**eng**

# Feature Papers in Eng

**MDPI**

# Feature Papers in Eng

# Feature Papers in Eng

Editor

**Antonio Gil Bravo**

**MDPI**

*Editor*
Antonio Gil Bravo
Science Department
Public University of Navarra
Pamplona
Spain

This is a reprint of articles from the Special Issue published online in the open access journal *Eng-Advances in Engineering* (ISSN 2673-4117) (available at: www.mdpi.com/journal/eng/special_issues/Feature_Papers_Eng).

For citation purposes, cite each article independently as indicated on the article page online and as indicated below:

LastName, A.A.; LastName, B.B.; LastName, C.C. Article Title. *Journal Name* **Year**, *Volume Number*, Page Range.

# Contents

*Editorial*

# Special Issue: Feature Papers in *Eng* 2021

Antonio Gil Bravo (ID)

INAMAT2-Science Department, Campus of Arrosadia, Public University of Navarra, Building Los Acebos, E-31006 Pamplona, Spain; andoni@unavarra.es

The aim of this first *Eng* Special Issue is to collect experimental and theoretical research relating to engineering science and technology. The general topics of *Eng* are as follows: Electrical, Electronic and Information Engineering; Chemical and Materials Engineering; Energy Engineering; Mechanical and Automotive Engineering; Industrial and Manufacturing Engineering; Civil and Structural Engineering; Aerospace Engineering; Biomedical Engineering; Geotechnical Engineering and Engineering Geology; Ocean and Environmental Engineering. Therefore, the following editorial is a selection of representative works of these topics.

This book contains up to 14 papers, including 3 paper reviews, published by several authors interested in new cutting-edge developments in the field of engineering. Solid oxide fuel cells (SOFC) are considered as one of the most promising fuel cell types for application as high-efficiency power generators by Tonekabonimoghaddam and Shamiri [1]. In this interesting work, the authors review the use of computational fluid dynamics (CFD) to maximize SOFC performance and life and minimize cost by considering numerous configurations and designs. A critical analysis of available literature proves that detailed research on the simulation of thermal stress and its damaging impact on the SOFC is still in its early stage of development. Numerical simulation is expected to help optimize the design, operating parameters and fuel cell materials. Therefore, sensitivity analysis of fuel cell parameters using simulation models is analyzed to address the issue. Finally, the present status of the SOFC optimization efforts is summarized by the authors so that unresolved problems can be identified and solved.

Aras and Diaconeasa [2] report that Probabilistic Risk Assessment (PRA) is one of the technologies that is used to inform the design, licensing, operation and maintenance activities of nuclear power plants (NPP). They indicate that PRA can be performed by considering the single hazard (e.g., earthquake, flood, high wind, landslide) or by considering multi-hazards (e.g., earthquake and tsunami, high wind and internal fire). Single-hazard PRA was thought sufficient to cover the analysis of a severe accident until the Fukushima Daiichi NPP accident in 2011. Since then, efforts were made to consider multi-hazards as well; thus, multi-hazard PRA are starting to be seen as being indispensable for NPP. However, a general framework for multi-hazard PRA is still missing. In this paper, the authors argue that the starting point for any multi-hazard PRA general framework should be the Advanced Non-LWR Licensing Basis Event Selection (LBE) Approach and Probabilistic Risk Assessment Standard for Non-LightWater Reactor (non-LWR) Nuclear Power Plants. For Probabilistic Risk Assessment (PRA), history has shown the path forward before, with the Three Mile Accident being seen as one milestone to understand the necessity of PRA. The Fukushima Daiichi NPP Accident is another milestone in the development of PRA, showing the need for performing multi-hazard PRA for the current and future NPP.

Metal phosphates represent an important group of materials with established industrial applications that are still attracting special scientific interest, owing to their outstanding physical and chemical properties. In this review, Amghouz et al. [3] provide an account on the synthetic routes and applications of zirconium and titanium phosphates, with a special focus on their application in the medicinal field. While zirconium phosphate has been extensively studied and explored with several reported industrial and medicinal applications,

especially for drug delivery applications, titanium phosphates have not yet attracted the deserved attention regarding their established applications. However, titanium phosphates have been the focus of several structural studies with their different polymorphic forms, varied chemical structures, and morphologies. These variations introduce titanium phosphates as a strong candidate for technological and, particularly, biomedical applications.

In the following work, El Sheshtawy et al. [4] develop a method to perform shape optimization of a tidal stream turbine hydrofoil using a multi-objective genetic algorithm. A Bezier curve parameterized the reference hydrofoil profile NACA 63815. Shape optimization of this hydrofoil maximized its lift-to-drag ratio and minimized its pressure coefficient, thereby increasing the turbines power output power and improving its cavitation characteristics. The Elitist Non-dominated Sorting Genetic Algorithm (NSGA-II) is employed by the authors to perform the shape optimization. A comparative study of two- and three-dimensional optimizations is carried out. The effect of varying the angle of attack on the quality of optimized results is also studied. Predictions based on the two-dimensional panel method results are also performed. Predictions based on a two-dimensional panel method and on a computational fluid dynamics code are compared to experimental measurements.

Non-carbonated fruit juices often tend to foam over during bottling. The resulting foam height corresponds to the equilibrium of foam formation and decay. Therefore, the foam unexpectedly occupies more space in the bottle and carries parts of the juice out of the bottle, resulting in product loss under filled containers and hygienic problems in the plant. Recent ultrasonic defoamers are effective but only capable outside the container and after the filling. In this article, a lateral ultrasonication through the bottle wall with frequencies between 42 and 168 kHz is used in-line by the authors [5] for non-invasive foam prevention during filling. The comparably high frequencies have a particular effect on the fresh foams, where a large fraction of small resonant bubbles is still present. Foam volume reductions of up to 50% are reached in these experiments. A low power of 15 W was sufficient for changing the rise of entrained bubbles and minimizing the foam development from the start. The half-life of the remaining foam could be reduced by up to 45% from the reference case. The authors indicate that the main observed effects were a changed rise of entrained bubbles and an increased drainage.

Aziz et al. [6] experimentally studied the surface roughening (Ra), martensitic phase transformation (MPT) and grain misorientation (GMO) behavior of stainless steel 304 and 316 in various grain sizes (Dg), including five cycles of sequential uniaxial tensile stress testing and Scanning Electron Microscope-Electron Back Scattered Diffraction (SEM-EBSD) investigation. The MPT and GMO characteristics were sequentially investigated by the authors using tensile testing and SEM-EBSD analysis. The correlation between MPT, GMO, martensitic volume fraction (Mf) and Ra behavior were also investigated. The experimental results showed that increasing the total strain from 5.0 to 25.0% increased the MPT, GMO, and Mf, which were transformed from the metastable austenitic phase in stainless steel (SUS) 304.

Sand control screens (SCD) have been widely installed in wells producing bitumen from unconsolidated formations. The screens are typically designed using general rules-of-thumb. The sand retention testing (SRT) technique has gained attention from the industry for the custom design and performance assessment of SCD. However, the success of SRT experimentation highly depends on the accuracy of the experimental design and variables. In this work, the authors examine the impact of the setup design, sample preparation, near-wellbore stress conditions, fluid flow rates, and brine chemistry on the testing results and, accordingly, screen design [7]. The SRT experiments were carried out using the replicated samples from the McMurray Formation at Long Lake Field. The results were compared with the test results on the original reservoir samples presented in the literature. Subsequently, a parametric study was performed by the authors changing one testing parameter at a test, gradually making the conditions more comparable to the actual wellbore conditions. The results indicated that the fluid flow rate is the most influential parameter on sand production, followed by the packing technique, stress magnitude and brine salinity level.

The next study involves the use of food-grade chemicals in the integrated pest management of dry-cured ham through the use of 100% polyester weft knitted mesh nets. Tubular mesh nets that are used to contain dry-ageing hams, commonly referred to as ham nets, were treated with a patent-pending food-grade chemical solution (40% Propylene Glycol + 1% Propylene Glycol Alginate + 1% Carrageenan) to control ham mites. Both treated and untreated ham nets were compared by the authors [8] for mechanical performance characteristics based on the following standards: abrasion resistance (ASTM D4966), elastic recovery (BS EN 14704-1:2005), breaking strength (ASTM D5034-09) and bursting strength (ASTM D3786). The results indicate that the chemical treatment had minimal to no impact on the mechanical performance characteristics of ham nets. The findings support the use of treated ham nets to increase the end-use functionality and provide ham producers an option for integrated pest management without compromising mechanical performance needs.

A compact artificial magnetic conductor (AMC) structure for the application of specific absorption rate (SAR) reduction is presented in this work by Karimian et al. [9]. A magneto-dielectric (MD) structure as a host of AMC substrate is used to miniaturize the AMC size. The magneto-dielectric has been designed with a low-profile spiral loop in a way to have a high permittivity and permeability for the desired center frequency of 3.5 GHz. Simulation results confirm the zero-degree reflection phase of the proposed AMC unit cell. Moreover, a 70% reduction has been achieved by the authors in comparison to the conventional AMC. To validate the simulation results, a prototype of the board is fabricated and measured with a coplanar waveguide (CPW) antenna for the reflection coefficient. The measurement results display an excellent agreement with the simulation ones.

Optical multi-band (OMB) systems exploit the low-loss spectrum of the single mode fiber (SMF) and are key enablers to increase the transportation capacity and node connectivity of already deployed systems. The realization of OMB systems is mainly based on the technological advances on the component and system level, and for this purpose, a broad gamut of various structural elements, such as transceivers, amplifiers, filters, etc. have been commercialized already or are close to commercialization. This wide range of options, which aid in unlocking the concurrent transmission in all amplification bands, is reviewed here by Uzunidis et al. [10] for the first time, whilst their pros and cons, as well as their limitations, are discussed. Furthermore, the needs for additional components in order to fully exploit the 390 nm low-loss wavelength range of SMF, which spans from 1260 to 1650 nm, are highlighted by the authors. Finally, based on a physical layer formalism, which incorporates the impact of the most important physical layer constraints for an OMB system, the attainable capacity and transparent reach of each amplification band are quantified.

Power electronics converters are traditionally designed regarding efficiency, power density, cost, and reliability figures of merit. Today, with the extreme spread of power electronic applications in our modern societies, together with the Earth's limits in terms of materials resources, it is important to consider the ecological impact of the converter not only during its usage but over its whole life cycle. In this article, de Freitas Lima et al. [11] introduce an eco-dimensioning methodology for analyzing and accounting for the energy consumption over the entire converter life. The analysis is applied on a small DC–DC converter considering the main components dual active bridge (DAB) converter. The planar transform is one of the key elements modeled in this article, including material and manufacturing conditions. The traditional and eco-dimensioning approaches are carried out and compared in order to emphasize the possible consequences on total energy cost.

Most commercially available hybrid electric vehicle (HEV) drivetrains are made of small internal combustion (IC) engines and large electric drives to improve fuel economy. They usually have higher cost than the conventional IC-engine-based vehicles because of the high costs of the electric drives. In this paper, Hu et al. [12] propose a hybridized powertrain composed of the original full-size engine of the vehicle and a universally optimum size parallel electric drive. The dynamic programming (DP) algorithm was used to obtain the sensitivity of the maximum miles per gallon (MPG) values versus

the power rating of the electric drive. This sensitivity was then analyzed by the authors to determine the optimal window of the electric drive power ratings. This was proven to be universal for all passenger cars of various masses and engine powers. The fuel economy and vehicle performance of this HEV was compared with those of the 2019 Toyota Corolla, a conventional IC-engine-based vehicle, and the 2019 Toyota Prius, a commercially available HEV. The results showed that the proposed universally optimized HEV powertrain achieved better fuel economy and vehicle performance than both the original ICE and HEV vehicles, at low additional vehicle cost.

When the concentration of a gas exceeds the equilibrium concentration in a liquid, the gas–liquid system is referred to as a supersaturated system. Supersaturation can be achieved by either changing the pressure and/or temperature of the system. The gas from a supersaturated liquid escapes either through bubble nucleation that usually occurs on solid surface and/or gas diffusion through the gas–liquid interface. The bubble nucleation requires a minimum threshold supersaturation. A waiting time is required to observe whether the applied supersaturation is sufficient to initiate bubble nucleation. When the supersaturation is not sufficient to cause bubble nucleation, some or all of the supersaturated gas may diffuse out from the liquid through the gas–liquid interface before further reducing the pressure in order to increase the supersaturation. In this work, using Fick's second law of diffusion and Henry's law, an analytical method is proposed by Pradhan and Bikkina [13] to estimate the level of supersaturations generated in three gas–liquid systems at several step-down pressures. Characteristic times of the gas–liquid systems were estimated by the authors to validate whether the waiting times used in this study are in accordance with the semi-infinite diffusion model used to estimate the supersaturations generated.

In the last manuscript, Zengerle et al. [14] investigate the damping phenomena acting on piezoelectrically driven MEMS oscillators. Three geometrical shapes of MEMS oscillators are presented, including cantilevers, bending oscillators and paddle oscillators. An analytical model for their resonance frequencies is derived. The bending modes of these micro-oscillator structures are characterized regarding their resonance frequency and their quality factor as a function of the ambient pressure in a nitrogen atmosphere as well as the dependence on the distance to a neighboring plate representing a geometrical boundary (e.g., to the package or to the mounting). The investigations cover a pressure range from $10^{-3}$ mbar up to 900 mbar and a gap width from 150 μm to 3500 μm. Consequently, a Knudsen number range over 6 orders of magnitude from 100 to $10^{-4}$ is covered. The measurement data are evaluated with a generalized damping model consisting of four parts representing the individual damping mechanisms (intrinsic, molecular, transitional and viscous). The evaluated parameters are analyzed as a function of the resonance frequency and the gap width. From the results found, the authors gain fundamental insights on the viscous and transitional damping mechanisms as well as on the intrinsic losses. In conclusion, a basic concept is provided to reduce the damping of micro-oscillator bending modes and thus increase the quality factor. Additionally, the results are supported by finite element simulations, revealing the temperature and pressure distribution within the gap.

**Conflicts of Interest:** The author declares no conflict of interest.

## References

1. Tonekabonimoghaddam, M.; Shamiri, A. Simulation and Sensitivity Analysis for Various Geometries and Optimization of Solid Oxide Fuel Cells: A Review, Greece. *Eng* **2021**, *2*, 386–415. [CrossRef]
2. Aras, E.M.; Diaconeasa, M.A. A Critical Look at the Need for Performing Multi-Hazard Probabilistic Risk Assessment for Nuclear Power Plants. *Eng* **2021**, *2*, 454–467. [CrossRef]
3. Amghouz, Z.; García, J.R.; Adawy, A. A Review on the Synthesis and Current and Prospective Applications of Zirconium and Titanium Phosphates. *Eng* **2022**, *3*, 161–174. [CrossRef]
4. El Sheshtawy, H.; El Moctar, O.; Natarajan, S. Multi-Point Shape Optimization of a Horizontal Axis Tidal Stream Turbine. *Eng* **2021**, *2*, 340–355. [CrossRef]

5. Thünnesen, J.; Gatternig, B.; Delgado, A. Ultrasonic Effects on Foam Formation of Fruit Juices during Bottling. *Eng* **2021**, *2*, 356–371. [CrossRef]

6. Aziz, A.; Yang, M.; Shimizu, T.; Furushima, T. Effect of Martensitic Transformation and Grain Misorientation on Surface Roughening Behavior of Stainless Steel Thin Foils. *Eng* **2021**, *2*, 372–385. [CrossRef]

7. Kotb, O.; Haftani, M.; Nouri, A. An Investigation into Current Sand Control Testing Practices for Steam Assisted Gravity Drainage Production Wells. *Eng* **2021**, *2*, 435–453. [CrossRef]

8. Al-Amin, M.; Freeman, C.; Schilling, W.; Black, C.; Campbell, Y.L.; Shao, W.; Kundu, S.; Varadajan, A. The Application of Food-Grade Chemical Treatment and Its Effect on the Mechanical Performance Characteristics of Ham Nets. *Eng* **2021**, *2*, 468–479. [CrossRef]

9. Karimian, R.; Ardakani, M.D.; Ahmadi, S.; Zaghloul, M. Human Body Specific Absorption Rate Reduction Employing a Compact Magneto-Dielectric AMC Structure for 5G Massive-MIMO Applications. *Eng* **2021**, *2*, 501–511. [CrossRef]

10. Uzunidis, D.; Apostolopoulou, F.; Pagiatakis, G.; Stavdas, A. Analysis of Available Components and Performance Estimation of Optical Multi-Band Systems. *Eng* **2021**, *2*, 531–543. [CrossRef]

11. de Freitas Lima, G.; Rahmani, B.; Rio, M.; Lembeye, Y.; Crebier, J.-C. Eco-Dimensioning Approach for Planar Transformer in a Dual Active Bridge (DAB) Application. *Eng* **2021**, *2*, 544–561. [CrossRef]

12. Hu, Z.; Mehrjardi, R.T.; Lai, L.; Ehsani, M. Optimal Hybridization of Conventional ICE Vehicles. *Eng* **2021**, *2*, 592–607. [CrossRef]

13. Pradhan, S.; Bikkina, P.K. An Analytical Method to Estimate Supersaturation in Gas–Liquid Systems as a Function of Pressure-Reduction Step and Waiting Time. *Eng* **2022**, *3*, 116–123. [CrossRef]

14. Zengerle, T.; Ababneh, A.; Seidel, H. Generalized Damping Model for MEMS Oscillators from Molecular to Viscous Flow Regime. *Eng* **2022**, *3*, 124–141. [CrossRef]

**Short Biography of Author**

**Antonio Gil Bravo** (Full Professor of Chemical Engineering, Universidad Pública de Navarra, Spain): Professor Gil earned his BS and MS in Chemistry at University of Basque Country (San Sebastián), and his PhD in Chemical Engineering at University of Basque Country (San Sebastián). He did postdoctoral research at the Université catholique de Louvain (Belgium) working on Spillover and Mobility of Species on Catalyst Surfaces. The research interests of Professor Gil can be summarized as: Evaluation of the porous and surface properties of solids. Pillared clays. Gas adsorption. Energy and $CO_2$ storage. Pollutants adsorption. Environmental technologies. Environmental management. Preparation, characterization and catalytic performance of metal supported nanocatalysts. Industrial waste valorization.

# A Review on the Synthesis and Current and Prospective Applications of Zirconium and Titanium Phosphates

**Zakariae Amghouz** [1] , **José R. García** [2] **and Alaa Adawy** [3],*

1    Department of Material Science and Metallurgical Engineering, University of Oviedo, 33203 Oviedo, Spain;
     amghouzzakariae@uniovi.es
2    Department of Organic and Inorganic Chemistry, University of Oviedo—CINN, 33006 Oviedo, Spain;
     jrgm@uniovi.es
3    Unit of Electron Microscopy and Nanotechnology, Institute for Scientific and Technological Resources (SCTs),
     University of Oviedo, 33006 Oviedo, Spain
*    Correspondence: a.adawy@outlook.com or uo263127@uniovi.es

**Abstract:** Metal phosphates represent an important group of materials with established industrial applications that are still attracting special scientific interest, owing to their outstanding physical and chemical properties. In this review, we account on the different synthetic routes and applications of zirconium and titanium phosphates, with a special focus on their application in the medicinal field. While zirconium phosphate has been extensively studied and explored with several reported industrial and medicinal applications, especially for drug delivery applications, titanium phosphates have not yet attracted the deserved attention regarding their established applications. However, titanium phosphates have been the focus of several structural studies with their different polymorphic forms, varied chemical structures, and morphologies. These variations introduce titanium phosphates as a strong candidate for technological and, particularly, biomedical applications.

## 1. Introduction

Two-dimensional compounds represent an extraordinary group of materials that could provide fascinating scientific applications. The term *laminar composite or lamellar solid* is often associated with the concept of *intercalation*, owing to the ability of this type of layered structures to accommodate guest molecules in between their layers, creating supramolecular materials. These intercalated guest molecules can be released accordingly and, therefore, these stratified solids have been conventionally considered as repositories (containers) for different functioning chemicals. Accordingly, intercalated laminar composites have characteristics of a dual nature and dual functions from both the host layered structures and their intercalated guest compounds. Herein, we focus on two of these laminar composites: zirconium phosphates and titanium phosphates, extending our review to their other variants and their established and prospective applications.

## 2. Zirconium Phosphates

Since the 1950s, many studies have focused on lamellar solids, particularly their properties and prospective applications. Special attention has been given to the development of innovative synthetic approaches for their preparation, and subsequent implementation in applications for different technological fields [1]. In this context, a group of laminar tetravalent metal phosphates is worth mentioning, with zirconium phosphates as remarkable structures that have attracted intensive investigation interests [2–4]. The great appeal

of zirconium phosphates (hereafter abbreviated as ZrP) is mainly attributed to their robustness, the versatility of their crystalline structure, and the tunability of their crystallinity degree. In addition, the textural properties of ZrP particles can be controlled by modifying the synthetic conditions, or chemistry of their layers, generally by introducing functional groups on their surface. In fact, the scientific interest in ZrP goes back to the beginning of the 20th century, when the first trials of using soluble zirconium salts, e.g., $ZrOCl_2 \cdot 8H_2O$, in the elimination of phosphate ions in solution were reported [5–7]. Since the 1920s, the thermogravimetric determination of zirconium, as ZrP, has been one of the most widely used quantitative analysis methods [8].

Although the cation exchange properties of ZrP began to be studied in the 1950s, the crystalline zirconium bis(monohydrogen phosphate) monohydrate, $\alpha$-$Zr(HPO_4)_2 \cdot H_2O$ ($\alpha$-ZrP), was not synthesized until 1964 [9]. In 1968, the description of its laminar structure was reported [10], for which more precise data appeared in 1969 [11]. The structural refinement of $\alpha$-ZrP was published in 1977 in two independent articles [12,13], making it possible to explain the behavior of the material as cation exchangers after determining the position of the hydrogen atoms in the crystalline phase. The $\alpha$-ZrP consists of a single plane of metal atoms linked through monohydrogen phosphate groups, located alternatively above and below each plane. Each zirconium atom is coordinated with six oxygen atoms of six different phosphates, and each phosphate tetrahedron shares three of its four oxygen atoms with three zirconium atoms, while the fourth is linked to a hydrogen atom, constituting the most active site of the material. The packing of the layers generates six-sided cavities that contain the water molecules, which interact through hydrogen bonds with the P–OH groups in the same plane (Figure 1).
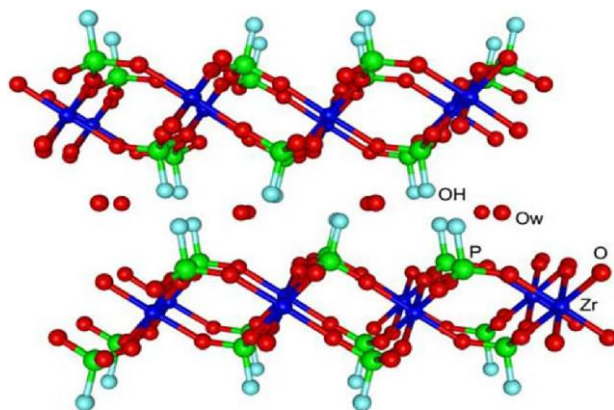


**Figure 1.** The crystalline structure of $\alpha$-ZrP. Reprinted with permission from reference [14].

*2.1. Reported Methods for $\alpha$-ZrP Synthesis*

Three main routes for the synthesis of crystalline $\alpha$-ZrP have been described: (i) the reflux methods; (ii) the metal complexing reagent-utilizing methods; and (iii) the hydrothermal methods.

The reflux method was first published by Clearfield and Stynes in 1964 [9]. In this method, an amorphous solid is obtained from the direct reaction between phosphoric acid ($H_3PO_4$) and zirconyl chloride, ($[Zr_4(OH)_8(H_2O)_{16}]Cl_8(H_2O)_{12}$), that is subsequently washed with distilled water to eliminate all chloride ions, before being refluxed in concentrated phosphoric acid until obtaining crystalline $\alpha$-ZrP. The degree of crystallinity varies directly with the refluxing duration and the concentration of the used phosphoric acid solutions. Worth noting that almost four decades later, it was demonstrated in our laboratories that the formation of amorphous zirconium phosphate is actually a consequence of washing the precipitate with distilled water [15]. This led to the conclusion that crystalline $\alpha$-ZrP can be directly obtained after the reaction between acid phosphoric and zirconyl chloride at room temperature, although the obtained $\alpha$-ZrP microcrystals easily hydrolyze in water [15].

On the other hand, the use of complexing agents for the synthesis of highly crystalline α-ZrP was first proposed in 1968 by Alberti and Torracca [16]. This approach, dubbed the HF method, includes the complexation of the Zr(IV) cation with fluoride anions, to form with $ZrF_6^{2-}$ species, and its slow subsequent decomposition in the presence of phosphoric acid. The degree of α-ZrP crystallinity is controlled through the rate of HF elimination. García-Rosales et al. revisited this approach to study the precipitation of α-ZrP from solutions of fluoro complexes of zirconium (IV) using Mexican zirconium sand as raw material [17]. The oxalate anion ($C_2O_4^{2-}$) was also reported to be another complexing agent for zirconium (IV). Horsley et al. [18] and Capitani et al. [19] explored the synthesis of α-ZrP using oxalic acid ($HO_2C-CO_2H$). Their studies showed that the obtained material had similar structural characteristics and properties to those prepared in the presence of hydrofluoric acid (HF), with the additional advantage of providing more efficient reaction yields in shorter times.

As a third alternative, Shuai et al. proposed the use of a hydrothermal method based on the slow addition (under stirring) of a zirconyl chloride ($[Zr_4(OH)_8(H_2O)_{16}]Cl_8(H_2O)_{12}$) solution to phosphoric acid ($H_3PO_4$) solutions, followed by subsequent heating of the reaction mixture to 200 °C [20]. They observed that the length and thickness of the crystalline α-ZrP [$Zr(HPO_4)_2·H_2O$] plates are directly proportional to the concentration of phosphoric acid, whereas their length/thickness varies inversely. They also studied the growth rate of α-ZrP crystals as a function of reaction time and phosphoric acid concentration. They found that the process starts with a rapid homogeneous nucleation in the presence of an amorphous phase, in a regime similar to liquid–liquid phase separation that was first reported by Vekilov et al. as the very first step of nucleation [21]. This is followed by slow progress, especially at low concentrations of phosphoric acid. This slowness in the metastable stage allows for autoselection of better nuclei and the formation of secondary clusters, from which crystals grow with an observable homogeneity and regularity in size and morphology of the crystals, after dissolving poorly ordered crystals and their reprecipitation (Figure 2) [22]. This nucleation and crystallization mechanism resembles what was known later in the literature as the two-step nucleation mechanism that was mainly reported for protein crystallization, and very recently, for metal–organic frameworks [21,23,24].
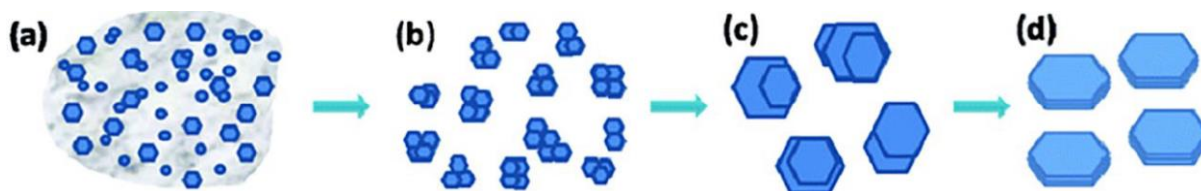


**Figure 2.** Representation of α-ZrP crystal formation. (**a**) Precipitation of nanocrystals and amorphous particles from solution. (**b**) Lattice orientation of nanocrystals with formation of secondary clusters. (**c**) Growth of microcrystals from secondary clusters. (**d**) Formation of homogeneous and regular crystals by dissolving poorly ordered areas and subsequent re-precipitation. Reprinted with permission from reference [20], License number: 1173274-1.

Later, Cheng et al. described the synthesis of α-ZrP, but rather in the form of rods, in addition to the usual morphology of plates, as a function of the presence or absence of small amounts of fluoride ions [25]. They showed that in the absence of fluoride ions, mixing zirconyl chloride and concentrated phosphoric acid solutions, that is followed by heating the mixture at 100–120 °C, resulted in the formation of plate-shaped α-ZrP, with a notable increase in crystallinity with increasing temperature. On the other hand, the presence of small amounts of fluoride ions led to the formation of rod-shaped α-ZrP together with the plate-shaped α-ZrP crystals. They attributed the growth of crystals along the crystallographic *a*-axis to be due to the selective adsorption of fluoride ions on the surface of the plates.

Other unconventional strategies for the synthesis of α-ZrP were reported to have resulted in the formation of microcrystals with different morphologies and textural properties. As an example, Benhamza et al. used a sol–gel synthesis in which a concentrated

phosphoric acid solution was added to zirconium (IV) propoxide [$Zr(OCH_2CH_2CH_3)_4$] in 1-propanol ($CH_3CH_2CH_2OH$) [26]. The aging in the mother liquor increased the degree of crystallinity of the resultant α-ZrP, which varied in structure from amorphous to semi-crystalline materials. Another approach was reported by Hajipour et al. who prepared hexagonal shaped α-ZrP nanoparticles using polyvinyl alcohol [-$CH_2CHOH$-]$_n$ (PVA) or polyvinylpyrrolidone [($C_6H_9NO)_n$] (PVP) in an aqueous solution [27]. PVA and PVP acted as complexing organic matrices for Zr(IV) that facilitated the dispersion of α-ZrP particles and controlled their sizes. Afterwards, their (PVA/PVP) elimination through the calcination led to the growth of hexagonally-shaped plates with an average length of 60 nm.

On the other hand, mesoporous zirconium phosphate that bears relatively high specific surface areas and proton conduction capacity [28,29], are usually synthesized by calcination of the reaction product of soluble phosphate and/or phosphonate salts with zirconium oxychloride ($ZrOCl_2$) or zirconium alkoxide [$Zr(OR)_4$] [30–33]. Moreover, Zhao et al. used the reaction between phosphoric acid and zirconium propoxide, in the presence of non-ionic surfactants, to obtain gels that transformed to ZrP with a vermicular morphology after being thermally activated [34]. Furthermore, Tarafdar et al. synthesized spherical zirconium phosphate particles (micro- and meso-porous) with zirconium carbonate [$Zr(CO_3)_2$] as the source of the metal [35]. In addition to the above, Zhu et al. obtained α-ZrP with unusual textural properties, using an approach consisting of a sol–gel process followed by supercritical drying [36]. For this process, two polymers: polyethylene glycol [$H-(O-CH_2-CH_2)_n-OH$] (PEO) and polyacrylamide [(-$CH_2CHCONH_2$-)](PAM), were used that led to the formation of solids with macropores of controllable size (0.5–5 μm). These solids can be modified through altering the relative proportion of both polymers. These materials in monolithic form maintain high porosity, with specific surface areas close to 600 $m^2 \cdot g^{-1}$, and good mechanical properties in uniaxial compression.

### 2.2. Functionally-Adapted α-ZrP Synthesis Routes

The increasing interest in nanotechnology demands the development of new synthetic approaches for laminar materials as α-ZrP, in order to prepare nanoparticles of nanometric dimensions, both in their planes (length/width) as well as in their thickness, until reaching colloidal dispersions of quasi-individual sheets.

In that respective, Bellezza et al. proposed the synthesis of α-ZrP nanoparticles from water-in-oil microemulsions containing zirconyl chloride and phosphoric acid [37]. In addition, Alberti et al. synthesized α-ZrP of low crystallinity in a one-step reaction, where a solution of an organic zirconyl salt such as zirconyl propionate [$Zr(CH_3CH_2COO)_4$], in a polar aprotic solvent (e.g., *N,N*-dimethylformamide [$(CH_3)_2NC(O)H$], DMF) is mixed at room temperature with a solution of phosphoric acid in the same solvent, inducing the formation of a transparent dispersion from which the solvent is removed by heating to dryness [38]. Although this procedure is chemically effective, the use of organic solvents deprives it from being an ecologically friendly alternative and, thus, biocompatible for medical applications [39]. In order to develop synthetic procedures with less environmental drawbacks by avoiding the usage of toxic solvents, ethanol ($C_2H_5OH$) was used to dissolve zirconyl propionate [40]. This led to the formation of gel-like materials containing crystalline domains of α-ZrP sheets intercalated with ethanol. Heat treatment of these gels generates irregularly shaped α-ZrP nanoplates. As expected, the presence of low-coordination phosphate groups ($H_2PO_4^-$ and $H_3PO_4$) was detected in α-ZrP nanoparticles in a much higher proportion than in massive materials, which favors their reactivity compared to microcrystalline materials.

### 2.3. Technological Applications of α-ZrP and Their Introduction in Medicinal Chemistry

In addition to its well-known applications as ion exchangers that can assist in purification [41,42], and proton conductors [43,44], or catalysts and catalyst supports [14], α-ZrP and its derivatives have been explored to obtain different functionalities, such as anticorrosion coating and reinforced mechanical properties.

Owing to its barrier property, exfoliated α-ZrP, after being functionalized with polypyr-role, could be used for the fabrication of waterborne epoxy coatings (PPy-ZrP/WEC), that bear outstanding anticorrosion resistance [45]. In another context, despite the difficulty of introducing inorganic nanoparticles homogenously in a polymeric matrix, different amounts of bromine initiator could be introduced on the surface of α-ZrP nanoplatelets using a dual-epoxide-modified method [46] that facilitated their usage to graft polymethyl-methacrylate (PMMA) polymer. This resulted in enforced mechanical properties for the resultant matrix (ZrP-g-PMMA) that could present a new alternative to this PMMA polymer, widely used for biomedical applications [47].

α-ZrP and its derivatives have also been extensively investigated in biomedical appli-cations. The implementation of nanotechnology for drug delivery processes has become a very promising area of research, with a primary objective to realize the necessity to completely eliminate (or even inhibit) the inherent side effects for the compulsory administration of medi-cations. In this context, lamellar zirconium phosphate can serve as a good choice for being a carrier material in drug delivery systems owing to its excellent biocompatibility [48–51]. This has allowed for its commercial use in peritoneal dialysis as a sequestrant of ammonium cations [52].

The intercalation capacity of α-ZrP facilitates the incorporation of functional biomolecules into its structure, avoiding their unnecessary interaction with the biological environment, preventing their protein denaturation, and prolonging their lifetime and thus, their biofunc-tionality. The acidic character of the interlaminar hydrogen phosphate groups favors the reversible intercalation of basic species and, therefore, the ability to sequester and release molecules of biological interest. In general, the expulsion of the molecules contained in the lamellar inorganic matrices is achieved by alterations in the chemical environment, such as a concentration gradient or changes in pH, and other biological stimuli [48–51].

Ding et al. were the first to intercalate proteins in microcrystalline α-ZrP [53]. Af-terwards, Kumar and McLendon described the intercalation of heme proteins with redox activity [54]. Thereafter, Kim et al. obtained multilayer thin films comprising previously exfoliated α-ZrP sheets and polycations, such as cytochrome c (a monomeric protein that acts as a mitochondrial electronic transporter) [55]. In addition, α-ZrP was also used to immobilize myoglobin, lysozyme, hemoglobin, chymotrypsin, and glucose oxidase [56]. Furthermore, Bellezza et al. immobilized lipase on pristine and functionalized α-ZrP, ob-serving that the amount of lipase retained increases as the hydrophobicity of the surface increases [57]. Moreover, it was observed that the presence of DNA or urea stabilizes the activity of enzymes and proteins in α-ZrP based nanomaterials. For example, the activity of hemoglobin inserted in α-ZrP is greater when it is co-intercalated with DNA [58], while moderate concentrations of urea stabilize heme proteins (methemoglobin and metmyo-globin) retained in α-ZrP, thus increasing its activity as well as its average lifetime by alterations in the chemical environment [59].

A quite different biologically relevant application is the one that has been reported by Xu et al., who selectively captured phosphopeptides in complex mixtures using α-ZrP nanoplates, in order to analyze tryptic digestions of mouse liver and leukemia cell phosphoproteomics, identifying 158 phosphopeptides (209 phosphorylation sites) of 101 phosphoproteins in mouse liver lysate, and 78 phosphopeptides (104 phosphorylation sites) of 59 phosphoproteins in leukemia cell extract [60]. On the other hand, Díaz et al. studied the controlled release of nanoencapsulated insulin in α-ZrP, using pH variations as a stimulus [48]. Afterwards, they reported the use of zirconium phosphate nanoplates for the encapsulation of cisplatin (a well-known drug in chemotherapy treatments against various types of cancer) and its administration to tumor cells [51]. Cisplatin, which was inserted using ion exchange, was tested in vitro for cytotoxicity against human breast cancer (MCF-7) with promising results. The intercalated drug was released only at low pHs, typical for tumor cells, which should minimize the side effects of cisplatin on noncancerous tissue. In addition, these researchers used α-ZrP to intercalate doxorubicin (DOX), a cytostatic from the anthracycline family, widely used in cancer chemotherapy, demonstrating its

effectiveness against MCF-7 cells compared to free DOX, in a process that leads to the uptake of nanoparticles by endocytosis, reaching and penetrating cancer cells to a greater extent than healthy cells [50]. In this mechanism, the hydrogen-mediated phospholipid cell membrane phosphates participate with P-OH groups on the surface of $\alpha$-ZrP. Studies by others on DOX@$\alpha$-ZrP demonstrated that the zirconium phosphate nanoparticles are highly hemocompatible and do not show hemolytic activity towards human red blood cells [49]. González et al. conducted in vitro studies showing that $\alpha$-ZrP nanoplates deliver DOX to cancer cells, while DOX retains its original anticancer activity [61]. Furthermore, while DOX induces the generation of oxygen radicals in both cancer cells and healthy mammary epithelial cells, DOX@$\alpha$-ZrP causes greater oxidative stress in MCF-7 breast cancer cells than in MCF-10A breast epithelial cells, thus showing selective toxicity towards cancer cells. Subsequent studies have reported the success of modifying the surface of the zirconium phosphate in DOX@$\alpha$-ZrP, improving the ability to administer DOX, and increasing its biocompatibility [62].

Another variety of $\alpha$-ZrP (3-aminopropyl)triethoxysilane [$H_2N(CH_2)_3Si(OCH_2CH_3)_3$] (APTES) functionalized zirconium phosphate nanoparticles has also been used for drug delivery. Li et al. used APTES@$\alpha$-ZrP to anchor hyaluronic acid (HA) to combine the carrying capacity of the drug in zirconium phosphate and the tumor targeting capacity of HA in the administration of paclitaxel (PTX), a drug used in the treatment for cancers of the lung, ovary, breast, and advanced forms of Kaposi's sarcoma [63]. In vivo experiments in A549 lung cancer showed good antineoplastic ability and minimization of the toxic side effects of PTX.

Donnadio et al. prepared materials with antimicrobial and antibiofilm activity using nanocrystalline $\alpha$-ZrP as a carrier for chlorhexidine (CLX), which was used to prepare carboxymethylcellulose (CMC)-based composite polymeric films as wound dressings with reduced cytotoxicity, in order to release the antiseptics in a prolonged way, showing good antibacterial activity against Gram-positive and Gram-negative bacteria [64]. CMC@$\alpha$-ZrP films are uncytotoxic to human dermal fibroblast and keratinocyte cell lines. When CLX was incorporated into CMC@$\alpha$-ZrP, its cytotoxicity was significantly reduced when compared to free CLX, probably because of the controlled release of CLX that kept its concentration at low levels.

Very recently, Adawy et al. implemented the hydrothermal methodology to synthesize monohydrated nanolaminar $\alpha$-ZrP and subsequently enriched their nanolayers with silver nanoparticles (AgNPs) of uniform dimensions and distribution [65]. The structural stability of $\alpha$-ZrP was confirmed before and after enrichment with AgNPs as well as the assessment of their antimicrobial activity and the cytocompatibility. They reported that the pristine $\alpha$-ZrP did not show any cytotoxic effects against Sarcoma osteogenic cells, with a small loss of its biocompatibility when enriched with silver, yet with outstanding antimicrobial effectiveness against *Escherichia coli*.

### 3. Titanium Phosphates

Although zirconium phosphates have so far been the most studied lamellar solids of all tetravalent metal phosphates since 1950s, titanium laminar phosphates have also attracted particular attention around a decade later (1960s). Alberti et al. [16,66], following the wake of the seminal investigations on zirconium phosphates by Clearfield et al. [9], were pioneers in the study of crystalline titanium phosphates. In addition, the contributions of the Soviet school are also noteworthy in this field, with special mention to Chernorukov et al. [67]. In this scientific environment, our working group at the University of Oviedo (Research Group of Synthesis, Structure and Technological Application of Materials [SYSTAM], https://systam.grupos.uniovi.es/) published its first articles on titanium phosphates in the early 1980s [68–70]. The established synthetic routes for the layered titanium phosphate [$\alpha$-Ti (HPO$_4$)$_2 \cdot$H$_2$O] are very similar to those of layered zirconium phosphate [$\alpha$-Zr(HPO$_4$)$_2 \cdot$H$_2$O]. Only recently new methodologies have been introduced, mostly based on the hydrothermal synthesis, but rather using other compounds as precursors, such as protonated layered lepidocrocite-type K$_{0.7}$(Ti)$_2$O$_4$ [71]. Unlike zirconium

phosphates, titanium phosphates can be also synthesized in another morphology, exhibiting the form of nanorods, and thus are known as nanofibrous titanium phosphates. The most popular synthetic route of nanofibrous titanium phosphates is also the hydrothermal method at temperatures 160–250 °C, using a titanium precursor such as titanium tetrachloride ($TiCl_4$) in aqueous phosphoric acid solutions [72]. Moreover, several studies have reported on the possibility for growing a nanofibrous titanium phosphate thin film on titanium substrates after hydrothermally treating the latter in phosphoric acid solutions at 180 °C [73] or 250 °C [74], but also at 120 °C in the presence of hydrogen peroxide as a catalyst [75]. Herein, we provide further details on the different titanium phosphate phases, categorizing them into nanolayered and nanofibrous phases.

### 3.1. Nanolayered Titanium Phosphates

So far, there are three nanolayered titanium phosphates reported in the literature. The crystalline titanium bis(monohydrogen phosphate) monohydrate, $\alpha$-Ti $(HPO_4)_2 \cdot H_2O$ ($\alpha$-TiP), is an iso-structure of its zirconium counterpart ($\alpha$-ZrP) [76]. $\gamma$-titanium phosphate ($\gamma$-TiP), formulated as $\gamma$-Ti$(HPO_4)_2 \cdot 2H_2O$ and later as $\gamma$-Ti$(PO_4)(H_2PO_4) \cdot 2H_2O$, is another interesting dihydrate layered material that is considered as an analogue to $\alpha$-TiP, with larger interlayer spacing than $\alpha$-TiP since it comprises two interlayer water molecules. $\gamma$ titanium phosphate ($\gamma$-TiP) is also an iso-structure of $\gamma$ zirconium phosphate, so do its dehydrated form known as $\beta$-TiP [$\beta$-Ti$(PO_4)(H_2PO_4)$] and $\beta$-ZrP [$\beta$-Zr$(PO_4)(H_2PO_4)$] [77–79]. As expected, $\alpha$-TiP and $\gamma$-TiP have many profound applications, such as ion exchange [80], catalysis [30], ionic conductivity [44,81], and electrodes for Li- and Na ion storage [82–84], in addition to their capability of being functionalized with organic species and also being exfoliated [85–87]. Nevertheless, there is almost no record for their biologically relevant applications.

### 3.2. Nanofibrous Titanium Phosphates

Several titanium phosphate compounds have been reported with different morphologies and thus different plausible applications [88–91]. With a nanofibrous morphology, two polymorphs with the chemical formula: $Ti_2O(PO_4)_2 \cdot 2H_2O$, called $\pi$-TiP and $\rho$-TiP, were synthesized around two decades ago [72]. Although $\pi$-TiP and $\rho$-TiP are compounds with neutral lattices, they exhibit an outstanding ability to incorporate intracrystalline metal cations through ion exchange processes [92,93]. Particularly, $\pi$-TiP reaction with alkali metal nitrates in the molten state results in the formation of fibrous phases charged with the reactant alkali metal [92]. In terms of their crystallographic structures, the structure of $\rho$-TiP was reported shortly after being solved ab initio from synchrotron X-ray and neutron powder diffraction data (Figure 3a) [94]. On the other hand, there were only some scientific hypotheses reported about the most probable structural ordering of $\pi$-TiP [92]. The crystal structure of the $\rho$-TiP phase has two channels that run parallel to the crystallographic *a*-axis in which the water molecules are housed. Recently, we succeeded to reveal the crystal structure of $\pi$-TiP [95]. The $\pi$-TiP phase, obtained under mild hydrothermal conditions, crystallizes in the monoclinic system (space group $P2_1/c$) in contrast to the previously described $\rho$-TiP (triclinic, *P*-1). The $\pi$-TiP structure is made up of $TiO_6$ and $TiO_4(H_2O)_2$ octahedra, connected with orthophosphate groups, defining the anisotropic three-dimensional packing, consistent with the existence of preferential crystal growth directions. The properties of both polymorphs have scarcely been explored, probably because they were just synthesized less than 3 decades ago. Although both $\pi$- and $\rho$-TiP exhibit low accessible porosity when hydrated (11–16 $m^2 \cdot g^{-1}$) [96], Amghouz et al. showed that this can be changed significantly if they are treated thermally. In such a process, the crystallized water molecules are eliminated, and the resultant dehydrated phases exhibit an unusual adsorption capacity for nitrogen a little above ambient temperature [97]. For the $\rho$-TiP, the thermal treatment at temperatures of 200–300 °C results in the generation of tetrahedral (ex-octahedral) titanium atoms (Figure 3b) that are responsible for the enhanced nitrogen adsorption capability [97]. On the other hand, this treatment for the $\pi$-TiP was shown to lead to a change in the coordination environment of one of the titanium atoms, which

goes from octahedral to tetrahedral, with the formation of the anhydrous Ti$_2$O(PO$_4$)$_2$ [97]. Beside their nitrogen adsorption properties, π-TiP has been recently investigated for its proton conductivity. π-TiP was doped in a chitosan matrix to study the proton conductivity of the resultant chitosan-based composite membranes (CS@π-TiP) [95]. This study showed that doping the CS matrix with π-TiP (5 *w/w*%) results in a 1.8-fold rise in the proton conductivity compared to the bare membrane [95]. Another explored property is their attraction to water. Yada et al. prepared super-hydrophilic thin films made up of micro- and nano-crystals of π-TiP with controlled morphologies that transform to superhydrophobic layers when π-TiP is modified with alkylamines [75]. Similarly, Cai et al. showed that TiP thin films grown on titanium substrates can possess a hydrophilicity that can be switched to a super-hydrophobicity through increasing the preparation temperature [98]. Recently, π-TiP has been reported to possess ultrastability as a Ca$^{2+}$ storage material with a minimal dimensional change and almost no transformation in the crystallographic structure upon the insertion or extraction of Ca$^{2+}$ [99]. This exceptional stability that stands for over 1700 cycles of the Ca$^{2+}$ insertion and extraction presents π-TiP as a viable electrode material for calcium-ion battery applications [99]. Therefore, we can estimate that many attempts could be made in the near future to utilize these nanofibrous phases for some of the other applications that have been established for the nanolayered titanium phosphates.
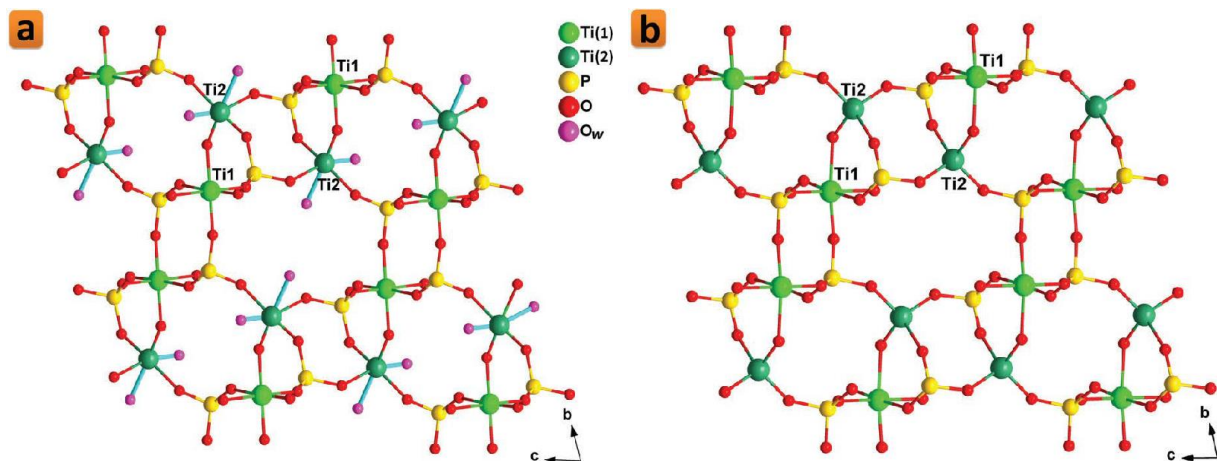


**Figure 3.** Crystalline structure of the dihydrate (**a**) and anhydrous (**b**) phases of ρ-TiP. Reprinted with permission from reference [86], License number: 1173277-1.

*3.3. Prospective Applications of TiP in the Biomaterials Research Field*

Since α-TiP is an iso-structure of α-ZrP, and has been reported to share many of its physical and chemical properties beside its non-spherical layered morphology, it should have potential applications for being used in biological applications [100]. Similar to α-ZrP, one of the most important aspects to consider for this inorganic laminar nanomaterial in biological applications lies in its ability to provide a controlled release of drugs or charged nanoparticles, while its non-spherical morphology favors its adherence and margination to biological tissue [101,102]. Therefore, just like α-ZrP, α-TiP could be an interesting alternative for drug delivery and bone cement applications, especially if it is enriched with biofunctional nanoparticles that could improve its antimicrobial properties [103,104]. In this context, Adawy et al. recently used the hydrothermal methodology to synthesize monohydrated nanolayered α-TiP, whose surfaces were, afterwards, enriched with AgNPs of uniform dimensions and distribution [65]. After the structural evaluation of the resultant Ag-enriched α-TiP phase, assessments of the cyto-viability and antimicrobial efficacy were performed. Surprisingly, it was reported that α-TiP bears an outstanding cytocompatibility that even surpasses that of α-ZrP, particularly after being enriched with AgNPs [65]. More interestingly, α-TiP appeared to possess higher intercalation capacity for silver ions compared to that of α-ZrP, while, as predicted, AgNPs-enriched α-TiP possessed a perfect

antimicrobial activity [65]. Therefore, these recent results highlight the potential capabilities of α-TiP, especially under stressful biological conditions, in which the decrease in pH values can stimulate more release of the functional silver ions that are richly intercalated in the interlayers of α-TiP.

On the other hand, nanostructured inorganic compounds, such as nanofibrous ρ-TiP and π-TiP, have generated increasing interest in the research directed towards the synthesis of biomaterials. These compounds are usually prepared in advance and used as bone cement or applied as a surface layer covering the implant [105–109], although it is also possible to induce its growth on the substrate surface [110–112]. In the latter case, the thin layer that constitutes the coating acts as a messenger through which interaction with the surrounding biological system occurs [113–115]. The outer layer of the device should be compromised of compounds naturally present in the biological environment, commonly calcium phosphate phases [116–119]. However, nanostructured systems are also used for this application using the metal of the (metallic) substrate as the source of the inorganic matter [120,121]. A good example of these metals is titanium, with high osseointegration capacity and ease of physical attachment to bone, owing to the spontaneously formed oxide layer on its surface, which does not cause denaturation of proteins in the proximity of the implant [122]. Thus, the formation of coating layers of nanostructured compounds containing titanium on the metallic surface of titanium has been described with positive results [123,124], including its use, together with its widely used alloy, titanium–aluminum–vanadium alloy (Ti–6Al–4V), in implant applications [125,126]. Modifications of the surfaces of titanium and its alloys by coating them with a layer of nanostructured materials could promote their hydrophilicity, surface roughness, biocompatibility, and bioactivity, by improving both their functionality and their osseointegration and, therefore, their long-term fixation of the implant. In this respect, the two polymorphs π- and ρ-TiP are nanostructured compounds that obviously contain titanium, the former of which could be grown successfully on titanium surfaces [73–75].

Recently, Adawy et al. synthesized nanorods of the polymorphs π-TiP and ρ-TiP and enriched them with AgNPs that provided outstanding antimicrobial activity without much of a compromise to their cytocompatibility [127]. Additionally, they hydrothermally treated the commercially available Ti–6Al–4V alloy in aqueous phosphoric acid solution in attempts that led to the growth of a nanofibrous coating layer of π-TiP on the surface of Ti–6Al–4V alloy discs, where the size as well as the surface roughness of this nanofibrous π-TiP coating layer can be controlled [127]. Although intercalation is not an option for the nanofibrous morphology of π-TiP, Adawy et al. reported the enrichment of the coated surfaces with AgNPs and their subsequent doping with strontium ions, which in turn played an outstanding role in controlling the release of the antimicrobial AgNPs, in a mechanism that improved the cytocompatibility and long-term biofunctionality of the resultant coated biomedical alloy as they described in their long-term ion release assessments [127].

## 4. Conclusions

More than half a century of research on metal phosphates has provided multiple synthetic approaches for α-ZrP, some of which were later adopted for the synthesis of α-TiP. The different available synthetic routes allowed for controlling the crystallinity, particle size, textural properties, functionalization and, consequently, the reactivity of these materials. In addition to the traditional methods, based on the use of zirconium (IV) complexing agents, such as fluoride and oxalate, new routes were developed, mainly aimed at obtaining particles with varied morphologies, such as nanoplates, nanofibers, and mesoporous compounds. The research also offered several methods for the effective functionalization of the α-ZrP layers, and thus obtaining tailored materials for specific applications, many of them targeted, among many profound technological applications, in the field of biomedicine research. Although α-ZrP has attracted substantial research attention for biomedical application, mainly in drug delivery applications, α-TiP, as an iso-structure of α-ZrP, has been scarcely investigated in this regard until recently. α-TiP, which shares with its zirconium counterpart many of

the technological applications, could also be the basis for many prospective medicinal applications, particularly after the recent reports on its better biocompatibility compared to $\alpha$-ZrP. On the other hand, nanofibrous TiP, which has attracted research interests in the latest decade, could also be useful for many prospective technological applications. In addition, it can be considered as an easily functionalized nanostructured material for use in biocement applications, as well as coating for biomedical alloys, thus enhancing their bio-functionality and, ultimately, their long-term success.

## References

1. Pica, M.; Donnadio, A.; Casciola, M. From microcrystalline to nanosized $\alpha$ zirconium phosphate: Synthetic approaches and applications of an old material with a bright future. *Coord. Chem. Rev.* **2018**, *374*, 218–235. [CrossRef]
2. Kraus, K.A.; Phillips, H.O. Adsorption on inorganic materials. I. Cation exchange properties of zirconium phosphate. *J. Am. Chem. Soc.* **1956**, *78*, 694. [CrossRef]
3. Amphlett, C.B.; McDonald, L.A.; Redman, M.J. Synthetic inorganic ion-exchange materials—I. Zirconium phosphate. *J. Inorg. Nucl. Chem.* **1958**, *6*, 220–235. [CrossRef]
4. Amphlett, C.B.; McDonald, L.A.; Burgess, J.S.; Maynard, J.C. Synthetic inorganic ion-exchange materials—III. The separation of rubidium and cesium on zirconium phosphate. *J. Inorg. Nucl. Chem.* **1959**, *10*, 69–73. [CrossRef]
5. Curtman, L.J.; Greenslade, T.B. Removal of phosphate ion. *J. Chem. Educ.* **1936**, *13*, 238–239. [CrossRef]
6. Holness, H.; Mattock, G. Phosphate removal in qualitative analysis. *Analyst* **1949**, *74*, 43–46. [CrossRef]
7. Friend, J.N.; Vallance, R.H.; Challis, H.J.G. Phosphate separation in qualitative analysis. *Nature* **1940**, *146*, 63. [CrossRef]
8. Kiefer, E.W.; Boltz, D.F. Spectrophotometric determination of zirconium. *Anal. Chem.* **1952**, *24*, 542–544. [CrossRef]
9. Clearfield, A.; Stynes, J.A. The preparation of crystalline zirconium phosphate and some observations on its ion exchange behaviour. *J. Inorg. Nucl. Chem.* **1964**, *26*, 117–129. [CrossRef]
10. Clearfield, A.; Smith, S.D. The crystal structure of zirconium phosphate and the mechanism of its ion exchange behaviour. *J. Colloid Interface Sci.* **1968**, *28*, 325–330. [CrossRef]
11. Clearfield, A.; Smith, G.D. Crystallography and structure of $\alpha$-zirconium bis(monohydrogenorthophosphate) monohydrate. *Inorg. Chem.* **1969**, *8*, 431–436. [CrossRef]
12. Troup, J.M.; Clearfield, A. Mechanism of ion exchange in zirconium phosphates. 20. Refinement of the crystal structure of $\alpha$-zirconium phosphate. *Inorg. Chem.* **1977**, *16*, 3311–3314. [CrossRef]
13. Albertsson, J.; Oskarsson, A.; Tellgren, R.; Thomas, J.O. Inorganic ion exchangers. 10. A neutron powder diffraction study of the hydrogen bond geometry in $\alpha$-zirconium bis(monohydrogenorthophosphate) monohydrate. A model for the ion exchange. *J. Phys. Chem.* **1977**, *81*, 1574–1578. [CrossRef]
14. Pica, M. Zirconium phosphate catalysts in the XXI century: State of the art from 2010 to date. *Catalysts* **2017**, *7*, 190. [CrossRef]
15. Trobajo, C.; Khainakov, S.A.; Espina, A.; García, J.R. On the synthesis of $\alpha$-zirconium phosphate. *Chem. Mater.* **2000**, *12*, 1787–1790. [CrossRef]
16. Alberti, G.; Torracca, E. Crystalline insoluble salts of polybasic metals—II. Synthesis of crystalline zirconium or titanium phosphate by direct precipitation. *J. Inorg. Nucl. Chem.* **1968**, *30*, 317–318. [CrossRef]
17. García-Rosales, G.; Ordóñez-Regil, E.; Romero-Guzmán, E.T.; Ordóñez-Regil, E. The influence of agitation speed on the morphology and size particle synthesis of $Zr(HPO_4)_2 \cdot H_2O$ from mexican sand. *J. Mineral. Mater. Charact. Eng.* **2007**, *6*, 39–51.
18. Horsley, S.E.; Nowell, D.V. The preparation and characterization of crystalline $\alpha$-zirconium phosphate. *J. Appl. Chem. Biotechnol.* **1973**, *23*, 215–224. [CrossRef]
19. Capitani, D.; Casciola, M.; Donnadio, A.; Vivani, R. High yield precipitation of crystalline $\alpha$-zirconium phosphate from oxalic acid solutions. *Inorg. Chem.* **2010**, *49*, 9409–9415. [CrossRef]
20. Shuai, M.; Mejia, A.F.; Chang, Y.W.; Cheng, Z. Hydrothermal synthesis of layered zirconium phosphate disks: Control of aspect ratio and polydispersity for nano-architecture. *Cryst. Eng. Comm.* **2013**, *15*, 1970–1977. [CrossRef]

21. Vekilov, P.G. Two-step mechanism for the nucleation of crystals from solution. *J. Cryst. Growth* **2005**, *275*, 65–76. [CrossRef]

22. Adawy, A. The Ceiling Method for the Growth of High-Resolution Protein Crystals. Ph.D. Thesis, Radboud University, Nijmegen, The Netherlands, 2014.

23. Adawy, A.; Amghouz, Z.; van Hest, J.C.M.; Daniela, A.W. Sub-micron polymeric stomatocytes as promising templates for confined crystallization and diffraction experiments. *Small* **2017**, *13*, 1700642. [CrossRef] [PubMed]

24. Liu, X.; Chee, S.W.; Raj, S.; Sawczyk, M.; Král, P.; Mirsaidov, U. Three-step nucleation of metal–organic framework nanocrystals. *Proc. Nat. Acad. Sci. USA* **2021**, *118*, e2008880118. [CrossRef] [PubMed]

25. Cheng, Y.; Wang, X.; Jaenicke, S.; Chuah, G.K. Minimalistic liquid assisted route to highly crystalline $\alpha$-zirconium phosphate. *Chem. Sus. Chem.* **2017**, *10*, 3235–3242. [CrossRef]

26. Benhamza, H.; Barboux, P.; Bouhaouss, A.; Josien, F.A.; Livage, J. Sol–gel synthesis of $Zr(HPO_4)_2 \cdot H_2O$. *J. Mater. Chem.* **1991**, *1*, 681–684. [CrossRef]

27. Hajipour, A.R.; Karimi, H. Synthesis and characterization of hexagonal zirconium phosphate nanoparticles. *Mater. Lett.* **2014**, *116*, 356–358. [CrossRef]

28. Alberti, G.; Casciola, M.; Marmottini, F.; Vivani, R. Preparation of mesoporous zirconium phosphate-pyrophosphate with a large amount of thermally stable acid groups on the pore surface. *J. Porous Mater.* **1999**, *6*, 299305. [CrossRef]

29. Alberti, G.; Casciola, M.; Cavalaglio, S.; Vivani, R. Proton conductivity of mesoporous zirconium phosphate pyrophosphate. *Solid State Ion.* **1999**, *125*, 91–97. [CrossRef]

30. Clearfield, A.; Thakur, D.S. Zirconium and titanium phosphates as catalysts: A review. *Appl. Catal.* **1986**, *26*, 1–26. [CrossRef]

31. Jiménez-Jiménez, J.; Maireles-Torres, P.; Olivera-Pastor, P.; Rodríguez-Castellón, E.; Jiménez-López, A.; Jones, D.J.; Roziere, J. Surfactant-assisted synthesis of a mesoporous form of zirconium phosphate with acidic properties. *Adv. Mater.* **1998**, *10*, 812–815. [CrossRef]

32. Zhang, F.; Xie, Y.; Lu, W.; Wang, X.; Xu, S.; Lei, X. Preparation of microspherical $\alpha$-zirconium phosphate catalysts for conversion of fatty acid methyl esters to monoethanolamides. *J. Colloid Interface Sci.* **2010**, *349*, 571–577. [CrossRef] [PubMed]

33. Sinhamahapatra, A.; Sutradhar, N.; Roy, B.; Tarafdar, A.; Bajaj, H.C.; Panda, A.B. Mesoporous zirconium phosphate catalyzed reactions: Synthesis of industrially important chemicals in solvent-free conditions. *Appl. Catal. A* **2010**, *385*, 22–30. [CrossRef]

34. Zhao, G.L.; Yuan, Z.Y.; Chen, T.H. Synthesis of amorphous supermicroporous zirconium phosphate materials by nonionic surfactant templating. *Mater. Res. Bull.* **2005**, *40*, 1922–1928. [CrossRef]

35. Tarafdar, A.; Panda, A.B.; Pradhan, N.C.; Pramanik, P. Synthesis of spherical mesostructured zirconium phosphate with acidic properties. *Microporous Mesoporous Mater.* **2006**, *95*, 360–365. [CrossRef]

36. Zhu, Y.; Shimizu, T.; Kitajima, T.; Morisato, K.; Moitra, N.; Brun, N.; Kiyomura, T.; Kanamori, K.; Takeda, K.; Kurata, H.; et al. Synthesis of robust hierarchically porous zirconium phosphate monolith for efficient ion adsorption. *New J. Chem.* **2015**, *39*, 2444–2450. [CrossRef]

37. Bellezza, F.; Cipiciani, A.; Costantino, U.; Marmottini, F.; Quotadamo, M.A. Zirconium phosphate nanoparticles from water-in-oil microemulsions. *Colloid Polym. Sci.* **2006**, *285*, 19–25. [CrossRef]

38. Alberti, G.; Casciola, M.; Capitani, D.; Donnadio, A.; Narducci, R.; Pica, M.; Sganappa, M. Novel Nafion–zirconium phosphate nanocomposite membranes with enhanced stability of proton conductivity at medium temperature and high relative humidity. *Electrochim. Acta* **2007**, *52*, 8125–8132. [CrossRef]

39. Gescher, A. Metabolism of N,N-dimethylformamide: Key to the understanding of its toxicity. *Chem. Res. Toxicol.* **1993**, *6*, 245–251. [CrossRef]

40. Pica, M.; Donnadio, A.; Capitani, D.; Vivani, R.; Troni, E.; Casciola, M. Advances in the chemistry of nanosized zirconium phosphates: A new mild and quick route to the synthesis of nanocrystals. *Inorg. Chem.* **2011**, *50*, 11623–11630. [CrossRef]

41. Li, Z.; Vivas, E.L.; Suh, Y.J.; Cho, K. Highly efficient and selective removal of $Sr^{2+}$ from aqueous solutions using ammoniated zirconium phosphate. *J. Environ. Chem. Eng.* **2022**, *10*, 107333. [CrossRef]

42. Pan, S.; Shen, J.; Deng, Z.; Zhang, X.; Pan, B. Metastable nano-zirconium phosphate inside gel-type ion exchanger for enhanced removal of heavy metals. *J. Hazard. Mater.* **2022**, *423*, 127158. [CrossRef] [PubMed]

43. Bashir, A.; Ahad, S.; Malik, L.A.; Qureashi, A.; Manzoor, T.; Dar, G.N.; Pandith, A.H. Revisiting the old and golden inorganic material, zirconium phosphate: Synthesis, intercalation, surface functionalization, and metal ion uptake. *Ind. Eng. Chem. Res.* **2020**, *59*, 22353–22397. [CrossRef]

44. Colodrero, R.M.P.; Olivera-Pastor, P.; Cabeza, A.; Bazaga-García, M. Properties and Applications of Metal Phosphates and Pyrophosphates as Proton Conductors. *Materials* **2022**, *15*, 1292. [CrossRef] [PubMed]

45. Zhao, Y.; Yan, S.; He, Y.; Li, Z.; Li, C.; Li, H. Synthesis of ultrathin $\alpha$-zirconium phosphate functionalized with polypyrrole for reinforcing the anticorrosive property of waterborne epoxy coating. *Colloids Surf. A Physicochem. Eng. Asp.* **2022**, *635*, 128052. [CrossRef]

46. Zhao, M.; Baker, J.; Jiang, Z.; Zhu, Z.; Wu, H.-M.; Wu, J.-L.; Kang, W.-H.; Sue, H.-J. Preparation of Well-Exfoliated Poly(ethyleneco-vinyl acetate)/alpha-Zirconium Phosphate Nanocomposites. *Langmuir* **2021**, *37*, 4550–4561. [CrossRef]

47. Zhu, Z.; Tsai, C.-Y.; Zhao, M.; Baker, J.; Sue, H.-J. PMMA Nanocomposites Based on PMMA-Grafted $\alpha$-Zirconium Phosphate Nanoplatelets. *Macromolecules* **2022**, *55*, 1165–1177. [CrossRef]

48. Díaz, A.; David, A.; Pérez, R.; González, M.L.; Báez, A.; Wark, S.E.; Zhang, P.; Clearfield, A.; Colón, J.L. Nanoencapsulation of insulin into zirconium phosphate for oral delivery applications. *Biomacromolecules* **2010**, *11*, 2465–2470. [CrossRef]

49. Saxena, V.; Díaz, A.; Clearfield, A.; Batteas, J.D.; Hussain, M.D. Zirconium phosphate nanoplatelets: A biocompatible nanomaterial for drug delivery to cancer. *Nanoscale* **2013**, *5*, 2328–2336. [CrossRef]

50. Díaz, A.; Saxena, V.; González, J.; David, A.; Casanas, B.; Carpenter, C.; Batteas, J.D.; Colón, J.L.; Clearfield, A.; Hussain, M.D. Zirconium phosphate nanoplatelets: A novel platform for drug delivery in cancer therapy. *Chem. Commun.* **2012**, *48*, 1754–1756. [CrossRef]

51. Díaz, A.; González, M.L.; Pérez, R.J.; David, A.; Mukherjee, A.; Báez, A.; Clearfield, A.; Colón, J.L. Direct intercalation of cisplatin into zirconium phosphate nanoplatelets for potential cancer nanotherapy. *Nanoscale* **2013**, *5*, 11456–11463. [CrossRef]

52. Wester, M.; Simonis, F.; Gerritsen, K.G.; Boer, W.H.; Wodzig, W.K.; Kooman, J.P.; Joles, J.A. A regenerable potassium and phosphate sorbent system to enhance dialysis efficacy and device portability: An in vitro study. *Nephrol. Dyalysis Transp.* **2013**, *28*, 2364–2371. [CrossRef] [PubMed]

53. Ding, Y.; Jones, D.J.; Maireles-Torres, P.; Roziere, J. Two dimensional nanocomposites: Alternating inorganic-organic polymer layers in zirconium phosphate. *Chem. Mater.* **1995**, *7*, 562–571. [CrossRef]

54. Kumar, C.; Mclendon, G. Nanoencapsulation of cytochrome c and horseradish peroxidase at the galleries of α-zirconium phosphate. *Chem. Mater.* **1997**, *9*, 863–870. [CrossRef]

55. Kim, H.N.; Keller, S.W.; Mallouk, T.E. Characterization of zirconium phosphate/polycation thin films grown by sequential adsorption reactions. *Chem. Mater.* **1997**, *9*, 1414–1421. [CrossRef]

56. Kumar, C.V.; Chaudhuri, A. Proteins immobilized at the galleries of layered α-zirconium phosphate: Structure and activity studies. *J. Am. Chem. Soc.* **2000**, *122*, 830–837. [CrossRef]

57. Bellezza, F.; Cipiciani, A.; Costantino, U.; Negozio, M.E. Zirconium phosphate and modified zirconium phosphates as supports of lipase. Preparation of the composites and activity of the supported enzyme. *Langmuir* **2002**, *18*, 8737–8742. [CrossRef]

58. Mudhivarthi, V.K.; Bhambhani, A.; Kumar, C.V. Novel enzyme/DNA/inorganic nanomaterials: A new generation of biocatalysts. *Dalton Trans.* **2007**, *36*, 5483–5497. [CrossRef]

59. Bhambhani, A.; Kumar, C.V. Enzyme-inorganic nanoporous materials: Stabilization of proteins intercalated in α-zirconium phosphate by a denaturant. *Micropor. Mesopor. Mater.* **2008**, *110*, 517–527. [CrossRef]

60. Xu, S.; Whitin, J.C.; Yu, T.T.; Zhou, H.; Sun, D.; Sue, H.S.; Zou, H.; Cohen, H.J.; Zare, R.N. Capture of phosphopeptides using α-zirconium phosphate nanoplatelets. *Anal. Chem.* **2008**, *80*, 5542–5549. [CrossRef]

61. González, M.L.; Ortiz, M.; Hernández, C.; Cabán, J.; Rodríguez, A.; Colón, J.L.; Báez, A. Zirconium phosphate nanoplatelet potential for anticancer drug delivery applications. *J. Nanosci. Nanotechnol.* **2016**, *16*, 117–129. [CrossRef]

62. González-Villegas, J.; Kan, Y.; Bakhmutov, V.I.; García-Vargas, A.; Martínez, M.; Clearfield, A.; Colón, J.L. Poly(ethylene glycol)-modified zirconium phosphate nanoplatelets for improved doxorubicin delivery. *Inorg. Chim. Acta* **2017**, *468*, 270–279. [CrossRef]

63. Li, R.; Liu, T.; Wang, K. Hyaluronic acid-modified zirconium phosphate nanoparticles for potential lung cancer therapy. *Biomed. Eng. Biomed. Tech.* **2016**, *62*, 67–73. [CrossRef] [PubMed]

64. Donnadio, A.; Ambrogi, V.; Pietrella, D.; Pica, M.; Sorrentino, G.; Casciola, M. Carboxymethylcellulose films containing chlorhexidine–zirconium phosphate nanoparticles: Antibiofilm activity and cytotoxicity. *RSC Adv.* **2016**, *6*, 46249. [CrossRef]

65. García, I.; Trobajo, C.; Amghouz, Z.; Adawy, A. Nanolayered metal phosphates as biocompatible reservoirs for antimicrobial silver nanoparticles. *Materials* **2021**, *14*, 1481. [CrossRef]

66. Alberti, G.; Cardini-Galli, P.; Costantino, U.; Torracca, E. Crystalline insoluble salts of polybasic metals. 1. Ion-exchange properties of crystalline titanium phosphate. *J. Inorg. Nucl. Chem.* **1967**, *29*, 571–578. [CrossRef]

67. Chernorukov, N.G.; Moskvich, E.P.; Zhuk, M.I. Crystallographic characteristics of phosphates and arsenates of tetravalent elements. *Kristallograf.* **1974**, *19*, 1084–1085.

68. Suárez, M.; García, J.R.; Rodríguez, J. Calorimetric determination of $H^+/Na^+$ ion-exchange on α-titanium phosphate. *J. Phys. Chem.* **1984**, *88*, 157–159. [CrossRef]

69. Suárez, M.; García, J.R.; Rodríguez, J. Thermodynamic treatment of $H^+/Na^+$ ion-exchange on α-titanium phosphate. *J. Phys. Chem.* **1984**, *88*, 159–162. [CrossRef]

70. García, J.R.; Suárez, M.; Guarido, C.G.; Rodríguez, J. X-Ray diffraction spectrometry for the analysis of crystalline solid phases. *Anal. Chem.* **1984**, *56*, 193–196. [CrossRef]

71. Kumada, N.; Imase, A.; Yanagida, S.; Takei, T.; Miura, A.; Itoi, N.; Goto, T. Hydrothermal synthesis of $KTi_2(PO_4)_3$, α-$Ti(HPO_4)_2 \cdot H_2O$ and γ-$Ti(PO_4)(H_2PO_4) \cdot 2H_2O$ from a lepidocrocite-type titanate. *J. Asian Ceram. Soc.* **2019**, *7*, 361–367. [CrossRef]

72. Bortun, A.I.; Khainakov, S.A.; Bortun, L.N.; Poojary, D.M.; Rodriguez, J.; Garcia, J.R.; Clearfield, A. Synthesis and Characterization of Two Novel Fibrous Titanium Phosphates $Ti_2O(PO_4)_2 \cdot 2H_2O$. *Chem. Mater.* **1997**, *9*, 1805–1811. [CrossRef]

73. Park, J.-W.; Jang, J.-H.; Lee, C.S.; Hanawa, T. Osteoconductivity of hydrophilic microstructured titanium implants with phosphate ion chemistry. *Acta Biomater.* **2009**, *5*, 2311–2321. [CrossRef] [PubMed]

74. Lu, J.-S. Corrosion of titanium in phosphoric acid at 250 °C. *Trans. Nonferrous Met. Soc. China* **2009**, *19*, 552–556. [CrossRef]

75. Yada, M.; Inoue, Y.; Sakamoto, A.; Torikai, T.; Watari, T. Synthesis and controllable wettability of micro-and nanostructured titanium phosphate thin films formed on titanium plates. *ACS Appl. Mater. Interfaces* **2014**, *6*, 7695–7704. [CrossRef] [PubMed]

76. Salvadó, M.A.; Pertierra, P.; García-Granda, S.; García, J.R.; Rodríguez, J. Fernández-Díaz, M.T. Neutron powder diffraction study of α-$Ti(HPO_4)_2 \cdot H_2O$ and α-$Hf(HPO_4)_2 \cdot H_2O$; H-atom positions. *Acta Cryst. B* **1996**, *52*, 896–898. [CrossRef]

77. Christensen, A.N.; Andersen, E.K.; Andersen, I.G.K.; Alberti, G.; Nielsen, M.; Lehmann, M.S. X-Ray powder diffraction study of layer compounds. The crystal structure of alpha-Ti(HPO$_4$)$_2$·H$_2$O and a proposed structure for gamma-Ti(H$_2$PO$_4$)(PO$_4$)·2H$_2$O. *Acta. Chem. Scand.* **1990**, *44*, 865–872. [CrossRef]

78. Poojary, D.M.; Shpeizer, B.; Clearfield, A. X-Ray powder structure and Rietveld refinement of γ-zirconium phosphate, Zr(PO$_4$)(H$_2$PO$_4$)·2H$_2$O. *J. Chem. Soc. Dalton Trans.* **1995**, *19*, 111–113. [CrossRef]

79. Krogh Andersen, A.M.; Norby, P.; Vogt, T. Determination of formation regions of titanium phosphates; determination of the crystal structure of β-titanium phosphate, Ti(PO$_4$)(H$_2$PO$_4$), from neutron powder data. *J. Solid State Chem.* **1998**, *140*, 266–271. [CrossRef]

80. Llavona, R.; Suárez, M.; García, J.R.; Rodríguez, J. Lamellar inorganic ion exchangers. Alkali metal ion exchange on α- and γ-titanium phosphate. *Inorg. Chem.* **1989**, *28*, 2863–2868. [CrossRef]

81. Zhou, G.F.; Wang, Q.; Zeng, R.Q.; Fu, X.K.; Yang, X.B. Preparation and application of zirconium phosphate and its derivatives. *Progress Chem.* **2014**, *26*, 87–99.

82. Kishore, M.S.; Pralong, V.; Caignaert, V.; Varadaraju, U.V.; Raveau, B. Electrochemical intercalation of lithium in the titanium hydrogeno phosphate Ti(HPO$_4$)$_2$·H$_2$O. *J. Power Sources* **2007**, *169*, 355–360. [CrossRef]

83. Lee, G.; Zhang, X.; Zhang, H.B.; Varanasi, C.V.; Liu, J. Effect of interlayer spacing on sodium ion insertion in nanostructured titanium hydrogeno phosphates/carbon nanotube composite. *RSC Adv.* **2016**, *6*, 60015–60021. [CrossRef]

84. Xiang, X.; Li, X.; Chen, K.; Tang, Y.; Wan, M.; Ding, X.; Xue, L.; Zhang, W.; Huang, Y. Gamma titanium phosphate as an electrode material for Li-ion and Na-ion storage: Performance and mechanism. *J. Mater. Chem. A* **2016**, *4*, 18084–18090. [CrossRef]

85. García-Glez, J.; Trobajo, C.; Khainakov, S.A.; Amghouz, Z. α-Titanium phosphate intercalated with propylamine: An alternative pathway for efficient europium (III) uptake into layered tetravalent metal phosphates. *Arab. J. Chem.* **2017**, *10*, 885–894. [CrossRef]

86. García-Glez, J.; Trobajo, C.; Adawy, A.; Amghouz, Z. Exfoliation and europium (III)-functionalization of α-titanium phosphate via propylamine intercalation: From multilayer assemblies to single nanosheets. *Adsorption* **2020**, *26*, 241–250. [CrossRef]

87. Albitres, G.A.V.; Cestari, S.P.; Freitas, D.F.S.; Rodrigues, D.C.; Mendes, L.C.; Neumann, R. Intercalation of α-titanium phosphate with longchain amine aided by short-chain amine. *Appl. Nanosci.* **2020**, *10*, 907–916. [CrossRef]

88. Bortun, A.I.; Bortun, L.N.; Clearfied, A.; Villa-García, M.A.; García, J.R.; Rodríguez, J. Synthesis and characterization of a novel layered titanium phosphate. *J. Mater. Res.* **1996**, *11*, 2490–2498. [CrossRef]

89. García-Granda, S.; Salvadó, M.A.; Pertierra, P.; Bortun, A.I.; Khainakov, S.A.; Trobajo, C.; Espina, A.; García, J.R. Hydrothermal synthesis and characterization of an ammoniumtitanium (IV) phosphate with pyrochlore-type structure. *Inorg. Chem. Comm.* **2001**, *4*, 555–557. [CrossRef]

90. Blanco, J.A.; Khainakov, S.A.; Khainakova, O.; García, J.R.; García-Granda, S. An old material in the nanoworld: Organic-inorganic hybrid nanotubes based on γ-titanium phosphate layered crystal structure. *Phys. Status Solidi C* **2009**, *6*, 2190–2194. [CrossRef]

91. García-Granda, S.; Khainakov, S.A.; Espina, A.; García, J.R.; Castro, G.R.; Rocha, J.; Mafra, L. Revisiting the thermal decomposition of layered γ-titanium phosphate and structural elucidation of its intermediate phases. *Inorg. Chem.* **2010**, *49*, 2630–2638. [CrossRef]

92. Espina, A.; Trobajo, C.; Khainakov, S.A.; García, J.R.; Salvadó, M.A.; Pertierra, P.; García-Granda, S. Reaction of π-Ti$_2$O(PO$_4$)$_2$·2H$_2$O with molten alkali nitrates. Synthesis of the fibrous materials π-M$_{0.5}$H$_{0.5}$TiOPO$_4$ (M = Na, K). *Mater. Res. Bull.* **2002**, *37*, 1381–1392. [CrossRef]

93. Espina, A.; Trobajo, C.; Khainakov, S.A.; García, J.R. Synthesis of the fibrous materials ρ-Na$_{0.50}$H$_{0.50}$TiOPO$_4$ and ρ-K$_{0.67}$H$_{0.33}$TiOPO$_4$ by reaction of ρ-Ti$_2$O(PO$_4$)$_2$(H$_2$O)$_2$ with molten alkali nitrates. *Mater. Res. Bull.* **2001**, *36*, 2531–2541. [CrossRef]

94. Salvadó, M.A.; Pertierra, P.; García-Granda, S.; García, J.R.; Fernández- Díaz, M.T.; Dooryhee, E. Crystal structure, including H-atom positions, of Ti$_2$O(PO$_4$)$_2$(H$_2$O)$_2$ determined from synchrotron X-ray and neutron powder data. *Eur. J. Solid State Inorg. Chem.* **1997**, *34*, 1237–1247. [CrossRef]

95. Babaryk, A.A.; Adawy, A.; García, I.; Trobajo, C.; Amghouz, Z.; Colodrero, R.M.P.; Cabeza, A.; Olivera-Pastor, P.; Bazaga-García, M.; dos Santos-Gómez, L. Structural and proton conductivity studies of fibrous π-Ti$_2$O(PO$_4$)$_2$·2H$_2$O: Application in chitosan-based composite membranes. *Dalton Trans.* **2021**, *50*, 7667–7677. [CrossRef]

96. Bereznitski, Y.; Jaroniec, M.; Bortun, A.I.; Poojary, D.M.; Clearfield, A. Surface and structural properties of novel titanium phosphates. *J. Colloid Interface Sci.* **1997**, *191*, 442–448. [CrossRef]

97. García-Glez, J.; Amghouz, Z.; da Silva, I.; Ania, C.O.; Parra, J.B.; Trobajo, C.; García-Granda, S. The ability of a fibrous titanium oxophosphate for nitrogen-adsorption above room temperature. *Chem. Commun.* **2017**, *53*, 2249–2251. [CrossRef]

98. Cai, B.; Jiang, N.; Tan, P.; Hou, Y.; Li, Y.B.; Zhang, L. The custom making of hierarchical micro/nanoscaled titanium phosphate coatings and their formation mechanism analysis. *RSC Adv.* **2019**, *9*, 41311–41318. [CrossRef]

99. Prabakar, S.J.R.; Park, W.-B.; Seo, J.Y.; Singh, S.P.; Ahn, D.; Sohn, K.-S.; Pyo, M. Ultra-stable Ti$_2$O(PO$_4$)$_2$(H$_2$O) as a viable new Ca$^{2+}$ storage electrode material for calcium-ion batteries. *Energy Storage Mater.* **2021**, *43*, 85–96. [CrossRef]

100. Korneikov, R.I.; Aksenova, S.V.; Ivanenko, V.I.; Lokshin, E.P. Stability of titanyl hydrogen phosphates in aqueous media. *Inorg. Mater.* **2018**, *54*, 689–693. [CrossRef]

101. Geng, Y.; Dalhaimer, P.; Cai, S.; Tsai, R.; Tewari, M.; Minko, T.; Discheret, D.E. Shape effects of filaments versus spherical particles in flow and drug delivery. *Nat. Nanotech.* **2007**, *2*, 249–255. [CrossRef]

102. Decuzzi, P.; Ferrari, M. The receptor-mediated endocytosis of nonspherical particles. *Biophy. J.* **2008**, *94*, 3790–3797. [CrossRef]

103. Prabhu, S.; Poulose, E.K. Silver nanoparticles: Mechanism of antimicrobial action, synthesis, medical applications, and toxicity effects. *Int. Nano Lett.* **2012**, *2*, 32. [CrossRef]

104. Wypij, M.; Jędrzejewski, T.; Ostrowski, M.; Trzcińska, J.; Rai, M.; Golińska, P. Biogenic silver nanoparticles: Assessment of their cytotoxicity, genotoxicity and study of capping proteins. *Molecules* **2020**, *25*, 3022. [CrossRef] [PubMed]

105. Roşu, R.A.; Şerban, V.A.; Bucur, A.I.; Dragoş, U. Deposition of titanium nitride and hydroxyapatite-based biocompatible composite by reactive plasma spraying. *App. Surf. Sci.* **2011**, *258*, 3871–3876. [CrossRef]

106. Esteban-Tejeda, L.; Díaz, L.A.; Cabal, B.; Prado, C.; López-Piriz, R.; Torrecillas, R.; Moya, J.S. Biocide glass–ceramic coating on titanium alloy and zirconium oxide for dental applications. *Mater. Lett.* **2013**, *111*, 59–62. [CrossRef]

107. Mondal, J.; Aarik, L.; Kozlova, J.; Niilisk, A.; Mändar, H.; Mäeorg, U.; Simões, A.; Sammelselg, V. Functionalization of Titanium Alloy Surface by Graphene Nanoplatelets and Metal Oxides: Corrosion Inhibition. *J. Nanosci. Nanotechnol.* **2015**, *15*, 6533–6540. [CrossRef]

108. Castellini, I.; Andreani, L.; Parchi, P.D.; Bonicoli, E.; Piolanti, N.; Risoli, F.; Lisanti, M. Hydroxyapatite in total hip arthroplasty. Our experience with a plasma spray porous titanium alloy/hydroxyapatite double-coated cementless stem. *Clin. Cases Miner. Bone Metab.* **2016**, *13*, 221–227. [CrossRef]

109. Li, K.; Yan, J.; Wang, C.; Bi, L.; Zhang, Q.; Han, Y. Graphene modified titanium alloy promote the adhesion, proliferation and osteogenic differentiation of bone marrow stromal cells. *Biochem. Biophys. Res. Comm.* **2017**, *489*, 187–192. [CrossRef]

110. Gnedenkov, S.V.; Sinebryukhov, S.L.; Puz, A.V.; Egorkin, V.S.; Kostiv, R.E. In vivo study of osteogenerating properties of calciumphosphate coating on titanium alloy Ti-6Al-4V. *Biomed. Mater. Eng.* **2016**, *27*, 551–560.

111. Trybuś, B.; Zieliński, A.; Beutner, R.; Seramak, T.; Scharnweber, D. Deposition of phosphate coatings on titanium within scaffold structure. *Acta Bioeng. Biomech.* **2017**, *19*, 65–72.

112. Chellappa, M.; Vijayalakshmi, U. Electrophoretic deposition of silica and its composite coatings on Ti-6Al-4V, and its in vitro corrosion behaviour for biomedical applications. *Mater. Sci. Eng. C Mater. Biol. Appl.* **2017**, *71*, 879–890. [CrossRef] [PubMed]

113. Meyer, U.; Joos, U.; Mythili, J.; Stamm, T.; Hohoff, A.; Fillies, T.; Stratmann, U.; Wiesmann, H.P. Ultrastructural characterization of the implant/bone interface of immediately loaded dental implants. *Biomaterials* **2004**, *25*, 1959–1967. [CrossRef] [PubMed]

114. Sharan, J.; Koul, V.; Dinda, A.K.; Kharbanda, O.P.; Lale, S.V.; Duggal, R.; Mishra, M.; Gupta, G.; Singh, M.P. Bio-functionalization of grade V titanium alloy with type I human collagen for enhancing and promoting human periodontal fibroblast cell adhesion: An in-vitro study. *Colloids Surf. B Biointerface* **2018**, *161*, 1–9. [CrossRef] [PubMed]

115. AlFarraj, A.A.; Sukumaran, A.; Al Amri, M.D.; Van Oirschot, A.B.; Jansen, J.A. A comparative study of the bone contact to zirconium and titanium implants after 8 weeks of implantation in rabbit femoral condyles. *Odontology* **2018**, *106*, 37–44. [CrossRef]

116. Hassan, A.A.M. Surface Modification and Biophysical Characterization of a Prosthetic Alloy. Master's Thesis, Ain Shams University, Cairo, Egypt, 2008.

117. Adawy, A.; Abdel-Fattah, W.I.; Talaat, M.S.; El-Sayed, M.E.-S. Biomimetic coating of a precalcified Ti-6Al-4V alloy. *Open Med. Devices* **2009**, *1*, 19–28. [CrossRef]

118. Abdel-Fattah, W.I.; El-Sayed, M.E.-S.; Talaat, M.S.; Adawy, A. Comparative Study of $Sr^{2+}$ and $Zn^{2+}$ incorporation in the biomimetic coating of a prosthetic alloy. *Open Biomater. J.* **2011**, *3*, 4–13. [CrossRef]

119. Adawy, A.; Abdel-Fattah, W.I. An efficient biomimetic coating methodology for a prosthetic alloy. *Mater. Sci. Eng. C Mater. Biolog. App.* **2013**, *33*, 1813–1818. [CrossRef]

120. Jiang, N.; Guo, Z.; Sun, D.; Li, Y.; Yang, Y.; Chen, C.; Zhang, L.; Zhu, S. Promoting osseointegration of Ti implants through micro/nanoscaled hierarchical Ti phosphate/Ti oxide hybrid coating. *ACS Nano* **2018**, *12*, 7883–7891. [CrossRef]

121. Leelanarathiwat, K.; Minato, K.; Katsuta, Y.; Otsuka, Y.; Katsuragi, H.; Watanabe, F. Cytotoxicity of hydroxyapatite-tyrosine complex with gray titania coating on titanium alloy surface to L929 mouse fibroblasts. *Dent. Mater. J.* **2019**, *38*, 573–578. [CrossRef]

122. Raines, A.L.; Olivares-Navarrete, R.; Wieland, M.; Cochran, D.L.; Schwartz, Z.; Boyan, B.D. Regulation of angiogenesis during osseointegration by titanium surface microstructure and energy. *Biomaterials* **2010**, *31*, 4909–4917. [CrossRef]

123. Tschernitschek, H.; Borchers, L.; Geurtsen, W. Non alloyed titanium as a bioinert metal: A review. *Quintessence Int.* **2005**, *36*, 523–530. [PubMed]

124. Yada, M.; Inoue, Y.; Akihito, G.; Noda, I.; Torikai, T.; Watari, T.; Hotokebuchi, T. Apatite-forming ability of titanium compound nanotube thin films formed on a titanium metal plate in a simulated body fluid. *Colloids Surf. B Biointerfaces* **2010**, *80*, 116–124. [CrossRef] [PubMed]

125. Liu, X.; Chu, P.K.; Ding, C. Surface modification of titanium, titanium alloys, and related materials for biomedical applications. *Mater. Sci. Eng. Res.* **2004**, *92*, 49–121. [CrossRef]

126. Gil, F.J.; Manzanares, N.; Badet, A.; Aparicio, C.; Ginebra, M.P. Biomimetic treatment on dental implants for short-term bone regeneration. *Clin. Oral Investig.* **2014**, *18*, 59–66. [CrossRef] [PubMed]

127. García, I.; Trobajo, C.; Amghouz, Z.; Alonso-Guervos, M.; Díaz, R.; Mendoza, R.; Mauvezín-Quevedo, M.; Adawy, A. Ag- and Sr-enriched nanofibrous titanium phosphate phases as potential antimicrobial cement and coating for a biomedical alloy. *Mater. Sci. Eng. C* **2021**, *126*, 112168. [CrossRef] [PubMed]

# Generalized Damping Model for MEMS Oscillators from Molecular to Viscous Flow Regime

Tobias Zengerle [1,*], Abdallah Ababneh [2] and Helmut Seidel [1]

1   Department of Physics, Saarland University, 66123 Saarbrücken, Germany; seidel@lmm.uni-saarland.de
2   Electronic Engineering Departement, Hijjawi Faculty for Engineering Technology, Yarmouk University, Irbid 21163, Jordan; a.ababneh@yu.edu.jo
*   Correspondence: t.zengerle@lmm.uni-saarland.de

**Abstract:** In this study, we investigate the damping phenomena acting on piezoelectrically driven MEMS oscillators. Three different geometrical shapes of MEMS oscillators are presented, including cantilevers, bending oscillators, and paddle oscillators. An analytical model for their resonance frequencies is derived. The bending modes of these micro-oscillator structures are characterized regarding their resonance frequency and their quality factor as a function of the ambient pressure in a nitrogen atmosphere as well as the dependence on the distance to a neighboring plate representing a geometrical boundary (e.g., to the package or to the mounting). The investigations cover a pressure range from $10^{-3}$ mbar up to 900 mbar and a gap width from 150 μm to 3500 μm. Consequently, a Knudsen number range over six orders of magnitude from 100 to $10^{-4}$ is covered. The measurement data are evaluated with a generalized damping model consisting of four parts representing the individual damping mechanisms (intrinsic, molecular, transitional, and viscous). The evaluated parameters are analyzed as a function of the resonance frequency and the gap width. The data reveal an exponential growing saturation behavior, which is determined by two characteristic lengths, being correlated with the viscous and the thermal boundary layer thickness, respectively. This leads to an estimation of the strength and of the range of the damping effect just by calculating the boundary layer thicknesses given by the resonance frequency and the gas properties. From these results, we gain fundamental insights on the viscous and transitional damping mechanisms as well as on the intrinsic losses. In conclusion, a basic concept is provided to reduce the damping of micro-oscillator bending modes and thus increase the quality factor. Additionally, the results are supported by finite element simulations revealing the temperature and pressure distribution within the gap.

**Keywords:** $Q$ factor; damping model; transitional flow regime; viscous flow regime; MEMS oscillator; bending modes

## 1. Introduction

MEMS oscillators are widely used in our technological world and are found in a variety of applications [1–5]. Based on their operating principle, the oscillators can be divided into two classes: electrostatic excitation and piezoelectric excitation. The advantages of electrostatic excitation include a high-quality factor $Q$ and good frequency stability, especially for frequency timing applications [6]. However, disadvantages are a high operating voltage, occurring nonlinearities, and difficult processing due to the narrow gaps that are required between the two capacitor plates. Piezoelectric oscillators, on the other hand, are characterized by low operating voltage, high electromechanical coupling, and simple CMOS-compatible batch processing since there is no requirement for a gap. However, a disadvantage compared to electrostatic excitation is the lower $Q$ factor and the associated impact on oscillator performance, such as resolution and frequency stability. For this reason, an important area of research is to increase the quality factor of piezoelectric oscillators [7]. This includes an enhancement of the material-related intrinsic quality factor, as well as the

reduction of extrinsic damping effects, being related to environmental ambient pressure and geometrical boundaries [8,9]. In many applications, MEMS oscillators are operated under various ambient conditions, e.g., in a liquid, in a gas atmosphere, or encapsulated under vacuum conditions. Furthermore, the distance of the micro-oscillator to a mounting or to a package is an important factor [10–12]. Therefore, we investigate the influence of an encapsulation package under various ambient pressure conditions and experimentally correlate them with the damping effects.

The overall quality factor of a micro-oscillator is determined by four different damping mechanisms. Therefore, in the first step, the main loss factor must be determined, which subsequently has to be minimized to increase the overall quality factor of the MEMS oscillator. At low pressure ($p = 0.01$ mbar), the intrinsic damping losses are dominant and the extrinsic losses due to the gas atmosphere can be neglected. The extrinsic damping mechanisms are subdivided further, corresponding to the Knudsen number, which is defined by the ambient pressure and the gap width. In the low-pressure range, with a respectively high Knudsen number ($Kn > 1$), molecular flow damping dominates, whereas in the high-pressure range up to atmospheric pressure and beyond viscous flow damping is defined ($Kn < 0.01$) [13]. In between, the transitional flow regime is encountered, where the damping is described by thermal wave resonance effects [14]. Within this paper, we examine a generalized damping model for various oscillator structure shapes with different coverage sizes of the piezoelectric thin film layer. The oscillators are characterized over a large pressure range and over a varying gap width. The generalized damping model shows good agreement with the experimental data, covering six orders of magnitude for the Knudsen number. Additionally, numerical simulations are performed to model the pressure and temperature distribution supporting the found damping fundamentals. With the help of the results, design rules for an optimization of the quality factor on the dependence of the gap width to the package are derived, based only on the resonance frequency.

## 2. Generalized Damping Model

### 2.1. MEMS-Oscillator

The investigated MEMS oscillator structures are shown in Figure 1 and are subdivided into three classes regarding their shape: paddle oscillator (PO), bending oscillator (BO), and cantilever oscillator (CO). The resonance frequency of such MEMS oscillators can be derived by the theory of an externally driven damped harmonic oscillator [15]. This results in the fundamental resonance frequency $f_r$:

$$f_r = \frac{1}{2\pi}\sqrt{\frac{k}{m}},\tag{1}$$

where $k$ and $m$ denote the spring constant, and the mass of the MEMS oscillator, respectively.

The mass of the MEMS oscillator can be calculated easily by the density–volume product, and the spring constant is given by solid mechanics [16]:

$$k_B = \frac{3E_{Si}I}{L^3} = \frac{E_{Si}BT^3}{4L^3},\tag{2}$$

with Young's modulus $E_{Si}$ of the MEMS oscillator, as well as the moment of inertia $I$, the length $L_B$, the width $B_B$ and the thickness $T$ of the beam.

For the PO and BO structures, we assume two parallel-connected beam spring constants ($k = k_B \| k_B = 2k_B$) with the length $L = L_B + L_O$, and width $B = B_B$. For the CO structures, the length $L$ and the width $B$ are directly given by the oscillator dimension $L_O$ and $B_O$.
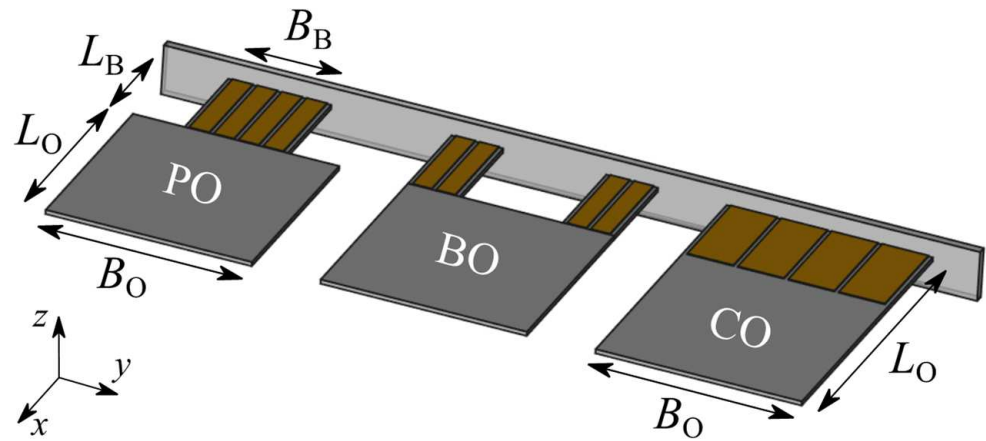
**Figure 1.** Schematic 3D view on the investigated MEMS oscillator structures. From left to right: paddle oscillator PO, bending oscillator BO, and cantilever oscillator CO. The oscillators with thickness *T* are actuated with piezoelectric AlN elements, which are covered with gold electrodes. Microscope images of the fabricated micro-oscillators can be found in Appendix A.

Moreover, we scale the mass *m* of the MEMS oscillator by a factor *α* attributing the effect that not the entire structure is vibrating [17]. Therefore, the effective mass $m_{\text{eff}} = \alpha \cdot m$ is used in the following calculations.

Taking everything into account, this leads to the resonance frequency of the MEMS oscillator as a function of its geometry and material properties:

$$f_{\text{r}} = \frac{1}{2\pi} \begin{cases} T\sqrt{\dfrac{E_{\text{Si}}}{2\alpha\rho_{\text{Si}}}\dfrac{B_{\text{B}}}{(L_{\text{B}}+L_{\text{O}})^3 (2B_{\text{B}}L_{\text{B}}+B_{\text{O}}L_{\text{O}})}} & \text{for PO/BO} \\[2ex] \dfrac{T}{2L_{\text{O}}^2}\sqrt{\dfrac{E_{Si}}{\alpha\rho_{Si}}} & \text{for CO} \end{cases}. \tag{3}$$

This calculation is valid for the resonance frequency of the first fundamental bending mode (BM) of the MEMS oscillator. The resonance frequency of the higher bending modes $f_n$ with mode number *n* can be calculated as a multiple of the squared eigenvalue $\lambda_n$ [18]:

$$f_n = \lambda_n^2 f_{\text{r}}. \tag{4}$$

The eigenvalue $\lambda_n$ is given by 1.88; 4.69; 7.86; 11.00 for the first four bending modes [17]. This analytical approach will be discussed in the experimental section, regarding the accuracy of the description as well as the influence of the different geometric shapes on the scale factor *α*, which is experimentally investigated.

### 2.2. Quality Factor Q

The second characteristic value besides the resonance frequency of a MEMS oscillator is the quality factor *Q*. This is a measure for the damping and is defined by the ratio of the total energy of an oscillator *E* to the energy dissipation per cycle *ΔE*. Its value can be measured from a resonance curve plot using the 3dB bandwidth $\Delta f_{\text{3dB}}$ [19]:

$$Q = \frac{E}{\Delta E} = \frac{f_n}{\Delta f_{\text{3dB}}}. \tag{5}$$

The experimentally measured quality factor always represents a combination of several individual damping mechanisms, which can be combined via the superposition principle.

The damping phenomena can be divided into two classes, the material related intrinsic ones, and the extrinsic ones from the surrounding gas atmosphere:

$$Q_{\text{exp}} = \left( \frac{1}{Q_{\text{int}}} + \frac{1}{Q_{\text{ext}}} \right)^{-1} = \left( \frac{1}{Q_{\text{int}}} + \underbrace{\frac{1}{Q_{\text{mol}}} + \frac{1}{Q_{\text{trans}}} + \frac{1}{Q_{\text{vis}}}}_{\frac{1}{Q_{\text{gas}}}} \right)^{-1}. \tag{6}$$

The intrinsic damping losses are independent of the external pressure and are determined by anchor losses [20], by the residual stress of the piezoelectric layer [21], and by thermoelastic damping effects [22]. Which of the three types of damping is dominant depends on the resonance frequency and on the bending mode.

The extrinsic damping influences can be further subdivided according to several flow regimes with respect to the Knudsen number *Kn*. The Knudsen number is given by the ratio of the mean free path $l_{\text{mfp}}$ to a representative length scale *l\**. In the context of the investigations, the Knudsen number is changed on the one hand by the pressure which influences the mean free path, and on the other hand by the representative length scale which can be adjusted by a variable gap width *h*.

For the different individual flow regimes, descriptions of the damping behavior in the literature are available which are summarized in a more generalized damping model. In the molecular flow regime (*Kn* > 1), the quality factor can be described by a reciprocal dependency on the external pressure which was expanded by Bao et al. with a linear gap width *h* dependency [23,24] (valid for small gap widths [25]):

$$Q_{\text{mol}} = \frac{\zeta_{\text{mol}}(h)}{p}. \tag{7}$$

The other dependencies on the geometric shape, the size of the oscillator, the gas atmosphere, and the ambient temperature are summarized in the fit parameter $\zeta_{\text{mol}}$ [bar].

For low Knudsen numbers (*Kn* < 0.01), the quality factor can be calculated for the viscous flow regime based on the Navier–Stokes equations under some simplifications (assumption of the MEMS oscillator as a chain of spheres/discs [26,27]). The result of these theories is a viscous quality factor exhibiting a dependency to the reciprocal square root of the external pressure:

$$Q_{\text{vis}} = \frac{\zeta_{\text{vis}}(h)}{\sqrt{p}}. \tag{8}$$

The fit parameter $\zeta_{\text{vis}}$ [bar$^{0.5}$] combines the influence of the geometry, the gas atmosphere, and temperature on the viscous quality factor. Dependence of the quality factor on the gap width is not covered by any of these works, except for the squeeze film regime, and is therefore specifically investigated in our work.

The transitional flow regime in-between the two previously presented flow regimes can be described by the resonance effect of thermal waves. The quality factor is described by a parabola-shaped function with a vertex located at the thermal wave resonance point where the mechanical resonance frequency $f_n$, matches with the thermal frequency of the gas locked within the gap width $f_{\text{th}}$. Here, the thermal frequency of the gas depends on its thermal diffusivity, which is changed by the external pressure, as well as on the adjusted gap width. The mathematical description of the quality factor plot of the transitional flow regime is as follows:

$$Q_{\text{trans}} = \zeta_{\text{trans}}(h) \frac{f_n \cdot f_{\text{th}}}{f_n{}^2 + f_{\text{th}}{}^2}, \text{ with } f_{\text{th}} = \frac{B_{\text{trans}}}{\sqrt{p}}. \tag{9}$$

Further elaboration on this theory including the fit parameters $\zeta_{\text{trans}}$ [1] and $B_{\text{trans}}$ [Hz·mbar$^{0.5}$] can be found in [14]. In previous works [14,25] and in Appendix B, the process

steps of the sensor fabrication and the experimental procedure of recording resonance curves for the extraction of the quality factor and resonance frequency are described.

## 3. FEM Simulations

In addition to the analytical description and the experimental investigation of the MEMS oscillators, finite element (FEM) simulations were performed for modeling the temperature and pressure distribution within the gap caused by the oscillator movement and also for determining the thickness of the oscillator.

In the oscillator manufacturing process, the thickness of the oscillators is determined by a time-controlled wet chemical etching process in a KOH solution. This process is affected by variations on the order of $\pm 3$ μm in the targeted oscillator thickness. For this reason, the thickness is determined numerically in COMSOL Multiphysics 5.4 with the solid mechanics' module using an eigenfrequency study of the MEMS oscillator [28]. The exact thickness of one micro-oscillator structure was measured by scanning electron microscopy (SEM). For this method of investigation, however, a thin Au layer is necessary for the conduction of the electrons, after which the oscillator is no longer suitable for further measurements. As an example, an SEM image is attached in Appendix A (see Figure A1d) and confirms the range of thickness of the micro-oscillators.

The pressure and temperature distribution within the gap are modeled by coupling the oscillator movement with the surrounding fluid using an adiabatic and no-slip boundary condition. The geometrical boundary at the top is a fixed plate placed in a defined spacing from the micro-oscillator. It possesses a nearly infinite heat capacity (relative to the amount of heat produced by the micro-oscillator movement) and is therefore modeled by an isothermal wall. Towards the bottom, the micro-oscillator is not confined, which is represented by the spherical wave radiation boundary condition. The behavior of the fluid is described by a linearized Navier–Stokes equation, which is valid for the continuum regime ($Kn < 0.1$). Furthermore, we use the continuity equation and the energy equation for fluids. The complete set of equations can be found in [14].

As a result, we obtain the pressure distribution and temperature distribution within the gap for the individual bending modes and geometric shapes. No significant difference between the distributions of the individual geometric shapes PO, BO, and CO is observed. Therefore, we only illustrate the distributions of a CO structure for the first three bending modes.

Figure 2a indicates a pressure increase $\Delta p$ as well as the resulting temperature rise $\Delta T$ within the gap due to the compression by the MEMS oscillator. Correspondingly, the same behavior in the opposite direction (expansion leading to pressure and temperature drop) can be observed on the bottom side. Furthermore, an asymmetry between the upper and the lower side is significant due to the spatial limitation of the gap towards the top. This asymmetry in the distributions decreases with increasing mode number and disappears almost completely up to the 3rd BM. From this, it can be concluded that the range of an interaction of the oscillator with a spatial boundary decreases with the mode number, respectively, with the resonance frequency.

**Pressure**  **Temperature**
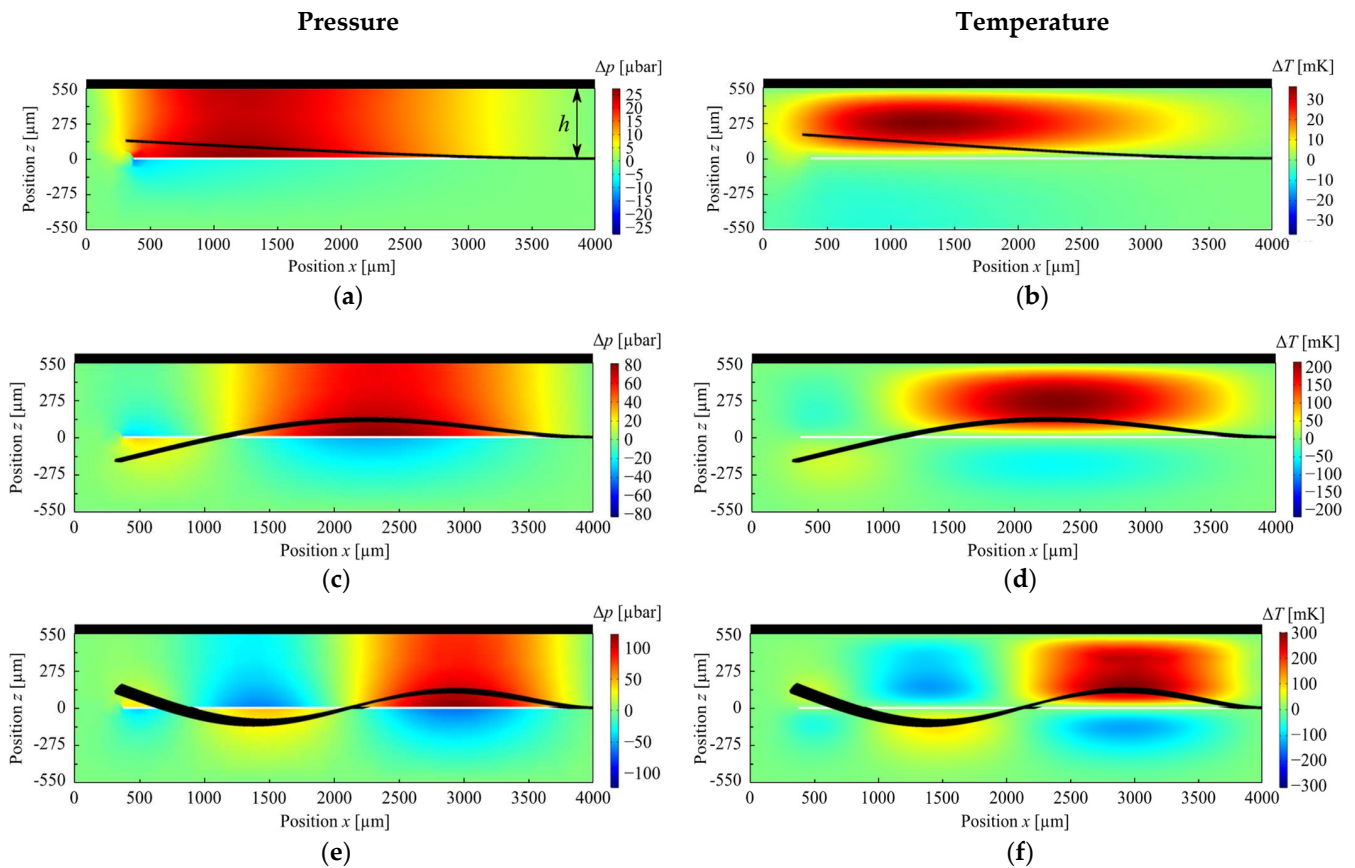


**Figure 2.** FEM-simulated pressure and temperature distribution for CO1 (*cf.* Table 1) at a gap width of 550 μm. In (**a**,**c**,**e**), the pressure distribution of the 1st, 2nd, and 3rd BM is displayed, respectively. In (**b**,**d**,**f**), the corresponding temperature distribution of the 1st, 2nd, and 3rd BM is shown. The ambient values are indicated by a green color. Deviations to the top are shown by red, or to the bottom by blue. The oscillator deflection (black) of the bending mode is shown enlarged and not to scale.

## 4. Experimental Results

Figure 3 shows a composition of the experimentally-measured quality factor $Q_{exp}$ consisting of the individual damping components, according to the theory section. The data show the quality factor plots for the 1st BM of the three different geometries and the 3rd BM of sensor BO for a gap width of 550 μm in a nitrogen atmosphere.

The fitting of the composition of the four different damping mechanisms was performed by a LabView nonlinear fitting tool. The quality of the fitting procedure was ensured by the coefficient of determination $R^2$ which was exceeding the value of 0.99 for all cases ($R^2 = 1$ is representing a perfect match between the experimental data and the theoretical fit.) The relative deviation of the fit model to the experimental data is in the range of 5%. Consequently, a very good agreement between the fitted quality factor $Q_{fit}$ and the experimental data over the entire pressure range can be achieved. The extrinsic gas part of the quality factor $Q_{gas}$ is calculated by subtracting the intrinsic damping and then is inserted additionally. Subsequently, for each gap width and mode, a set of parameters is obtained.
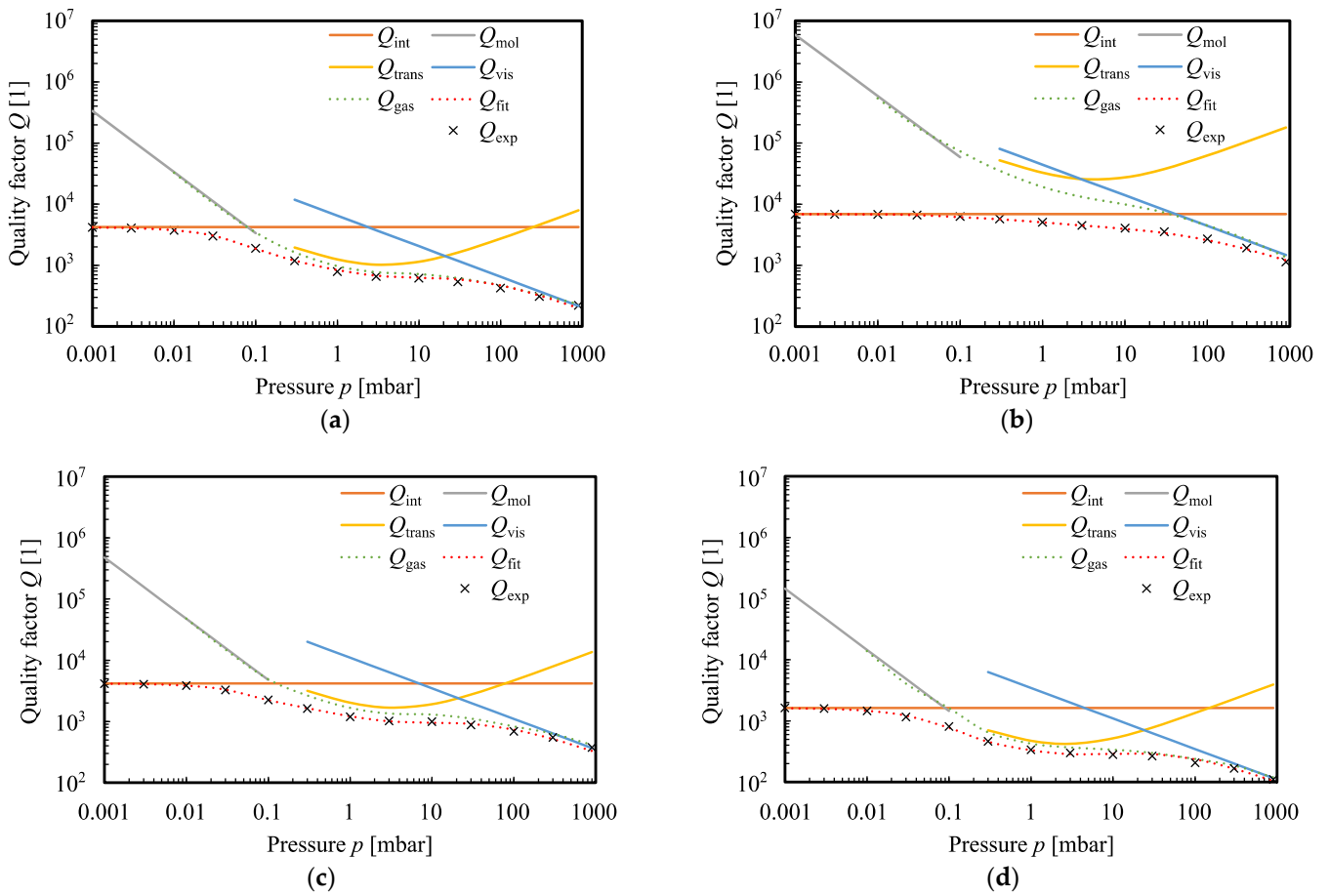
**Figure 3.** Reconstruction of the experimentally-measured quality factor plot $Q_{exp}$ as function of the ambient pressure. The total quality factor $Q_{fit}$ is composed of the individual damping mechanisms and shows very good agreement over the entire range from molecular to viscous flow regime for the 1st BM (**a**) and for the 3rd BM (**b**) of MEMS oscillator BO3 as well as for the 1st BM of the micro-oscillators PO1 (**c**) and CO1_33 (**d**). The data are measured at a gap width $h = 550$ μm in an $N_2$ atmosphere.

In Figure 4a, the quality factor plot of micro-oscillator BO3 is depicted for the fundamental BM and the next higher bending modes up to a mode number of $n = 5$. The strong decrease in the quality factor of the first BM in the medium pressure range is particularly noticeable. This is due to the stronger damping in the molecular flow regime (up to $p = 0.1$ mbar) and due to the high thermal losses in the transitional flow regime caused by standing thermal waves [14]. The first bending mode also has the largest oscillation amplitude, which results in a stronger interaction with the surrounding gas atmosphere and, accordingly, to larger damping losses. The strength of the damping decreases with the mode number and is mainly visible in the viscous flow regime. For mode 5, the intrinsic damping is very pronounced, which leads to a masking of the extrinsic damping losses.
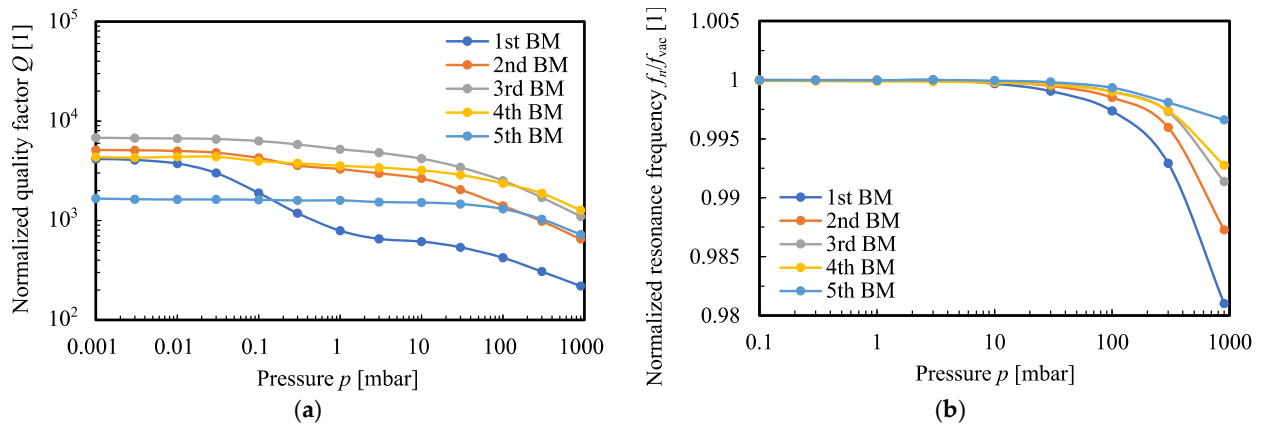
**Figure 4.** Quality factor plot (**a**) and normalized resonance frequency plot (**b**) as function of the ambient pressure of MEMS oscillator BO3 for the first five bending modes. The resonance frequency is normalized on the high vacuum value (at $p = 0.001$ mbar). All data are recorded for a gap width of $h = 550$ μm in an $N_2$ atmosphere.

For all bending modes, the strong damping in the viscous flow regime is common, which leads to a decrease in the quality factor as well as to a reduction of the resonance frequency (see Figure 4b). The highest shift of the resonance frequency due to the strongest damping can also be seen in the 1st BM. The other modes are sorted according to the mode number.

In addition to the quality factor, the resonance frequency of the fundamental mode, as well as the higher bending modes of the MEMS oscillators, are investigated. The frequencies are measured in the vacuum range at 0.001 mbar and then associated with the bending modes using a laser Doppler vibrometer. Besides the bending modes there are other modes, such as torsional oscillations or more complex roof-tile-shaped modes, which are not considered in this work. All electrically-measured bending modes are summarized in Appendix C. Based on the fundamental mode, the thickness of the MEMS oscillator is derived by means of FEM simulation. The results are shown in Table 1. The MEMS oscillators BO1, BO2, BO3, PO1, and PO2 were fabricated on a wafer with a target thickness of 20 μm and the CO structures on a second wafer with a target thickness of 10 μm. These thickness values could be found numerically and are in the range of the variations of the etching process.

**Table 1.** Listing of all investigated MEMS oscillator geometries and their dimensions. The thickness T is obtained from numerical simulations and the scale factor $\alpha$ is determined by applying best fit method of Equation (3) to resonance frequency plot shown in Figure 5. The Young's modulus $E_{Si} = 130$ GPa and density $\rho_{Si} = 2.330$ g/cm$^3$ are taken from [29].

| Name | $L_O$ [μm] | $B_O$ [μm] | $L_B$ [μm] | $B_B$ [μm] | $T$ [μm] | $\alpha$ [1] |
|---|---|---|---|---|---|---|
| BO1 | 1200 | 2400 | 2400 | 600 | 24.5 | 0.36 |
| BO2 | 1800 | 2400 | 1800 | 600 | 24 | 0.27 |
| BO3 | 2400 | 1800 | 1200 | 600 | 17 | 0.47 |
| PO1 | 1200 | 2400 | 2400 | 600 | 24 | 0.45 |
| PO1_100 [1] | 1200 | 2400 | 2400 | 600 | 21 | 0.30 |
| PO2 | 1800 | 2400 | 1800 | 600 | 27 | 0.34 |
| PO2_100 [1] | 1800 | 2400 | 1800 | 600 | 27 | 0.30 |
| CO1_33 [1] | 3600 | 2400 | - | - | 12 | 0.84 |
| CO1_50 [1] | 3600 | 2400 | - | - | 12 | 0.58 |
| CO1_100 [1] | 3600 | 2400 | - | - | 8.5 | 0.61 |
| CO2_33 [1] | 3600 | 1800 | - | - | 12 | 0.95 |
| CO2_100 [1] | 3600 | 1800 | - | - | 9 | 0.67 |

[1] The value is indicating the percentage of coverage of the oscillator surface [30].
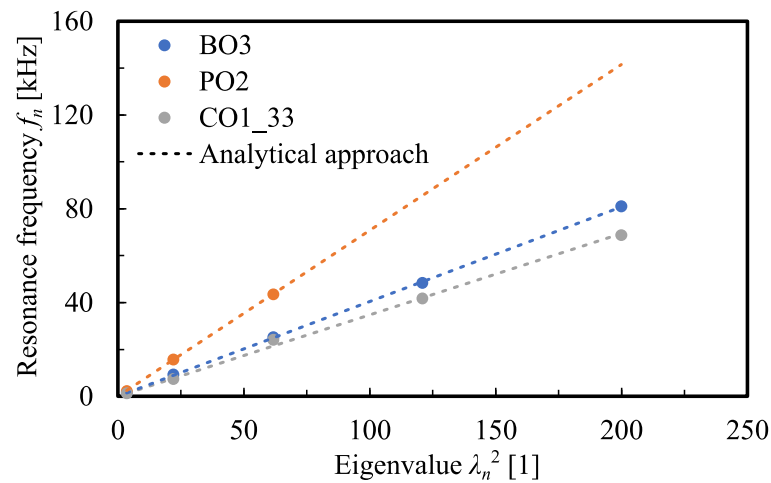
**Figure 5.** Exemplary resonance frequency plot of three oscillator geometries dependent on the eigenvalue. Additionally, the analytical calculation according to Equation (3) is included, which shows a very good agreement. Mode 4 and 5 of micro-oscillator PO2 were electrically not measurable with the piezoelectric elements because the deflection is not in the area of the sensing patches.

In the following, the resonance frequencies of the fundamental mode and the higher bending modes are plotted against the squared eigenvalue $\lambda_n$. By means of the scaling factor $\alpha$, Equation (3) is fitted to the measurement points using the best fit method. Figure 5 shows this exemplarily for a MEMS oscillator of each geometric shape and indicates a very good agreement with the analytical approach. The determined scaling values are listed in Table 1. It is noticeable that the CO structures have the largest values. This is since the oscillator and the bending beam are congruent to each other. The dependence can also be observed on the coverage size. A larger actuator area counteracts the oscillation because the stress of the piezoelectric layer has an additional restoring effect. The smallest possible actuator area thus represents the ideal case of a fully oscillating structure.

For the BO and PO structures, on the other hand, the oscillating mass is mainly given by the oscillator plate dimension itself which is actuated via the beam structure. The larger the beam structure, the smaller the fraction of oscillating mass, after which the scale factor $\alpha$ decreases. Regarding the size of the coverage of the oscillator with the actuator surface, the same observation can be seen for the CO structures. The completely covered structures PO1_100 and PO2_100 show a lower scaling factor $\alpha$ than the regular structures PO1 and PO2.

In the following section, the determined parameters obtained by the fitting procedure shown in Figure 3 will be discussed. In Figure 6, the fitting parameters are depicted as a function of the gap width. The parameters of the individual damping mechanisms show an increase with the gap width which leads to saturation for larger values of the gap width. The molecular fit parameter $\xi_{mol}$ has already been discussed in a previous paper [25] and is excluded in the following. The parameters of the transitional flow regime ($\xi_{trans}$, $B_{trans}$) and the viscous flow regime ($\xi_{vis}$) are fitted with the following function which describes an upward limited growth function:

$$\xi(h) = \xi_{sat} \cdot \left(1 - e^{-\frac{h}{l^*}}\right). \tag{10}$$

**(a)**

**(b)**

**(c)**
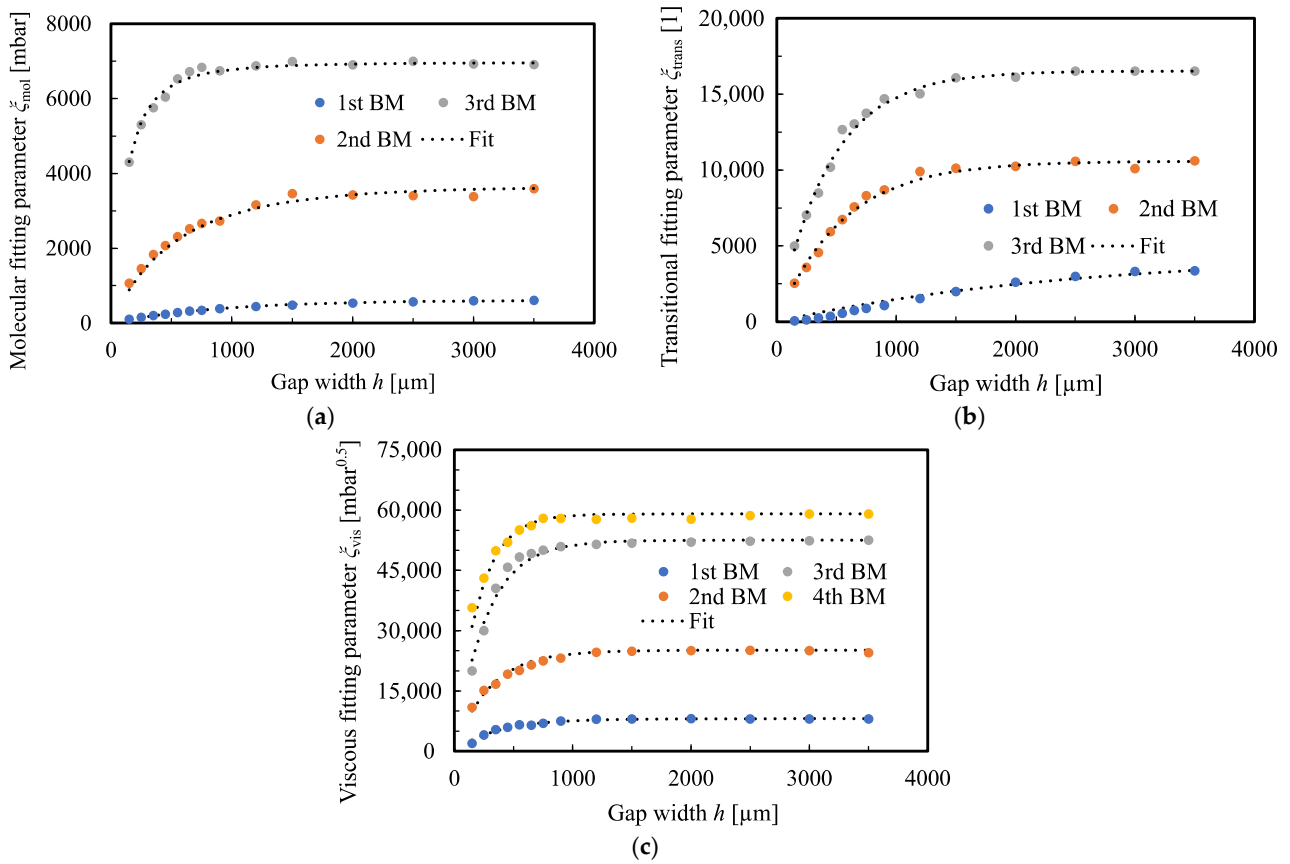
**Figure 6.** The determined fit parameters for MEMS oscillator BO3 of the higher bending modes as a function of the gap width. In (**a**) the curve is shown for the molecular fitting parameter, in (**b**) for the transitional fitting parameter, and in (**c**) for the viscous fitting parameter. Each plot was fitted by an upward limited exponential growth function to determine the characteristic length (*cf.* Equation (10)).

The course is determined by the saturation value $\tilde{\zeta}_{\text{sat}}$, which represents the horizontal asymptote of the function and by the characteristic length $l^*$, which describes the speed of the increase of the gap width.

## 5. Discussion

By applying the generalized fit model, the quality factor plots of three different geometric shapes with different dimensions were fitted from the molecular to the viscous flow range and simultaneously for gap widths to a geometric boundary varying from 150 µm up to 3500 µm. As shown in the previous section, fit parameters were obtained that exhibit a gap width dependence, which can be described according to Equation (10). The determined quantities, the saturation value, and the characteristic length are analyzed in more detail in the following.

In Figure 7, the characteristic lengths for the viscous and thermal damping are depicted. These characteristic lengths represent the range of the interaction with the geometrical boundary. A gap width in the range of the characteristic length thereby means an increase in the quality factor by 63% ($=1 - e^{-1}$) of the saturation value of a freely oscillating structure.
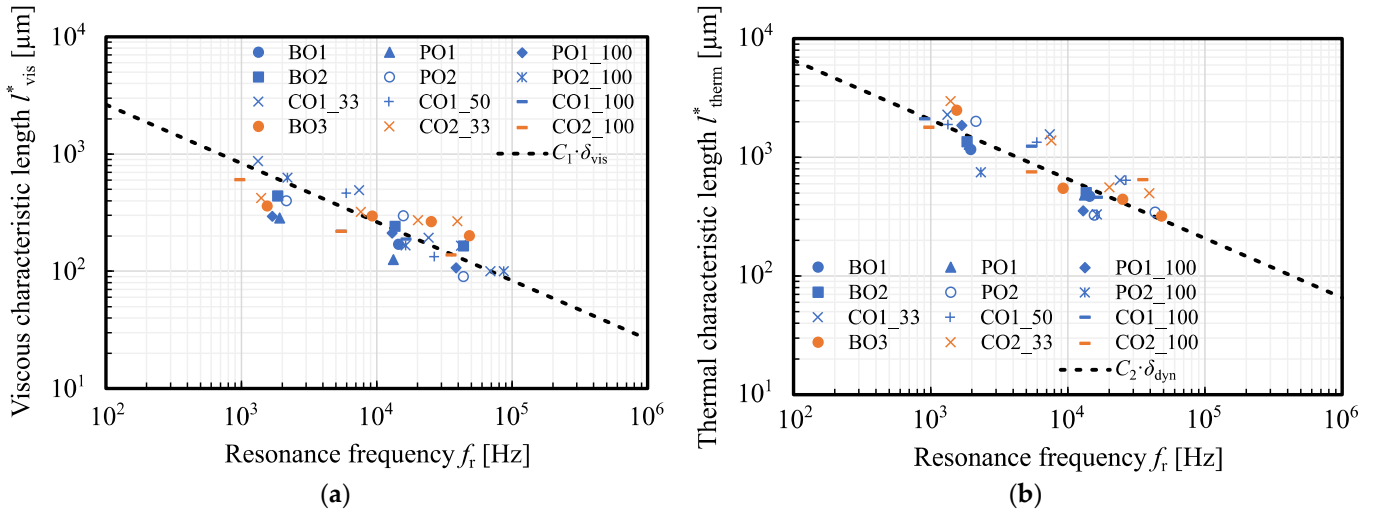
**Figure 7.** Determined characteristic length of the viscous damping (**a**) as well as of the thermodynamic damping of the transitional flow regime (**b**). In both cases, a multiple of the viscous or thermal boundary layer thickness is included ($C_1$ = 12, $C_2$ = 25).

At a distance larger than three times the characteristic length, the damping has dropped to 5% of the original value. Thus, knowledge of the ratio of the distance characteristic can be taken as a measure of the strength of the damping.

In Figure 7, a multiple of the viscous $\delta_{vis}$ and the thermal $\delta_{dyn}$ boundary layer thickness is inserted, which are given as a function of the frequency $f$ by the following equations [31]:

$$\delta_{vis}(f) = \frac{1}{f} \sqrt{\frac{\mu}{\pi \rho f}} \tag{11}$$

$$\delta_{dyn}(f) = \frac{1}{f} \sqrt{\frac{k}{\pi \rho C_p f}}, \tag{12}$$

with $\mu$, $\rho$, $k$, $C_p$ the dynamic viscosity, the density, the thermal conductivity, and the heat capacity at constant pressure, respectively. The gas properties for the fitting of the boundary layer thicknesses were taken from NIST [32].

For the viscous damping, we find a characteristic length $l^*_{vis}$ in the order of $C_1$ = 8–15 times the viscous boundary layer thickness depending on the MEMS oscillator and a mean value of 12. A similar consideration of the damping of the thermal resonance effects in the transitional flow regime yields to a thermal characteristic length $l^*_{therm}$ in the range of $C_2$ = 20–30 times the thermal boundary layer thickness also depending on the exact oscillator geometry with a mean value of about 25. Thus, the thermal effects are generally more long-range than the viscous losses. Furthermore, the decrease in the characteristic length with frequency is observed which agrees with the FEM simulations (*cf.* Figure 2). A higher mode or frequency interacts more short-range with a geometric boundary than the fundamental mode at a lower frequency.

At large distances, a saturation value is obtained, which represents the ideal case of free oscillatory movement without the influence of spatial constraints. In Figure 8, the saturation value for the thermal damping is illustrated. The measurement values clearly distinguish a difference between the MEMS oscillators regarding their width (Blue $B_O$ = 2400 µm; Orange $B_O$ = 1800 µm). The overall trend can be described by a logarithmic function which correlates with a strong decrease in the thermal damping with the frequency. Thermal resonance effects are therefore particularly pronounced at low frequencies or in the fundamental mode.
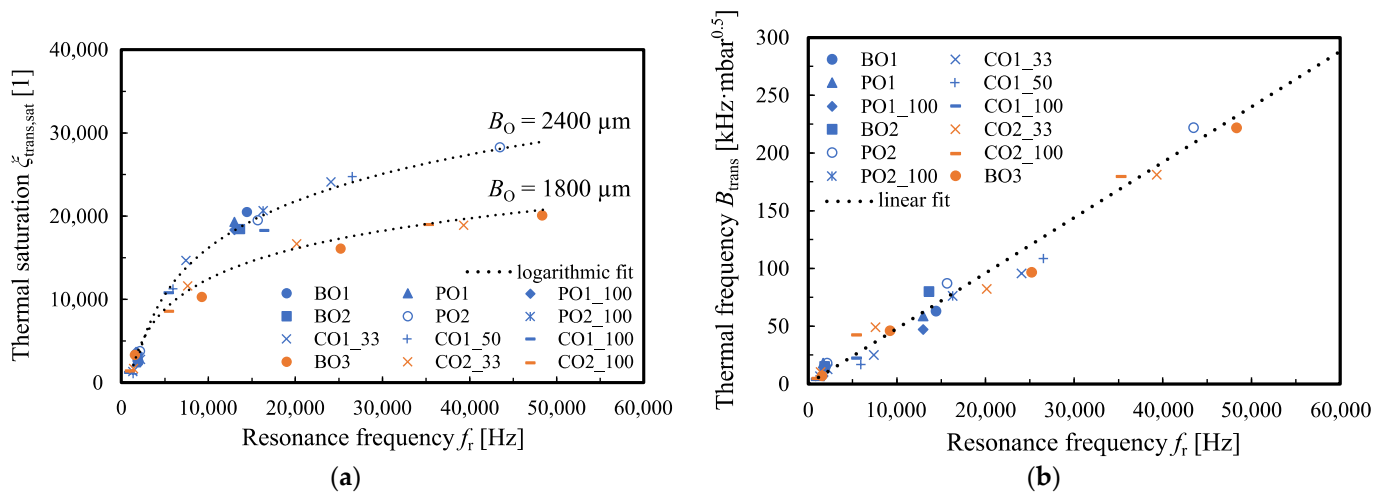
**Figure 8.** Saturation values of the thermodynamic damping for large gap widths determined according to Equation (10) (**a**). The data of the MEMS oscillators show a good agreement with a logarithmic function. The graph reveals two groups that differ in the width of the oscillators $B_O$. The thermodynamic resonance frequency is determined by the fit parameter $B_{dyn}$ which shows an almost linear trend to the resonance frequency of the oscillator (**b**).

The pressure-dependent factor of the thermal resonance frequency shows a linear dependence on the mechanical resonance frequency (see Figure 8b). This derives from the fact that the thermal and mechanical resonance frequencies match at a specific pressure, which we define as the thermal resonance pressure.

The viscous saturation value is predicted by various theories of a freely vibrating structure in a viscous medium. The resulting dependence of the viscous quality on the reciprocal root of the resonant frequency is plotted in Figure 9 and reflects the general trend of the measurements. The individual MEMS oscillators differed in strength from each other due to the different geometries. It is noticeable that the cantilever structures have the highest damping. The further sorting is carried out according to the size of the surface of the MEMS oscillators after which only the absolute area of interaction with the viscous medium is decisive. This is different from damping in liquids where the perimeter of the edges and thus additional slits and edges are counterproductive. According to Basak et al., additional edges lead to turbulence effects which decrease the quality factor [33].

The last point to be investigated is the intrinsic quality factor. All data are summarized in Appendix C and plotted in Figure 9b as a function of the resonance frequency. Several points are noticeable as a result. First, a larger beam structure leads to higher losses via the anchor. For this reason, the cantilever structures have the lowest intrinsic values. The BO and PO structures are sorted by their beam width. On the other hand, larger piezoelectric actuator films also have a negative effect on intrinsic quality. Up to a frequency of about 30 kHz, the quality factor increases approximately linearly with resonance frequency for all structures. At higher frequencies, the thermoelastic damping mechanism is dominant, which leads to a decrease in the intrinsic Q-factor [34]. In general, it can be concluded that the intrinsic quality factor can be improved with the smallest possible actuator area and minimal anchor structures.
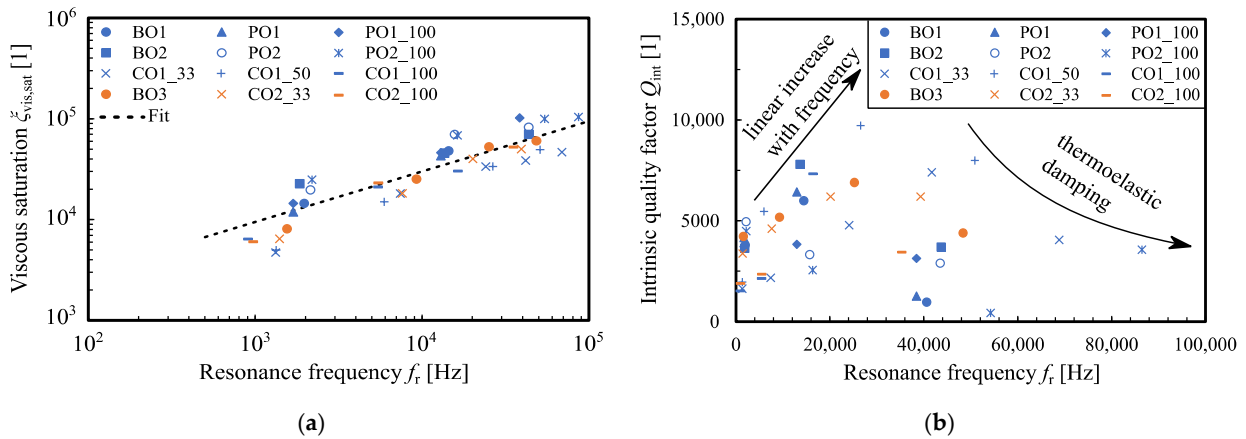
(**a**)                                                                  (**b**)

**Figure 9.** Viscous saturation values for large gaps of all investigated MEMS oscillators as a function of resonance frequency (**a**). The plotted fit line shows the dependence of the quality factor on the frequency in the viscous flow regime according to Equation (8). In (**b**), the intrinsic quality factors of all MEMS oscillators as well as all determined bending modes measured in the high vacuum range at a pressure level of 0.001 mbar are shown.

## 6. Conclusions

A generalized damping model based on the literature results for individual damping mechanisms was presented. First, an analytical description of the resonance frequency was presented, which is in good agreement with the experimental results for all considered geometries and allows numerical determination of the thickness $T$. The quality factor could be reproduced by superposition of four components for the three different oscillator geometries (BO, PO, CO) over six orders of magnitude of the Knudsen number $Kn$. The results of the fitting procedure were further analyzed and interpreted using FEM simulations. From these results, the pressure and temperature distributions within the gap caused by the micro-oscillator movement were obtained. The thermal and viscous interaction with the boundary was visualized and a decrease in the range of the effects with the resonance frequency and the mode number was observed. The decrease in the influence of the damping phenomena was also found experimentally and described with a saturation behavior, from which a characteristic length was obtained, which was used for the estimation of the range of the effects. From these results, indications for reducing the damping were found, thus giving opportunities for increasing the quality factor. Furthermore, the influence of a geometrical boundary due to the package or the mounting can be predicted. The results are briefly summarized below:

- The range of viscous damping is about 8–15 times the viscous boundary layer thickness.
- The measure of the strength is the surface which interacts with the viscous fluid and decreases with the resonance frequency.
- The range of thermal effects in the transitional flow regime is about 20–30 times the thermal boundary layer thickness.
- Exponential decreasing of thermal damping effects occurs with an increase in resonance frequency.
- The thermal resonance frequency of the gas depends linearly on the mechanical resonance frequency of the micro-oscillator.
- The intrinsic quality factor of the fundamental mode is limited by anchor losses and residual stress of the piezoelectric thin film. Therefore, the minimization of an anchor structure and the actuator area are recommended to maximize the intrinsic quality factor.
- At higher modes, thermoelastic damping is crucial, which leads to a lowering of the intrinsic quality of higher modes.

With this model, we give an approach to describing the resonance frequency and quality factor of various MEMS oscillators as well as optimization possibilities to improve the performance of piezoelectric-actuated oscillators.

## Appendix A



**Figure A1.** Microscopic image of the three different micro-oscillator structures: PO (**a**), BO (**b**), and CO (**c**) and a scanning electron microscope image illustrating the side view on a micro-oscillator revealing a thickness of 27 μm (**d**).

## Appendix B

The micro-oscillators were fabricated as shown in the cross-sectional view depicted in Figure A2a. A highly p-doped 4″ Si-wafer with a thickness of 525 μm was used as substrate. In the first step, $SiO_2$ was grown with a thickness of 120 nm on both sides for electrical insulation and as a bonding agent of the $Si_3N_4$ layer. The $Si_3N_4$ layer was deposited in a PECVD process in a thickness of 550 nm and was used as a passivation layer for the anisotropic KOH etching step to release the micro-oscillators. Both layers were structured simultaneously in a 6% HF solution. Next, the piezoelectric thin film was deposited in a reactive sputter deposition process in a thickness of 1100 nm and etched with phosphoric acid solution after a photolithographic structuring step. Subsequently, the piezoelectric elements are connected with a 500 nm thick Au layer using a DC sputtering process. The Au film is structured by an etching step with aqua regia. Afterward, the micro-oscillators

are released in a two-step process. Firstly, a thin membrane is etched anisotropically from the backside in a KOH etching solution, and released by a DRIE step (Bosch-process) from the top side. Before the DRIE process, the KOH cavities are filled with photoresists to mechanically stabilize the oscillator and to stop the DRIE etching process. The photoresist is removed after the dicing of the wafer within the cleaning process of the individual chips. Finally, the micro-oscillators are glued on PCB-boards to contact them electrically via Au bond wires. The PCB boards are drilled in advance at the locations of the micro-oscillator so that they are not confined to the bottom.
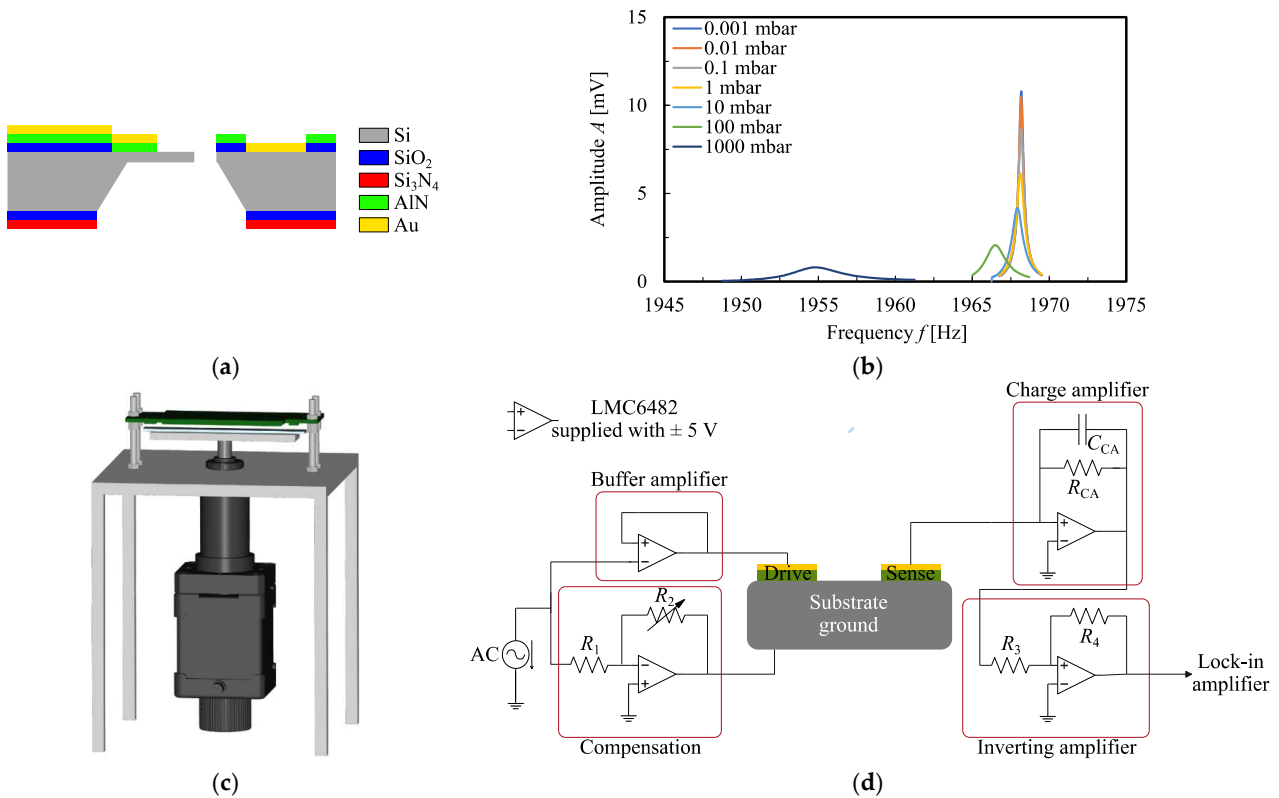


**Figure A2.** Cross-sectional view on a micro-oscillator (layer thickness are not to scale) (**a**), and recorded resonance curves of BO1 (**b**) for various ambient pressure levels. In (**c**), the construction for adjusting the gap width to the neighboring plate is depicted, and in (**d**) the electric circuit for the actuation and detection of the micro-oscillators.

The quality factor and resonance frequency of the micro-oscillators as shown in Figures 3–5 are measured by recording the resonance curves sown in Figure A2b for various ambient pressure conditions. Therefore, the micro-oscillators are placed in a vacuum chamber ranging from $10^{-3}$ mbar up to 900 mbar. Additionally, the PCB board containing four micro-oscillators is built in a construction shown schematically in Figure A2c to adjust the gap width to the neighboring plate. The distance between the plate and the micro-oscillator is initially calibrated with two cooper gauge blocks with a defined thickness of 800 μm. The actuation of the micro-oscillator and the sensing of the oscillation are achieved with the electric circuit illustrated in Figure A2d. The micro-oscillators are actuated with a sinusoidal signal with an amplitude of 1 V which is decoupled by a buffer amplifier. The electrical crosstalk between actuation and sensing is compensated by a shift of the substrate ground [35]. The amplitude of the oscillation movement is measured by collecting the charge of the sensing elements with a charge amplifier which amplifies the signal and converts it in the range of a few μV. Finally, this signal is lifted to several mV by an inverting amplifier. Then, the actuation and sensing signal is compared by a lock-in amplifier and evaluated regarding their amplitude and phase.

**Appendix C**

| Name | $L_O$ [μm] | $B_O$ [μm] | $L_B$ [μm] | $B_B$ [μm] | $T$ [μm] | $\alpha$ [1] | Mode | $f_n$ [kHz] | $Q_{int}$ [1] |
|---|---|---|---|---|---|---|---|---|---|
| BO1 | 1200 | 2400 | 2400 | 600 | 24.5 | 0.36 | 1st BM | 1.97 | 3805 |
| | | | | | | | 2nd BM | 14.3 | 6002 |
| | | | | | | | 3rd BM | 40.6 | 965 |
| BO2 | 1800 | 2400 | 1800 | 600 | 24 | 0.27 | 1st BM | 1.85 | 3640 |
| | | | | | | | 2nd BM | 13.6 | 7800 |
| | | | | | | | 3rd BM | 43.7 | 3704 |
| BO3 | 2400 | 1800 | 1200 | 600 | 17 | 0.47 | 1st BM | 1.55 | 4230 |
| | | | | | | | 2nd BM | 9.24 | 5175 |
| | | | | | | | 3rd BM | 25.2 | 6905 |
| | | | | | | | 4th BM | 48.3 | 4402 |
| PO1 | 1200 | 2400 | 2400 | 600 | 24 | 0.45 | 1st BM | 1.92 | 4170 |
| | | | | | | | 2nd BM | 13.2 | 6436 |
| | | | | | | | 4th BM | 69.7 | 1272 |
| PO1_100 | 1200 | 2400 | 2400 | 600 | 21 | 0.30 | 1st BM | 1.69 | 3722 |
| | | | | | | | 2nd BM | 13.0 | 3832 |
| | | | | | | | 3rd BM | 38.4 | 3133 |
| PO2 | 1800 | 2400 | 1800 | 600 | 27 | 0.34 | 1st BM | 2.15 | 4951 |
| | | | | | | | 2nd BM | 15.7 | 3326 |
| | | | | | | | 3rd BM | 43.4 | 2898 |
| PO2_100 | 1800 | 2400 | 1800 | 600 | 27 | 0.30 | 1st BM | 2.11 | 4488 |
| | | | | | | | 2nd BM | 16.3 | 2555 |
| | | | | | | | 3rd BM | 54.2 | 431 |
| | | | | | | | 4th BM | 86.4 | 3571 |
| CO1_33 | 3600 | 2400 | - | - | 12 | 0.84 | 1st BM | 1.33 | 1635 |
| | | | | | | | 2nd BM | 7.40 | 2174 |
| | | | | | | | 3rd BM | 24.1 | 4781 |
| | | | | | | | 4th BM | 41.7 | 7404 |
| | | | | | | | 5th BM | 68.8 | 4050 |
| CO1_50 | 3600 | 2400 | - | - | 12 | 0.58 | 1st BM | 1.34 | 1956 |
| | | | | | | | 2nd BM | 5.94 | 5459 |
| | | | | | | | 3rd BM | 26.5 | 9718 |
| | | | | | | | 4th BM | 50.9 | 7995 |
| CO1_100 | 3600 | 2400 | - | - | 8.5 | 0.61 | 1st BM | 0.90 | 1519 |
| | | | | | | | 2nd BM | 5.45 | 2140 |
| | | | | | | | 3rd BM | 16.4 | 7337 |
| | | | | | | | 4th BM | 34.8 | 4002 |
| CO2_33 | 3600 | 1800 | - | - | 12 | 0.95 | 1st BM | 1.40 | 3371 |
| | | | | | | | 2nd BM | 7.59 | 4614 |
| | | | | | | | 3rd BM | 20.2 | 6199 |
| | | | | | | | 4th BM | 39.3 | 6198 |
| CO2_100 | 3600 | 1800 | - | - | 9 | 0.67 | 1st BM | 0.97 | 1896 |
| | | | | | | | 2nd BM | 5.47 | 2350 |
| | | | | | | | 4th BM | 35.3 | 3448 |

## References

1. Pillai, G.; Li, S. Piezoelectric MEMS resonators: A review. *IEEE Sens. J.* **2021**, *21*, 12589–12605. [CrossRef]
2. Weng, C.; Pillai, G.; Li, S. A thin-film piezoelectric-on-Silicon MEMS oscillator for mass sensing applications. *IEEE Sens. J.* **2020**, *12*, 7001–7009. [CrossRef]
3. Zengerle, T.; Stopp, M.; Ababneh, A.; Seidel, H. Using the nonlinear Duffing effect of piezoelectric micro-oscillators for wide-range pressure sensing. *Actuators* **2021**, *10*, 172. [CrossRef]
4. Ruiz-Diez, V.; Toledo, J.; Hernando-Garcia, J.; Ababneh, A.; Seidel, H.; Sanchez-Rojas, J. A geometrical study on the roof tile-shaped modes in AlN-based piezoelectric microcantilevers as viscosity–density sensors. *Sensors* **2019**, *19*, 658. [CrossRef]
5. van Beek, J.; Puers, R. A review of MEMS oscillators for frequency reference and timing applications. *J. Micromech. Microeng.* **2012**, *22*, 1–35. [CrossRef]
6. Kim, B.; Melamud, R.; Candler, R.; Hopcroft, A.; Jha, C.; Chandorkar, S.; Kenny, T. Encapsulated MEMS resonators—A technology path for MEMS into frequency control applications. In Proceedings of the IEEE International Frequency Control Symposium, Newport Beach, CA, USA, 1–4 June 2010; pp. 1–4.
7. Tu, C.; Lee, J.; Zhang, X. Dissipation analysis methods and Q-enhancement strategies in piezoelectric MEMS laterally vibrating resonators: A review. *Sensors* **2020**, *20*, 4978. [CrossRef] [PubMed]
8. Xu, C.; Piazza, G. Active boost in the quality factor of an AlN MEMS resonator up to 165,000. In Proceedings of the IEEE 32nd International Conference on Micro Electro Mechanical Systems (MEMS), Seoul, South Korea, 27–31 January 2019; pp. 907–910.
9. Haider, S.; Saleem, M.; Ahmed, M. Effect of environmental conditions and geometric parameters on the squeeze film damping in RF-MEMS switches. *AICPEF* **2019**, *100*, 357–368. [CrossRef]
10. Patocka, F.; Schneidhofer, C.; Dörr, N.; Schneider, M.; Schmid, U. Novel resonant MEMS sensor for the detection of particles with dielectric properties in aged lubricating oils. *Sens. Actuators A* **2020**, *315*, 112290. [CrossRef]
11. Kim, B.; Melamud, R.; Candler, R.; Hopcroft, M.; Kenny, T. MEMS packaging for reliable resonators and oscillators. In Proceedings of the IEEE/MTT-S International Microwave Symposium, Montreal, QC, Canada, 17–22 June 2012; pp. 1–3.
12. Brown, J.; Lutz, M.; Partridge, A.; Gupta, P.; Radza, E. MEMS as low-cost high-volume semiconductor solutions: It's all in the packaging and assembly. *SPIE MOEMS-MEMS* **2008**, *6884*, 183–188.
13. Karniakadis, G.; Beskok, A.; Aluru, A. *Microflows and Nanoflows: Fundamentals and Simulation*, 1st ed.; Springer: New York, NY, USA, 2005.
14. Zengerle, T.; Joppich, J.; Schwarz, P.; Ababneh, A.; Seidel, H. Modeling the damping mechanism of MEMS oscillators in the transitional flow regime with thermal waves. *Sens. Actuators A* **2020**, *311*, 112068. [CrossRef]
15. Alrasheed, S. Oscillatory motion. In *Principles of Mechanics*, 1st ed.; Springer International Publishing: Cham, Switzerland, 2019; pp. 155–171.
16. Young, W.; Budynas, R. *Roark's Formulas for Stress and Strain*, 7th ed.; McGraw-Hill Professional: New York, NY, USA, 2001; pp. 799–812.
17. Boisen, A.; Dohn, S.; Keller, S.; Schmid, S.; Tenje, M. Cantilever-like micromechanical sensors. *Rep. Prog. Phys.* **2011**, *74*, 3. [CrossRef]
18. Timoshenko, S.; Young, D.; Weaver, W. *Vibration Problems in Engineering*, 4th ed.; Wiley: New York, NY, USA, 1974.
19. Petersan, P.; Anlage, S. Measurement of resonant frequency and quality factor of microwave resonators: Comparison of methods. *J. Appl. Phys.* **1998**, *84*, 3392–4002. [CrossRef]
20. Rodriguez, J.; Chandorkhar, S.; Glaze, G.; Gerrard, D.; Chen, Y.; Heinz, D.; Flader, I.; Kenny, T. Direct detection of anchor damping in MEMS tuning fork resonators. *J. Microelectromech. Syst.* **2018**, *27*, 800–809. [CrossRef]
21. Qiu, H.; Ababneh, A.; Feili, D.; Wu, X.; Seidel, H. Analysis of intrinsic damping in vibrating piezoelectric microcantilevers. *Microsyst. Technol.* **2016**, *22*, 2017–2025. [CrossRef]
22. Zener, C. Internal friction in solids: II. General theory of thermoelastic internal friction. *Phys. Rev.* **1937**, *53*, 90–99. [CrossRef]
23. Christian, R. The theory of oscillating-vane vacuum gauges. *Vacuum* **1966**, *16*, 175–178. [CrossRef]
24. Bao, M.; Yang, H.; Yin, H.; Sun, Y. Energy transfer model for squeeze-film air damping in low vacuum. *J. Micromech. Microeng.* **2002**, *12*, 341–346. [CrossRef]
25. Zengerle, T.; Joppich, J.; Schwarz, P.; Ababneh, A.; Seidel, H. Polyatomic degrees of freedom and their temporal evolution extracted from the damping of micro-oscillators. *Sens. Actuators A* **2019**, *297*, 111460. [CrossRef]
26. Kokubun, K.; Hirata, M.; Ono, M.; Murakami, H.; Toda, Y. Unified formula describing the impedance dependence of a quartz oscillator on gas pressure. *J. Vac. Sci. Technol. A* **1987**, *5*, 2450–2453. [CrossRef]
27. Blom, F.; Bouwstra, S.; Elwenspoek, M.; Fluitman, J. Dependence of the quality factor of micromachined silicon beam. *J. Vac. Sci. Technol. B* **1992**, *10*, 341–346. [CrossRef]
28. Kurmendra, S.; Kumar, R. Design analysis, modeling and simulation of novel rectangular cantilever beam for MEMS sensors and energy harvesting applications. *Int. J. Inf. Technol.* **2017**, *9*, 295–302. [CrossRef]
29. Hopcroft, M.; Nix, W.; Kenny, T. What is the Young's modulus of silicon? *J. Microelectromech. Syst.* **2010**, *3*, 229–238. [CrossRef]
30. Zengerle, T.; Stopp, M.; Ababneh, A.; Seidel, H. Investigations on nonlinearities of roof tile-shape modes for pressure measurement applications. In Proceedings of the 21st International Conference on Solid-State Sensors, Actuators and Microsystems (Transducers), Orlando, FL, USA, 20–24 June 2021; pp. 1343–1346.

31. Qiu, H.; Seidel, H. Hydrodynamic loading on vibrating piezoelectric microresonators. In *Piezoelectricity–Organic and Inorganic Materials and Applications*, 1st ed.; IntechOpen Limited: London, UK, 2018; pp. 43–64.

32. Lemmon, E.; McLinden, M.; Friend, D. Thermophysical Properties of Fluid Systems NIST. 1998. Available online: https://webbook.nist.gov/chemistry/fluid/ (accessed on 10 February 2022).

33. Basak, S.; Raman, A.; Garimella, S. Hydrodynamic loading of microcantilevers vibrating in viscous fluids. *J. Appl. Phys.* **2006**, *99*, 114906. [CrossRef]

34. Duwel, A.; Candler, R.; Kenny, T.; Varghese, M. Engineering MEMS resonators with low thermoelastic damping. *J. Microelectromech. Syst.* **2006**, *15*, 1437–1445. [CrossRef]

35. Qiu, H.; Schwarz, P.; Völlm, H.; Feili, D.; Wu, X.; Seidel, H. Electrical crosstalk in two-port piezoelectric resonators and compensation solutions. *J. Micromech. Microeng.* **2013**, *23*, 4. [CrossRef]

*Communication*

# An Analytical Method to Estimate Supersaturation in Gas–Liquid Systems as a Function of Pressure-Reduction Step and Waiting Time

Sushobhan Pradhan [ID] and Prem Kumar Bikkina *[ID]

School of Chemical Engineering, Oklahoma State University, Stillwater, OK 74078, USA;
sushobhan.pradhan@okstate.edu
* Correspondence: prem.bikkina@okstate.edu; Tel.: +1-(405)-744-9112

**Abstract:** When the concentration of a gas exceeds the equilibrium concentration in a liquid, the gas–liquid system is referred as a supersaturated system. The supersaturation can be achieved by either changing the pressure and/or temperature of the system. The gas from a supersaturated liquid escapes either through bubble nucleation that usually occurs on solid surface and/or gas diffusion through the gas–liquid interface. The bubble nucleation requires a minimum threshold supersaturation. A waiting time is required to observe whether the applied supersaturation is sufficient to initiate bubble nucleation. When the supersaturation is not sufficient to cause bubble nucleation, some or all of the supersaturated gas may diffuse out from the liquid through the gas–liquid interface before further reducing the pressure in order to increase the supersaturation. In this article, using Fick's second law of diffusion and Henry's law, an analytical method is proposed to estimate the level of supersaturations generated in three gas–liquid systems at different step-down pressures. Characteristic times of the gas–liquid systems were estimated to validate whether the waiting times used in this study are in accordance with the semi-infinite diffusion model used to estimate the supersaturations generated.

**Keywords:** bubble nucleation; gas-liquid systems; supersaturation; step-down pressure; Fick's second law diffusion; characteristic time

## 1. Introduction

A gas–liquid system is referred to as supersaturated when the concentration of the solute phase (gas) exceeds the equilibrium concentration in the solvent phase (liquid). A supersaturated solution can be achieved by changing either the temperature and/or pressure of the system [1,2]. Thermodynamically, the degree of supersaturation can be described by chemical potential ($\mu_i$), which represents the deviation from equilibrium at a given temperature and pressure for a component $i$.

For a component $i$ in a mixture, the chemical potential in an isothermal–isobaric ensemble is given by:

$$\mu_i(T,\, P,\, x_i) = \left( \frac{\partial G}{\partial N_i} \right)_{T, P, N_{j, j \neq i}}$$

where $G$ is the free energy of the system $j$, $N_i$ is the number of moles of the $i$th component, $T$ is the temperature, $P$ is the pressure, $j$ refers to the number of remaining components, and $x_i$ is the molar fraction of the $i$th component in the mixture [3]. For a supersaturated system, the change in chemical potential of the component $i$ ($\Delta \mu_i$) w.r.t. temperature and pressure is >0.

Dissolution and exsolution of $CO_2$ in and from water/brine are relevant in $CO_2$ sequestration. Depending on whether the $CO_2$ exsolution proceeds with or without bubble nucleation, the pressure response and relative permeability characteristics of the fluids in

the porous media would be influenced [4]. Bubble nucleation and growth are important in sparkling beverages that are weakly supersaturated with $CO_2$ [5]. During limnic eruptions, which are natural calamities, high concentrations of gas slowly built up in the water column of a lake suddenly erupt after a trigger mechanism leads to local supersaturation [6]. For example, disproportionate accumulations of magmatic $CO_2$ in the bottom layers of Lakes Monoun and Nyos in Cameroon and their eruptions in 1984 and 1986, respectively, caused a sudden release of huge quantities of $CO_2$ into the atmosphere causing the loss of thousands of lives. In addition, Lake Kivu located in the Democratic Republic of the Congo is known to have about 2 and 10 trillion cubic feet of high concentrations of $CH_4$ and $CO_2$ gas, respectively, in its deep water [7–9]. The presence of high concentrations of such gases in the lake and the high population density in the nearby area pose a similar risk. The $CH_4$-water system is also relevant to nucleation of methane gas hydrates [10]. Nitrogen removal from wastewater streams is necessary to limit eutrophication and algal blooms [11].

When the pressure of a gas-saturated liquid is reduced at a constant temperature, the liquid becomes supersaturated with the gas and it may lead to bubble nucleation [12]. In our study on the influence of wettability on pressure-driven bubble nucleation of sparingly soluble gases $CH_4$ and $N_2$ in water, we observed that when a supersaturation is created with a higher (500 mbar) step-down pressure, the liberation of dissolved gas occurred through bubble nucleation on a hydrophobic solid surface. However, with a smaller (100 mbar) step-down pressure, the dissolved gas separation occurred only via diffusion of gas molecules from the liquid to the free gas through the gas–liquid interface, as there is not enough supersaturation to cause bubble nucleation even when the pressure is reduced to atmospheric pressure. In the case where bubble nucleation occurs with the available supersaturation, a certain waiting time is required for the gas bubbles to nucleate and grow in order to be observed using a microscope or naked eye, before applying further step-down pressure for generating higher supersaturation. Therefore, in order to know whether a given supersaturation is sufficient for bubble nucleation or not, it is important to estimate the supersaturation level at different step-down pressures [2]. In this study, we present a simple analytical method developed based on Fick's second law of diffusion and Henry's law to estimate the supersaturations generated in three gas–liquid systems ($CO_2$-water, $CH_4$-water, and $N_2$-water) with two step-down pressures (100 and 500 mbar). The three gas–liquid systems were chosen based on their importance in many natural and industrial applications. Moreover, we also had conducted experimental investigations to study the influence of wettability and step-down pressure on bubble nucleation of these gases in water [2,13].

Gas liberation from its dissolved state (in the liquid) to free gas state requires supersaturation. In our previous experimental study, we observed that bubble nucleation did not occur in a $CO_2$-supersaturated water in a hydrophilic container, even when the pressure was reduced in one step from the saturation pressure (6000 mbar) to atmospheric pressure (0 mbar) [13]; whereas in the case of hydrophobic container, the incipiation pressure for bubble nucleation and hence, the required supersaturation was observed to be influenced by the surface wettability of the container for all the three gases ($CO_2$, $CH_4$, and $N_2$) and the step-down pressure for the sparingly soluble gases ($CH_4$ and $N_2$). A certain waiting time is required at each pressure reduction step to ensure either enough time is given for the bubbles to nucleate and grow, or to conclude that the supersaturation is not sufficient to initiate the bubble nucleation. However, during the waiting time, some (or all) of the supersaturated gas may diffuse out before initiating the next step-down pressure. Therefore, it is important to know the actual supersaturation (which is the difference in the gas saturations at the initial and current pressures minus the gas diffused out until that point of time) in the system as a function of time during the waiting time in order to quantify the actual supersaturation required to initiate the bubble nucleation. Therefore, in this study, we investigated the influence of step-down pressure, and the waiting time between the pressure reduction steps on supersaturation.

## 2. Methodology

### 2.1. Derivation of Supersaturation Equation

The Fick's second law of diffusion given below describes the one-dimensional diffusion equation for binary mixtures of species A and B under transient condition [14,15]:

$$\frac{\partial C}{\partial t} = D_{AB}\frac{\partial^2 C}{\partial x^2} \tag{1}$$

where $D_{AB}$ is the diffusion coefficient of species A (gas) in species B (liquid), $C$ is the solute concentration, $t$ is the time, and $x$ is the position.

The derivation to solve Equation (1) is briefly demonstrated below; however, the detailed derivation of the differential equation can be found out in the supplementary section provided with this paper.

$y = f(x, t)$ defined by the following dimensionless characteristic equation:

$$y = \frac{x}{2\sqrt{D_{AB}t}} \tag{2}$$

$$\frac{\partial y}{\partial x} = \frac{1}{2\sqrt{D_{AB}t}}, \text{ and } \frac{\partial y}{\partial t} = \frac{-x}{4\sqrt{D_{AB}t^3}} \tag{3}$$

$$\frac{\partial C}{\partial t} = \frac{\partial C}{\partial y} \times \frac{\partial y}{\partial t} \tag{4}$$

$$\frac{\partial C}{\partial t} = \frac{-x}{4\sqrt{D_{AB}t^3}} \times \frac{\partial C}{\partial y} \tag{5}$$

$$\frac{\partial^2 C}{\partial x^2} = \frac{1}{4D_{AB}t} \times \frac{\partial^2 C}{\partial y^2} \tag{6}$$

From the Gaussian integral [16], we have

$$\int_0^\infty \exp\left(-y^2\right) dy = \frac{\sqrt{\pi}}{2} \tag{7}$$

The following error function and the boundary conditions were used to solve Fick's second law of diffusion [14]:

$$\text{erf}(y) = \frac{2}{\sqrt{\pi}} \int_0^y \exp\left(-y^2\right) dy \tag{8}$$

The initial and boundary conditions for the gas–liquid system shown in Figure 1 are given below:

$$C_A = C_{Aeq}, \text{ at x = 0 and for any t}$$

$$C_A = C_{AS}, \text{ at t = 0 and for all x}$$

where $C_{AS}$ is the saturated gas concentration in mol/L (after complete saturation), $C_A$ is the gas concentration after any time $t$, $C_{Aeq}$ is the gas concentration at the gas–liquid interface at equilibrium (with the current gas pressure) in mol/L. As mentioned earlier, $y = f(x, t)$; therefore, it is required to redefine $x$ and $y$ in Figure 1. In Figure 1, $x$ is an independent variable and it is a position in the liquid column, which is taken as 5 mm (total height of the liquid column) for estimating $C_A$ in Tables S2–S7 in the Supplementary Material. $y$ is a dependent variable that varies with position in the liquid column, and the waiting time after each pressure reduction step.
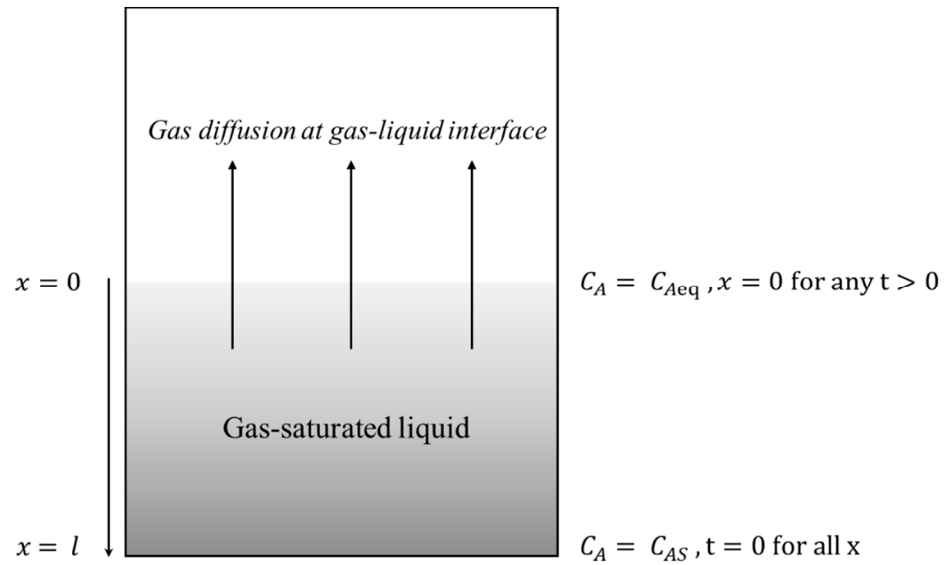
**Figure 1.** Schematic of bubble nucleation system used to study gas diffusion and saturation in a liquid.

Solving the differential equation mentioned in Equation (1) and applying the boundary conditions, we get the following:

$$\frac{C_A - C_{Aeq}}{C_{AS} - C_{Aeq}} = \text{erf}\left(\frac{x}{2\sqrt{D_{AB}t}}\right) \tag{9}$$

Equation (11) is the solution to the problem for time-dependent diffusion assuming a semi-infinite system.

For a given $x$, $D_{AB}$, and $t$, assuming $erf\left(\frac{x}{2\sqrt{D_{AB}t}}\right)$ as $\eta$, Equation (11) reduces to:

$$\frac{C_A - C_{Aeq}}{C_{AS} - C_{Aeq}} = \eta \tag{10}$$

The gas concentration $C_A$ in water at any given time can be determined using the following equation:

$$C_A = \eta \times \left(C_{AS} - C_{Aeq}\right) + C_{Aeq} \tag{11}$$

The equilibrium concentration of gas can be calculated by using Henry's law for low-pressure systems, which is given by:

$$C_{Aeq} = H_g P \tag{12}$$

where $H_g$ is the Henry's constant in mol/L/Pa and $P$ is the partial pressure of the gas [17].

The saturation time for carbon dioxide, methane and nitrogen gases in water can be estimated using the one-dimensional bounded diffusion equation, given below [14].

$$t_s = \frac{\tau l^2}{D_{AB}} \tag{13}$$

where $\tau$ is the dimensionless time, $D_{AB}$ is the diffusion coefficient in mm$^2$/s, $t_s$ is the diffusion time in seconds, and $l$ is the height of the water column in mm.

For this study, $\tau$ is taken as 4 to achieve 99.99% gas saturation, and the height of the water column $l$ is 5 mm. The saturation times required to achieve 99.99% saturation of carbon dioxide, methane, and nitrogen gases in water at the initial pressure, and constant ($\eta$) in the diffusion equation, are given in Table S1, in the Supplementary Material provided with this paper.

*2.2. Validation of Waiting Time after Each Pressure Reduction Step with 'Characteristic Time'*

Since semi-infinite diffusion solution is used to estimate the supersaturations in this study, therefore, it is important to validate whether the waiting time, i.e., 15 min, after each pressure reduction step satisfies the semi-infinite diffusion model. The characteristic diffusion time ($\tau_s$) for a gas to diffuse on a liquid column of height ($L$) and diffusion co-efficient ($D$) is given by [18]:

$$\tau_s = \frac{L^2}{4D_{AB}} \tag{14}$$

**3. Results and Discussion**

The saturation times for $CO_2$-water, $CH_4$-water, and $N_2$-water systems are 62,500 s (17.4 h), 56,497 s (15.7 h), and 52,910 s (14.7 h), respectively (ref.: Table S1). The details of characteristic time for all the three gases used in this study are given in Table 1.

**Table 1.** Characteristic time and $\eta$ calculations for carbon dioxide, methane, and nitrogen gases.

| Type of Gas | Diffusion Coefficient $D_{AB}$ (mm$^2$/s) | $\eta$ | Characteristic Time (s) |
|---|---|---|---|
| Carbon dioxide | 0.0016 [19] | 0.99678 | 3906 (1.09 h) |
| Methane | 0.00177 [20] | 0.99491 | 3531 (0.98 h) |
| Nitrogen | 0.00189 [21] | 0.99329 | 3307 (0.92 h) |

Since the time allowed after each pressure reduction step (15 min) is much less than the characteristic time, a semi-infinite diffusion model is applicable for this study. For 15 min waiting time, the supersaturation generated, remaining $CO_2$ concentration, and amount of $CO_2$ diffused out from water after each pressure reduction step, for 500 mbar and 100 mbar step-down pressures, are given in Tables S2 and S3, respectively. The corresponding calculations for two sparingly soluble gases, methane and nitrogen, are given in Tables S4–S7, respectively, in the Supplementary Material provided with this paper.

*Effect of Waiting Time on Supersaturation after Each Pressure Reduction Step*

To study the effect of waiting time on supersaturation in $CO_2$-water, $CH_4$-water, and $N_2$-water systems, the supersaturations were estimated at four different waiting times (15, 30, 45, and 60 min) for 100 and 500 mbar step-down pressures. Figure 2a–c represent the effect of waiting time on supersaturation in $CO_2$-water, $CH_4$-water, and $N_2$-water systems, respectively. From the figures, it can be observed that as the waiting time increases, at a given pressure, the supersaturation decreases as the consequence of the decrease in the actual concentration of gas in the liquid phase for both the step-down pressures. In addition, from the figures, it can be observed that the supersaturations generated in the $CO_2$-water system are an order of magnitude higher compared to those of $CH_4$-water and $N_2$-water systems. This is due to fact that the Henry's law constants for $CO_2$ ($3.36 \times 10^{-7}$ mol/L/Pa) are approximately 25- and 56-times higher than $CH_4$ ($1.36 \times 10^{-8}$ mol/L/Pa) and $N_2$ ($6.02 \times 10^{-9}$ mol/L/Pa) gases, respectively. Hence, the dissolved gas concentration of $CO_2$ in water is much higher compared to the sparingly soluble gases $CH_4$ and $N_2$ in the pressure range considered in this work. The supersaturation calculations for both 100 and 500 mbar step-down pressures for the aforementioned gas–liquid systems are tabulated in Tables S8–S13 (100 mbar), in the Supplementary Material provided with this paper. It is important to note that the mathematical model presented in this study is not applicable at the onset nucleation pressure for bubble nucleation or any pressure below that.
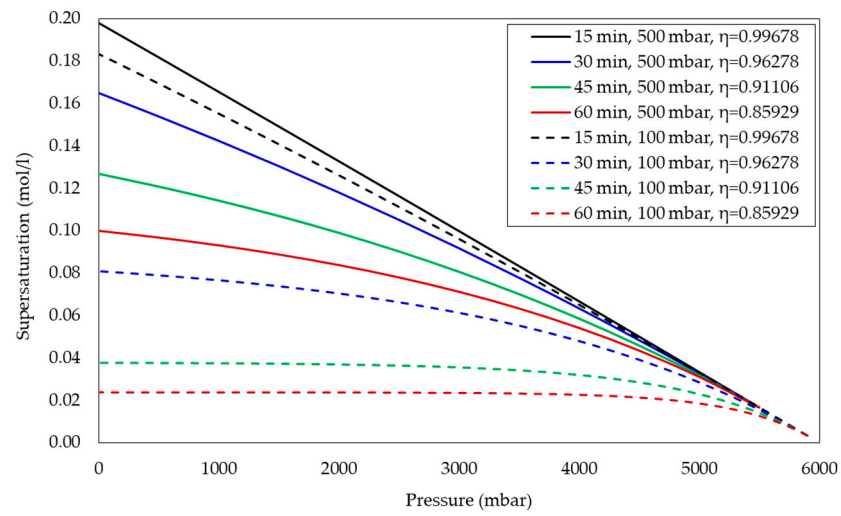
(**a**)
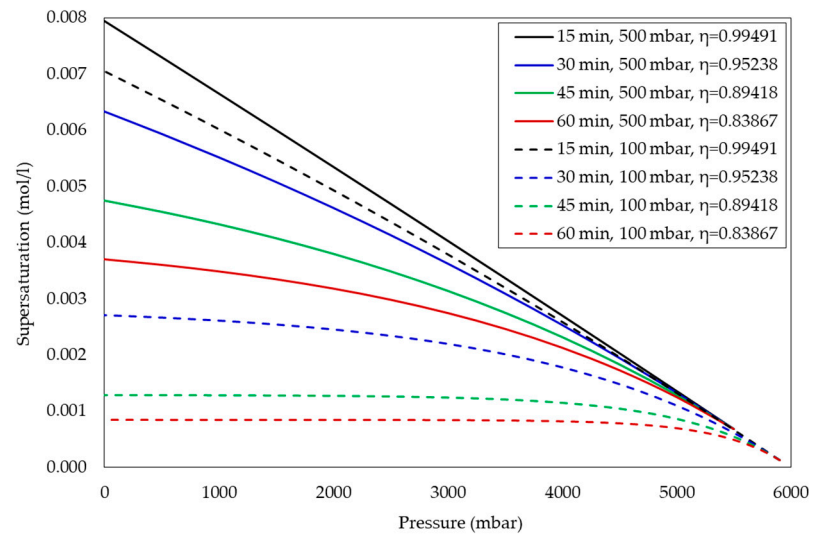


(**b**)
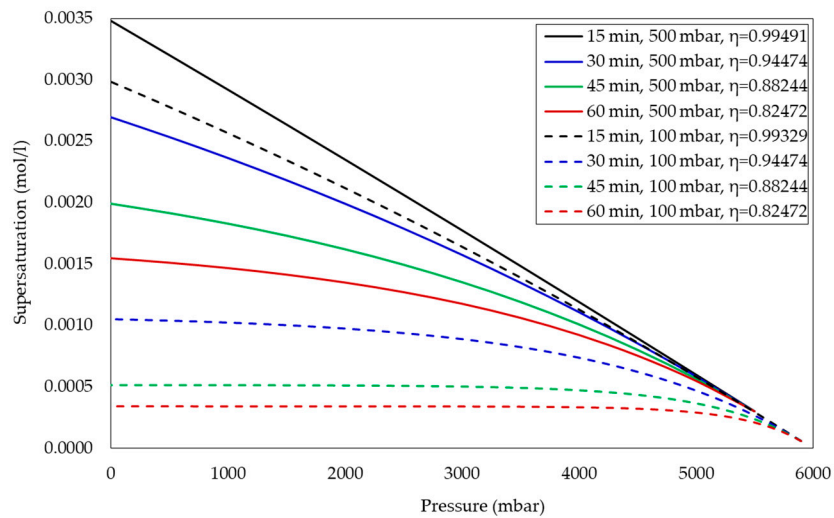


(**c**)

**Figure 2.** Effect of waiting time on supersaturation in the (**a**) $CO_2$-water system; (**b**) $CH_4$-water system; (**c**) $N_2$-water system.

## 4. Conclusions

In this study, an analytical method to estimate the supersaturations generated with two different step-down pressures from an initial saturation pressure is presented. The proposed method estimates supersaturation by integrating Fick's second law of diffusion and Henry's law. The semi-infinite diffusion model was used to estimate the supersaturations generated in gas–liquid systems. Since a certain waiting time is required to observe bubble nucleation after creating a supersaturation, the waiting times used in this study were validated with the characteristic times. In addition, we analyzed the effect of waiting time on supersaturation and found that, at a given pressure, as the waiting increases, the supersaturation decreases due to the decrease in the actual concentration of gas in the liquid. The supersaturation is higher with the higher step-down pressure. The dissolved gas concentration of $CO_2$ in water is much higher compared to $CH_4$ and $N_2$, in the pressure range of this work; therefore, the supersaturations generated in the $CO_2$-water system are correspondingly higher compared to $CH_4$-water and $N_2$-water systems.

**Supplementary Materials:** The following are available online at https://www.mdpi.com/article/10.3390/eng3010010/s1, Table S1: Saturation time calculations for carbon dioxide, methane, and nitrogen gases in water, Table S2: Supersaturation and diffusion calculations for carbon dioxide with 500 mbar (50,000 Pa) step-down pressure, Table S3: Supersaturation and diffusion calculations for carbon dioxide with 100 mbar (10,000 Pa) step-down pressure, Table S4: Supersaturation and diffusion calculations for methane gas with 500 mbar (50,000 Pa) step-down pressure, Table S5: Supersaturation and diffusion calculations for methane with 100 mbar (10,000 Pa) step-down pressure, Table S6: Supersaturation and diffusion calculations for nitrogen with 500 mbar (50,000 Pa) step-down pressure, Table S7: Supersaturation and diffusion calculations for nitrogen with 100 mbar (10,000 Pa) step-down pressure, Table S8: $CO_2$ supersaturations generated in water with 500 mbar (50,000 Pa) step-down pressure and different waiting times (saturation pressure: 6000 mbar or 600,000 Pa), Table S9: $CO_2$ supersaturations generated in water with 100 mbar (10,000 Pa) step-down pressure and different waiting times (saturation pressure: 6000 mbar or 600,000 Pa), Table S10: $CH_4$ supersaturations generated in water with 500 mbar (50,000 Pa) step-down pressure and different waiting times (saturation pressure: 6000 mbar or 600,000 Pa), Table S11: $CH_4$ supersaturations generated in water with 100 mbar (10,000 Pa) step-down pressure and different waiting times (saturation pressure: 6000 mbar or 600,000 Pa), Table S12: $N_2$ supersaturations generated in water with 500 mbar (50,000 Pa) step-down pressure and different waiting times (saturation pressure: 6000 mbar or 600,000 Pa), Table S13: $N_2$ supersaturations generated in water with 100 mbar (10,000 Pa) step-down pressure and different waiting times (saturation pressure: 6000 mbar or 600,000 Pa).

**Author Contributions:** Conceptualization, P.K.B. and S.P.; methodology, S.P.; validation, S.P. and P.K.B.; formal analysis, S.P. and P.K.B.; investigation, S.P.; resources, P.K.B.; data curation, S.P.; writing—original draft preparation, S.P.; writing—review and editing, P.K.B.; visualization, P.K.B.; supervision, P.K.B.; project administration, P.K.B.; funding acquisition, P.K.B. All authors have read and agreed to the published version of the manuscript.

**Institutional Review Board Statement:** Not relevant.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** In this research, there was no data availability statement.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Bikkina, P.K.; Pradhan, S. A potential solution for boiling crisis. In Proceedings of the 18th International Topical Meeting on Nuclear Reactor Thermal Hydraulics (NURETH 18), Portland, Oregon, OR, USA, 18–23 August 2019; American Nuclear Society: Portland, OR, USA, 2019; pp. 1383–1396.
2. Pradhan, S.; Qader, R.J.; Sedai, B.R.; Bikkina, P.K. Influence of wettability on pressure-driven bubble nucleation: A potential method for dissolved gas separation. *Sep. Purif. Technol.* **2019**, *217*, 31–39. [CrossRef]

3. Baillon, F.; Espitalier, F.; Cogne, C.; Peczalski, R.; Louisnard, O. *Crystallization and Freezing Processes Assisted by power Ultrasound, in Power Ultrasonics*; Elsevier: Amsterdam, The Netherlands, 2015; pp. 845–874.

4. Zuo, L.; Krevor, S.; Falta, R.W.; Benson, S.M. An experimental study of $CO_2$ exsolution and relative permeability measurements during $CO_2$ saturated water depressurization. *Transp. Porous Media* **2012**, *91*, 459–478. [CrossRef]

5. Liger-Belair, G.; Parmentier, M.; Jeandet, P. Modeling the kinetics of bubble nucleation in champagne and carbonated beverages. *J. Phys. Chem. B* **2006**, *110*, 21145–21151. [CrossRef] [PubMed]

6. Schmid, M.; Lorke, A.; Wuest, A.; Halbwachs, M.; Tanyileke, G. Development and sensitivity analysis of a model for assessing stratification and safety of lake nyos during artificial degassing. *Ocean Dyn.* **2003**, *53*, 288–301. [CrossRef]

7. Kusakabe, M. Lakes nyos and monoun gas disasters (Cameroon)—Limnic eruptions caused by excessive accumulation of magmatic $CO_2$ in crater lakes. *Geochem. Monogr. Ser.* **2017**, *1*, 1–50. [CrossRef]

8. Vaselli, O.; Tedesco, D.; Cuoco, E.; Tassi, F. Are limnic eruptions in the $CO_2$–$CH_4$-rich gas reservoir of lake kivu (Democratic Republic of the Congo and Rwanda) possible? Insights from physico-chemical and isotopic data. In *Volcanic Lakes*; Springer: New York, NY, USA, 2015; pp. 489–505.

9. Jones, N. How dangerous is Africa's explosive lake kivu? *Nature* **2021**, *597*, 466–469. [CrossRef] [PubMed]

10. Hawtin, R.W.; Quigley, D.; Rodger, P.M. Gas hydrate nucleation and cage formation at a water/methane interface. *Phys. Chem. Chem. Phys.* **2008**, *10*, 4853–4864. [CrossRef] [PubMed]

11. Sharp, R.; Khunjar, W.; Daly, D.; Perez-Terrero, J.; Chandran, K.; Niemiec, A.; Pace, G. Nitrogen removal from water resource recovery facilities using partial nitrification, denitratation-anaerobic ammonia oxidation (PANDA). *Sci. Total Environ.* **2020**, *724*, 138283. [CrossRef] [PubMed]

12. Bergeron, V.; Walstra, P. *Foams, in Fundamentals of Interface and Colloid Science*; Elsevier: Amsterdam, The Netherlands, 2005; pp. 7.1–7.38.

13. Pradhan, S. Influence of Wettability on Dissolved Gas Separation, Nucleate Boiling, and Enhanced Oil Recovery, in School of Chemical Engineering. Ph.D. Dissertation, Oklahoma State University, Stillwater, OK, USA, 2021.

14. Bird, R.B.; Stewart, W.E.; Lightfoot, E.N. *Transport Phenomena 2002*; JohnWiley & Sons: New York, NY, USA, 2004.

15. Crank, J. *The Mathematics of Diffusion*; Oxford University Press: Oxford, UK, 1979.

16. Conrad, K. *The Gaussian Integral*; University of Connecticut: Storrs, CT, USA, 2016; pp. 1–2.

17. Smith, J.M. *Introduction to Chemical Engineering Thermodynamics*; ACS Publications: Washington, DC, USA, 1950.

18. Berthier, J.; Silberzan, P. *Microfluidics for Biotechnology*; Artech House: Norwood, MA, USA, 2010.

19. Himmelblau, D. Diffusion of dissolved gases in liquids. *Chem. Rev.* **1964**, *64*, 527–550. [CrossRef]

20. Witherspoon, P.; Saraf, D. Diffusion of methane, ethane, propane, and n-butane in water from 25 to 43. *J. Phys. Chem.* **1965**, *69*, 3752–3755. [CrossRef]

21. Ferrell, R.T.; Himmelblau, D.M. Diffusion coefficients of nitrogen and oxygen in water. *J. Chem. Eng. Data* **1967**, *12*, 111–115. [CrossRef]

*Article*

# Optimal Hybridization of Conventional ICE Vehicles

Zhemin Hu, Ramin Tafazzoli Mehrjardi [ID], Lin Lai and Mehrdad Ehsani *

Department of Electrical Engineering, Texas A&M University, College Station, TX 77840, USA;
zhemin37@tamu.edu (Z.H.); ramin.tafazzoli@tamu.edu (R.T.M.); lailinabc@gmail.com (L.L.)
* Correspondence: ehsani@ece.tamu.edu; Tel.: +1-979-845-7582

**Abstract:** Most commercially available hybrid electric vehicle (HEV) drivetrains are made of small internal combustion (IC) engines and large electric drives to improve fuel economy. They usually have higher cost than the conventional IC-engine-based vehicles because of the high costs of the electric drives. This paper proposes a hybridized powertrain composed of the original full-size engine of the vehicle and a universally optimum size parallel electric drive. The dynamic programming (DP) algorithm was used to obtain the sensitivity of the maximum miles per gallon (MPG) values versus the power rating of the electric drive. This sensitivity was then analyzed to determine the optimal window of the electric drive power ratings. This was proven to be universal for all passenger cars of various masses and engine powers. The fuel economy and vehicle performance of this HEV was compared with those of the 2019 Toyota Corolla, a conventional IC-engine-based vehicle, and the 2019 Toyota Prius, a commercially available HEV. The results showed that the proposed universally optimized HEV powertrain achieved better fuel economy and vehicle performance than both the original ICE and HEV vehicles, at low additional vehicle cost.

**Keywords:** fuel optimization; low cost HEV; optimum hybridization

## 1. Introduction

Conventional vehicles, powered by internal combustion (IC) engines, are a major source of carbon dioxide emission, causing global warming [1]. They also pollute the air with significant emissions of toxic gases such as nitrogen oxides (NOx), carbon monoxide (CO), and unburned hydrocarbons [1]. Another drawback of the IC engine is its low efficiency. Its typical efficiency is around 20% which is much lower than an electric motor whose efficiency is around 85% [2]. This low efficiency leads to the poor fuel economy of IC engine based vehicles, especially in urban driving cycles.

It is now known that electric vehicles (EV) have certain advantages over IC engine based vehicles, such as higher efficiency, no tailpipe emissions, smoother operation and less noise [1]. However, they also have several disadvantages, such as short travel range, long battery recharging time, and high comparative costs. For example, the 2020 Chevrolet Bolt, a commercially available EV, can have an additional travel range of only 90 miles after its battery is recharged for 30 min at a Level 3 charging station [3]. This time is much longer than the time needed for filling a gasoline tank [2]. In addition, the travel range of EV can be even shorter under lower ambient temperatures. For instance, it can be shown that the travel distance of the Mitsubishi i-MiEV, a commercial EV, decreases at a rate of 2.5 km per 1 °C temperature drop in the ambient temperature range of +20 °C to −15 °C [4]. Furthermore, the electric drive and battery in the EV have a significantly higher total cost than an IC engine and its gasoline tank, leading to a higher cost for the EV.

To combine the advantages of the IC engine based vehicle and the EV, the hybrid electric vehicle (HEV) typically contains an IC engine to deliver the average tractive power and an electric motor to supply the peak power. In this way, the power rating of the IC engine can be reduced to less than half of that in an equivalent conventional vehicle with the same mass and dynamic specifications [1,5]. However, the HEV must use this

small IC engine alone if the state of charge (SOC) of the battery drops to its lower limit, negatively affecting the vehicle performance. In addition, due to its large electric drive and battery bank, the cost of the HEV is usually much higher than its equivalent conventional vehicle. Although the fuel economy is improved in such HEV, the cost savings on the fuel consumption cannot offset its high initial cost in a reasonable time period. For instance, it can be shown that the Toyota Prius, a conventional HEV, has a cost payback mileage of about 150,600 miles which is close to the vehicle's lifespan [2]. Furthermore, control strategies for HEV powertrains usually have to be adjusted according to different driving conditions [6,7]. For these reasons, the market share of HEVs in the US has remained small over the past two decades. To further improve the fuel economy, plug-in hybrid electric vehicles (PHEV) were introduced. These use even larger battery banks with energy capacity of over 4 kWh and large electric drives to supply adequate power for pure electric drive [8]. In the PHEV, more of the gasoline energy is replaced with the battery energy, obtained from the power grid [9]. Thus, the fuel economy of the PHEV can be increased to 95 miles per gallon (MPG) while in its pure electric range of about 37 miles [10]. However, the MPG value of the PHEV will drop to the same level as that of Toyota Prius when traveling beyond this pure electric range. In the meantime, the PHEV has two times larger battery bank, which significantly raises its cost, with a payback mileage as long as 117,600 miles even with the highest 95 MPG [2]. Reference [11] shows that the PHEV can only bring a fuel savings of $10 to $120 annually, which cannot offset its high cost in a short term.

To overcome the drawbacks of the conventional HEVs and PHEVs, this paper proposes a new approach to HEV drivetrain design. The new design approach keeps the full-size engine as the primary vehicle power plant, with its associated longer travel range, better performance and much higher energy density of the fossil, compared to batteries. The full-size engine is defined as an engine that has the capability of propelling the vehicle alone during typical driving conditions. Then, a relatively small electric drive is added in parallel with this full-size engine to optimally improve the fuel economy and vehicle performance, beyond the IC engine alone. Here, the additional cost and mass of the electric drive will be offset by the savings in fuel economy improvement. For this, the optimal window of the electric drive power ratings, in which the fuel economy is maximized, has to be determined. This paper presents the results of this optimization using the dynamic programming (DP) algorithm to analyze the sensitivity of the fuel consumption to the electric drive power ratings. The sensitivity analysis was applied to hybridization of three conventional vehicles with different IC engine power ratings and different vehicle masses in order to demonstrate that the optimal window of the electric drive power ratings is applicable to all common types of passenger cars. This means all sizes of conventional IC engine vehicles can be optimally hybridized by a specific electric drive power rating in the optimal window. This universally sized electric drive package can potentially reduce the production costs of the electric drives in all full-size engine HEVs. To illustrate the advantages of these HEVs over the conventional IC engine vehicles and commercially available HEVs, their vehicle performances, fuel economies and the payback mileages were compared.

## 2. Basic Structure of a Full-Size Engine HEV

The full-size engine has been used in a commercially available vehicle propelled solely by the IC engine (hereinafter called an IC engine vehicle). It has been commercially proven to have the capability of propelling the vehicle alone during all common driving conditions. This paper will take three typical IC engine vehicles, Toyota Corolla 2019, Toyota Camry 2016, and Chrysler 300 2016, as examples to demonstrate the transformation of IC engine vehicles into full-size engine HEVs by adding a small electric drive in parallel with the full-size engine of the IC engine vehicles. The three IC engine vehicles have different curb masses and IC engine power ratings, as shown in Table 1, which can represent compact, medium-sized and large-sized passenger cars on the market.

The drivetrain configuration of a torque-coupling parallel HEV with a full-size engine is shown in Figure 1. It contains two power plants, an IC engine and an electric motor. The engine is connected to a torque-coupler via a multi-gear transmission, and the motor is connected to the same torque-coupler via a single-gear transmission. The torque-coupler enables the torque on the driving shaft to be equal to the sum of the torques on the engine shaft and the motor shaft, whereas the speed of the driving shaft is proportional to the speeds of both the engine shaft and the motor shaft. A final drive is used to distribute the driving shaft torque to each driving wheel through a differential. The load power on the driving wheels, $P_l$, has a relation with powers of engine and motor expressed as

$$P_1 = P_{e\_w} + P_{m\_w}, \tag{1}$$

where $P_{e\_w}$ and $P_{m\_w}$ are the engine power and the motor power transmitted onto the driving wheels, respectively.
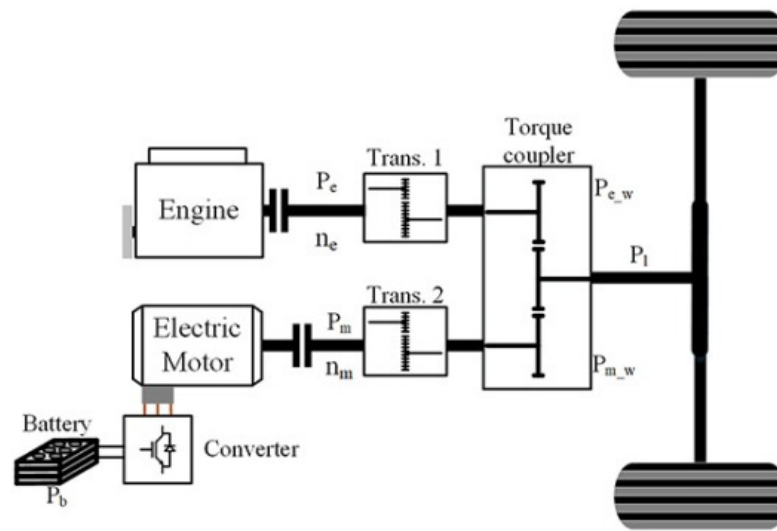


**Figure 1.** Drivetrain configuration of a torque-coupling parallel HEV. Adapted from Ref. [2].

The parameters of the three IC engine vehicles and of the designed motors used in this study are listed in Tables 1 and 2, where $\delta_1$ and $\delta_2$ are used to calculate the rotational inertia factor which converts the inertia of all rotating components to the equivalent mass and can be expressed as

$$\delta = 1 + \delta_1 + \delta_2 i_0^2 \left( i_{ge}^2 + i_{gm}^2 \right). \tag{2}$$

where, $i_{ge}$ and $i_{gm}$ are the gear ratios of the transmissions connected to the engine and the motor, respectively, and $i_0$ is the gear ratio of the final drive. For an eight-gear transmission, $i_{ge}$ can be any one of the eight gear ratios of the engine transmission listed in Table 1, depending on which gear is used at a time. To demonstrate the universality associated with the power rating of the electric drive (discussed later in Section 4), the motor parameters in Table 2 are designed to be the same for all the three IC engine vehicles, except $i_{gm}$ which is designed to keep the product of $i_{gm}$ and $i_0$ the same for all the three IC engine vehicles.

**Table 1.** Vehicle parameters of 3 typical IC engine vehicles [1,2,12–14].

| Vehicle | Toyota Corolla 2019 | Toyota Camry 2016 | Chrysler 300 2016 |
|---|---|---|---|
| Curb mass ($M$) | 1285 kg | 1497 kg | 1828 kg |
| IC engine power rating ($P_{e\text{rated}}$) | 98 kW | 133 kW | 218 kW |
| Number of transmission gears | 6 | 6 | 8 |
| 1st gear ratio ($i_{ge1}$) | 3.54 | 3.3 | 4.71 |
| 2nd gear ratio ($i_{ge2}$) | 1.91 | 1.9 | 3.14 |
| 3rd gear ratio ($i_{ge3}$) | 1.31 | 1.42 | 2.1 |
| 4th gear ratio ($i_{ge4}$) | 0.97 | 1 | 1.67 |
| 5th gear ratio ($i_{ge5}$) | 0.71 | 0.713 | 1.29 |
| 6th gear ratio ($i_{ge6}$) | 0.62 | 0.608 | 1 |
| 7th gear ratio ($i_{ge7}$) | - | - | 0.84 |
| 8th gear ratio ($i_{ge8}$) | - | - | 0.67 |
| Gear ratio of final drive ($i_0$) | 4.21 | 3.634 | 2.62 |
| Gravitational acceleration ($g$) | 9.8 | 9.8 | 9.8 |
| Rolling resistance coefficient ($f_r$) | 0.01 | 0.01 | 0.01 |
| Air density ($\rho_a$) | 1.205 kg/m$^3$ | 1.205 kg/m$^3$ | 1.205 kg/m$^3$ |
| Frontal area ($A_f$) | 2.2 m$^2$ | 2.3 m$^2$ | 2.4 m$^2$ |
| Aerodynamic drag coefficient ($C_D$) | 0.3 | 0.3 | 0.3 |
| Wheel radius ($r_d$) | 0.31 m | 0.33 m | 0.36 m |
| Wheel mass factor ($\delta_1$) | 0.04 | 0.04 | 0.04 |
| Power plant rotor mass factor ($\delta_2$) | 0.0025 | 0.0025 | 0.0025 |
| Transmission efficiency from engine to driving wheels ($\eta_{t,e}$) | 0.9 | 0.9 | 0.9 |

**Table 2.** Parameters of designed electric motors for hybridization of three typical IC engine vehicles.

| Vehicle | Toyota Corolla 2019 | Toyota Camry 2016 | Chrysler 300 2016 |
|---|---|---|---|
| Maximum speed of motor ($n_{m\max}$) | 6000 rpm | 6000 rpm | 6000 rpm |
| Base speed of motor ($n_{mb}$) | 1500 rpm | 1500 rpm | 1500 rpm |
| Gear ratio of motor transmission ($i_{gm}$) | 1.48 | 1.72 | 2.39 |
| Power efficiency of motor ($\eta_m$) | 0.88 | 0.88 | 0.88 |
| Transmission efficiency from motor to driving wheels ($\eta_{t,m}$) | 0.95 | 0.95 | 0.95 |

The mechanical resistances opposing the vehicle movement consist of rolling resistance, aerodynamic drag, and grading resistance. Taking the vehicle acceleration into consideration, we can express the load power $P_l$ as

$$P_l = V \left( Mgf_r + \frac{1}{2}\rho_a C_D A_f V^2 + Mgi \right) + M\delta V \frac{dV}{dt} \tag{3}$$

where, $V$ is the vehicle speed, and $i$ is the road grade.

The electric drive consists of the motor, power electronic converters, and a battery pack. The maximum torque of the motor is constant when the motor speed is below the base speed, whereas the maximum power of the motor is constant when the motor speed is above the base speed. That is, the motor has nearly ideal torque-speed and power-speed profiles. Therefore, only a single-gear transmission is needed for the motor. The widely used battery packs for HEVs and EVs are nickel metal hydride (NiMH) batteries and Li-ion batteries [15]. Research in [16–19] shows that the energy density and efficiency of Li-ion batteries are significantly higher than those of NiMH batteries. Hence, the Li-ion battery pack is used here to deliver electrical power to the motor. The total mass of the electric drive should be considered during the optimization process of the electric drive power rating because it has an adverse effect on the fuel economy of the HEV. We assume that motors with power ratings from 3 kW to 40 kW have a constant mass per kW factor which is 2.5 in this paper. We also assume that the specific power of the battery pack is 1.28 kW/kg, and

the energy/power ratio of the battery pack is 0.03 h. Based on these assumptions, the masses of electric drives with various power ratings from 3 kW to 45 kW are estimated and listed in Table 3.

**Table 3.** Mass estimation of the electric drives. Adapted from Ref. [2].

| Electric Drive Power Rating (kW) | Mass (kg) |
|:---:|:---:|
| 45 | 147 |
| 40 | 130 |
| 35 | 114.5 |
| 30 | 99 |
| 25 | 82.5 |
| 20 | 66.3 |
| 15 | 49.8 |
| 10 | 33.3 |
| 5 | 16.8 |
| 3 | 10.6 |

### 3. Dynamic Programming Algorithm in HEV

The dynamic programming (DP) algorithm optimizes a cost function and finds the optimal sequences of control variables in the whole process by making control decisions, step-by-step, in time domain [20,21]. The cost function in HEV drivetrain control is the total fuel consumption of the engine during a certain driving cycle.

Research in [5,22] uses the DP algorithm whose control variable is the output power or torque of the engine. This algorithm will either lead to the depletion of the battery or select the pure electric mode for only a short distance and switch to the engine-alone mode after the battery SOC drops to its lower bound, which is not an optimal control for an HEV. Although this situation can be avoided by adding electric energy usage penalties (i.e., electric penalties) to the cost function, the weights of the gasoline usage cost and the electric penalties will highly affect the control sequences, leading to the non-optimality of the obtained solution.

To overcome the disadvantages of algorithms applied in previous research, the DP algorithm used in this paper selects the battery SOC as well as the number of gears of the engine transmission as the control variables. This selection can avoid battery depletion without adding the electric penalties, and thus the optimality of the obtained solution is guaranteed. In this algorithm, the output powers of the engine and the motor are calculated according to the battery SOCs, SOC($k$) and SOC($k$ + 1), at two adjacent time steps, $k$ and $k$ + 1, in a driving cycle. If SOC($k$) > SOC($k$ + 1), the battery is discharged, and the motor power on the driving wheels is expressed as

$$P_{m\_w} = \frac{(\text{SOC}(k) - \text{SOC}(k+1))Q_n \eta_{b\_w}}{\Delta t} \tag{4}$$

where $Q_n$ is the battery energy capacity, $\eta_{b\_w}$ is the power efficiency from the battery to the driving wheels, and $\Delta t$ is the time interval between time steps $k$ and $k$ + 1. Otherwise, the battery is charged if SOC($k$) < SOC($k$ + 1), and the motor power on the driving wheels is expressed as

$$P_{m\_w} = \frac{(\text{SOC}(k) - \text{SOC}(k+1))Q_n}{\eta_{w\_b} \Delta t} \tag{5}$$

where $\eta_{w\_b}$ is the power efficiency from the driving wheels to the battery. If SOC($k$) = SOC($k$ + 1), there is no energy delivered from or to the battery, and thus $P_{m\_w}$ = 0. The battery energy capacity, $Q_n$, is related to the output power rating of the electric drive, $P_{m\_rated}$, which is expressed as

$$Q_n = \frac{P_{m\_rated} R_{e/p}}{\eta_m} \tag{6}$$

where, $R_{e/p}$ is the energy/power ratio of the battery and is assumed to be 0.03 h in this paper.

Using (1), the engine power on the driving wheels, $P_{e\_w}$, can also be calculated, with the load power, $P_l$, obtained from the driving cycle as expressed in (3). The output power, $P_e$, and speed, $n_e$, of the engine on its own shaft are calculated as

$$P_e = \frac{P_{e\_w}}{\eta_{t,e}} \qquad (7)$$

$$n_e = \frac{30 i_{ge} i_0 V}{\pi r_d} \qquad (8)$$

where, $i_{ge}$ is determined by the gear number of the engine transmission, as shown in Table 1.

The fuel consumption, $F_c$, in the time interval $\Delta t$ is calculated as

$$F_c = \frac{P_e g_e \Delta t}{\rho_f} \qquad (9)$$

where $\rho_f$ is the mass density of gasoline, which is assumed to be 0.75 kg/L in this paper, and $g_e$ is the brake specific fuel consumption (BSFC) of the engine, which is determined by $P_e$ and $n_e$ in the engine fuel consumption map.

Suppose the driving cycle has $N + 1$ discrete time steps from time step 0 to time step $N$. Because the engine delivers the average tractive power and the motor only supplies the peak power, all the net energy used during the entire driving cycle is only delivered by the engine. Therefore, the battery SOC must be the same value, $SOC_{init}$, at both time steps 0 and $N$. At each of other time steps, the battery SOC can be any value within its full usable range, $SOC_{min} \leq SOC \leq SOC_{max}$. Suppose the range is discretized into $i_m$ values, that is,

$$SOC_{min} = SOC_1 < SOC_2 < \ldots < SOC_{i_m} = SOC_{max} \qquad (10)$$

Suppose the value of the battery SOC is $SOC_i$ at time step $k$ and $SOC_r$ at time step $k + 1$, where $SOC_i$ and $SOC_r$ are within the SOC range, $1 \leq i, r \leq i_m$, and $1 \leq k \leq N-2$. Suppose the $j$th gear of the engine transmission, $G_j$, is selected at time step $k$, where $1 \leq j \leq j_m$, and $j_m = 6$ for a six-gear transmission. For the time interval $\Delta t_k$ from time step $k$ to time step $k + 1$, we have a recursive relation of fuel consumption expressed as

$$J^*_{SOC_r}(k+1) = \min_{1 \leq i \leq i_m} \left\{ \min_{1 \leq j \leq j_m} \left[ F_c(SOC_{i,r}(k), G_j(k)) \right] + J^*_{SOC_i}(k) \right\} \qquad (11)$$

where $J^*_{SOC_i}(k)$ and $J^*_{SOC_r}(k+1)$ are the minimum total fuel consumptions from time step 0 to time step $k$ with $SOC_i$ at time step $k$ and from time step 0 to time step $k + 1$ with $SOC_r$ at time step $k + 1$, respectively, and $F_c(SOC_{i,r}(k), G_j(k))$ is the fuel consumption in the time interval $\Delta t_k$ with $SOC_i$ at time step $k$, $SOC_r$ at time step $k + 1$, and the $j$th gear selected at time step $k$. Using (11), we can obtain the minimum total fuel consumption from time step 0 to time step $k + 1$ for every $SOC_r$ within $SOC_{min} \leq SOC_r \leq SOC_{max}$ at time step $k + 1$ and every $k$ within $1 \leq k \leq N-2$.

In the initial time interval $\Delta t_0$ from time step 0 to time step 1 and the terminal time interval $\Delta t_{N-1}$ from time step $N-1$ to time step $N$, the recursive relations of fuel consumption are expressed as

$$J^*_{SOC_r}(1) = \min_{1 \leq j \leq j_m} \left[ F_c(SOC_{init,r}(0), G_j(0)) \right] + J^*_{SOC_{init}}(0) \qquad (12)$$

$$J^*_{SOC_{init}}(N) = \min_{1 \leq i \leq i_m} \left\{ \min_{1 \leq j \leq j_m} \left[ F_c(SOC_{i,init}(N - 1), G_j(N - 1)) \right] + J^*_{SOC_i}(N - 1) \right\} \qquad (13)$$

where, $F_c(SOC_{init,r}(0), G_j(0))$ is the fuel consumption in the time interval $\Delta t_0$ with $SOC_r$ at time step 1 and the $j$th gear selected at time step 0, $F_c(SOC_{i,init}(N - 1), G_j(N - 1))$ is the fuel consumption in the time interval $\Delta t_{N-1}$ with $SOC_i$ at time step $N-1$ and the $j$th gear

selected at time step $N-1$, and $J^*_{SOC_{init}}(N)$ is the minimum total fuel consumption during the entire driving cycle. The initial fuel consumption at time 0, $J^*_{SOC_{init}}(0)$, is assumed to be zero. Using (12), the minimum total fuel consumption from time step 0 to time step 1, $J^*_{SOC_r}(1)$, is obtained for every $SOC_r$ within $SOC_{min} \leq SOC_r \leq SOC_{max}$ at time step 1, whereas $J^*_{SOC_{init}}(N)$ obtained in (13) is a unique value solely for the single initial and terminal SOC value, $SOC_{init}$, which is also within the SOC usable range.

Using (1)–(3) and (4)–(9), we can obtain $P_{e\_w}$, $P_{m\_w}$ and $F_c$ in every time interval $\Delta t_k$ for every set of $SOC_i(k)$, $SOC_r(k+1)$ and $G_j(k)$ values. However, not all of the discrete values within the SOC usable range can be assigned to $SOC_i$ and $SOC_r$ in practical conditions because $P_{e\_w}$ and $P_{m\_w}$ must satisfy the following constraint conditions expressed as

$$P_{m\_min}(n_m(k)) \leq P_{m\_w}(k) \leq P_{m\_max}(n_m(k)) \tag{14}$$

$$P_{e\_min}(n_e(k)) \leq P_{e\_w}(k) \leq P_{e\_max}(n_e(k)). \tag{15}$$

where the upper and lower bounds of $P_{e\_w}$ and $P_{m\_w}$ are functions of the engine speed, $n_e(k)$, and the motor speed, $n_m(k)$, respectively. The functions are determined by the power ratings and the power-speed characteristics of the engine and the motor. Every set of $SOC_i(k)$, $SOC_r(k+1)$ and $G_j(k)$ values which cause $P_{e\_w}$ or $P_{m\_w}$ to go beyond the corresponding constraint condition in (14) or (15) should be excluded from the optimization processes expressed in (11)–(13).

Besides the final minimum value of the cost function, $J^*_{SOC_{init}}(N)$, we can also obtain the optimal values of $SOC_i(k)$ and $G_j(k)$ that lead to $J^*_{SOC_{init}}(N)$ with $k = N-1$ from (13). Then, with $k = N-2$, we let $SOC_r(k+1)$ be equal to the optimal value of $SOC_i(N-1)$ and use (11) to find the optimal values of $SOC_i(k)$ and $G_j(k)$. Repeat this with $k = N-3, N-4, \ldots,$ 1. Finally, substituting the optimal value of $SOC_r(1) = SOC_i(1)$ into (14), we obtain the optimal value of $G_j(0)$. Letting $G_j(N) = G_j(0)$ and $SOC_i(N) = SOC_i(0) = SOC_{init}$, we obtain the optimal sequences of the control variables, $SOC_i(k)$ and $G_j(k)$ with $k = 0, 1, \ldots, N$. Using (1)–(3) and (4)–(8), we can also obtain the optimal sequences of $P_{e\_w}$, $P_{m\_w}$, $P_e$ and $n_e$ from those of the control variables to see how the total tractive power is distributed among the engine and the motor and to obtain the optimal engine operating points analyzed in the next section.

## 4. Optimization of Electric Drive Power Rating

To optimize the power rating of the electric drive, $P_{m\_rated}$, in a full-size engine HEV, the DP algorithm discussed in the previous section is used to calculate the minimum total fuel consumption, which is equivalent to the maximum MPG, during four typical driving cycles (FTP75 Urban, FTP75 Highway, LA92, and SC03) under various values of $P_{m\_rated}$. Then, the sensitivity of the maximum MPG to $P_{m\_rated}$ is analyzed.

Research in [23–27] has proposed an optimization methodology which fixes either the total power rating of the power plants or the acceleration performance of the vehicle and obtains the trend of maximum MPG by increasing the electric drive power rating. Under this methodology, the engine power rating has to be reduced when the electric drive power rating increases. This reduction will potentially worsen the vehicle performances because the smaller engine is the only available power plant to propel the vehicle if the battery SOC drops to its lower bound.

In this section, to ensure that the vehicle performance is not degraded in the optimization process, the power rating of the full-size engine in each of the three IC engine vehicles is kept constant. The performance and fuel economy are further improved with the addition of the electric drive. Therefore, the optimal window of $P_{m\_rated}$, obtained from the sensitivity analysis of the maximum MPG trend with the increasing $P_{m\_rated}$, will be practical for the full-size engine HEV.

The parameters of the DP algorithm used in the optimization of this section are listed in Table 4. The discrete SOC values expressed in (10) are uniformly distributed within the SOC usable range, and thus the difference between any two adjacent SOC values, $SOC_\Delta$, is

kept constant at $5 \times 10^{-4}$. The time interval between two adjacent time points, $\Delta t_k$, is one second in all of the four driving cycles, whereas the total time duration, $N$, is different in different driving cycles.

For hybridization of each IC engine vehicle with each of the electric drive power rating values listed in Table 3, the maximum MPG during each driving cycle and the average maximum MPG of the four driving cycles are calculated and plotted in Figure 2. The maximum MPG of each IC engine vehicle without the electric drive, i.e., $P_{m\_rated}$ = 0 kW, is also calculated and plotted for comparison and analysis.

**Table 4.** Parameters of the DP algorithm.

| Parameter | Value |
|---|---|
| $SOC_{init}$ | 0.5 |
| $SOC_{min}$ | 0.2 |
| $SOC_{max}$ | 0.8 |
| $SOC_{\Delta}$ | $5 \times 10^{-4}$ |
| $\Delta t_k$ | 1 s |
| $i_m$ | 1201 |
| $j_m$ | 6 (for Corolla 2019 and Camry 2016) 8 (for Chrysler 300 2016) |
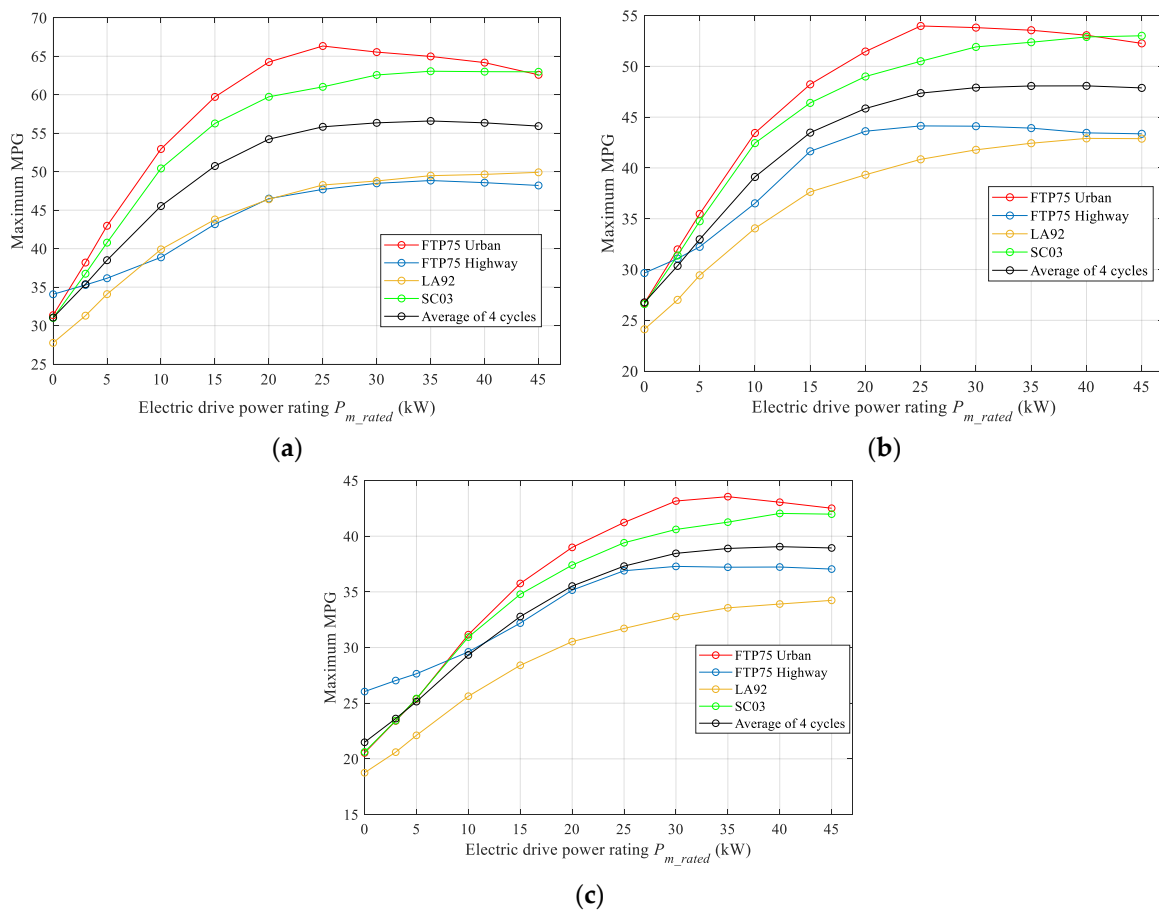


**Figure 2.** Maximum MPG versus electric drive power rating in typical driving cycles for hybridization of three typical vehicles: (**a**) Toyota Corolla 2019; (**b**) Toyota Camry 2016; (**c**) Chrysler 300 2016.

It is shown in Figure 2 that a vehicle with large mass and high engine power rating (e.g., Chrysler 300 2016) has lower maximum MPG than that with small mass and low engine power rating (e.g., Toyota Corolla 2019) after hybridized with electric drives of the same power rating. However, Figure 2 shows that all the three vehicles during all

the four driving cycles have the similar trend of sensitivity of MPG to the electric drive power rating $P_{m\_rated}$. Specifically, the maximum MPG starts to increase rapidly as $P_{m\_rated}$ increases from 0 kW to 10 kW, which means the fuel economy is significantly improved with the help of the electric drive. As $P_{m\_rated}$ keeps on increasing from 10 kW to 30 kW, the rate of increase in the maximum MPG is still positive and significant but starts to decline because the effect of fuel economy improvement is partially offset by the increasing resistance brought by the additional mass of the electric drive. After $P_{m\_rated}$ exceeds 30 kW, the maximum MPG shows negligible improvement or even starts to decrease as $P_{m\_rated}$ increases, which indicates that there is little room for further improvement on the fuel economy, whereas the significant mass increase of the electric drive starts to dominate the MPG trend. Although for different vehicles and during different driving cycles, there are minor differences in the $P_{m\_rated}$ value where the maximum MPG reaches its peak and starts to decline, the differences do not play a significant role in the optimization of $P_{m\_rated}$. This is because the additional cost brought by the increase of $P_{m\_rated}$ overwhelms the little, if any, improvement on the fuel economy after $P_{m\_rated}$ exceeds 30 kW. Therefore, the upper bound of the optimal window of $P_{m\_rated}$ is determined to be 30 kW, and the lower bound of this window is determined to be 10 kW because 10 kW is the upper limit of $P_{m\_rated}$ with the largest rate of increase in MPG. This optimal window of $P_{m\_rated}$ between 10 kW and 30 kW is universal regardless of vehicle mass, engine power rating and driving cycle, i.e., it is applicable to all the vehicles with different masses and engine power ratings during all the driving cycles, as illustrated in Figure 2. Therefore, it is possible to produce the same motor-battery package for all different sizes of passenger cars. For example, if the 30 kW electric drive package is produced for the purpose of optimizing fuel economy, it will be suitable for hybridizing all passenger cars, which will significantly decrease the production cost of the electric drive.

The MPG trends shown in Figure 2 can be also explained from the perspective of the distribution of operating points of the engine in the fuel consumption map. Here, each operating point of the engine refers to a set of the engine speed $n_e$ and the engine output power $P_e$ at a certain time point in the driving cycle. For example, for the hybridization of Toyota Corolla 2019 during the FTP75 Highway driving cycle, the engine operating points, overlapping the fuel consumption map with various $P_{m\_rated}$ values, are shown in Figure 3. The operating points in Figure 3 are classified into seven operating modes (hybrid traction, engine-alone traction, charging battery from engine, motor-alone traction, regenerative braking alone, hybrid braking, and mechanical braking alone), which determine the flow direction of motor power $P_m$ and load power $P_l$ at each operating point, as shown in Table 5. For load power $P_l$, positive power means traction, and negative power means braking. For motor power $P_m$, positive power means the battery is discharged so that the motor provides propulsion force to the vehicle, and negative power means the battery is charged so that the motor provides regenerative braking force to the vehicle. For engine power $P_e$, positive power means the engine provides propulsion force to the vehicle, and zero power means the engine is shut down. $P_e$ cannot be negative because the engine cannot be used for regenerative braking purpose. Thus, mechanical braking power $P_b$ is needed for braking if the motor cannot supply sufficient braking power to meet the braking demand of the vehicle load.
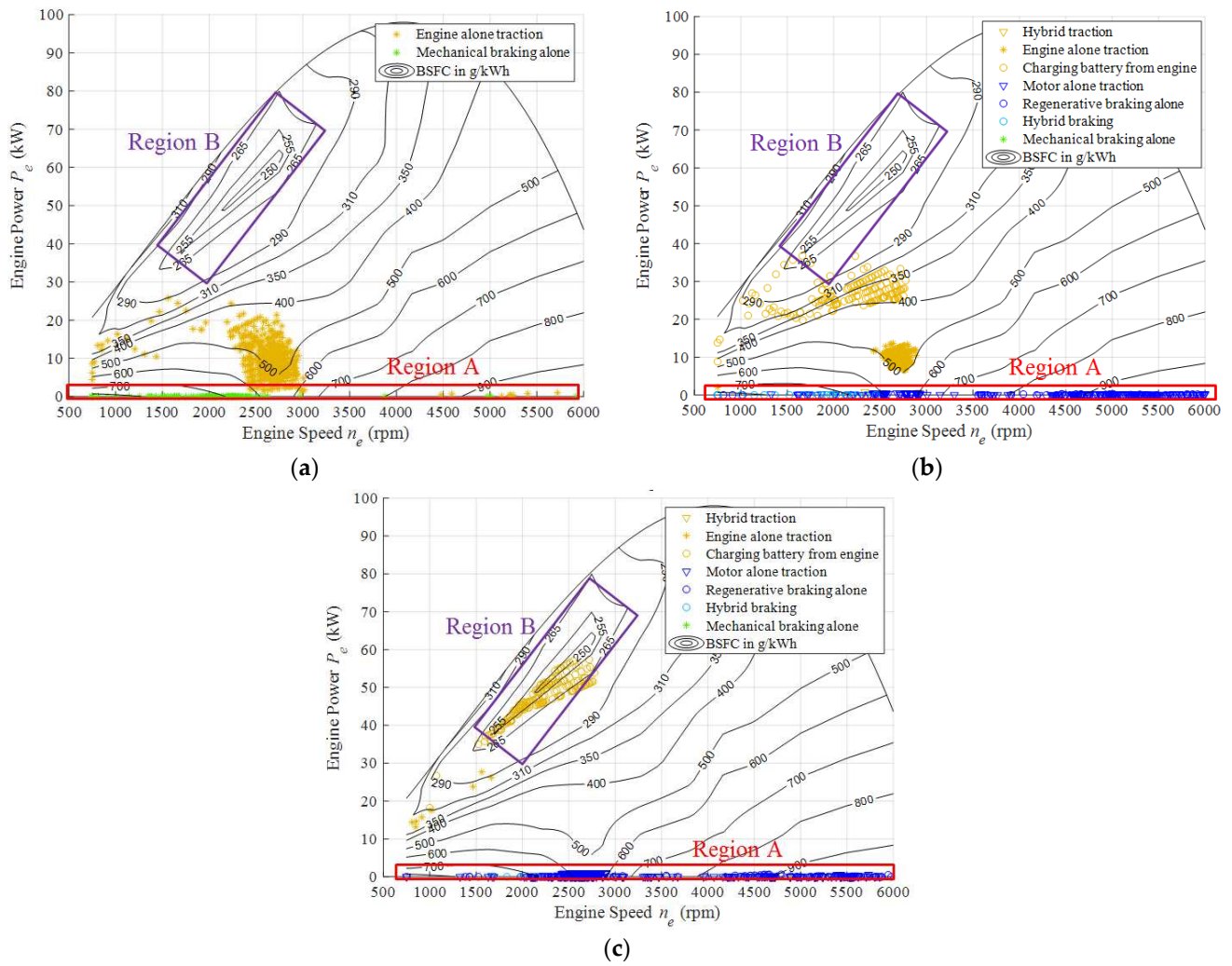
**(a)**



**(b)**



**(c)**

**Figure 3.** Optimal engine operating points overlapping the fuel consumption map [1,2] under various operating modes during the FTP75 Highway driving cycle for hybridization of Toyota Corolla 2019 with three $P_{m\_rated}$ values: (**a**) $P_{m\_rated}$ = 0 kW; (**b**) $P_{m\_rated}$ = 10 kW; (**c**) $P_{m\_rated}$ = 30 kW.

**Table 5.** Polarities of powers of motor, engine and load in each operating modes.

| Operating Modes | Polarities of Powers |
|---|---|
| Hybrid traction | $P_e > 0, P_m > 0, P_l > 0$ |
| Engine-alone traction | $P_e > 0, P_m = 0, P_l > 0$ |
| Charging battery from engine | $P_e > 0, P_m < 0, P_l > 0$ |
| Motor-alone traction | $P_e = 0, P_m > 0, P_l > 0$ |
| Regenerative braking alone | $P_e = 0, P_m < 0, P_l < 0, P_b = 0$ |
| Hybrid braking | $P_e = 0, P_m < 0, P_l < 0, P_b < 0$ |
| Mechanical braking alone | $P_e = 0, P_m = 0, P_l < 0, P_b < 0$ |

In Figure 3a with $P_{m\_rated}$ = 0 kW, most operating points of the engine lie in the region with significant engine power and high BSFC. According to (9), the fuel consumption $F_c$ is large, leading to poor fuel economy. The reason for this undesirable point distribution is that the absence of the electric drive forces the engine power to follow the load power which usually corresponds to points in the fuel consuming region of the map.

To improve the fuel economy, the engine operating points should move away from this fuel consuming region. According to (9), this can be accomplished by either lowering the points towards the region with negligible engine power (Region A), or lifting the points towards the region with low BSFC (Region B). In Region A, most of the operating points

are under the motor-alone traction mode or the regenerative braking mode because the negligible engine power is insufficient to meet the load power requirement. In Region B, the operating points are under the charging battery from engine mode because the engine power exceeds the load power requirement. Thus, the electric drive is needed in both desirable regions to let the motor power offset the power gap between the engine and the load. The larger rated power $P_{m\_rated}$ the electric drive has, the more motor power it can deliver to offset the power gap, and thus the more operating points it can move towards the two desirable regions, leading to better fuel economy. This principle is shown in Figure 3b,c.

With $P_{m\_rated}$ = 10 kW in Figure 3b, all the lowered operating points lie in Region A, leading to significant MPG improvement. Whereas, most of the lifted operating points are only closer to but do not lie in Region B, and a small portion of the operating points still stay in the fuel consuming region, indicating that there is still some room for further improvement on MPG. With $P_{m\_rated}$ = 30 kW in Figure 3c, almost all the operating points are moved into either Region A or Region B, indicating that the MPG is close to its peak value shown in Figure 2, and there is little room for further improvement on MPG if $P_{m\_rated}$ keeps on increasing. Similar trends in the distribution of engine operating points can also be seen for the other driving cycles. Therefore, we can conclude that the optimal window of the electric drive power rating is from 10 kW to 30 kW.

The independence of the optimal window of the electric drive power rating from vehicle mass and engine power rating can be explained from the distribution of operating points of the vehicle load with respect to the power-speed characteristics of the motor. Here, each operating point of the vehicle load refers to a set of the vehicle speed $V$ and the load power $P_l$ on the driving wheels at a certain time point in a driving cycle. For example, for each of the three IC engine vehicles without the electric drive, the distribution of the load operating points during the FTP75 Urban driving cycle with respect to the power-speed characteristics of the 10 kW rated motor and of the 30-kW-rated motor are shown in Figure 4. To make the comparisons between the load power and the motor power more convenient, the motor power-speed characteristics in Figure 4 are also transmitted onto the driving wheels.

Note that the engine traction mode in Figure 4 is divided into two submodes, one with full-linkage of the engine to the transmission and another with semi-linkage of the engine to the transmission. The engine usually operates in the full-linkage submode, where almost all of the engine power can be transmitted onto the driving wheels with high efficiency. However, the engine must switch to the semi-linkage submode if the speed of the engine operating in the full-linkage submode falls below the minimum permissible speed of the engine. This usually happens when the vehicle speed is very low. The semi-linkage submode will lead to low efficiency of the engine power transmission because there is a significant friction between the engine shaft and the transmission gearbox if they are only partially linked.

Figure 4a shows that the 30-kW-rated motor has the capacity to supply all the load powers of a compact vehicle like Toyota Corolla 2019. As the vehicle mass increases, the load powers also increase, as shown in Figure 4b,c for a medium-sized vehicle like Toyota Camry 2016 and for a large-sized vehicle like Chrysler 300 2016. However, the 30-kW-rated motor can still meet most of the load power demands for both heavier vehicles. Furthermore, for all the three vehicles, the 30-kW-rated motor can supply the powers of all the loads at which the engine has to operate in the semi-linkage submode. This means the 30-kW-rated motor alone can be used for propulsion to prevent the engine from inefficiently operating in the semi-linkage submode, which can significantly help improve the fuel efficiency. On the other hand, there are few or no load operating points lying outside the power range of the 30-kW-rated motor, indicating that there will be little or no further improvement on meeting load power demands as the motor power rating keeps on increasing beyond 30 kW. This explains why 30 kW is the universal upper bound of the optimal window of the electric drive power rating.
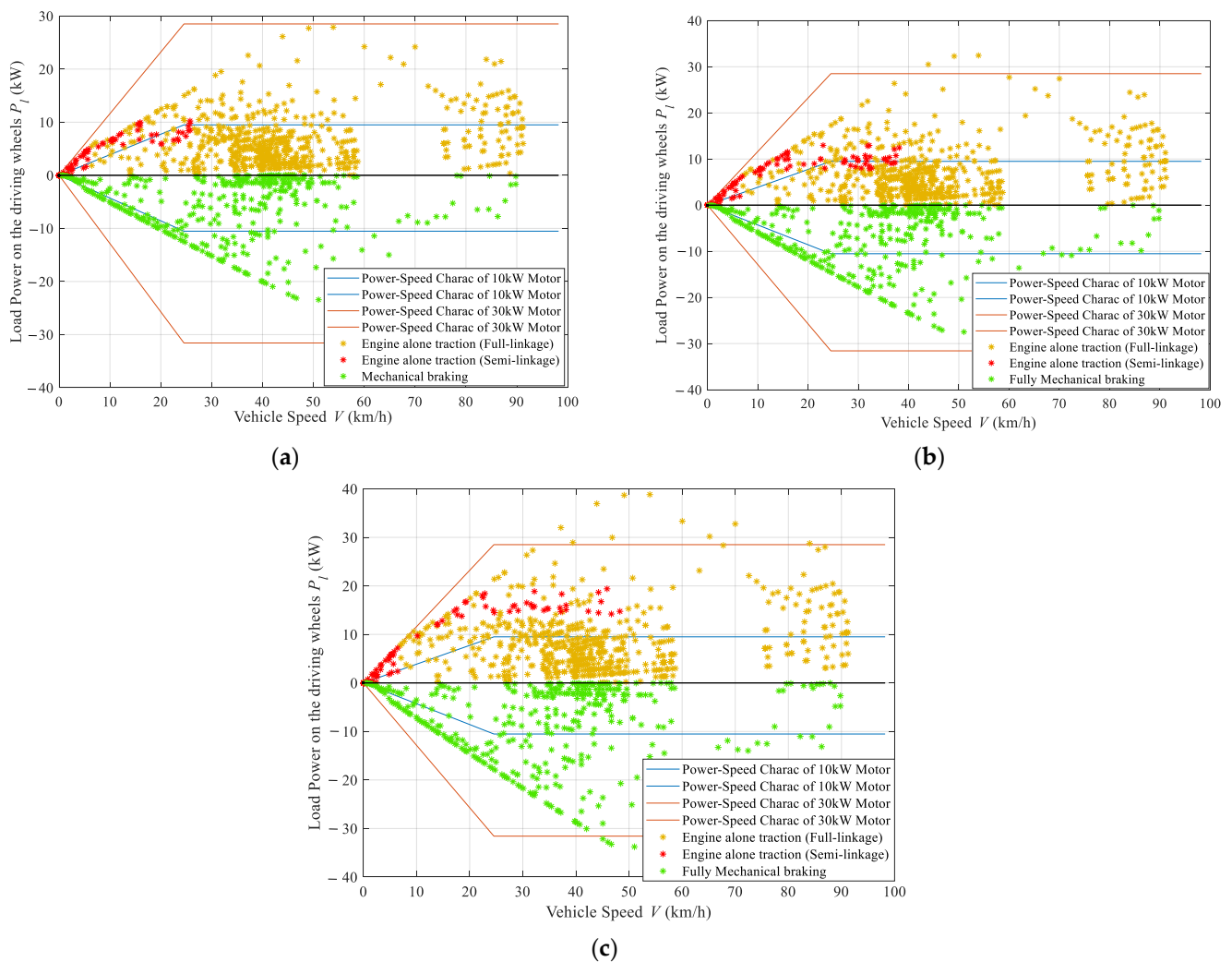
**Figure 4.** Load operating points during the FTP75 Urban driving cycle with respect to the power-speed characteristics of the 10 kW rated motor and of the 30-kW-rated motor for three typical IC engine vehicles: (**a**) Toyota Corolla 2019; (**b**) Toyota Camry 2016; (**c**) Chrysler 300 2016.

It is also shown in Figure 4 that for all the three vehicles, the density of load operating points within the power range of the 10 kW rated motor is significantly higher than that of the points beyond 10 kW rated motor power range. Therefore, the capability of meeting load power demands can be rapidly enhanced as the motor power rating increases from 0 to 10 kW, whereas the rate of this enhancement slows down as the motor power rating goes beyond 10 kW. This explains why 10 kW is the universal lower bound of the optimal window of the electric drive power rating.

## 5. Comparison with Commercially Available Vehicles

To demonstrate the advantages of the full-size engine HEVs over IC engine vehicles and commercially available HEVs, the fuel economy, acceleration performance, cost of power plants and cost payback mileage of the full-size engine HEVs are compared with those of Toyota Corolla 2019, an IC engine vehicle without hybridization, and with those of 2019 Toyota Prius, a commercially available HEV consisting of a large electric drive and a relatively small IC engine.

A model of Corolla-based full-size engine HEV is constructed by adding a relatively small electric drive in parallel to the full-size (98 kW) engine of Corolla 2019. Because 10 kW and 30 kW are the lower and upper bounds of the optimal window of the electric

drive power rating $P_{m\_rated}$, we make two models of the Corolla-based full-size engine HEVs, one with $P_{m\_rated}$ = 10 kW and another with $P_{m\_rated}$ = 30 kW.

Using the DP algorithm introduced in this paper, we calculate the maximum MPGs of the two HEV models during FTP 75 Urban and FTP 75 Highway driving cycles, and compare them with those of the Corolla 2019. To obtain the overall fuel economy during the two driving cycles, a combined MPG taking 55% of the urban MPG and 45% of the highway MPG is also calculated and compared. To make fair comparisons under the same conditions, the maximum MPGs of the Corolla 2019 are also calculated using the same DP algorithm, rather than directly copied from its commercial MPG data. The maximum MPGs of the three vehicles are listed in Table 6, which shows that both full-size engine HEVs have significantly better fuel economy than the Corolla 2019.

**Table 6.** Maximum MPGs of Corolla-based vehicles.

| Vehicles | Urban | Highway | Combined |
|---|---|---|---|
| Corolla 2019 | 31.35 | 34.07 | 32.57 |
| HEV ($P_{m\_rated}$ = 10 kW) | 52.95 | 38.88 | 46.62 |
| HEV ($P_{m\_rated}$ = 30 kW) | 65.54 | 48.49 | 57.87 |

The acceleration performance is specified in this study as the acceleration time for the vehicle speed from 0 to 60 mph (26.8 m/s) on a flat road, which can be calculated as

$$t_a = \int_{V_1}^{V_2} \frac{M\delta}{\frac{T_{e\_w}+T_{m\_w}}{r_d} - Mg\,f_r - \frac{1}{2}\rho_a C_D A_f V^2}\,dV \tag{16}$$

where $V_1$ = 0, $V_2$ = 26.8 m/s, $T_{e\_w}$ and $T_{m\_w}$ are the engine torque and the motor torque transmitted onto the driving wheels, respectively.

Using (15), we can calculate the acceleration time from 0 to 60 mph for the three vehicles, as listed in Table 7, which shows that both full-size engine HEVs have better acceleration performance than the Corolla 2019.

**Table 7.** Acceleration time from 0 to 60 mph for Corolla-based vehicles.

| Vehicles | Acceleration Time (s) |
|---|---|
| Corolla 2019 | 7.53 |
| HEV ($P_{m\_rated}$ = 10 kW) | 7.14 |
| HEV ($P_{m\_rated}$ = 30 kW) | 6.58 |

The Toyota Prius 2019 is a series-parallel HEV whose maximum MPGs cannot be calculated using the DP algorithm introduced in this paper because this DP algorithm can only be used to calculate the maximum MPGs of parallel HEVs. To estimate the fuel economy of the Prius 2019 under optimal control, an equivalent parallel HEV model (Modified Prius) is made. The electric drive power rating $P_{m\_rated}$ of the Modified Prius is designed to be the same as the Prius 2019, which is 53 kW, so that the Modified Prius will have roughly the same weight as Prius 2019, and its initial cost will also be close to Prius 2019. The reason for this design is that the fuel economy is affected by the vehicle weight, and the initial cost is largely determined by $P_{m\_rated}$ because the price per kW of the motor is much higher than that of the engine. The engine power rating of the Modified Prius is determined to be 31 kW, which ensures that the Modified Prius has the same acceleration performance (acceleration time from 0 to 60 mph) as the Prius 2019.

To construct full-size engine HEV models based on the Prius 2019, an IC engine vehicle model (Engine-only Prius) is made first. The engine power rating of the Engine-only Prius is determined to be 73 kW, which ensures that the Engine-only Prius has the same performance (acceleration time from 0 to 60 mph) as the Prius 2019. Then, a relatively small electric drive is added in parallel to the full-size engine of the Engine-only Prius,

making a Prius-based full-size engine HEV. Similar to the Corolla-based HEVs, two models of the Prius-based full-size engine HEVs are made, one with $P_{m\_rated}$ = 10 kW and another with $P_{m\_rated}$ = 30 kW.

The maximum MPGs of all the Prius-based vehicles are listed in Table 8. They are all calculated using the DP algorithm introduced in this paper, except for those of the Prius 2019, which is estimated to be lower than those of the Modified Prius. The reason for such estimation is that the Prius 2019 has a much larger engine (71 kW) than the Modified Prius (31 kW).

**Table 8.** Maximum MPGs of Prius-based vehicles.

| Vehicles | Urban | Highway | Combined |
|---|---|---|---|
| Engine-only Prius | 35.75 | 36.57 | 36.11 |
| HEV ($P_{m\_rated}$ = 10 kW) | 57.77 | 46.36 | 52.63 |
| HEV ($P_{m\_rated}$ = 30 kW) | 71.40 | 52.07 | 62.70 |
| Modified Prius | 73.31 | 58.91 | 66.83 |
| Prius 2019 [1] | <73.31 | <58.91 | <66.83 |

[1] Maximum MPGs of the Prius 2019 are estimated based on those of the Modified Prius.

To compare the costs of power plants for all the Prius-based vehicles, we assume that the average prices of the induction motor, the IC engine, and the battery pack are $110 per kW, $35 per kW, and $137 per kWh, respectively [2,28,29]. To calculate the cost payback mileages, we assume that the price of gasoline is $4 per gallon [2]. The costs of power plants and payback mileages for the Prius-based vehicles are listed in Table 9.

**Table 9.** Costs of power plants and cost payback mileages for Prius-based vehicles.

| Vehicles | Cost ($) | Payback (mile) |
|---|---|---|
| Engine-only Prius | 2553 | NaN |
| HEV ($P_{m\_rated}$ = 10 kW) | 3699 | 32,990 |
| HEV ($P_{m\_rated}$ = 30 kW) | 5993 | 73,249 |
| Modified Prius | 7163 | 90,562 |
| Prius 2019 [1] | 8563 | >118,065 |

[1] Payback mileage of the Prius 2019 is estimated to be longer than 118 k miles, based on the Prius 2019's cost of power plants and the Modified Prius's fuel economy.

Table 8 shows both the full-size engine HEVs have significantly better fuel economy than the Engine-only Prius, and the fuel economy of the full-size engine HEV with $P_{m\_rated}$ = 30 kW is close to that of the Modified Prius and may be better than that of the Prius 2019. In the meantime, the costs of power plants in both full-size engine HEVs are significantly lower than those in the Modified Prius and the Prius 2019, as shown in Table 9. Therefore, the cost payback mileages of both full-size engine HEVs are significantly shorter than those of the Modified Prius and the Prius 2019, as shown in Table 9.

The acceleration time from 0 to 60 mph for all the Prius-based vehicles are listed in Table 10, which shows that both full-size engine HEVs have better acceleration performance than the Engine-only Prius, the Modified Prius, and the Prius 2019. Note that the acceleration time for the Modified Prius and the Engine-only Prius is by design the same as that for the Prius 2019, which is found on [30].

**Table 10.** Acceleration time from 0 to 60 mph for Prius-based vehicles.

| Vehicles | Acceleration Time (s) |
|---|---|
| Engine-only Prius | 9.79 |
| HEV ($P_{m\_rated}$ = 10 kW) | 9.11 |
| HEV ($P_{m\_rated}$ = 30 kW) | 7.61 |
| Modified Prius | 9.79 |
| Prius 2019 | 9.79 |

## 6. Conclusions

This paper proposed a universally optimized parallel HEV drivetrain, which uses the full-size IC engine as the primary power plant to ensure that the vehicle performance and the travel range of the hybridized vehicle are at least the same as the original IC engine vehicle. To improve the fuel economy and performance, an optimal size electric drive, including an electric motor, a battery bank, and power electronic converters, was added in parallel with this full-size engine.

Three commercially available IC engine vehicles with different curb masses and engine power ratings were chosen as representative passenger cars on the market. The original power ratings of the engines were kept for this hybridization.

To find the optimal power rating of the electric drive, we developed a DP algorithm to calculate the maximum MPG and find the optimal sequence of the engine operating points during the driving cycle. The optimality of the solution obtained from the algorithm was guaranteed by selecting the battery SOC and the number of gears in the engine transmission as the control variables of the algorithm.

The optimal window of electric drive power rating was found to be between 10 kW and 30 kW. This was shown to be independent of IC engine power rating and vehicle mass by analyzing the sensitivity of the maximum MPG to the electric drive power rating during each test driving cycle. Therefore, all passenger cars, regardless of their masses and engine sizes, can be universally optimally hybridized with one size of electric drive packages, further reducing the hybridization cost. The sensitivity of the maximum MPG to the electric drive power rating and the universality of the optimal window were further explained in terms of the distribution of the optimal engine operating points in the engine fuel consumption map and the distribution of the load operating points with respect to the power-speed profiles of the motor.

To show the advantages of the full-size engine HEVs proposed in this paper over the conventional IC-engine based vehicles and commercially available HEVs, comparisons were made with two types of popular passenger cars for fuel economy, acceleration performance, cost of power plants and cost payback mileage. The proposed full-size engine HEV has significantly better fuel economy as well as better performance than the equivalent IC engine vehicle. This HEV was also shown to have better performance and shorter cost payback than conventional HEVs.

Further research may be done to find a real-time control strategy for the full-size engine HEV. This will be independent of the driving cycle.

**Author Contributions:** Conceptualization, L.L. and M.E.; methodology, Z.H. and L.L.; software, Z.H.; validation, Z.H., R.T.M. and M.E.; formal analysis, Z.H.; investigation, R.T.M.; resources, L.L., R.T.M. and M.E.; data curation, Z.H.; writing—original draft preparation, Z.H.; writing—review and editing, Z.H., R.T.M. and M.E.; visualization, Z.H.; supervision, M.E.; project administration, M.E.; funding acquisition, N/A. All authors have read and agreed to the published version of the manuscript.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Publicly available datasets were analyzed in this study. The data can be found here: [12–14,30].

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Ehsani, M.; Gao, Y.; Longo, S.; Ebrahimi, K. *Modern Electric, Hybrid Electric, and Fuel Cell Vehicles—Fundamentals, Theory, and Design*, 3rd ed.; CRC Press: Boca Raton, FL, USA, 2018; pp. 1–16.
2. Lai, L. A Development of Design and Control Methodology for Next Generation Parallel Hybrid Electric Vehicle. Ph.D. Thesis, Texas A&M University, College Station, TX, USA, 2012.
3. Car and Driver. Available online: https://www.caranddriver.com/chevrolet/bolt-ev (accessed on 5 January 2021).

4. Reyes, J.; Parsons, R.; Hoemsen, R. Winter happens: The effect of ambient temperature on the travel range of electric vehicles. *IEEE Trans. Veh. Technol.* **2016**, *65*, 4016–4022. [CrossRef]

5. Lin, C.; Peng, H.; Grizzle, J.; Kang, J. Power management strategy for a parallel hybrid electric truck. *IEEE Trans. Control Syst. Technol.* **2003**, *11*, 839–849.

6. Cieslik, W.; Zawartowski, J.; Fuc, P. *The Impact of the Drive Mode of a Hybrid Drive System on the Share of Electric Mode in the RDC Test*; SAE Technical Paper; SAE Powertrains, Fuels & Lubricants Meeting: Krakow, Poland, 2020. [CrossRef]

7. Pielecha, I.; Cieślik, W.; Merkisz, J. Analysis of the electric drive mode use and energy flow in hybrid drives of SUVs in urban and extra-urban traffic conditions. *J. Mech. Sci. Technol.* **2019**, *33*, 5043–5050. [CrossRef]

8. Wirasingha, S.; Schofield, N.; Emadi, A. Plug-in hybrid electric vehicle developments in the US: Trends, barriers, and economic feasibility. In Proceedings of the 2008 IEEE Vehicle Power and Propulsion Conference, Harbin, China, 3–5 September 2008.

9. Gao, Y.; Ehsani, M. Design and control methodology of plug-in hybrid electric vehicles. *IEEE Trans. Ind. Electron.* **2010**, *57*, 633–640.

10. Hardman, S.; Plotz, P.; Tal, G.; Axsen, J.; Figenbaum, E.; Karlsson, S. Exploring the Role of Plug-in Hybrid Electric Vehicles in Electrifying Passenger Transportation. Available online: https://escholarship.org/uc/item/3w53q2h9#author (accessed on 17 October 2021).

11. Peterson, S.; Whitacre, J.; Apt, J. The economics of using plug-in hybrid electric vehicle battery packs for grid storage. *J. Power Source* **2010**, *195*, 2377–2384. [CrossRef]

12. Car and Driver. Available online: https://www.caranddriver.com/toyota/corolla/specs/2019/toyota_corolla_toyota-corolla_2019/399278 (accessed on 17 October 2021).

13. New-Cars. Available online: http://www.new-cars.com/2016/toyota/specs/camry.html (accessed on 17 October 2021).

14. New-Cars. Available online: http://www.new-cars.com/2016/chrysler/specs/300.html (accessed on 17 October 2021).

15. Kelly, K.J.; Mihalic, M.; Zolot, M. Battery usage and thermal performance of the Toyota prius and Honda insight during chassis dynamometer testing. In Proceedings of the 17th Annual Battery Conference on Applications and Advances, Long Beach, CA, USA, 18 January 2002.

16. Kuhn, B.; Pitel, G.; Krein, P. Electrical properties and equalization of lithium-ion cells in automotive application. In Proceedings of the IEEE Vehicle Power and Propulsion Conference, Chicago, IL, USA, 7–9 September 2005.

17. Valoen, L.; Shoesmith, M. The effect of PHEV and HEV duty cycles on battery and battery pack performance. In Proceedings of the Plug-In Hybrid Electric Vehicle 2007 Conference, Winnipeg, MB, Canada, 1–2 November 2007.

18. Zaghib, K.; Charest, P.; Guerfi, A.; Shim, J.; Perier, M.; Striebel, K. Safe li-ion polymer batteries for HEV applications. *J. Power Source* **2004**, *134*, 124–129. [CrossRef]

19. Affanni, A.; Bellini, A.; Franceschini, G.; Guglielmi, P.; Tassoni, C. Battery choice and management for new-generation electric vehicles. *IEEE Trans. Ind. Electron.* **2005**, *52*, 1343–1349. [CrossRef]

20. Bertsekas, D.P. *Dynamic Programming and Optimal Control*, 3rd ed.; Athena Scientific: Belmont, TN, USA, 2005; pp. 18–34.

21. Cormen, T.H.; Leiserson, C.E.; Rivest, R.L.; Stein, C. *Introduction to Algorithms*, 3rd ed.; The MIT Press: Cambridge, MA, USA, 2009; pp. 359–413.

22. Perez, L.V.; Bossio, G.R.; Moitre, D.; Garcia, G.O. Optimization of power management in an hybrid electric vehicle using dynamic programming. *Math. Comput. Simul.* **2006**, *73*, 244–254. [CrossRef]

23. Zoelch, U.; Schroeder, D. Dynamic optimization method for design and rating of the components of a hybrid vehicle. *Int. J. Veh. Des.* **1998**, *19*, 1–13.

24. Varesi, K.; Radan, A. A novel methodology proposed for optimizing the degree of hybridization in parallel HEVs using genetic algorithm. *WASET* **2011**, *5*, 915–920.

25. Vinot, E.; Trigui, R.; Jcanneret, B.; Scordia, J.; Badin, F. HEVs comparison and components sizing using dynamic programming. In Proceedings of the 2007 IEEE Vehicle Power and Propulsion Conference, Arlington, TX, USA, 9–12 September 2007.

26. Lukic, S.M.; Emadi, A. Effects of drivetrain hybridization on fuel economy and dynamic performance of parallel hybrid electric vehicles. *IEEE Trans. Veh. Technol.* **2004**, *53*, 385–389. [CrossRef]

27. Kim, M.; Peng, H. Power management and design optimization of fuel cell/battery hybrid vehicles. *J. Power Source* **2007**, *165*, 819–832. [CrossRef]

28. Lowe, M.; Tokuoka, S.; Trigg, T.; Gereffi, G. *Lithium-Ion Batteries for Electric Vehicles: The U.S. Value Chain*; Duke University CGGC: Durham, NC, USA, 2010.

29. Renewable Energy World. Available online: https://www.renewableenergyworld.com/storage/annual-survey-finds-battery-prices-dropped-13-in-2020/#gref (accessed on 5 October 2021).

30. Zero To 60 Times. Available online: https://www.zeroto60times.com/vehicle-make/toyota-0-60-mph-times/?yearMin=2019&yearMax=2020&engineType%5B%5D=Gas+Electric+Hybrid (accessed on 5 January 2021).

# Eco-Dimensioning Approach for Planar Transformer in a Dual Active Bridge (DAB) Application

**Glauber de Freitas Lima** [1,*] **, Boubakr Rahmani** [1,2] **, Maud Rio** [2] **, Yves Lembeye** [1] **and Jean-Christophe Crebier** [1]

1   G2Elab, Grenoble INP, CNRS, University Grenoble Alpes, F-38000 Grenoble, France;
    boubakr.rahmani@grenoble-inp.fr (B.R.); Yves.Lembeye@g2elab.grenoble-inp.fr (Y.L.);
    jean-christophe.crebier@g2elab.grenoble-inp.fr (J.-C.C.)
2   GSCOP, Grenoble INP, CNRS, University Grenoble Alpes, F-38000 Grenoble, France; maud.rio@g-scop.eu
*   Correspondence: glauber.de-freitas-lima@g2elab.grenoble-inp.fr; Tel.: +33-(0)-74916-5697

**Abstract:** Power electronics converters are traditionally designed regarding efficiency, power density, cost, and reliability figures of merit. Today, with the extreme spread of power electronic applications in our modern societies, together with the earth limits in terms of materials resources, it is important to consider the ecological impact of the converter not only during its usage, but over its whole life cycle. This article introduces an eco-dimensioning methodology for analyzing and accounting for the energy consumption over the entire converter life. The analysis is applied on a small DC-DC converter considering the main components dual active bridge (DAB) converter. The planar transform is one of the key elements modeled in this article, including material and manufacturing conditions. The traditional and eco-dimensioning approaches are carried out and compared in order to emphasize the possible consequences on total energy cost.

**Keywords:** planar transformer; eco-design; environmental impact evaluation; product efficiency; DC-DC converter; dual active bridge (DAB); life cycle optimization

## 1. Power Electronic Design

### 1.1. Conventional (Efficiency vs. Power Density Focused)

Up to now, to achieve cost reductions due to the growing need for information and telecommunication technologies with the development of power electronic systems, it has been popular to increase the power density (up to 35 kW/dm$^3$ in a DC-DC converter [1]). Such an achievement is mainly performed by means of switching frequency increase and the use of forced cooling systems with the drawback of global efficiency reduction (<95%).

According to [2,3], nowadays, 40% of the total used energy is enabled by electric power, and this figure could reach 50% by 2030 [3] and 60% [2] by 2040. Such important growth implies that the role of power electronics is not only power density focused, but rather related to efficiently and ecologically processing the energy required by different specific applications. The applications can vary from small power (W) such as mobile phone chargers to higher power applications such as electric cars (kW), small and large renewable power plants, trains, metros (MW), and DC transmission lines (GW), etc.

Nowadays, with the advance in technologies for semiconductors (increase in junction temperature, ON/OFF states and rise/fall time improvement and optimization) and magnetic materials (higher operating temperatures, reduction of hysteresis losses) as well as by allowing power density reduction [2] (e.g., 2.2 kW/dm$^3$ [4], 7kW/dm$^3$ [3]), design with so-called extreme efficiency converters [5] (e.g., 98% [4], 99.2% [3,6]) have come to represent the major success of power electronics realization.

With the use of active and passive components as well as auxiliary circuity and cooling systems (when needed) [1], power electronic system design is conceived to process electrical specifications of desired voltages ($V_i$, $V_o$, $\Delta V_o$), currents ($I_o$, $\Delta I_o$), and power ($P_o$)

following the mandatory and safe standards (e.g., EMI (electro-magnetic interference), safety standards among others).

To proceed with the traditional power electronic design, a multi-objective optimization [3,6–8] is performed to guarantee optimum performances through Pareto fronts (e.g., Efficiency $\times$ Power Densities (W/dm$^3$ and W/kg) $\times$ Costs ($\text{€}$/W) $\times$ MTBF (Mean Time Between Failure) (h) $\eta - \rho_{box} - \sigma_P$ [3,8]). Figure 1 presents a conventional multi-objective flowchart based on [7] for applying and designing any power converter.
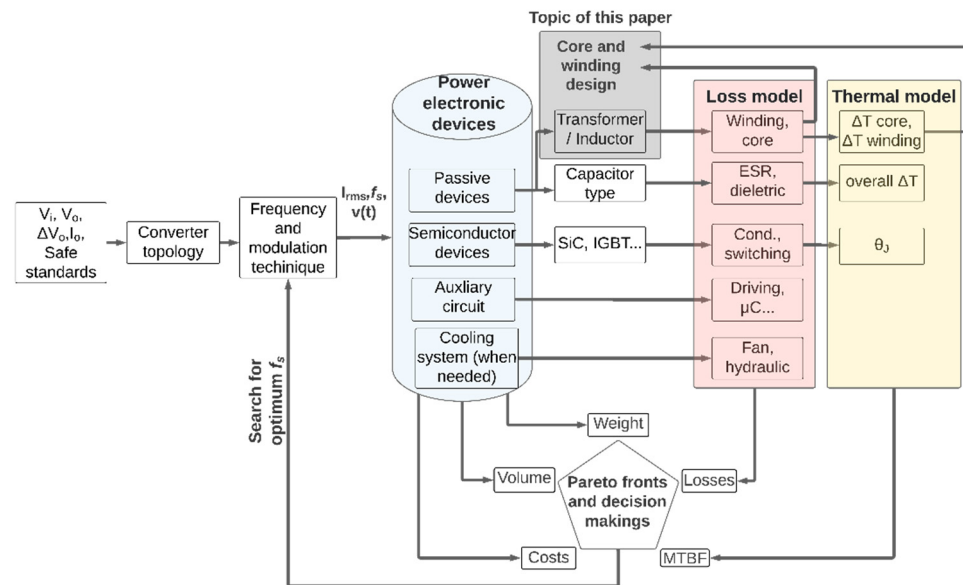


**Figure 1.** Summarized flowchart of a conventional design with global optimization of the main needs in a power electronic systems.

As can be seen, conventionally, power electronic choices are based on many criteria, such as failure rate, volume, weight, power losses, and costs. When focusing on this multi-objective conventional approach, extremely highly efficient (~96–99%) and/or power density systems (~2 kW/dm$^3$–~6 kW/dm$^3$) are conceived [3,7,8] in a Pareto front behavior over a wide load charge profile.

As covered in [3,7], designing a very efficient system (over 90%) while assuring high power density is a challenging work, as there is always a compromise between peak efficiency and power over total volume. The reason for this cognitive dissonance regarding "efficiency vs. power density", linking the need for miniaturization and maintenance of high efficiency, is mainly related to the switching frequency ($f_s$) in which power converters operate. By designing and operating at high frequencies, e.g., 300 kHz and above, the passive elements tend to decrease in volume and material, making the prototype more compact. Meanwhile, the increase of the $f_s$ may impose an increase of the power losses all over the converter (switching, skin effect, magnetic hysteresis losses, EMI filter effort), as well as increasing the design complexity (driving circuits' interconnects, EMI compliance, dynamic response management, $\mu$C bandwidth).

Therefore, such an important choice of $f_s$ is one of the main design parameters which justifies the design of the close loop introduced in Figure 1. Such a parameter will determine the power losses and the volume, which is directly related to the global life cycle energy cost, as covered in Section 1.2.

### 1.2. Eco-Design

When considering eco-design, it is necessary to represent and to take into account complementary parameters, such as the entire energy dissipation over the converter life cycle from the material extraction, and refine them to usages and finally end-of-life management. While the traditional design mainly focuses on optimizing the converter for its usage, looking after best efficiencies, the eco-design also considers the two other life

stages in order to prevent the converter from only being optimized for its usage, leaving for example a fully non-recyclable waste at the end of life or relying on extremely polluting materials. Today, with the pressure that our modern societies put on earth materials and energies, it is urgent to eco-design power converters.

In Figure 2, the complementary design blocks to the schematic introduced in Figure 1 are proposed to account for eco-dimensioning. In such a way, the weight, volume, efficiency, and MTBF are used to feed another set of models seeking to evaluate the total energy losses, the design model functionalities, the potential recycling fraction, and the waste and pollution generated during the life cycle.
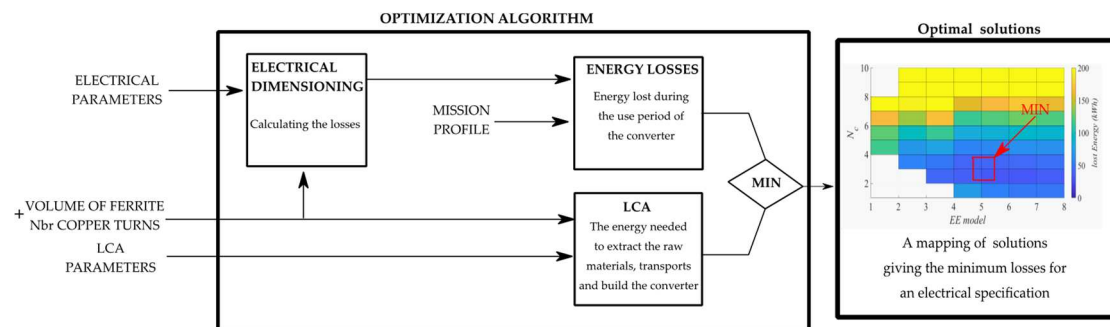


**Figure 2.** Summarized flowchart of a conventional design mixed with a Life Cycle Assessment vision (LCA) for global optimization of the main needs in a power electronic systems.

In the traditional concept, the weight, volume, and efficiency are chosen according to subjective specifications, e.g., costumer needs and available space constraints, whereas in eco-design, they have a direct impact on the objective.

In such a way, if it is easy for everyone to link converter efficiency and usage energy losses. An important difference can be related to the efficiency to be optimized. Whereas the peak efficiency could be optimized to outline a key characteristic on the converter, in eco-design, the mission profile and operating conditions are strongly impacting the energy losses under usage, bringing the design to consider not only the peak efficiency or the efficiency at maximum power efficiency over a wider power and operating condition range.

Similarly, regarding the active and passive components, and moreover their materials, the optimization of the component losses and power density may lead to the use of "exotic" materials, i.e., difficult to produce. This optimization process could be challenged by the environmental impact of the materials and/or technologies used to manufacture them.

One very tricky point is related to the integration process which receives a large plebiscite, increasing at best power densities, as well as performances under high switching frequencies thanks to reduced interconnects. The integration brings, on one side, collective processes, higher reliability, and cost reduction, but on the other side, it produces very heterogeneous systems that are difficult to disassemble and recycle. The processes used to extract the various materials are usually very polluting and energy consuming and most of the time only part of the materials remains fully useful, with the others being at best down-cycled. In this context, the life cycle assessment (ISO 14040:2006) of the system designed enables to model alternative scenarios to identify the potential environmental impact sources, support decision-making during the design process, and converge to an present an efficient life cycle system.

## 2. Dual Active Bridge (DAB)

There are different classes and topologies of power converters throughout industry and literature, e.g., Phase-shift DC-DC converters, Resonant DC-DC converters, DC link AC-AC converters, Matrix AC-AC converters, etc. Regarding topology, its selection is based on criteria such as efficiency, profile charge characteristics, power density, galvanic isolation, power directionality, costs, dynamic behavior and controllability, expertise, and complexity, etc. The scope of this section does not intend to impose the most optimum

topology choice, but rather describe one of the most popular topologies, the DAB converter, as a representative user case. This will be considered using single-phase-shift (SPS) modulation characteristics [9,10] and benefits found in our laboratory in previous research works [9,11,12].

The DAB has become a popular DC-DC converter over the last ten years that allows galvanic isolation, bidirectionality in power, symmetricity, controllable output, high power density, and efficiency. Due to such exceptional performance and features, it is mostly used in renewable energy systems [13,14], energy storage systems [15], DC distribution systems [16], and solid-state transformer (SST) [17]. The topology is presented in Figure 3.
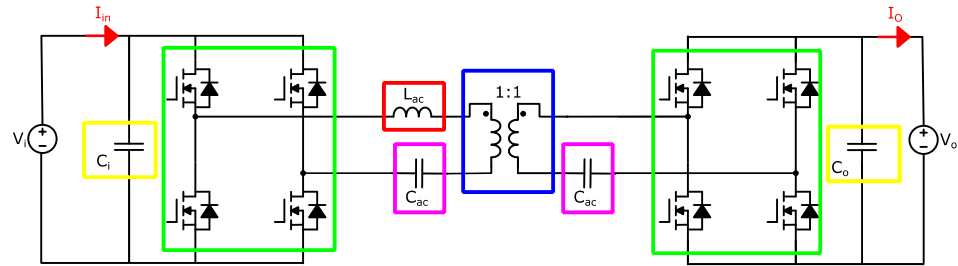


**Figure 3.** Dual Active Bridge (DAB) topology used as scenario study. Colored scares are highlighting the main components' types.

Even if ZVS (zero voltage switching) is a driving feature to reduce switching losses, the choice and optimization of the switching frequency remains a key parameter in the DAB, impacting on the design of the transformer (highlighted in blue), the choice of the AC inductance (highlighted in red), the AC capacitor (highlighted in purple), the active devices, here selected as semiconductor MOSFETs (metal-oxide-semiconductor field-effect transistor) (highlighted in green), and finally, the input and output capacitor (highlighted in yellow).

The DAB allows the implementation of different modulation techniques [18–21] that increase ZVS and ZCS (zero current switching) range, decrease RMS (root mean square) current value, and therefore improve efficiency and power density throughout the charge profile of a given application. In this work, the ordinary SPS modulation [9,10] was chosen due to its simplicity and due to an unpretentious profile charge that allows good performance: a constant input voltage $V_i$ = 20 V, a constant output voltage $V_o$ = 20 V and an output current $I_o$ of 2 A during half of the time and 1 A during the other half.

As in [9], in order to have a comprehensive analysis, the important parameters of design are portrayed in a dimensionless and meaningful plane, as a function of the static gain $M$ versus a normalized output current $\gamma$, given by (1) and (2), respectively.

$$M = \frac{V_o}{V_i} \tag{1}$$

$$\gamma = \frac{2 f_s L_{ac} I_o}{V_i} \tag{2}$$

Figure 4 presents the normalized RMS current (3) with respect to the output current in the parametrized output plane including the ZVS regions

$$\bar{I}_{rms} = \frac{I_{rms}}{I_o} \tag{3}$$

It is possible to conclude from Figure 4 that a safe design guarantying ZVS is located for the design criteria $\gamma_{min}$ = 0.05 and $\gamma_{max}$ = 0.1. This will allow, for example, minimization of conducting losses up to 5% with a circulating current value $I_{rms}$ of around 15% larger than the maximum output current $I_o$.
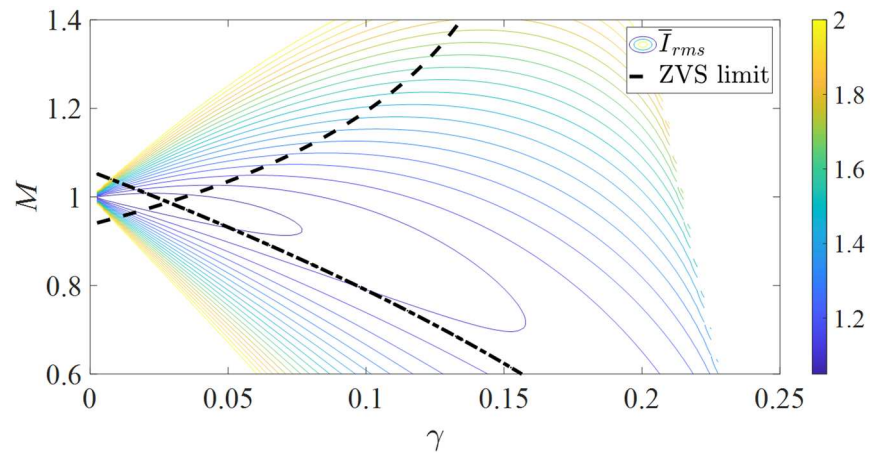
**Figure 4.** Normalized RMS current mapping of DAB under single phase-shift modulation including ZVS regions.

As it is possible to notice in Figure 5, all the devices are selected in a loop together with the choice of the switching frequency, which is an output of the minimizing function result. In this work, the selection of MOSFET, capacitors and inductor family are already given from previous work, which is presented in Table 1. Once they are selected, their losses must be computed and analyzed.

**Table 1.** Devices or family used for loss analysis.

| Main Devices | | |
|---|---|---|
| MOSFET | SiSA10DN, N-Channel 30 V | 5 mΩ; Qg = 18.75 nC |
| Inductor $L_{ac}$ | Coilcraft XEL6060 | |
| Capacitors | Ceramic | $R_{ac}$ = 2 mΩ |
| Transformer cores | EE Planar Ferroxcube [22] | Section 3. |

For the MOSFET, as the ZVS is assumed, the only losses considered are due to $Q_g$ and the conducting losses as presented in (4) thanks to datasheets. $R_{mosfet}$ can be expressed with respect to the junction temperature estimate and gate voltage, while $Q_g$ and $V_i$ are respectively derived from operating conditions.

$$P_{Mosfet} = R_{mosfet}I_{rms}{}^2 + Q_g V_i f_s \tag{4}$$

The losses of the $C_{ac}$ are simply computed as in:

$$P_{Ca} = R_{Ca}I_{rms}{}^2 \tag{5}$$

Then, the value of the inductance is chosen according to the load profile to respect ZVS condition and maximize efficiency, as already discussed and presented in Figure 4. By considering $\gamma_{max}$ = 0.1, the value of the inductance should go with the switching frequency as presented in (6).

$$f_s L_{ac} = \frac{V_i \gamma_{max}}{2 I_{o_{max}}} = 0.5 \Rightarrow L_{ac} = \frac{0.5}{f_s} \tag{6}$$
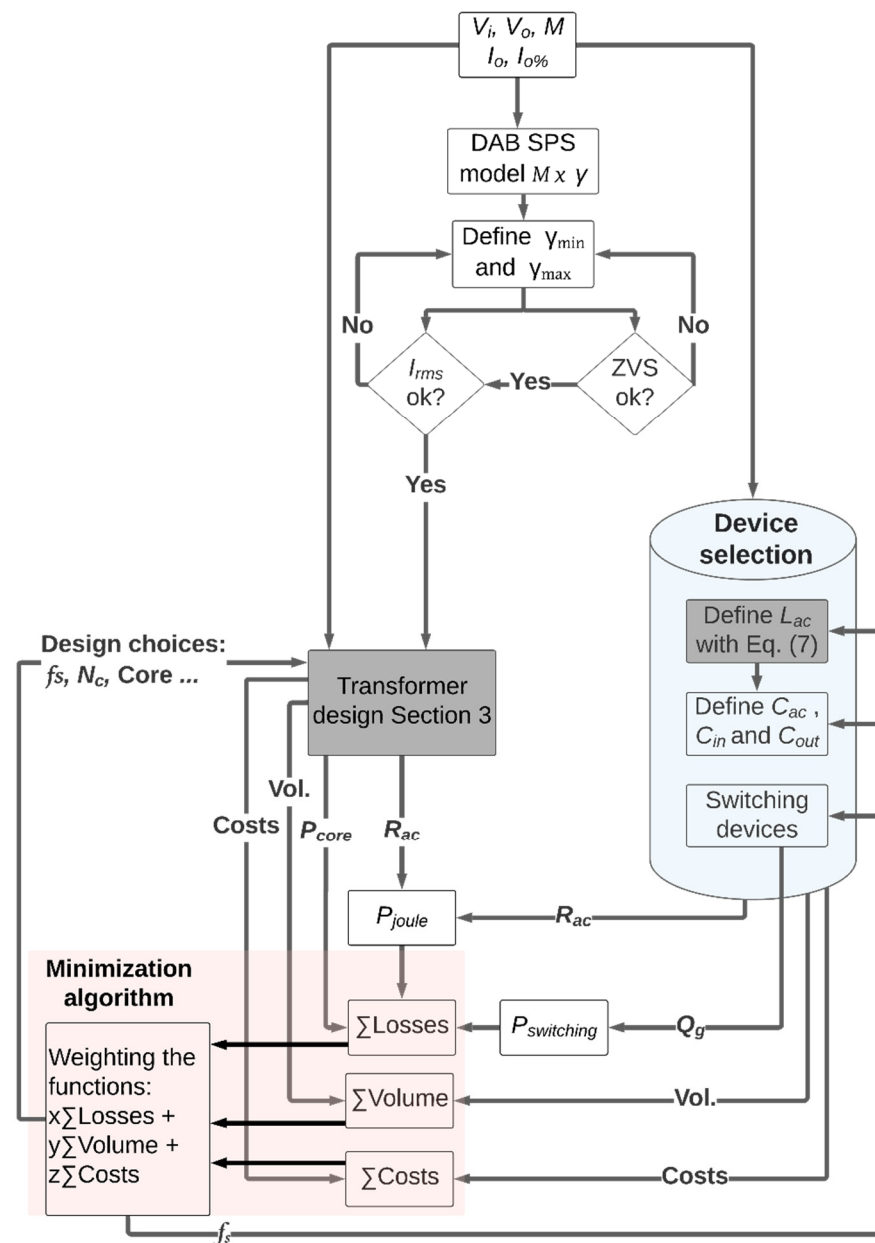
**Figure 5.** Optimizing routine for minimization of losses, volume or costs considering the main aspects in a DAB with SPS modulation.

Finally, considering current ratings, size, cost, and AC losses, the choice of the inductor can be selected from manufacturers, e.g., Coil Craft XEL6060. To take into account the DC resistance present in the inductors with respect to the inductance value, a linearization of its $R_{dc}$ vs. $L_{ac}$ is performed, as shown in Figure 6. The linear function is presented in (7).

$$R_{Lac_{dc}} = 2.9 \times 10^3 \times L_{ac} + 0.00024 \qquad (7)$$

Following from its spice model given by the manufacturer, the series AC resistance can be expressed as in (8).

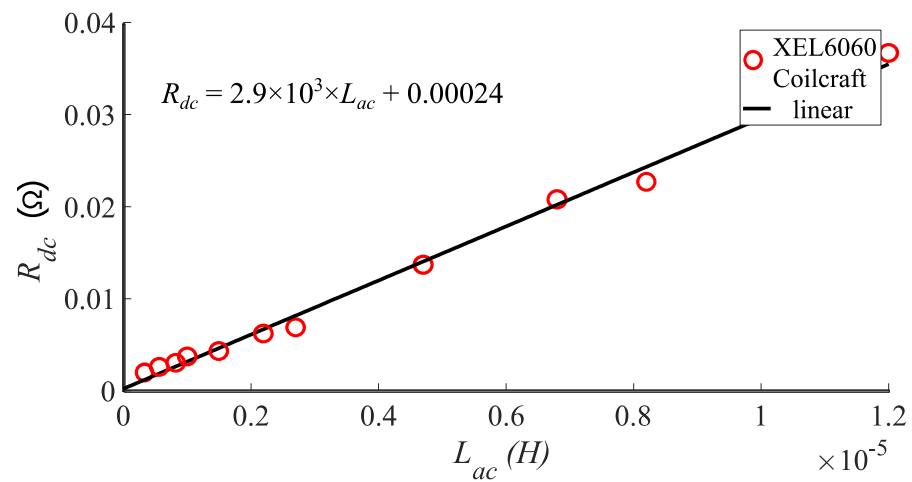$$R_{Lac_{ac}}(f_s) = R_{Lac_{dc}} + 10^{-6}\sqrt{f_s} \qquad (8)$$

**Figure 6.** Linearization of the DC resistance with respect to inductance found on Coilcraft's XEL6060 series.

Finally, substituting (8) in (7), and imposing the condition as in (6), the value of the AC resistance can be estimated with respect to the switching frequency in (9). The result is presented in Figure 7. Notice that this result depends on the technology of the inductor, that is, its DC and AC performances.

$$R_{Lac_{ac}}\left(L_{ac} = \frac{0.5}{f_s}\right) = \left(2.9 * 10^3 \frac{0.5}{f_s} + 0.00024\right) + 10^{-6}\sqrt{f_s} \tag{9}$$
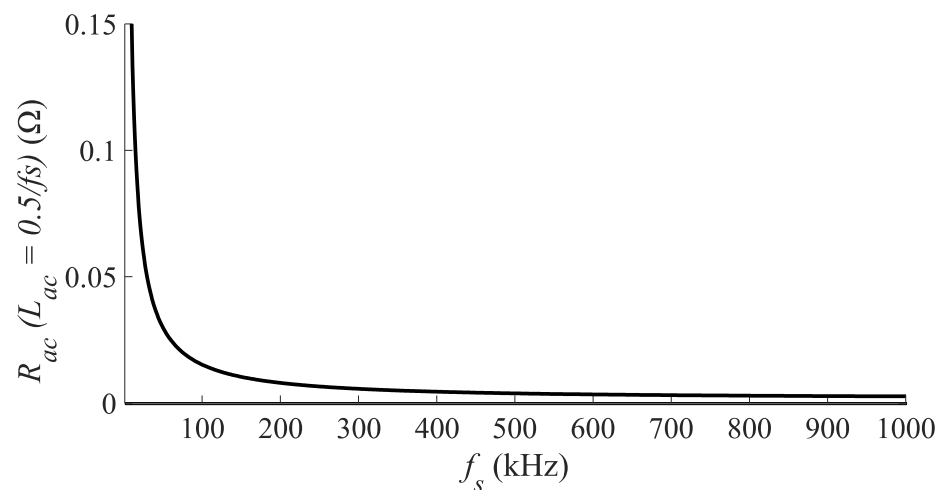


**Figure 7.** Estimated AC resistance considering the XEL 6060 series from Coil Craft by imposing $L_{ac} = 0.5/f_s$.

Then, after summing the losses on the MOSFETs, $C_{ac}$, and $L_{ac}$, the results are plotted with respect to the switching frequency as presented in Figure 8. Notice that an optimum switching frequency is found around 100 kHz. However, the losses on the transformer will be presented and included in Section 3.
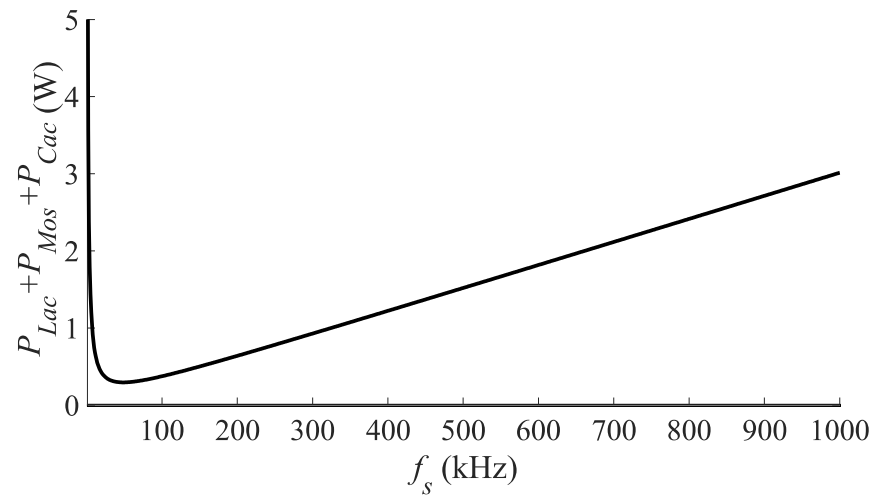
**Figure 8.** Estimated total losses with respect to switching frequency fs regardless the transformer.

## 3. Planar Transformer

In complement to the extra losses, the transformer plays a key role when it comes to weight and volume [23] in the design of the DAB. Even though loss sharing is not as significant, in an optimized design presented in [8], it occupies around 41% of the volume shares. In the work presented here, its volume represents about 45%, while its mass around 80% in the existing hardware prototype available in the laboratory from previous work [11,12]. In [10], a methodology is presented for designing a high frequency wire wound transformer for a DAB with SPS modulation 2.2 kW, by previously fixing one switching frequency design at 40 kHz. Therefore, the figure of merit of design is not clear with respect to power density and efficiency. The advantage of planar transformers [24,25] when compared to wire wound transformers is good thermal characteristics (due to higher surface area), easy assembly in power converters, low profile (e.g., by 37.3% as well as loss reduction [26]), predictability, repeatability, and cost reduction under certain conditions. Some disadvantages, however, include a larger footprint and higher winding capacitances.

Two different winding technologies may be used when prototyping a planar transformer: one made of PCB (printed circuit board), and another one, named copper foil. In Figure 9, the main aspects concerning the PCB winding technology are presented, which will be used through this work for result evaluation.
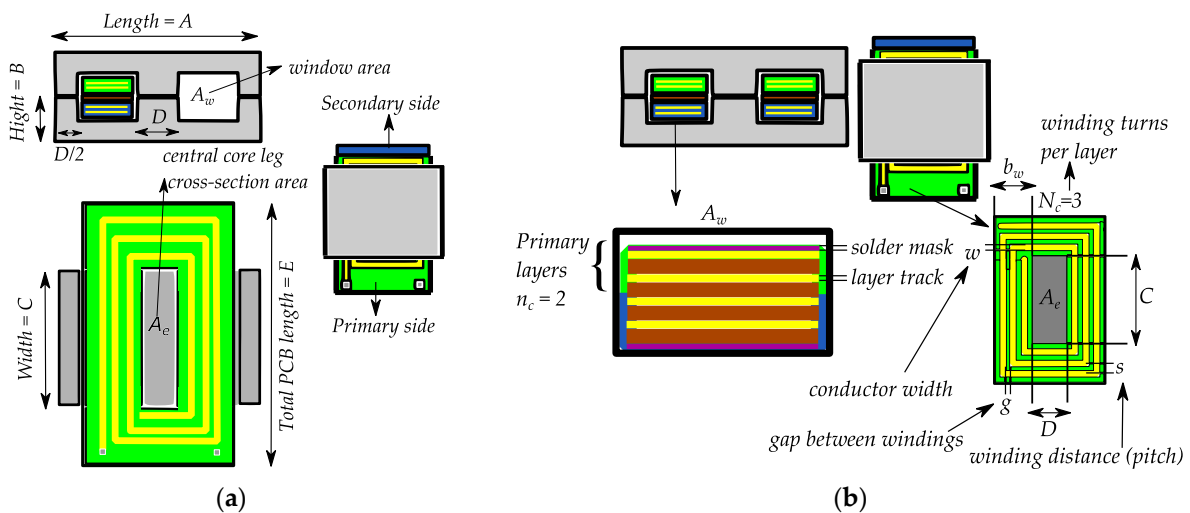


**Figure 9.** Important aspects in a EE planar transformer: (**a**) Core geometry; (**b**) Winding aspects (example of a 6:6 planar transformer).

*3.1. Classical Transformer Design*

A classical equation, used throughout industry and academia for choosing in a practical way the right ferrite core is based on the area product [27], which is the product between the window area $A_w$ and the magnetic cross section area $A_e$: $A_e A_w$ for forward converter types, as shown in (10). Such an equation allows to easily find the core provided on the manufacturer datasheets.

$$A_e A_{w_{\min}} = \frac{S_{max}}{K_r K_f J_{max} B_{peak} f_s} \tag{10}$$

where:

- $S_{max}$ is the total apparent power
- $B_{peak}$ is the magnetic induction (T), usually saturates at 0.3~0.4 T, but is often limited by the power losses;
- $J_{max}$ is the current density (A/cm$^2$)
- $K_r$ is the fulfill factor = 0.1~0.2 defined by the winding technology and isolation needs. As in [24], it is adopted 0.14 and can be later verified at the end of the design;
- $K_f$ is the form factor defined by the waveform in the input of the transformer (equal to 4 due to DAB a square waveform and 4.44 for sinusoidal one);

Notice that, depending on the choice of the current density $J_{max}$, magnetic induction $B_{max}$, and switching frequency $f_s$, different area products can be chosen, predefining the geometry and volume of the transformer, which are important parameters for efficiency, and power density. To calculate the efficiency, two losses occur inside the transformer, namely core and copper losses.

Normally, the core losses are calculated using the Steinmetz method, in which the manufacturer provides empirical parameters to be used in the Steinmetz equation adapted for a square waveform with 50% duty cycle commonly found on full-bridge DC-DC converter [28]. In this work, the employed parameters α, β, and *c* of a 3F3 ferrite were provided by the manufacturer, FerroxCube [22], and are mostly dependent on the core temperature and $f_s$. Commonly, for ferrite materials, lower temperatures imply higher losses, and for guaranteeing high efficiency, a temperature of 25 °C [23] was used throughout this equation. The total core loss is then presented in (12) by multiplying by the core volume $V_e$.

$$P_{core/V_e} = 2^{2\alpha-1} 0.5^{\beta-\alpha+1} c f_s{}^{\alpha} B_{peak}{}^{\beta} \tag{11}$$

$$P_{coreTotal} = P_{cor/Ve} \times V_e \ [\text{W}] \tag{12}$$

Regarding copper losses, the eddy losses are calculated based on Dowell factor and the general field solutions [25] defined in (13) and (14).

$$P_{winding} = 2 I_{RMS}{}^2 R_{dc} F_r [\text{W}] \tag{13}$$

$$F_r = \frac{R_{ac}}{R_{dc}} = \frac{\zeta}{2} \left[ \frac{\sinh\zeta + \sin\zeta}{\cosh\zeta - \cos\zeta} + (2m-1)^2 \frac{\sinh\zeta - \sin\zeta}{\cosh\zeta + \cos\zeta} \right] \tag{14}$$

where $\zeta$ is the ratio $h/\delta$, $\delta$ is the skin effect defined in (15), and *m* is the ratio defined by the MMFs at the limits of the layers in which will depend on the interleaving arrangement technique. The value of *m* = 1 is the choice for this work.

$$\delta = \sqrt{\frac{\rho_{cu}}{\pi f_s \mu}} \tag{15}$$

The DC resistance basic equation is presented in (16). However, besides depending on the number of turns and core geometry, it will also depend on the winding technology. For example, in Tables 2 and 3, the presented design guide and rules for PCB fabrication are from Wurth Elektronik [29]. The choice of them will impact on the efficiency as well as manufacturing costs.

**Table 2.** PCB main guideline and rules available in [29].

| Outer layers/Inner Layers Spacing (μm) | Min. Finished Copper (μm) ($h_c$ = 18 μm) | Min. Finished Copper (μm) ($h_c$ = 35 μm) | Min. Finished Copper (μm) ($h_c$ = 70 μm) | Min. Finished Copper (μm) ($h_c$ = 105 μm) |
|---|---|---|---|---|
| Track-Track = $g_{min}$ | ≥85 | ≥100 | ≥192 | ≥250 |
| Track width = $w_{min}$ | ≥85 | ≥100 | ≥192 | ≥250 |

**Table 3.** Multilayer PCB technology available in [29].

| Layer Quantity $2n_c$ | PCB Thickness Possibility 0.80 mm | PCB Thickness Possibility 1.00 mm | PCB Thickness Possibility 1.55 mm | PCB Thickness Possibility 2.40 mm |
|---|---|---|---|---|
| 1 | Yes | Yes | Yes | Yes |
| 2 | Yes | Yes | Yes | Yes |
| 4 | Yes | Yes | Yes | Yes |
| 6 | No | Yes | Yes | Yes |
| 8 | No | No | Yes | Yes |
| 10 | No | No | Yes | Yes |

Next, the mean total length *MLT* winding length per layer is applied in the expression (16). The cupper surface area is presented in (17).

$$R_{dc} = \frac{\rho_{cu} MLT . N_c n_c}{S_{Cu}} \ [\Omega] \tag{16}$$

$$S_{Cu} = w \times h_c \tag{17}$$

where:

$$N_c n_c = N_p = \left( \frac{V_i}{4 f_s A_e B_{peak}} \right) \tag{18}$$

$$w = \frac{I_{rms}}{J_{max} h_c} = \frac{b_w - (N_c + 1) g_{min}}{N_c} \geq w_{min} \tag{19}$$

According to [30], the temperature rise in the core can be related to the exterior surface, as presented in (20), relating the thermal resistance $R_{th}$ and the total core loss. For the temperature rise in the winding, the IPC-2152 could be used. However, as shown in [31], large deviation on the winding temperature rises from IPC-2152 compared to FEA (finite element analysis) simulations can be expected. Two conclusions should be pointed out: (i) IPC-2152 is too conservative and should not be used for optimization; (ii) the maximum temperature in the core and the winding are almost the same when the losses are evenly distributed. This provides evidence that planar transformer presents elevated heat conduction between the core and the PCB winding. Therefore, keeping the same power losses on winding and core is a good practice in terms of thermal prediction.

$$\Delta T_{Core} = R_{th} P_{coreTotal} = \frac{1}{h_{ext} S_{ext}} P_{coreTotal} \tag{20}$$

*3.2. Algorithm for Classical Transformer Design*

Based on the previous section, a classical transformer design is implemented. The first step is to identify the variables present in the system. Fixed electrical variables: ($V_i$ = 20, $V_o$ = 20, $I_o$ = 2A, $I_{rms}$ = 1.1 $I_o$)

- Fixed PCB variables ($h_c$ = 105 μm, $n_c$ = 4, $H_c$ = 2 mm, g = $g_{min}$ = 192 μm, $w_{min}$ = 192 μm)
- Design manufacture variables of magnetic cores (EE $A/B/C$) are presented in Table 4.

**Table 4.** *EE model* from Ferroxcube [22].

| EE model | A [mm] | B [mm] | C [mm] |
|----------|--------|--------|--------|
| 1 | 14 | 3.5 | 5 |
| 2 | 18 | 4 | 10 |
| 3 | 22 | 6 | 16 |
| 4 | 32 | 6 | 20 |
| 5 | 38 | 10 | 25 |
| 6 | 43 | 11 | 28 |
| 7 | 58 | 11 | 38 |
| 8 | 64 | 10 | 50 |

Design variables ($\Delta T_{Core} = 50°, fs, N_c = 1,2,3 \dots 10$)
Then, the domain, objective, and constraint functions:

- Variable domain:

$$
\begin{cases}
EE\ 14/3.5/5 \dots EE\ 64/10/50 \Rightarrow 300 < V_e < 40,700 [\text{mm}^3] \\
100 \leq f_s \leq 1000\ [\text{kHz}] \\
B_{peak} \leq B_{Sat} = 0.4\ [\text{T}]
\end{cases}
\tag{21}
$$

- Objective function:

$$
\min\left\{\sum\left(P_{Lac} + P_{Mos} + P_{Cac} + P_{core} + P_{winding}\right)\right\}
\tag{22}
$$

- Constraint

$$
\begin{cases}
\text{a) Temperature}: R_{th}P_{coreTotal} \leq \Delta T_{CoreMax} = 50 \Rightarrow \text{This might eventually limit } N_{c_{min}} \\
\text{b) Magnetic saturation}: B_{peak} \leq B_{Sat} = 0.4\ T \Rightarrow N_{c_{min}} \geq \frac{V_i}{4f_s A_e n_c B_{Sat}} \\
\text{c) Winding PCB}: w \geq w_{min} \Rightarrow \frac{b_w - (N_c+1)g}{N_c} \geq w_{min} \Rightarrow N_{c_{max}} \leq \frac{b_w - g_{min}}{w_{min} + g_{min}}
\end{cases}
\tag{23}
$$

To support a fast convergence, the parameter $B_{peak}$ was expressed as a function of discrete values: number of turns $N_c$, and $Ae$ from EE A/B/C, leaving only $fs$ as a variable to be optimized in an outer loop containing $N_c$ iterations. From (23), it is possible to delimit from b) and c) the range of $N_c$. However, one must pay attention that such aa range can be limited by temperature constraints a), while the Joule losses will normally limit the maximum amount of $N_c$.

In addition, the algorithm intends to let the designer manage the dimensioning process, and therefore visualization and comprehension of many designs is possible rather than letting the machine produce one single result without important physical and engineering understanding, which can be essential in terms of other aspects, such as costs or potential environmental impact generation (LCA), as will be demonstrated in Section 3.1.

The results of the total power losses, and optimum $fs$ could then be expressed in $N_c$ vs. *EE model*, as presented in Figures 10 and 11.
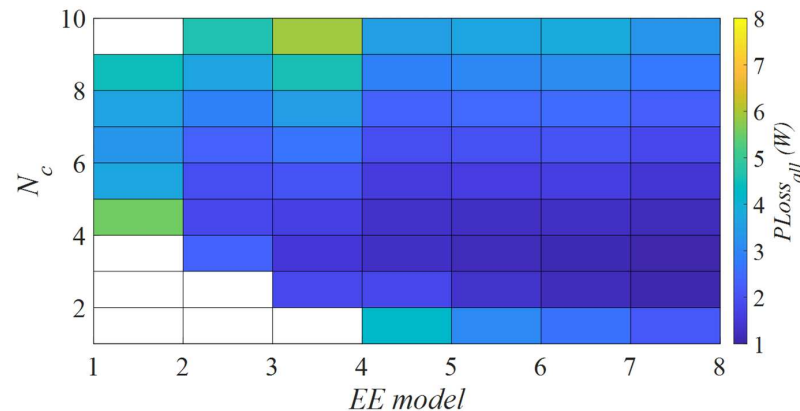


**Figure 10.** Total power losses resulted by the optimizing algorithm as function of $N_c$ vs. *EE model*.
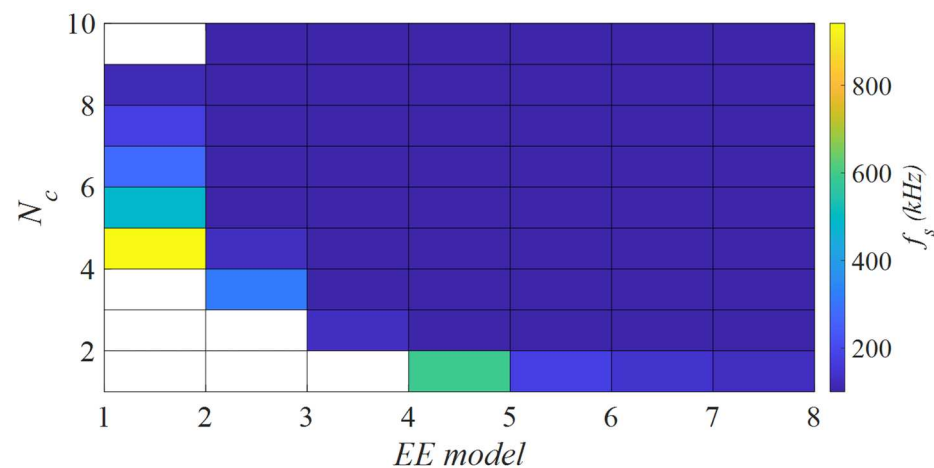
**Figure 11.** Resulting optimum *fs* as function of $N_c$ vs. *EE model*.

The regions in blank mean that no solution is possible for the constraints presented in (23), which is normal considering the small $A_e A_w$ that they present. From each *EE model*, it is possible to choose the $N_c$ that results in the minimum total power loss. It is possible to notice that, as presented in the literature, larger transformers will result in fewer power losses with the drawback of rapidly increasing volume (from 300 mm$^3$ to 40,700 mm$^3$). It also confirms that the discussion concerning efficiency versus power density is also found in a planar transformer, which goes along with the results in [32]. Based on this conclusion, the model 8 Ferroxcube E58/11/28 would be chosen in an "efficiency oriented" design. However, considering subjective Pareto analysis between efficiency and power density, it could result in smaller models.

As shown on Figure 11 the lower frequencies will mainly dominate in larger models, as the total loss is increasing with *fs*. The frequencies are therefore limited to 100 kHz. However, the inner borders will result in different frequencies due to important constraints that force the frequency to be larger than the minimum required one. The choice of PCB technology ($n_c$, $h_c$, $H_c$) makes the results vary as well.

### 3.3. Towards a Whole Life Cycle Energy and Material Impact Assessment of the Transformer

This section addresses the importance of evaluating the transformer on its whole life cycle. To perform this task, a simplified environmental impact assessment has been conducted. This streamlined approach presented aims at providing some first estimations based on a formula adapted for a transformer, to prepare the full life cycle assessment (not presented in this section). The life cycle assessment (LCA) is a method for quantifying the environmental impact of products, introduced in 1997 and framed by ISO 14040 [33]. The idea is to evaluate each process unit included in the life cycle scope of the converter to multiple environmental impacts (local and global). Global impacts include, e.g., global warming (in $CO_2$ eq.) and the ozone layer depletion, while local impacts include some local pollution in the air/water/soil (acidification, eutrophication, etc.), land use, etc. LCA tools are based on the ISO standard 14040&44 and integrate important databases that enable the aggregation of the primary flows (material, energy) included in the unit processes modeling the product life cycle (based on a functional unit, reference flows, in a defined perimeter) and the resulting potential environmental impact estimations. The environmental impact calculation is supported by several environmental impact methods (e.g., CML, Ecoindicator 99, Impact 2002+) that are discussed and improved by the worldwide LCA community.

For the perimeter of this study, the streamlined life cycle of the transformer has been divided into:

- The extraction of raw materials used for the electronic components.
- The production processes of transformer components (including the sub-systems).
- The assembly processes.

- The use stage.
- The separation process of the components; with reuse and recycling fraction end-of-life treatment scenarios (based on existing electronic waste stream states).
- The main transport stages (between each previously mentioned step).

In this study, the streamlined environmental impact assessment in this perimeter is focused on providing a first estimation of the potential embodied energy, approached through an energy cost analysis. This input flow analysis compiles the cost flow associated to the life cycle inventory to estimate the resulting embodied energy. In a line of similar reasoning to input-output flow analysis methods used in LCA, but streamlined at first, every step of the life cycle of the transformer included in the defined perimeter consumes an amount of energy, as expressed in (24). The equation is split in two terms. The term on the left is related to the energy cost of manufacturing, transport, and end of use scenarios, while the term on the right is the energy loss related to a certain given task (the energy spent during usage). This proposition allows the designers of transformer to estimate the embodied energy in regard to the technical functions and performances of the system. A cause to consequence reasoning is therefore made possible for the electronic designer. The embodied energy estimation factors may require some information exchanges with the LCA expert that has access to the full LCA software, i.e., databases (including data uncertainty) and calculations methods (including the characterization factors to estimate the environmental impacts indicator contributions, and uncertainty life cycle analysis).

$$E_{acv} = \sum M_m.E_{p,m} + \int_{use} P_{fct} dt \tag{24}$$

The $E_{p,m}$ are the energy costs of each unit process of transformation of a given kilogram of material. The latter is obtained from LCA inventory data and databases such as Ecoinvent©, ELCD (European Reference Life Cycle Database), electronic and electrical, associated with LCA software from the BaseImpact® of the French agency ADEME, including the end-of-life current scenarios and waste stream evaluation reports, and the European Union from the EUROSTAT annual report. The $P_{fct}$ are the energy losses of the converter during its use.

The manufacturing process of the transformer is divided into three main stages plus transport:

- Extraction of the raw material.
- Manufacture of the components.
- Assembly processes.
- Transport

The end of life is divided into several scenarios including transport again:

- Recycling (material processing to obtain the same or lower quality).
- Upcycling (partial reuse of part of the converter or some components in functional condition).
- Re-use (re-use of the components depending on its condition in a new converter system).

Note that reuse and upcycling would potentially be possible on a large scale because the conversion bricks are standardized in this modular design. The conversion bricks would retain their own value for a new use. However, this re-use scenario is prospective. This model cannot be based on existing data, as it is not yet operated by the current electronic repair-reuse scheme (as in most of Europe for instance). Again, establishing some design scenario related to other life cycle stage (usage/end-of-life) supports the electronic designers' capacity to estimate (roughly) the potential embodied energy losses in its design choices.

The embodied energy of each unit modeled in the life cycle perimeter proposed in this research follows:

$$E_i = M_m.E_{unit} \tag{25}$$

where $E_i$ is the energy per kg of a transformation process of a material stream of mass $M_m$.

The materials extracted and used in the transformer are included in Figure 12. They are all taken into account in the quantification of the energy cost of extraction of the raw material. Hence, minimizing the usage losses and the embodied energy during the whole life cycle can be expressed through the following objective function:

- Objective function:

$$\min\left\{\sum\left(P_{joule} + P_{core} + P_{winding}\right) \times years + E_{LCA}\right. \tag{26}$$
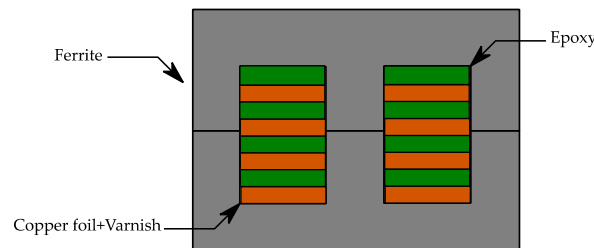


**Figure 12.** A simplified representation of planar transformer for Bill Of Material (BOM), or life cycle inventory in the streamlined perimeter of this study.

The parameter of the "streamlined LCA" is calculated over a simplified design version of the transformer (that can evolve during the design process development of the transformer to reach the best design choices compromise):

To perform the streamlined life cycle environmental impact, essentially based on an energy efficiency analysis, including the embodied energy associated to the material flow and the process included at the stages included in the study perimeter, the following parameters are taken into account, extracted from the databases presented in Table 5:

$$\begin{cases} E_{LCA} = E_{Ferrite} + E_{Copper} + E_{Epoxy+Varnish} \\ E_{Copper} = E_{Extraction+Production(FeNi)} + E_{Transport} \\ E_{Ferrite} = E_{Extraction+Production(Cu)} + E_{Transport} \\ E_{Epoxy+Varnish} = 31.4\,\text{kWh/kg} \\ E_{Transport} = 0.6\,\text{kWh/kg} \\ E_{Extraction+Production(Cu)} = 8.7\,\text{kWh/kg} \\ E_{Extraction+Production(FeNi)} = 37\,\text{kWh/kg} \\ E_{EndOfLife(WEEE)} = 31.4\,\text{kWh/kg} \\ E_{Assembly} = 27\,\text{kWh/kg} \end{cases}$$

**Table 5.** References for the streamlined calculation of the embodied energy: factors are extracted from calculation methods (IDEMAT datasheet, from the Idematapp 2021 cvs, cf. www.ecocostsvalue.com, and www.idematapp.com (accessed on 13 September 2021). databases used for the life cycle inventory stage (Bill of Material).

| Raw Materials' Extraction | Eurostat, Ecoinvent V3.6©, [34,35] |
|---|---|
| Transports | EcoTransit from the website (accessed Sept. 2021): https://www.ecotransit.org/fr (accessed on 13 September 2021), datasheet MURATA, [36,37] |
| Manufacturing | Ecoinvent V3.6© databases. |
| Usage-Maintenance | MIL-HDBK, EIME LCA software [38], MURATA database |
| Recycling and reuse | Ecoinvent©, Base impact® of the ADEME online tool (available on: www.base-impacts.ademe.fr, accessed on 13 September 2021) |

The algorithm is then run with respect to expected total converter's lifetime. Depending on components' MTBF, but also converter expected usage duration depending on application, the total lifetime can be set. For the purpose of the design optimization, a

total lifetime of ten years is selected at first. The results of the total energy losses over ten years and optimum of life cycle energy cost is expressed in $N_c$ vs. *EE model*, as presented in Figure 13.
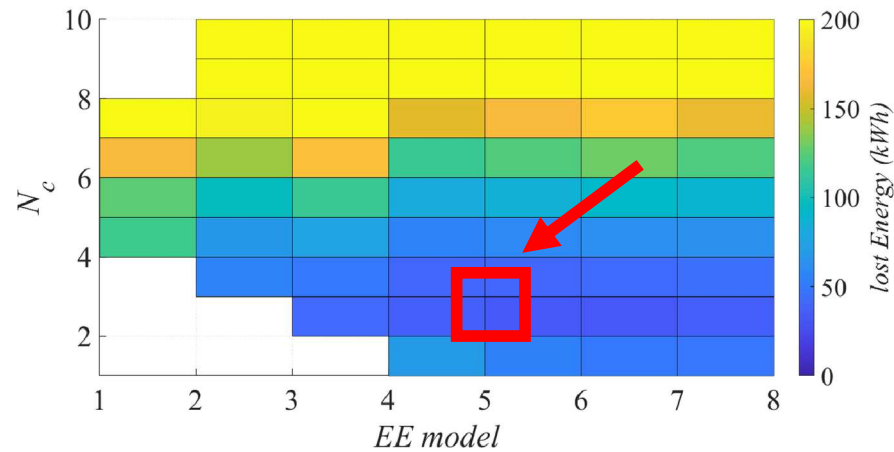


**Figure 13.** Usage energy losses and life cycle energy cost depending on the number of turns and the ferrite model.

It is noticeable that the optimum design for the transformer taking into account its whole life cycle shifts more to the left to a smaller ferrite from the Ferroxcube E58/11/28 to the E38/10/28.

An eco-efficient analysis aims to minimize the cost of energy by reducing the volume, but the electrical constraints and the design rules of the transformer push the optimum more to the right (wider transformers), which results in a possible minimum that satisfies both the energy efficiency performances and the minimization of the whole life cycle energy. The full LCA is necessary to evaluate scientifically the best life cycle design scenarios in regard to energy efficiency and other environmental impacts.

While the other indicators, such as the toxicity, ozone depletion, or land use (etc.), are important to take into account to have a global view over all the environmental impacts, it is still very difficult to implement a full LCA and more importantly to have a common indicator be easily understood by the electrical engineering community. The formulated model can be used for a first streamlined estimation in the early stage of the design process of such power electronics.

## 4. Results and Discussion

When plotting the optimized total energy lost during the usage and the embodied energy related to the manufacture, transport, and end-of-life treatments of the transformer with respect to the total lifetime, the two curves cross for a lifetime of seven years, as presented in Figure 14. Choosing a smaller transformer model E32/6/20 can therefore be a solution to reduce the energy losses, when the components around it have a short life or the whole device has a shorter life than the transformer. Nevertheless, if the device is robust and can be guaranteed to function without failure for approximately 15 years or more, then it would be interesting to over-size the transformer performance to limit the energy losses during the usage of the converter.

In our modern societies where many of our products are designed to last between five and 10 years (electric mobility such as electric bikes and scooters) or because those products will be obsolete within 2–5 years (computers, mobiles), the eco dimensioning could drive the design "naturally" toward more compact solutions, toward a higher global power density, together with a higher total energy efficiency. In this eco-design optimization, both efficiency and power densities are heading in the same direction and based on factual criteria. It is interesting to underline that the eco-design efforts could provide a complementary justification to push further power electronics designs toward

higher power densities and efficiencies, providing additional motivation for the community to keep improving both aspects in power converters.
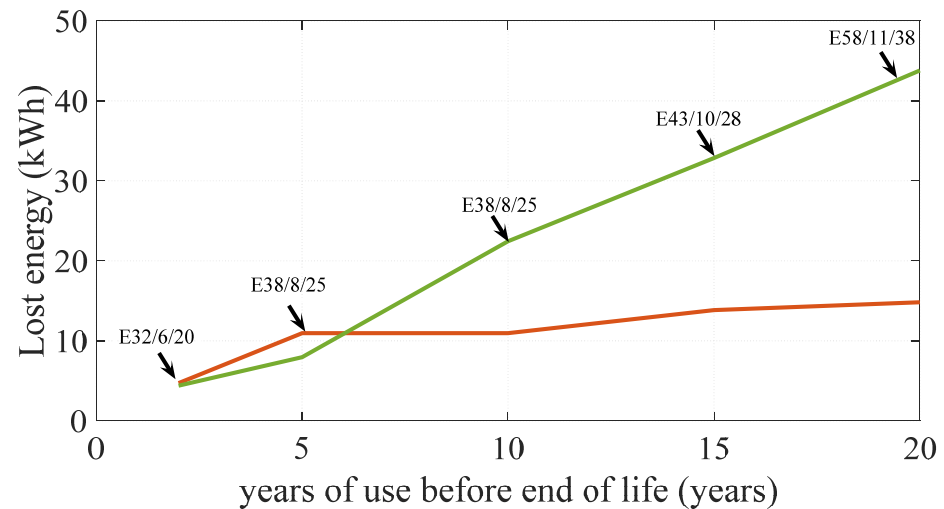


**Figure 14.** In Green the energy losses during the usage, in orange the energy losses during the other steps of the life cycle of the transformer in kWh, for example if the device has a lifetime of 5 years, the E38/8/25 Ferrite model is optimum when it comes to the whole life cycle cost.

It is important to point out here that the LCA analysis and the eco-design optimization have been partially carried out, leaving important additional work to be done taking into account the other components. Nonetheless, in DAB based converters, the transformer remains a key element to be optimized with respect to environmental issues since it represents the largest parts of the volume and weight of the whole converter. Further work will have to be carried out to check if this assumption remains valid if a full LCA analysis is done.

## 5. Conclusions

A methodology for easily designing the main magnetic devices found in a DAB with SPS modulation was proposed. This methodology allows an efficiency and power density maximization process, while ZVS is rather guaranteed for the sake of efficiency and loss of model accuracy. The transformer becomes the central key device, while other elements are designed around it.

In a conventional method, an "only" energy efficiency oriented would result in targeting the largest transformer size, while in a power density-oriented analysis the smallest transformer size would be best selected. The conventional method for determining the transformer size and the switching frequency is, therefore, most of the time subjective. On the other hand, the methodology proposed in this research allowed finding, objectively, the optimum switching frequency from the planar transformer design optimization, that can be materially and energy efficiency oriented, or even better, based on a LCA results orientation.

By following the methodology of a planar transformer for a DAB, the conventional and the proposed, environmentally (energy–material) oriented, optimized life cycle were compared. Results suggest that an increase in volume becomes interesting when the whole converter is robust enough and when its usage is guaranteed over a long period. If the case differs, the designer should choose a smaller transformer responding to the electrical constraints considered.

## References

1. Kolar, J.W.; Drofenik, U.; Biela, J.; Heldwein, M.L.; Ertl, H.; Friedli, T. Round SD PWM Converter Power Density Barriers. In Proceedings of the 2007 Power Conversion Conference, Nagoya, Japan, 2–5 April 2007; pp. P-9–P-29. [CrossRef]

2. Friedrichs, P.; Millan, J.; Harder, T.; Kaminski, N.; Lindemann, A.; Lorenz, L.; Schindele, L.; Ward, P. *ECPE Position Paper on Next Generation Power Electronics Based on Wide Bandgap Devices*; ECPE: Nuremberg, Germany, 16 May 2016; pp. 1–17.

3. Kolar, J.W.; Biela, J. Mini Book, Exploring the pareto front of multi-objective single-phase PFC rectifier design optimization—99.2% efficiency vs. 7 kW/dm$^3$ power density. In Proceedings of the 2009 IEEE 6th International Power Electronics and Motion Control Conference IPMC, Wuhan, China, 17–20 May 2009; pp. 1–21. [CrossRef]

4. Kasper, M.; Chen, C.W.; Bortis, D.; Kolar, J.W.; Deboy, G. Hardware verification of a hyper-efficient (98%) and super-compact (2.2 kW/dm$^3$) isolated AC/DC telecom power supply module based on multi-cell converter approach. In Proceedings of the 2015 IEEE Applied Power Electronics Conference and Exposition (APEC), Charlotte, NC, USA, 15–19 March 2015; pp. 65–71. [CrossRef]

5. Kolar, J.W.; Krismer, F.; Lobsiger, Y.; Muhlethaler, J.; Nussbaumer, T.; Minibock, J. Extreme efficiency power electronics. In Proceedings of the 2012 7th International Conference on Integrated Power Electronics Systems (CIPS), Nuremberg, Germany, 6–8 March 2012; pp. 1–22.

6. Schrittwieser, L.; Leibl, M.; Kolar, J.W. 99% Efficient Isolated Three-Phase Matrix-Type DAB Buck–Boost PFC Rectifier. *IEEE Trans. Power Electron.* **2020**, *35*, 138–157. [CrossRef]

7. Kolar, J.W.; Biela, J.; Waffler, S.; Friedli, T.; Badstuebner, U. Performance trends and limitations of power electronic systems. In Proceedings of the 2010 6th International Conference on Integrated Power Electronics Systems, Nuremberg, Germany, 16–18 March 2010; IEEE: Piscataway, NJ, USA, 2010.

8. Burkart, R.M.; Kolar, J.W. Comparative η–ρ–σ Pareto Optimization of Si and SiC Multilevel Dual-Active-Bridge Topologies With Wide Input Voltage Range. *IEEE Trans. Power Electron.* **2017**, *32*, 5258–5270. [CrossRef]

9. De Freitas Lima, G.; Lembeye, Y.; Ndagijimana, F.; Crebier, J.-C. Modeling of a DAB under phase-shift modulation for design and DM input current filter optimization. In Proceedings of the 2020 22nd European Conference on Power Electronics and Applications (EPE'20 ECCE Europe), Lyon, France, 7–11 September 2020; pp. P.1–P.10. [CrossRef]

10. Hoang, K.D.; Wang, J. Design optimization of high frequency transformer for dual active bridge DC-DC converter. In Proceedings of the 2012 XXth International Conference on Electrical Machines, Marseille, France, 2–5 September 2012; pp. 2311–2317. [CrossRef]

11. Crébier, J.C.; Phung, T.H.; Nguyen, V.S.; Lamorelle, T.; Andreta, A.; Kĕachev, L.; Lembeye, Y. DC-AC Isolated Power Converter Array. Focus on Differential Mode Conducted EMI. *Electronics* **2019**, *8*, 999. [CrossRef]

12. Andreta, A.; Lavado Villa, L.F.; Lembeye, Y.; Crebier, J.C. A Novel Automated Design Methodology for Power Electronics Converters. *Electronics* **2021**, *10*, 271. [CrossRef]

13. Herrera-Jaramillo, D.A.; Henao-Bravo, E.E.; González Montoya, D.; Ramos-Paja, C.A.; Saavedra-Montes, A.J. Control-Oriented Model of Photovoltaic Systems Based on a Dual Active Bridge Converter. *Sustainability* **2021**, *13*, 7689. [CrossRef]

14. Henao-Bravo, E.E.; Ramos-Paja, C.A.; Saavedra-Montes, A.J.; González-Montoya, D.; Sierra-Pérez, J. Design Method of Dual Active Bridge Converters for Photovoltaic Systems with High Voltage Gain. *Energies* **2020**, *13*, 1711. [CrossRef]

15. Carvalho, E.L.; Meneghetti, L.H.; Carati, E.G.; da Costa, J.P.; de Oliveira Stein, C.M.; Cardoso, R. Asymmetrical Pulse-Width Modulation Strategy for Current-Fed Dual Active Bridge Bidirectional Isolated Converter Applied to Energy Storage Systems. *Energies* **2020**, *13*, 3475. [CrossRef]

16. Choi, H.-j.; Lee, W.-b.; Jung, J.-h. Practical Controller Design of Three-Phase Dual Active Bridge Converter for Low Voltage DC Distribution System. *Electronics* **2020**, *9*, 2101. [CrossRef]

17. Shamshuddin, M.A.; Rojas, F.; Cardenas, R.; Pereda, J.; Diaz, M.; Kennel, R. Solid State Transformers: Concepts, Classification, and Control. *Energies* **2020**, *13*, 2319. [CrossRef]

18. Calderon, C.; Barrado, A.; Rodriguez, A.; Alou, P.; Lazaro, A.; Fernandez, C.; Zumel, P. General Analysis of Switching Modes in a Dual Active Bridge with Triple Phase Shift Modulation. *Energies* **2018**, *11*, 2419. [CrossRef]

19. Lei, T.; Wu, C.; Liu, X. Multi-Objective Optimization Control for the Aerospace Dual-Active Bridge Power Converter. *Energies* **2018**, *11*, 1168. [CrossRef]

20. Li, J.; Luo, Q.; Mou, D.; Wei, Y.; Sun, P.; Du, X. A Hybrid Five-variable Modulation scheme for Dual Active Bridge Converter with Minimal RMS Current. *IEEE Trans. Ind. Electron.* **2021**, *69*, 336–346. [CrossRef]

21. Krismer, F.; Kolar, J.W. Efficiency-Optimized High-Current Dual Active Bridge Converter for Automotive Applications. *IEEE Trans. Ind. Electron.* **2012**, *59*, 2745–2760. [CrossRef]

22. Ferroxcube Soft ferrites and accessories data handbook 2013 and Application Note: Design of Planar Power Transformers. Available online: http://ferroxcube.home.pl/ (accessed on 8 September 2021).

23. Friedemann, R.A.; Krismer, F.; Kolar, J.W. Design of a minimum weight dual active bridge converter for an Airborne Wind Turbine system. In Proceedings of the 2012 Twenty-Seventh Annual IEEE Applied Power Electronics Conference and Exposition (APEC), Orlando, FL, USA, 5–9 February 2012; pp. 509–516. [CrossRef]

24. Magambo, J.S.; Bakri, R.; Margueron, X.; Le Moigne, P.; Mahe, A.; Guguen, S.; Bensalah, T. Planar Magnetic Components in More Electric Aircraft: Review of Technology and Key Parameters for DC–DC Power Electronic Converter. *IEEE Trans. Transp. Electrif.* **2017**, *3*, 831–842. [CrossRef]

25. Ouyang, Z.; Andersen, M.A.E. Overview of Planar Magnetic Technology—Fundamental Properties. *IEEE Trans. Power Electron.* **2014**, *29*, 4888–4900. [CrossRef]

26. Zhang, K.; Wu, T.X.; Kutkut, N.; Shen, J.; Woodburn, D.; Chow, L.; Wu, W.; Mustain, H.; Batarseh, I. Modeling and design optimization of planar power transformer for aerospace application. In Proceedings of the IEEE 2009 National Aerospace & Electronics Conference (NAECON), Dayton, OH, USA, 21–23 July 2009; pp. 116–120. [CrossRef]

27. Gu, W.J.; Liu, R. A study of volume and weight vs. frequency for high-frequency transformers. In Proceedings of the IEEE Power Electronics Specialist Conference (PESC '93), Seattle, WA, USA, 20–24 June 1993; pp. 1123–1129. [CrossRef]

28. Ouyang, Z.; Thomsen, O.C.; Andersen, M.A. Optimal Design and Tradeoff Analysis of Planar Transformer in High-Power DC–DC Converters. *IEEE Trans. Ind. Electron.* **2012**, *59*, 2800–2810. [CrossRef]

29. WEdirekt Design Guide for PCBs from the Online Shop. Available online: https://www.wedirekt.com/web/media/16_wedirekt/spezifikationen/leiterplatten_daten_downloads/WEdirekt_Design_Guide_EN.pdf (accessed on 15 September 2021).

30. Bakri, R.; Ngoua Teu, J.S.; Margueron, X.; Le Moigne, P.; Idir, N. Planar transformer equivalent thermal resistance variation with ambient temperature and power losses. In Proceedings of the 2016 18th European Conference on Power Electronics and Applications (EPE'16 ECCE Europe), Karlsruhe, Germany, 5–9 September 2016; pp. 1–9. [CrossRef]

31. Tria, L.A.; Zhang, D.; Fletcher, J.E. Electromagnetic and thermal characterisation of PCB planar transformer. In Proceedings of the 2015 IEEE 11th International Conference on Power Electronics and Drive Systems, Sydney, Australia, 9–12 June 2015; pp. 1024–1028. [CrossRef]

32. Bakri, R.; Margueron, X.; Le Moigne, P.; Idir, N. Dimensionnement de transformateurs planar sous contraintes thermiques. In Proceedings of the Symposium de Génie Electrique 2021 (SGE 2021), Nantes, France, 6–8 July 2021; Available online: https://uncloud.univ-nantes.fr/index.php/s/GbiMj8AZxCxjDPR?dir=undefined&path=%2FSP-2%20-%20Mercredi%207_7&openfile=551212549 (accessed on 11 September 2021).

33. ISO 14040:2006 Environmental Management—Life Cycle Assessment—Principles and Framework. Available online: https://www.iso.org/standard/37456.html (accessed on 16 September 2021).

34. Oguchi, M.; Murakami, S.; Sakanakura, H.; Kida, A.; Kameya, T. A preliminary categorization of end-of-life electrical and electronic equipment as secondary metal resources. *Waste Manag.* **2011**, *31*, 2150–2160. [CrossRef] [PubMed]

35. Reller, A. Criticality of metal resources for functional materials used in electronics and microelectronics. *Phys. Status Solidi (RRL)-Rapid Res. Lett.* **2011**, *5*, 309–311. [CrossRef]

36. Ninlawan, C.; Seksan, P.; Tossapol, K.; Pilada, W. The implementation of green supply chain management practices in electronics industry. In Proceedings of the World Congress on Engineering 2012, London, UK, 4–6 July 2012; Volume 2182, pp. 1563–1568.

37. Berry, D.; Towill, D.R.; Wadsley, N. Supply Chain Management in the Electronics Products Industry. *Int. J. Phys. Distrib. Logist. Manag.* **1994**, *24*, 20–32. [CrossRef]

38. Jaouen, C.; Multon, B.; Barruel, F. Design optimization methodology for power converters based on global energy requirement criteria. Application to a DC-DC flyback structure. In Proceedings of the 2011 14th European Conference on Power Electronics and Applications, Birmingham, UK, 30 August–1 September 2011; pp. 1–10.

# Analysis of Available Components and Performance Estimation of Optical Multi-Band Systems

**Dimitris Uzunidis** [1], **Fotini Apostolopoulou** [2], **Gerasimos Pagiatakis** [2,*] **and Alexandros Stavdas** [1]

1   OpenLightComm Ltd., Suffolk IP5 3RE, UK; duzunidis@openlightcomm.uk (D.U.); astavdas@openlightcomm.uk (A.S.)

2   Department of Electrical and Electronic Engineering Educators, School of Pedagogical & Technological Education (ASPETE), 14121 Athens, Greece; fot.apo@elec.aspete.gr

\*   Correspondence: pagiatakis@aspete.gr; Tel.: +30-210-289-6781

**Abstract:** Optical multi-band (OMB) systems exploit the low-loss spectrum of the single mode fiber (SMF) and are key enablers to increase the transportation capacity and node connectivity of already deployed systems. The realization of OMB systems is mainly based on the technological advances on the component and system level, and for this purpose, a broad gamut of various structural elements, such as transceivers, amplifiers, filters, etc. have been commercialized already or are close to commercialization. This wide range of options, which aid in unlocking the concurrent transmission in all amplification bands, is reviewed here for the first time, whilst their pros and cons as well as their limitations are discussed. Furthermore, the needs for additional components in order to fully exploit the ≈390 nm low-loss wavelength range of SMF, which spans from 1260 to 1650 nm, are highlighted. Finally, based on a physical layer formalism, which incorporates the impact of the most important physical layer constraints for an OMB system, the attainable capacity and transparent reach of each amplification band are quantified.

**Keywords:** optical multi-band systems; physical layer modeling; fiber-optic communications; commercially available components

## 1. Introduction

To meet the increasing demand for higher bandwidth and more dynamic traffic patterns of the emerging Enhanced Fiber Broad Band (eFBB) services, such as ultra-high-definition video streaming, cloud office, and virtual/augmented reality [1], the ultimate bandwidth of the SMF has to be exploited [2–4]. A potential solution to this issue is the optical multi-band (OMB) system, which can fully exploit the second and third transmission windows of SMFs (practically, the 1260 to 1650 nm wavelength range) and increase the capacity of currently deployed systems by as much as an order of magnitude. However, research on OMB systems is still in its infancy, and the related work is mainly experimental, focusing more on C and L-bands and lesser on O, E, and S-bands [5–9].

The realization of an OMB system depends on the availability of optical components, such as optical transceivers, amplifiers, filters etc., for a number of optical transmission bands. In this work, we tabulate the state-of-the-art details for these commercially and experimentally available structural elements, and we discuss their specifications and their impact on the overall system performance. We show that using the currently available technology, the deployment of an OMB system is feasible only for a confined spectral range, and an important amount of work has to be carried out, along with the introduction of novel optical components, to unlock the entire low attenuation range of the single-mode optical fiber.

Next, we employ a novel physical layer formalism, which calculates the performance of an OMB transmission system considering an amplification scheme based on commercially available rare-earth doped fiber amplifiers for the S, C and L-bands. The proposed

expression is closed-form and incorporates the impact of the main transmission impairments of an OMB system, which are the Amplified Spontaneous Emission (ASE) noise accumulation, the Four-Wave-Mixing (FWM) crosstalk, and the Stimulated Raman Scattering (SRS). Using this formula, we calculate the attainable capacity and transparent reach for each band and for three different modulation formats, namely polarization multiplexing (PM) quadrature phase shift keying (QPSK), PM-16 quadrature amplitude modulation (QAM), and 64QAM, discussing their performance trade-offs.

The paper is organized as follows: Section 2 includes a review of commercially available state-of-the-art components (such as transceivers, amplifiers, and filters), Section 3 attempts a quantification of the capacity and the transparent reach of an OMB system employing commercially available equipment, Section 4 discusses the main findings of our work, and Section 5 concludes the paper.

## 2. Review of Commercially Available Components

### 2.1. Transceivers

In this section, we decouple the types of transceivers into two categories based on their target applications. The first category incorporates transceivers for long-reach purposes, e.g., core networks, while the second one includes transceivers for short-reach applications, e.g., access and intra-data center applications. The details of the transceivers for the first category are tabulated in Table 1. To compose this table, we are based on both commercially available solutions and experimental components such as [10,11]. The first observation is that a 200 G line rate can be attained by increasing the modulation format, from PM-QPSK to 16QAM, simply doubling the number of bits per symbol. However, this comes with the cost of a higher value of required optical to signal plus interference ratio (OSNIR) in order to attain the same Bit Error Rate (BER), which will directly result in a smaller attainable reach. Secondly, the 400 G and 800 G line rates can be realized using a number of different options. Each alternative has its own trade-offs, e.g., lower modulation formats such as 16QAM, require a lower OSNIR in order to attain the same BER as 64QAM. This will lead to a higher reach for the lower cardinality modulation format; however, this reach is traded for a lower number of channels, as they consume a higher channel bandwidth. In this case, capacity is traded with connectivity, and the optimal solution is based solely on the high-level design set by the network operator. The OSNIR shown in the last column of Table 1 is calculated over the whole channel bandwidth while designating the target pre-Forward Error Correction (FEC) value. The existence of a FEC, e.g., [12], is considered, which can make a pre-FEC BER of $10^{-3}$ shrink to a post-FEC BER of less than $10^{-12}$. The exact resulting performance after the FEC depends on the employed algorithm and its complexity; in commercially available transceivers, a common option is soft-decision FEC.

**Table 1.** Commercially available transceivers for regional and long-haul transmission.

| Type | Modulation Format | Symbol Rate | Channel Spacing | Data Rate | Required OSNIR (for BER = $10^{-3}$) |
|---|---|---|---|---|---|
| 100 G | PM-QPSK | 32 Gbaud | 37.5 GHz | 100 Gb/s | 9.8 dB |
| 200 G | PM-16QAM | 32 Gbaud | 37.5 GHz | 200 Gb/s | 16.55 dB |
| 400 G | PM-16QAM | 63 Gbaud | 75 GHz | 400 Gb/s | 16.55 dB |
| | PM-64QAM | 42 Gbaud | 50 GHz | 400 Gb/s | 22.5 dB |
| | PCS-16QAM | 80–95 Gbaud | 100 GHz | 400 Gb/s | varies |
| 800 G | PM-16QAM | 128 Gbaud | 150 GHz | 800 Gb/s | 16.55 dB |
| | PM-32QAM | 96 Gbaud | 112.5 GHz | 800 Gb/s | 19.5 dB |
| | PM-64QAM | 80 Gbaud | 100 GHz | 800 Gb/s | 22.5 dB |
| | PCS-64QAM | 90 Gbaud | 100 GHz | 800 Gb/s | N/A |
| 200–800 G | Probabilistic Shaping | 60–95 Gbaud | 75–100 GHz | 200–800 Gb/s | varies |

Next, we tabulate the details of indicative commercially available transceivers for short-reach applications (Table 2). These transceivers are part of the small form-factor pluggable (SFP) category of transceivers, which are mainly dedicated to access and/or shorter reach applications. SFP transceivers are hot-pluggable network interface modules which show some important features such as compactness, flexibility, cost-effectiveness, and high data rates. Based on their characteristics and application, they can be delineated in various categories such as simple SFP, SFP+, XFP, QSPF, and CFP. For example, XFP, where X stands for "10", and SFP+ are improved versions of simple SPF that can attain significantly higher data rates, e.g., 10 Gb/s compared with the 1 Gb/s rate of the simple SFP. Furthermore, QSFP, where Q stands for "Quad", and CFP, where C stands for "100", can extend further the data rate to 100 Gb/s and beyond, as they can employ multiple optical channels.

**Table 2.** Commercially available transceivers for data center interconnection and access applications.

| Type | Band | Line Rate | Distance | Output Power | Sensitivity |
|---|---|---|---|---|---|
| CFP8 Transceiver | O | 400 Gb/s | <2 km | N/A | N/A |
| CFP4/CFP2Transceiver | O | 100 Gb/s | <10 km | N/A | N/A |
| 200 G and 400 G Coherent Transceiver CFP2-DCO | C | 400 Gb/s | N/A | >−10 dBm | −30/>−20 dBm |
| 100 G Coherent Transceiver CFP-DCO | C | 100 Gb/s | N/A | >−5 dBm | >−28 dBm |
| 400 G QSFP-DCO Coherent Transceiver | C | 400 Gb/s | <120 km | −10dBm | >−12 dBm |
| 400 G Transceiver DD-QSFP-DCO | C | 400 Gb/s | <120 km | −10dBm | >−12 dBm |
| 16G-CWDM SFP+ Transceiver | L | 16 Gb/s | <40 km | >−1 dBm | −16 dBm |
| XFP Transceiver | O | 10 Gb/s | <40 km | >−5 dBm | −16 dBm |
| SFP Transceiver | O, S | 1 Gb/s | <40 km | >−6 dBm | −23 dBm |

From Table 2, it is evident that the main operational bands for these commercially available components are the O and C-band. This is expected, as the transmission bands for the upstream direction of important access standards, such as Gigabit-PON (GPON), 10G-EPON, and XG(S)-PON is O-band, while the C-band is also the premium band in current transmission systems due to its low fiber attenuation. Furthermore, the output power is an important parameter that can increase the overall power budget allowing (a) to support a larger number of users, as the number of power splits in a Passive Optical Network (PON) is higher and (b) to extend the transmission reach, as the tolerable link loss can be greater. Finally, a low sensitivity receiver is desirable in an optical system as it can detect an even smaller signal, leading to an improved power budget.

## 2.2. Amplifiers

Optical amplification is a critical procedure in optical transmission as it can rectify the signal power attenuated within the optical fiber. There are numerous commercially available optical amplifiers that can be exploited for an OMB system, the most indicative of which are illustrated in Table 3. In particular, the Doped Fiber Amplifier (DFA) is a family of amplification components that shows desirable characteristics for an OMB system, such as the sufficient flat gain over a large spectral extent, the low noise figure, and the high output power. In Figure 1, we illustrate the attainable gain and the noise figure of indicative commercially available amplifiers for the O, S, C, and L-bands, while for the E-band, the experimental neodymium (N) DFA of [13] is considered. Another considerable DFA solution for the E-band can be the bismuth (B) DFA of [14]. As it is evident from Figure 1, the amplification technology for the S, C, and L-bands is more mature compared with the O and E-bands, as it can provide a higher overall gain and a lower noise figure

compared with praseodymium (P) DFA and NDFA. These are important drawbacks for the exploitation of the lower bands, as a network operator may resort to a higher number of PDFAs and NDFAs to restore the optical power compared with thulium (T) and erbium (E) DFAs, which will directly lead to higher CapEx and OpEx costs. Furthermore, due to the increased noise figure, the overall physical layer performance of the channels in the O and E-band is expected to be lower compared with channels in the S, C, and L-bands.

**Table 3.** Commercially available amplifiers and their characteristics.

| Amplifier Type | Operating Bands | Flat Gain Wavelength Range | Gain | Gain Flatness | Noise Figure | Max Output Power |
|---|---|---|---|---|---|---|
| EDFA | mainly L | 1560–1610 nm | <35 dB | ±2 dB | <6 dB | >+22 dBm |
| | mainly C | 1525–1565 nm | <38 dB | ±2.5 dB | <5 dB | >+17 dBm |
| | L | 1575–1605 nm | <35 dB | ±2 dB | ≤7 dB | ≥+27 dBm |
| SOA | S, C, L | 1520–1580 nm | <20 dB | ±0.3 dB | 7 dB | 16 dBm |
| | 1050–1080 nm | 1050–1080 nm | <40 dB | 0.3 dB | 8 dB | 18 dBm |
| | O | 1280–1300 nm | <25 dB | ±1.5 dB | 7.5 dB | 12 dBm |
| RAMAN | mainly C, L | 1525–1605 nm | <18 dB | <±1.5 dB | N/A | 31.8 dBm |
| | L | 1565–1605 nm | <15 dB | <±1.5 dB | N/A | 28.7 dBm |
| YDFA | 1055–1075 nm | 1055–1075 nm | <30 dB | N/A | <6 dB | >+20 dBm/23 dBm on request |
| | 1025–1075 nm | 1025–1075 nm | <22 dB | ±5 dB | <8 dB | <+23 dBm |
| PDFA | O | 1280–1320nm | <35dB | ±2 dB | <7dB | >+20 dBm |
| TDFA | S | 1470–1520nm | <30dB | ±0.25 dB | <7dB | +19 dBm |
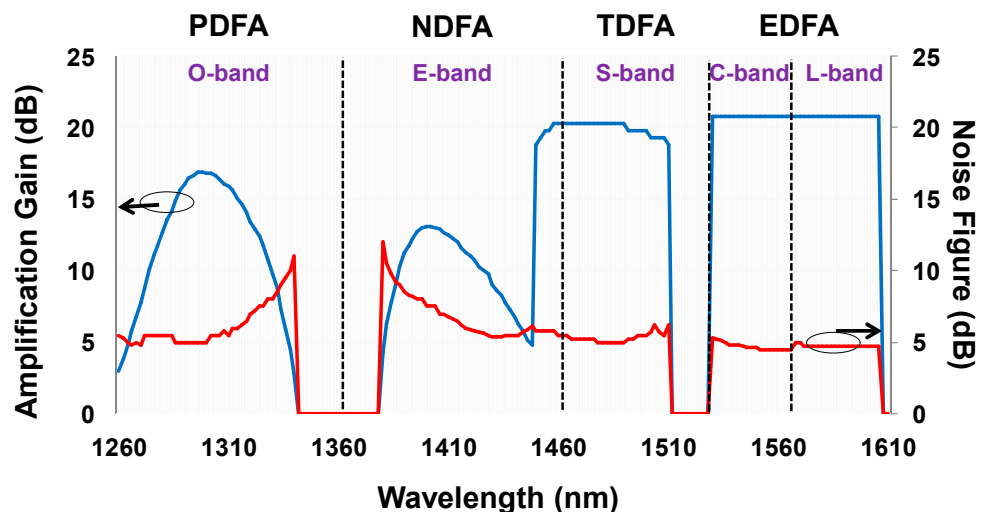


**Figure 1.** Gain and noise figure of indicative available amplifiers assuming a total input power in each amplifier equal to 0 dBm.

The second family of optical amplification components is the semiconductor optical amplifier (SOA). SOAs can be exploited in all bands, as shown in Table 3; however, they show some important drawbacks compared with the DFA family, such as the low saturated output power, the poor gain flatness, the significant transient effects, and the high-noise figure. On the other hand, SOAs are more compact and cheaper than EDFAs and a single SOA, as in [15], can restore the power of channels that are extended over 100 nm. Finally, SOAs as well as ytterbium (Y) DFAs can be exploited to unlock the beyond 1 μm spectral range of transmission.

The third category is the Raman amplification. This category can be exploited in all low attenuation bands similarly to SOAs. Raman-type amplifiers can achieve a sufficient flat gain activating several high-power pumps. An advantage of Raman amplifiers is that they can create arbitrary gain profiles at will, using machine learning methods, which can generate the desirable output power even in a channel basis. In this way, the gain can be tailored to each band to ensure an optimal physical layer performance in all bands of an OMB system concurrently [16]. Other significant advantages of the Raman-type amplifiers

are the distributed amplification and their high output power (about 30 dBm), although such an output power may not be always desirable due to non-linear phenomena and due to safety reasons. Furthermore, due to the smaller gain of Raman amplifiers compared with DFAs, usually a hybrid Raman/EDFA amplification can be considered [17]. Overall, an important advantage of Raman amplifiers and SOAs compared with DFAs is that they offer a seamless gain over a large spectral extend, avoiding a waste of bandwidths for the frequency areas between the bands, e.g., 1510–1530 nm in Figure 1. The selection procedure for the best amplification scheme depends on the system demands. In particular, if a large system capacity is required, amplifiers that can be extended to more than a single band need to be considered. Next, if a large amplification gain is targeted, DFAs seem to be a good option, as they can provide an adequate gain over a large amplification range. Furthermore, if the impact of ASE noise is an important constraint, low-noise figure amplifiers need to be selected.

### 2.3. Band Filters

Band filters are also key components for the realization of an OMB system as they can segregate and combine the optical bands at the node's ingress and egress, respectively, as well as at any amplification stage. Table 4 contains indicative band filters along with their key parameters. As it is evident, using a number of them e.g., in tandem, we can obtain a multi-band filter that can multiplex/demultiplex a number of bands. The most important parameters when selecting a band filter are (a) the wavelength range, as a cascade of different band filters may be required to split/combine the bands, (b) the insertion loss, as when a number of different filters are cascaded, the amplification gain should be higher to mitigate this loss, leading to a greater ASE noise, (c) the isolation, as an inadequate value can lead to an important channel crosstalk, degrading the physical layer performance, and (d) the return loss, which has to be as high as possible to avoid filter reflections. All in all, Table 4 reveals that a broad gamut of band filters is available, spanning from static to tunable and from filters with small spectral range of a few tens of nm to hundreds of nm.

**Table 4.** Commercially available filters for OMB transmission.

| Filter Type | Output Wavelength Range (Transmit) | Output Wavelength Range (Reflected) | Insertion Loss | Isolation Transmission Port/ Reflection Port | Return Loss |
|---|---|---|---|---|---|
| Band Filter | 1570–1610 nm (or 1500–1564 nm) | 1500–1564 nm (or 1570–1610 nm) | ≤0.5 dB (passband) ≤0.6 dB (reflect) | >30 dB/>12 dB | >50 dB |
| | 1460–1575 nm | 1610–1640 nm | <0.8 dB (passband) <0.5 dB (reflect) | 30 dB/15 dB | 50 dB |
| | 1500–1563 nm | 1570–1640 nm | <0.7 dB (passband) <0.6 dB (reflect) | 30 dB/15 dB | >50 dB |
| | 1410–1500 nm | 1510–1625 nm | <0.7 dB (passband) <0.6 dB (reflect) | 30 dB/15 dB | >50 dB |
| | 1310 nm > ±50 nm | 1470–1610 nm | ≤0.70 dB (passband) ≤0.45 dB (reflect) | ≥45 dB/≥15 dB | ≥45 dB |
| | 1310 nm ± 2 nm | 1270–1355 nm, 1317–1355 nm | ≤0.80 dB (passband) ≤0.45 dB (reflect) | ≥20 dB/≥12 dB | ≥45 dB |
| Passband | 1563–1568 nm | 1525–1561.5 nm, 1569.5–1620 nm | 0.5 dB | >30 dB/>20 dB | 50 dB |
| | 1500–1520 nm | 1530–1610 nm | 0.8 dB (passband) 0.4 dB (reflect) | 15 dB/40 dB | 50 dB |
| | 1280–1340 nm | 1350–1620 nm | <1.1 dB | >50 dB/>20 dB | 50 dB |
| Tunable | varies from 1260 to 1650 nm | varies from 1260 to 1650 nm | 4.5 dB | 60 dB | N/A |
| High-Isolation Filter | 1510–1590 nm | 1270–1350 nm | ≤0.8 dB (passband) ≤0.6 dB (reflect) | ≥40 dB/≥40 dB | ≥45 dB |
| Pluggable Filter | 1260–1360 nm, 1460–1581 nm | 1610–1660 nm | <1.2 dB | N/A | >18 dB (1260–1360 nm)/ >17 dB (1460–1581 nm) |
| Reflector | 1260–1570 nm | 1610–1680 nm | <0.60 dB | >15 dB/>40 dB | >45 dB |

### 3. Quantifying the Capacity and Transparent Reach of an OMB System

*3.1. System under Study*

In the physical layer study of this section, we consider only Doped Fiber Amplifiers (DFAs) to compensate for the fiber losses as the DFA family shows similar operational and performance features with C-band EDFAs, and at the same time, it is a widely deployed type of amplifier. In particular, the DFAs provide a sufficiently flat gain over a substantial spectral range, allowing the effective exploitation of the low-loss attenuation spectrum of the fiber. Furthermore, they provide a relatively low-noise figure, e.g., 5.5 dB, leading to an improved physical layer performance compared with SOAs. In our case, the DFAs for the new bands are collocated in the same amplifier duct with the EDFAs, and in this way, the operator has the flexibility to add new DFAs on an "as-needed" basis, preserving the investment made on the existing networking infrastructure. Moreover, the S-band is split into $S_1$ and $S_2$ sub-bands, since a single Thulium-Doped Fiber Amplifier (TDFA) cannot ensure sufficient power per channel at its output, e.g., higher than $-2$ dBm, due to the large amplification range (which is $\approx$55 nm).

The system under investigation is a core network that consists of nodes with an inter-node distance of 150 km and employs commercially available state-of-the-art components (Figure 2). The link consists of three fiber spans of 50 km, where a fiber span is set as the distance between two consecutive amplification stages. Each amplification stage comprises four DFAs to compensate for the fiber losses and the loss induced by the band filters, which in our case is set to 2 dB. For all amplifiers, we consider the presence of Gain-Flattening Filters (GFFs), which can offer an effectively flat gain with <1 dB ripple across the band. This is mandatory in order to maintain similar power levels between the channels of each band and avoid an unpredicted impact of FWM and SRS, which may lead to a significant degradation of the OSNIR on specific channels. Furthermore, the power equalization is performed at the optical nodes, which are Wavelength-Selective-Switch (WSS)-based, maintaining the power at the egress of one node equal to the power at the egress of every other node. These technologies are available today as in [18]. The band details and the corresponding DFAs are shown in Table 5.



**Figure 2.** The transmission link used to interconnect two consecutive optical nodes.

**Table 5.** Spectral partitioning used in our study.

| Band | Used Range (nm) | Range (nm) | Number of Channels | Noise Figure (dB) | Amplifier Type |
|------|-----------------|------------|--------------------|--------------------|----------------|
| $S_1$ band | 1455–1480 | 25 | 92 | 5.5 | TDFA |
| $S_2$ band | 1485–1510 | 25 | 89 | 5.5 | TDFA |
| C band | 1530–1565 | 35 | 116 | 5.5 | EDFA |
| L band | 1570–1615 | 45 | 141 | 6.0 | EDFA |

*3.2. Physical Layer formalism*

In an OMB system, the main physical layer impairments are ASE noise, FWM, and SRS [19–23]. ASE noise and FWM are intra-band effects, while SRS is an inter-band effect, as it leads to a power exchange from the lowest bands, such as the S-band, toward the highest ones, such as the L-band. By making the assumption that FWM is a Gaussian Noise (GN) source, statistically independent from ASE noise, then, we can calculate the OSNIR as follows [20]:

$$\text{OSNIR} = \frac{P_{ch} \cdot \prod\limits_{i=1}^{N_s} G_{SRS,i}}{P_{ASE} + P_{FWM}} \tag{1}$$

where $P_{ch}$ is the power of the examined channel, $P_{FWM}$ is the power of FWM interference, $G_{SRS,i}$ is the SRS gain/loss effect of the channel under observation at the $i$th fiber span, and $P_{ASE}$ is the power of ASE noise. The OSNIR can be directly used to calculate the BER for various modulation formats as follows [24]:

$$\begin{aligned}
\text{BER}_{\text{QPSK}} &= \tfrac{1}{2} erfc\left(\sqrt{\tfrac{\text{OSNIR}}{2}}\right) \\
\text{BER}_{\text{16QAM}} &= \tfrac{3}{8} erfc\left(\sqrt{\tfrac{\text{OSNIR}}{10}}\right) \\
\text{BER}_{\text{64QAM}} &= \tfrac{7}{24} erfc\left(\sqrt{\tfrac{\text{OSNIR}}{42}}\right)
\end{aligned} \tag{2}$$

Next, the power of ASE noise is given by

$$P_{ASE} = \sum_{i=1}^{N_s}\left[ hf(NF_i \cdot G_i - 1)B_0 \prod_{r=i+1}^{N_s} G_{SRS,r} \right] \tag{3}$$

where $G_i$ is the amplifier gain and $NF_i$ is the noise figure at the $i$th amplification stage. $N_s$ is the number of fiber spans a channel is traversing before its power is equalized using e.g., a WSS.

The SRS gain/loss effect for the $j$th wavelength in the $i$th fiber span, is given by [20,25]:

$$G_{SRS,i} = P_{tot,SRS} \frac{e^{\frac{g' \cdot B \cdot L_{eff}}{2A_e}(j-1)P_{tot,SRS}}}{\sum\limits_{m}\left[ P_{m,0} e^{\frac{g' \cdot B \cdot L_{eff}}{2A_e}(m-1)P_{tot,SRS}} \right]} \tag{4}$$

where $g'$ is the Raman gain slope equal to $4.9 \times 10^{-27}$ m/(W·Hz), $A_e$ is the effective cross-sectional area of the fiber equal to 80 μm², and $P_{m,0}$ is the power of the $m$th interfering channel at fiber input. The term $P_{tot,SRS}$ sums the power of channels that interact within the SRS gain bandwidth, which in our work is considered as 15 THz, since we use the triangular approximation [26]. This wide SRS gain bandwidth makes the OSNIR of one band a function of the power level of the channels in all other bands.

To estimate the impact of FWM, we use the expression of $P_{FWM}$ in [27], which was shown in [19] to be in a very good agreement with numerical results across the S, C, and L-bands:

$$\begin{aligned}
P_{FWM} &= \frac{32}{27} \frac{\gamma^2 L_{eff}{}^2 P_{ch}{}^3 N_s{}^2 c}{\lambda^2 B^2 D \sqrt{z_1}}\left(1 + \frac{4e^{-aL}}{(1-e^{-aL})^2}\right)\left(asinh\left(\frac{\pi\lambda^2 DB^2}{8c} N_{ch}{}^{\frac{2B}{B+GB}}\sqrt{z_2}\right) - \tfrac{5}{3}\Phi Log\left(N_{ch}{}^{\frac{B}{B+GB}}\right)\right) \\
&\quad - \frac{32}{27}\frac{\gamma^2 L_{eff}{}^2 P_{ch}{}^3 N_s{}^2 c}{\lambda^2 B^2 D\sqrt{z_1 + 12L^2}}\frac{4e^{-aL}}{(1-e^{-aL})^2}\left(asinh\left(\frac{\pi\lambda^2 DB^2}{8c} N_{ch}{}^{\frac{2B}{B+GB}}\sqrt{z_2 + 12L^2}\right) - \tfrac{5}{3}\Phi Log\left(N_{ch}{}^{\frac{B}{B+GB}}\right)\right)
\end{aligned} \tag{5}$$

where $z_1 = \left(\tfrac{2}{a}\right)^2 + 2L^2\left(N_s{}^2 - 1\right) / \left(\sum\limits_{k=x_1}^{x_2}\frac{1}{1+(2k\pi/(aL))^2}\right)^2$, $z_2 = \left(\tfrac{2}{a}\right)^2 + 2L^2\left(N_s{}^2 - 1\right)$, $x_1 = -\frac{\lambda^2 B^2 DL N_{ch}{}^2}{16c}$, $x_2 = \frac{\lambda^2 B^2 DL N_{ch}{}^2}{2c}$, $\gamma$ is the non-linear fiber coefficient, which takes the value of 1.31744 W⁻¹·km⁻¹, $D$ is the local dispersion, and $a$ is the fiber attenuation parameter, which are different in each band and can be extracted from Figure 2 of [20]. Table 6 illustrates the a, D, and λ for the central wavelength of each band. Next, $L$ is the span length, $N_{ch}$ is the number of channels of the band, $B$ is the channel bandwidth, and $B + GB$ is the channel spacing, which is equal to 37.5 GHz in this work. Finally, $\Phi$ is a

parameter that depends on the modulation format and takes the value of 1 for QPSK, 17/25 for 16QAM, and 13/21 for 64QAM.

**Table 6.** Operational parameters for each amplification band (extracted from [20]).

| Parameter/Band | $S_1$ | $S_2$ | C | L |
|---|---|---|---|---|
| $\lambda$ (nm) | 1467.5 | 1497.5 | 1547.5 | 1592.5 |
| $a$ (dB/km) | 0.246 | 0.23 | 0.211 | 0.209 |
| $D$ (ps/nm/km) | 12.1 | 14 | 16.9 | 19.4 |

The next step is to calculate the power of the channels in each band at the beginning of a link, in a way that ensures a balanced OSNIR performance across all channels at the end of the link. For this purpose, we employed closed-form expressions instead of more complex methods, e.g., Split Step Fourier Method, in order to perform power optimizations in a reasonable-time manner. The power arrangement is an important procedure, as it can minimize the impact of SRS, which as an inter-band effect can deplete the power of the channels in the lower bands degrading significantly their OSNIR. At the same time, it can also increase the channel power in the higher bands, stressing them toward the non-linear regime, where the increased FWM can severely degrade their OSNIR. As a result, a collective optimization of the OSNIR performance across all bands is required. For this purpose, we use the method of [23], which minimizes a suitably chosen merit function subject to specific power constraints $P_{min}$ and $P_{max}$ as follows:

$$
\begin{aligned}
A\left(P_{S_1}, P_{S_2}, P_C, P_L\right) &= \sum_b \left(\frac{1}{\text{OSNIR}_b\left(P_{S_1}, P_{S_2}, P_C, P_L\right)}\right)^2 \\
&\text{minimize } A\left(P_{S_1}, P_{S_2}, P_C, P_L\right) \\
&\text{subject to } P_{min} \leq P_{S_1}, P_{S_2}, P_C, P_L \leq +1 \text{ dBm}
\end{aligned}
\tag{6}
$$

where $P_{S_1}, P_{S_2}, P_C, P_L$ are the wanted quantities, which represent the power of the middle channel of $S_1$, $S_2$, C, and L-bands at the beginning of each link. $P_{min}$ is set in order to avoid a catastrophic OSNIR degradation due the small channel power, while $P_{max}$ is set to e.g., +1 dB, to avoid a DFA operation in its saturation regime, which can result in insufficient amplification gain and consequently to a severe OSNIR degradation. In essence, Equation (6) minimizes the OSNIR difference between the various exploited bands, providing a more balanced OSNIR performance across the entire spectrum. Using this method, the appropriate values of each band are calculated, allowing each band to swing between the ASE-limited and NL-limited regime, mitigating in this way the catastrophic SRS impact on the S-band. In this work, we calculate the transmission performance for a system where all bands are fully loaded with channels, which is the worst possible scenario in terms of physical layer performance.

### 3.3. Results

In this section, we employ the physical layer formalism presented in the previous section to investigate the attainable transparent reach and capacity for an OMB system with S, C, and L-bands. Using the power allocation method of Equation (6), the computed optimal power for the middle channel of each band is $-0.2$, $-1.4$, $-8.3$, and $-9$ dBm for $S_1$, $S_2$, C, and L-bands, respectively. From these power levels, it is evident that the S-band requires a significantly higher power compared with C and L in order to compensate for the power depletion due to SRS. Moreover, the power of C and L-bands stresses them to operate in a deeply linear regime, where ASE noise is the dominant effect. Furthermore, three modulation formats are examined, namely PM-QPSK, PM-16QAM, and PM-64QAM, where each one requires a different OSNIR value in order to attain the same BER, as it is obvious from Equation (2). In particular, we target a BER of $10^{-3}$ for which, from Equation (2), the required OSNIR is 9.8, 16.55, and 22.5 dB for PM-QPSK, PM-16QAM,

and PM-64QAM, respectively. Finally, the baud rate is set to 32 Gbaud, and the channel spacing equals 37.5 GHz.

The OSNIR when the signal traverses a different number of optical nodes is illustrated in Figure 3. The inter-node distance is set to 150 km, and this link comprises three fiber spans of 50 km. The target OSNIR, which leads to a BER equal to $10^{-3}$ for the three modulation formats, is also plotted with dashed lines. As it is evident, the C and L-bands, which are favored due to SRS, show a higher OSNIR performance compared with $S_1$ and $S_2$-bands. In particular, the OSNIR of the $S_1$-band is up to 3.4 dB lower compared with the other three bands, which in turn leads to a smaller attainable transparent reach. Evidently, when PM-QPSK is employed, a transparent reach of more than 2250 km can be attained in all bands. Next, when PM-16QAM is considered, the attainable reach varies significantly between bands, from 600 km for the $S_1$-band to 1350 km in the C-band. Finally, the PM-64QAM can be employed to interconnect only adjacent nodes, as the OSNIR can hardly clear the threshold of 22.5 dB after 150 km in $S_1$ and $S_2$-bands.
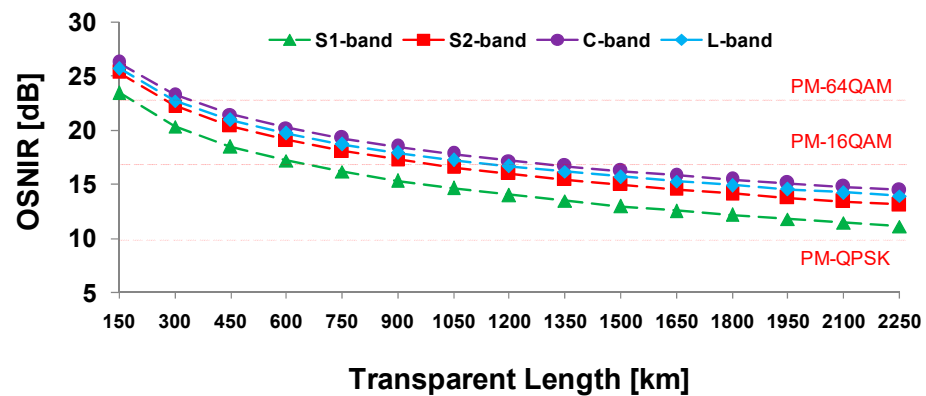


**Figure 3.** OSNIR for different transmission lengths. An inter-node distance of 150 km is considered.

As Figure 3 shows only an instance of OSNIR for the central channel of each band, Figure 4 illustrates the OSNIR for various channels per band. We calculate the OSNIR every 5 nm after 5, 10, and 15 optical nodes. A guard band of 5 nm is also considered between $S_1$ and $S_2$-bands and between C and L-bands, whilst a larger guard band of 20 nm is set between the $S_2$ and C-bands. The latter guard band is set as the commercially available DFAs fail to provide sufficient gain in this spectral region. We can observe that even within the same band, there is a significant OSNIR difference of up to 4 dB between the channels located at the band edges. This clearly indicates that Equation (6) is needed to incorporate a greater number of channels from each band, and not only the central one if a flatter OSNIR over the entire band is the target. Another observation is that in order to increase further the number of channels, the unexploited 20 nm guard band between the S and C-bands has to be filled. For this purpose, the existence of a commercially available amplifier for this amplification range is mandated.

Finally, we tabulate the overall attainable capacity and transmission reach when different modulation formats are employed (Table 7). In this table, we show the transmission reach for the channel with the lowest OSNIR in each band. For example, the worst channel of the $S_1$-band attains an OSNIR slightly above 10 dB after 2250 km. This OSNIR value allows supporting a PM-QPSK modulation format for up to 2250 km, which is the attainable reach for this band when PM-QPSK is employed. Next, in order to calculate the overall capacity, we consider line rates of 100 G, 200 G, and 300 G for PM-QPSK, PM-16QAM, and PM-64QAM, respectively. The attainable capacity in each band is the product of the number of channels that clear the OSNIR threshold for a BER of $10^{-3}$ times the channel line rate. This analysis reveals that when a higher modulation format is enabled, the overall capacity increases; however, the bit rate (C) times transmission reach (L) product decreases. This is expected, since a migration from PM-QPSK to PM-16QAM doubles the overall capacity; however, an about four times higher OSNIR is expected in order to attain the

same BER. Finally, with the proposed amplification scheme, a capacity of 43.8 Tb/s can be attained for up to 15 nodes when PM-QPSK is employed, whilst between adjacent nodes, the overall capacity can reach up to 131.4 Tb/s. A direct benchmarking of these results with relevant experimental works [5,8,15,28] is not possible, as we consider different system details, such as the amplification range, existence of WSSs, etc. However, our results are aligned with the general trends in terms of capacity, due to which more than 100 Tb/s can be attained only in short links, e.g., less than a few hundreds of km, while in longer links, a capacity of less than 100 Tb/s can be achieved. It is worth mentioning that in our study, we consider transmission only in the S, C, and L-bands, as there, the amplification technology is more mature compared with the O and E-bands. The overall capacity will be greatly enhanced as more amplifiers with the required features, e.g., flat gain over a spectral region, sufficient gain, low-noise figure, etc. for the O, the E, and even the U-band will become commercially available.
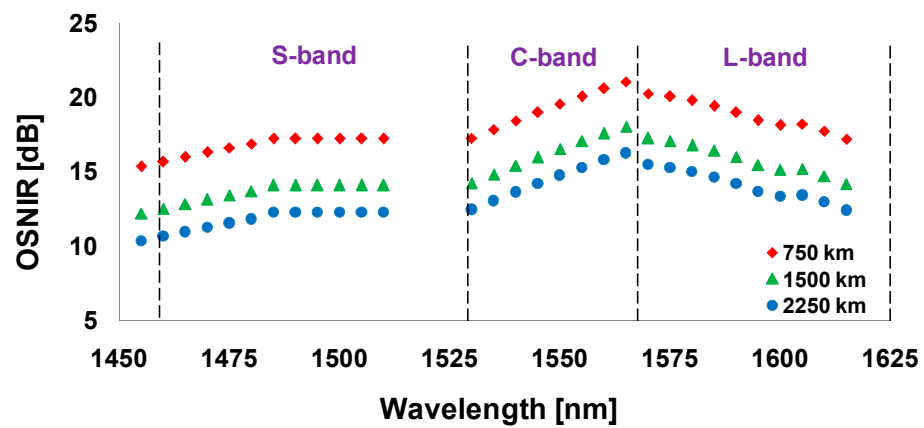


**Figure 4.** OSNIR for different wavelengths when S, C, and L-bands are employed.

**Table 7.** Attainable reach and capacity per band when PM-QPSK, PM-16QAM, and PM-64QAM are considered.

| Bands | PM-QPSK | PM-16QAM | PM-64QAM |
|---|---|---|---|
| $S_1$ band | L = 2250 km<br>C = 9.2 Tb/s | L = 450 km<br>C = 18.4 Tb/s | L = 150 km<br>C = 27.6 Tb/s |
| $S_2$ band | L = 2250 km<br>C = 8.9 Tb/s | L = 750 km<br>C = 17.8 Tb/s | L = 150 km<br>C = 26.7 Tb/s |
| C band | L = 2250 km<br>C = 11.6 Tb/s | L = 750 km<br>C = 23.2 Tb/s | L = 150 km<br>C = 34.8 Tb/s |
| L band | L = 2250 km<br>C = 14.1 Tb/s | L = 750 km<br>C = 28.2 Tb/s | L = 150 km<br>C = 42.3 Tb/s |
| Total (C × L) | 98.55 (Pb/s)·km | 60.18 (Pb/s)·km | 19.71 (Pb/s)·km |

## 4. Discussion

Our analysis revealed that the transceivers based on the application can be decoupled into two categories, the long-reach applications and access/intra-data center applications. In the first category, there are available transceivers for 100 G to 800 G line rates. However, in order to increase the line rate per channel, we can either consume more optical bandwidth, which will result in a lower number of channels, or migrate to a higher modulation format, which will require a higher OSNIR to attain the same BER, limiting the optical reach. The most suitable solution depends solely on the high-level objective set by the network operator. Furthermore, from the category of the SFP transceivers, there are various commercially available solutions mainly for O and C-bands, as these bands are the main bands for access and short-reach transmission. Finally, the most desirable characteristics

for optical transceivers are the high output power and the low receiver sensitivity, as both can lead to a longer transmission length and/or number of supported users.

Based on our study, the commercially available technology of band filters seems to be mature. By placing a number of these filters, e.g., in tandem, we can obtain a filter that can multiplex/demultiplex the desirable bands. The most important characteristics of a band filter are (a) the wavelength range, in order to exploit optimally the low-loss spectrum of the SMF, (b) the insertion loss, which has to be as low as possible to minimize the total losses, (c) the isolation, to be as high as possible, in order to avoid crosstalk, and (d) the return loss, which has to be kept in a high level to avoid the filter reflections.

Next, the amplification technology is mature in S, C, and L-bands, as there, the majority of the optical amplifiers exists. On the other hand, the technology at both the component and system level has to be developed further in O and E-bands in order to exploit the whole extent of the low-attenuation spectrum of the fiber. Our study revealed that SOAs can be employed to attain a wider amplification range; however, they suffer from low gain per channel, low gain flatness, transient effects, and the high noise figure. Furthermore, using Raman amplifiers, we can obtain the desirable gain profiles, which can tune the desirable output power in a channel basis. One significant disadvantage of Raman amplification is the low attainable gain, which compared with DFAs may require a larger number of amplifiers to restore the power of the optical channel. Finally, the rare earth amplification is designated as the winning amplification technology, as it (a) is a well-known technology due to the extensive use of EDFAs in the C-band, (b) allows for a modular engineering, allowing to introduce amplifiers on an as-needed basis, and (c) provides the desirable characteristics such as low noise figure, gain flatness, and high output power.

The proposed physical layer analysis revealed that by using the currently available components and in particular rare earth amplifiers, we can exceed a capacity of 130 Tb/s for shorter distances, e.g., 150 km, and 40 Tb/s for longer ones, e.g., >2000 km. Moreover, we highlighted a trade-off between the modulation format and the attainable reach, as the higher cardinality modulation formats require a higher OSNIR in order to attain the same BER. This resulted in a decrease in the bit rate times reach product, when a higher modulation format is employed. Finally, the adopted physical layer formalism balanced the impact of ASE, FWM, and SRS, leading to small differences of OSNIR between all bands; however, in order to attain a completely flat OSNIR within a band, a large number of channels has to be incorporated in the optimization procedure, which is the scope of an ongoing work.

## 5. Conclusions

In this work, we studied an OMB link with an amplification scheme based on commercially available rare-earth doped fiber amplifiers for the O, S, C, and L-bands. We derived a closed-form expression for the attainable transparent reach of each band of the link (taking into account the impact of the ASE accumulation and the FWM crosstalk), and then, we proceeded to the calculation of the link's capacity and the capacity-length product.

The proposed physical layer analysis showed that systems based on currently available components can exceed a capacity of 130 Tb/s for short distances such as 150 km and 40 Tb/s for longer ones, e.g., >2000 km. It also illustrated the fact that there is a trade-off between the modulation format and the attainable reach (as the higher cardinality modulation formats require a higher OSNIR in order to attain the same BER), which resulted in a decreased bit rate times reach product for the higher modulation formats.

## References

1. ETSI. *GR F5G 001: Fifth Generation Fixed Network (F5G)*; F5G Generation Definition Release #1; ETSI: Sophia Antipolis, France, 2020.
2. Lord, A.; Soppera, A.; Jacquet, A. The impact of capacity growth in national telecommunications networks. *Philos. Trans. R. Soc. A* **2016**, *374*, 20140431. [CrossRef]
3. Stavdas, A. Architectural solutions towards a 1000 channel ultra wideband WDM network. *Opt. Netw. Mag.* **2001**, 2, 51–60.
4. Bayvel, P.; Maher, R.; Xu, T.; Liga, G.; Shevchenko, N.A.; Lavery, D.; Alvarado, A.A.; Killey, R.I. Maximizing the optical network capacity. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **2016**, *374*, 20140440. [CrossRef] [PubMed]
5. Okamoto, S.; Horikoshi, K.; Hamaoka, F.; Minoguchi, K.; Hirano, A. 5-band (O, E, S, C, and L) WDM transmission with wavelength adaptive modulation format allocation. In Proceedings of the 2016 42nd European Conference on Optical Communication (ECOC), Dusseldorf, Germany, 18–22 September 2016.
6. Saavedra, G.; Tan, M.; Elson, D.J.; Galdino, L.; Semrau, D.; Iqbal, A.; Phillips, I.D.; Harper, P.; Ellis, A.; Thomsen, B.C.; et al. Experimental analysis of nonlinear impairments in fibre optic transmission systems up to 7.3 THz. *J. Lightwave Technol.* **2017**, *35*, 4809–4816. [CrossRef]
7. Cai, J.-X.; Batshon, H.G.; Mazurczyk, M.V.; Davidson, C.R.; Sinkin, O.V.; Wang, D.; Paskov, M.; Patterson, W.W.; Bolshtyansky, M.A.; Foursa, D.G. 94.9 Tb/s single mode capacity demonstration over 1900 km with C+L EDFAs and coded modulation. In Proceedings of the 2018 European Conference on Optical Communication (ECOC), Rome, Italy, 23–27 September 2018.
8. Hamaoka, F.; Nakamura, M.; Okamoto, S.; Minoguchi, K.; Sasai, T.; Matsushita, A.; Yamazaki, E.; Kisaka, Y. Ultra-wideband wdm transmission in S-, C-, and L-bands using signal power optimization scheme. *J. Lightwave Technol.* **2019**, *37*, 1764–1771. [CrossRef]
9. Iqbal, M.A.; Krzczanowicz, L.; Phillips, I.; Harper, P.; Forysiak, W. 150 nm SCL-Band Transmission through 70 km SMF using ultra-wideband dual-stage discrete raman amplifier. In Proceedings of the Optical Fiber Communication Conference (OFC), San Diego, CA, USA, 8–12 March 2020.
10. Arnould, A.; Mardoyan, H.; Pulka, F.; Ghazisaeidi, A.; Aref, V.; Bordez, B.; Tondo, P.; Cort, L.D.; Pincemin, E.; Brochier, N.; et al. Field trial demonstration over live traffic network of 400 Gb/s ultra-long haul and 600 Gb/s regional transmission. In Proceedings of the 2020 European Conference on Optical Communications (ECOC), Brussels, Belgium, 6–10 December 2020.
11. Sun, H.; Torbatian, M.; Karimi, M.; Maher, R.; Thomson, S.; Tehrani, M.; Gao, Y.; Kumpera, A.; Soliman, G.; Kakkar, A.; et al. 800G DSP ASIC design using probabilistic shaping and digital sub-carrier multiplexing. *J. Lightwave Technol.* **2020**, *38*, 4744–4756. [CrossRef]
12. Chang, F.; Onohara, K.; Mizuochi, T. Forward error correction for 100 G transport networks. *IEEE Commun. Mag.* **2010**, *48*, S48–S55. [CrossRef]
13. Dawson, J.W.; Kiani, L.S.; Pax, P.H.; Allen, G.S.; Drachenberg, D.R.; Khitrov, V.V.; Chen, D.; Schenkel, N.; Cook, M.J.; Crist, R.P.; et al. E-band $Nd^{3+}$ amplifier based on wavelength selection in an all-solid micro-structured fiber. *Opt. Express* **2017**, *25*, 6524–6538. [CrossRef] [PubMed]
14. Wang, Y.; Thipparapu, N.K.; Richardson, D.J.; Sahu, J.K. Ultra-broadband bismuth-doped fiber amplifier covering a 115-nm bandwidth in the O and E bands. *J. Lightwave Technol.* **2021**, *39*, 795–800. [CrossRef]
15. Renaudier, J.; Arnould, A.; Ghazisaeidi, A.; Le Gac, D.; Brindel, P.; Awwad, E.; Makhsiyan, M.; Mekhazni, K.; Blache, F.; Boutin, A.; et al. Recent advances in 100+ nm ultra-wideband fiber-optic transmission systems using semiconductor optical amplifiers. *J. Lightwave Technol.* **2020**, *38*, 1071–1079. [CrossRef]
16. De Moura, U.C.; Iqbal, M.A.; Kamalian, M.; Krzczanowicz, L.; Da Ros, F.; Brusin, A.M.R.; Carena, A.; Forysiak, W.; Turitsyn, S.k.; Sibar, D.; et al. Multi–band programmable gain raman amplifier. *J. Lightwave Technol.* **2020**, *39*, 429–438. [CrossRef]
17. Ionescu, M.; Galdino, L.; Edwards, A.; James, J.; Pelouch, W.; Sillekens, E.; Semrau, D.; Lavery, D.; Killey, R.I.; Barnes, S.; et al. 91 nm C+L hybrid distributed raman–erbium-doped fibre amplifier for high capacity subsea transmission. In Proceedings of the 2018 European Conference on Optical Communication (ECOC), Rome, Italy, 23–27 September 2018.
18. Kraemer, R.; Nakamura, F.; van den Hout, M.; van der Heide, S.; Okonkwo, C.; Tsuda, H.; Napoli, A.; Calabretta, N. Multi-Band Photonic Integrated Wavelength Selective Switch. *J. Lightwave Technol.* **2021**, *39*, 6023–6032. [CrossRef]
19. Uzunidis, D.; Matrakidis, C.; Stavdas, A.; Pagiatakis, G. On the attainable transparent length of multi-band optical systems employing rare-earth doped fiber amplifiers. In Proceedings of the 1st International Congress on Engineering Technologies (Electri-Tek), Irbid, Jordan, 16–18 June 2020.
20. Uzunidis, D.; Kosmatos, E.; Matrakidis, C.; Stavdas, A.; Lord, A. Strategies for upgrading an operator's backbone network beyond the C-band: Towards multi-band optical networks. *IEEE Photon-J.* **2021**, *13*, 1–18. [CrossRef]
21. Ferrari, A.; Napoli, A.; Fischer, J.K.; Costa, N.; D'Amico, A.; Pedro, J.; Forysiak, W.; Pincemin, E.; Lord, A.; Stavdas, A.; et al. Assessment on the achievable throughput of multi-band ITU-T G.652.D fiber transmission systems. *J. Lightwave Technol.* **2020**, *38*, 4279–4291. [CrossRef]
22. Semrau, D.; Killey, R.I.; Bayvel, P. The gaussian noise model in the presence of inter-channel stimulated raman scattering. *J. Lightwave Technol.* **2018**, *36*, 3046–3055. [CrossRef]
23. Uzunidis, D.; Matrakidis, C.; Stavdas, A.; Lord, A. Power optimization strategy for multi-band optical systems. In Proceedings of the 2020 European Conference on Optical Communications (ECOC), Brussels, Belgium, 6–10 December 2020.

24. Uzunidis, D.; Matrakidis, C.; Stavdas, A.A. Comparison of simplified FWM expressions for coherent optical systems in both dispersion managed and un-managed fiber links. In Proceedings of the 20th Pan-Hellenic Conference on Informatics, Patras, Greece, 10–12 November 2016.
25. Christodoulides, D.; Jander, R. Evolution of stimulated Raman crosstalk in wavelength division multiplexed systems. *IEEE Photon.-Technol. Lett.* **1996**, *8*, 1722–1724. [CrossRef]
26. Semrau, D.; Killey, R.; Bayvel, P. Achievable rate degradation of ultra-wideband coherent fiber communication systems due to stimulated Raman scattering. *Opt. Express* **2017**, *25*, 13024–13034. [CrossRef] [PubMed]
27. Uzunidis, D.; Matrakidis, C.; Stavdas, A. Closed-form FWM expressions accounting for the impact of modulation format. *Opt. Commun.* **2019**, *440*, 132–138. [CrossRef]
28. Hamaoka, F.; Minoguchi, K.; Sasai, T.; Matsushita, A.; Nakamura, M.; Okamoto, S.; Yamazaki, E.; Kisaka, Y. 150.3-Tb/s ultra-wideband (S, C, and L bands) single-mode fibre transmission over 40-km using >519 Gb/s/λ PDM-128QAM signals. In Proceedings of the European Conference on Optical Communication (ECOC), Rome, Italy, 23–27 September 2018.

# Human Body Specific Absorption Rate Reduction Employing a Compact Magneto-Dielectric AMC Structure for 5G Massive-MIMO Applications

Reza Karimian [1], Mansoor Dashti Ardakani [2,*], Shahrokh Ahmadi [1] and Mona Zaghloul [1]

[1] Department of Electrical Engineering, George Washington University (GWU), Washington, DC 20052, USA; karimian@gwu.edu (R.K.); Ahmadi@gwu.edu (S.A.); zaghloul@gwu.edu (M.Z.)
[2] Énergie Matériaux Télécommunications (EMT) Research Centre, Institut National de la Recherche Scientifique (INRS), Université du Québec, Montreal, QC H5A 1K6, Canada
* Correspondence: mansoor.dashti@inrs.ca

**Abstract:** A compact artificial magnetic conductor (AMC) structure for the application of specific absorption rate (SAR) reduction is presented in this paper. A magneto-dielectric (MD) structure as a host of AMC substrate is used to miniaturize the AMC size. The magneto-dielectric has been designed with a low-profile spiral loop in a way to have a high permittivity and permeability for the desired center frequency of 3.5 GHz. Simulation results confirm the zero-degree reflection phase of the proposed AMC unit cell. Moreover, a 70% reduction has been achieved in comparison to the conventional AMC. To validate the simulation results, a prototype of the board is fabricated and measured with a coplanar waveguide (CPW) antenna for the reflection coefficient. The measurement results display an excellent agreement with the simulation ones. A VOXEL model of a human body is utilized to determine the SAR value of the proposed structure. Considering the maximum SAR value for an average of 10 g human tissue, more than 70% SAR reduction is verified for the CPW antenna with the recommended MD-AMC structure compared to a conventional single CPW antenna.

**Keywords:** artificial magnetic conductor (AMC); specific absorption rate (SAR); coplanar waveguide (CPW) antenna; magneto-dielectric (MD); mobile phones; wireless communications

## 1. Introduction

The wireless systems revolution has generated a wave of innovative results that dramatically expand the availability of voice and data almost anywhere. While this change has significantly broadened the possibility for new and better wireless communication ends, it has proffered further considerations for antenna design concepts [1]. Control of specific absorption rate (SAR) is a crucial factor that should be equally considered alongside traditional antenna design parameters. The absorption of electromagnetic energy (EM) emitted from the cellular phone has been considered in recent years. The specific absorption rate is a defined parameter for evaluating power deposition in human tissue. For mobile phone compliance, the SAR value must not exceed the exposure guidelines, 2 watts per kilogram (W/kg). Mobile handset interaction with the human head has been investigated by numerous published papers providing two types of analysis [2–4].

First is the impact of the human head on the handset performance including the feed-point impedance, gain, and efficiency [5]. Second is the interference of the antenna electromagnetic radiation on the user's head due to the absorbed power, which is measured by predicting the induced SAR in the head tissue [6]. Therefore, many researchers are working on reducing SAR distribution in the human head.

In [7,8], a new platform with a study of attaching conductive materials to the cellular phone body for SAR reduction has been suggested. It is declared that the position of the shielding material is a significant factor for SAR reduction effectiveness [9]. There is a

requirement to endeavor to reduce the spatial peak SAR in the design phase of the metamaterial because the likelihood of a spatial peak SAR above the recommended exposure limit cannot be entirely ruled out. Another approach is the use of a directional or reflectional antenna [6]. Such an antenna structure sacrifices the availability of signals received from all directions to the phone model.

The further suggestion is the metamaterial structures with virtual control of the material's intrinsic characteristics were employed to reduce the SAR effects [10,11]. With proper design and arrangement of meta-cells, stopbands at desired mobile frequencies can be achieved. Although significant SAR reduction was observed with the use of Split Ring Resonators (SRR) meta-cells [12], the antenna profile and fabrication process are severe. The periodicity of SRR is aligned to the direction of mobile EM wave to achieve a low SAR effect. Therefore, the antenna profile size is large. In [13] a perfect electric conductor, the (PEC) reflector was realized between a human head and the driver of a folded loop antenna to achieve this purpose. The result shows that the radiation efficiency is improved, and the peak value of SAR is condensed. While the idea of inserting a PEC between the human head and the mobile antenna is a simple technique that can eliminate the EM wave of the antenna toward the human head, it faces a principal problem of the quarter of wavelength separation requirement.

Among all these approaches, Artificial Magnetic Conductor (AMC) can be an alternative for PEC to achieve our purposes, as it can act as a perfect magnetic conductor (PMC) at particular frequency bands. The AMC surface presents high impedance characteristics in some frequencies and changes the antenna's radiation pattern while maintaining antenna performances such as radiation efficiency and gain even when the antenna is too close to it [14]. In the aforementioned area, several research works have been proposed on AMC as a ground plane of the antenna to reduce RF exposure of mobile phones. However, the size of the designed antenna is relatively significant [12–15].

In this paper, a compact AMC surface, based on a magneto-dielectric substrate with a combination of a coplanar waveguide (CPW) antenna is proposed at 3.5 GHz band to reduce the specific absorption rate. The 3.5 GHz band is essential for Massive MIMO technology and 5G network strategy to deliver the best customer experience in urban areas [16]. Apart from numerous miniaturization techniques on the patch antenna design [17–19], in this paper, we focused on substrate miniaturization. For this reason, that magneto-dielectric (MD) substrate is a good candidate to achieve this purpose [8]. With the use of MD, an excellent dielectric property with high virtual permittivity can be performed, while the capacitive nature of the patch on the high dielectric substrate can be mitigated due to the magnetic properties (high permeability) of MD. However, most of the meta-substrate used for the antenna design has been employed with metamaterial spiral loops, which makes the whole structure high profile. In this configuration, a novel low-profile MD cell is used as a meta-substrate of a conventional AMC to reduce the total size. Simulation results show that 70% overall size reduction is obtained with the aforementioned design.

## 2. Analysis and Design

In this section, first, a small printed CPW-fed slot antenna for operation in the 3.5 GHz frequency band has been designed without any metallic layer on the bottom plane of the substrate. The antenna has a simple structure with only one layer of dielectric substrate and metallization. In the proceeding design, a novel AMC structure is described. The CPW antenna and AMC unit cell design simulations are carried out by the frequency solver of CST Microwave Studio and HFSS Software.

### 2.1. CPW Antenna Design

The configuration of the CPW-fed slotted patch antenna is shown in Figure 1. The antenna is printed on a 20-mil (0.508 mm) thickness RO4003 substrate with the relative permittivity of $\varepsilon_r = 3.55$. The basic structure of this antenna is a rectangular patch with a dimension of $35 \times 30$ mm$^2$. A narrow loop slot is embedded into the patch; the length of

this slot determines the resonant frequency, and two other parameters, the width of the slot and $l$t, adjust the impedance matching. As obtained by simulation results, the total size of the proposed ground plane has a critical role for impedance matching which is chosen based on HFSS optimization. As there is a linear relationship between the total slot-loop length and the dominant resonant mode, the slot loop length, $l$s is chosen slightly less than one-half of the slot wavelength ($l$s = 0.42 $\lambda$s); the slot-loop width, $w$s is kept less than one-quarter of lambda to minimize the magnetic current variation along the loop width.
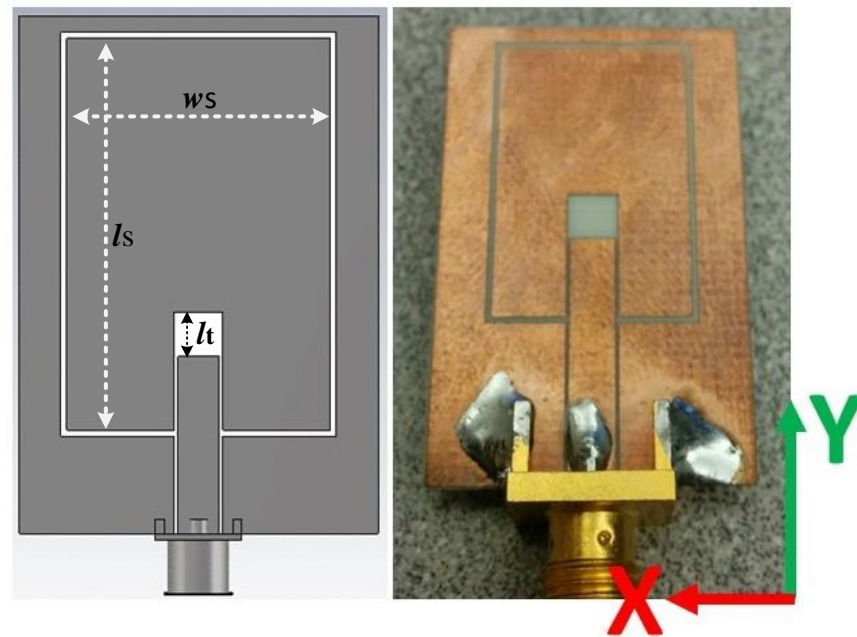


**Figure 1.** Configuration of the proposed CPW antenna.

However, the slot-loop width is critical for impedance matching. By fixing the optimum parameters of the CPW-fed antenna, proper impedance matching through the operation band for ISM application can be realized. These parameters are shown in Figure 1. The simulated and measured reflection coefficient characteristics of the intended antenna, obtained by CST Software and Agilent E8361C VNA, are shown in Figure 2. An acceptable agreement between simulation and measurement results is obtained.
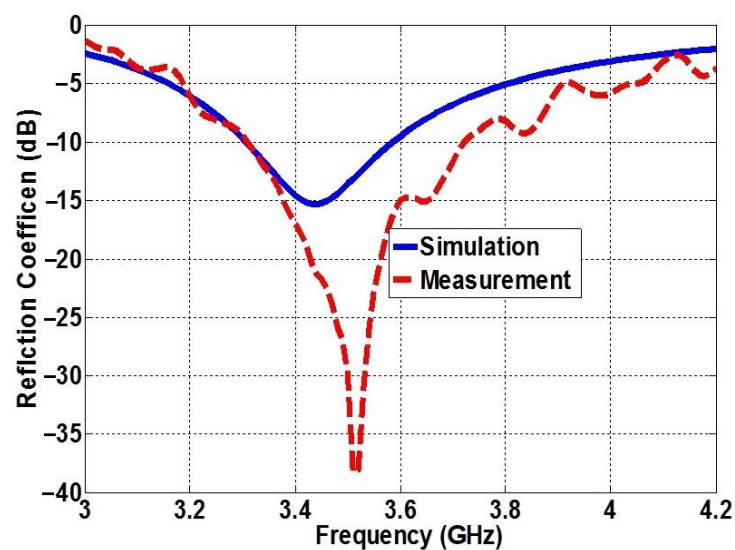


**Figure 2.** Simulation and measurement results of reflected coefficient of the CPW antenna.

## 2.2. AMC Design

By definition, an artificial magnetic conductor (AMC) is a structure, which shows a close-to-zero reflection coefficient phase, when it is subjected to the normal TEM plane wave illumination [9]. In their first appearance, AMCs were introduced as the periodic patches shorted to the ground planes via metallic pins [14]. However, they can also be implemented by loading a grounded dielectric substrate with different FSS-type periodic metallization [13]. Numerous techniques were carried out in recent years to compact the AMC unit cell. Some of these include lumped element loading [13] and engineered pattern [9]. However, by changing the permeability and permittivity of the substrate artificially, we have tried to compact the AMC unit cell size. According to [14] the approximate dimensions of the AMC patch can be calculated from the following equations:

$$L = \mu\, h \tag{1}$$

$$C = \frac{W\varepsilon_0(1+\varepsilon_r)}{\pi}\cosh^{-1}\left(\frac{2W+g}{g}\right) \tag{2}$$

In this model, $W$ and $h$ are the width of the AMC patch and the height of the substrate, respectively. Therefore, as it is cleared from the Equations (1) and (2) by increasing the $\varepsilon$ and $\mu$ of the substrate, $L$ and $C$ increase and the resonance frequency ($\omega_0 = 1/\sqrt{LC}$) moves toward lower frequencies. A magneto-dielectric (MD) structure has been used as a host substrate of a conventional AMC to increase $\varepsilon$ and $\mu$. As the other part of the proposed antenna, the effective medium theory was applied to our MD unit cell to obtain electric permittivity and magnetic permeability. This technique replaces the electromagnetic response of the complicated metamaterial structure with the electromagnetic response of a homogeneous isotropic or anisotropic slab [8]. Cell 1 in Figure 3 depicts the proposed magneto-dielectric substrate. The MD cell consists of metallic strips and vias, which make a loop normal to the *X*-direction. Both host and the upper (for the AMC patch) substrates are Rogers RO4003 with the dielectric permittivity $\varepsilon_r$ = 3.55 and loss tangent tan$\delta$ = 0.0027. The Thicknesses of MD and AMC substrates are 32 and 20 mils, respectively. In our proposed MD cell, sidewalls are PMC, and the top and bottom sides are PEC. It is declared that for a plane wave with normal incidence on a homogeneous slab, the wave impedance, and the refractive index are related to the S-parameters as follows:

$$S_{11} = \frac{R_{01}\left(1 - e^{i2nk_0d}\right)}{1 - R_{01}^2 e^{i2nk_0d}} \tag{3}$$

$$S_{21} = \frac{\left(1 - R_{01}^2\right)e^{i2nk_0d}}{1 - R_{01}^2 e^{i2nk_0d}} \tag{4}$$

where

$$R_{01} = \frac{z-1}{z+1}, \ \& z = \pm\sqrt{\frac{(1+S_{11}^2) - S_{21}^2}{(1-S_{11}^2) - S_{21}^2}};$$

And by employing the Kramers-Kronig relations according to [20], and by following the formula extraction procedure, the constitutive parameters can be extracted by:

$$\mu_{eff} = N_{eff}Z_{eff}, \ \& \varepsilon_{eff} = \frac{N_{eff}}{Z_{eff}} \tag{5}$$

The extracted constitutive parameters, which are effective permittivity, effective permeability and loss are shown in Figure 4a. In the vicinity of the resonant frequency, the product of the effective permittivity and effective permeability is around 41. More specifically at the frequency of 3.5 GHz, the equivalent, extracted, effective permittivity is $\varepsilon_{eff}$ = 7.48 and the equivalent, extracted, effective permeability is $\varepsilon_{eff}$ = 5.5. Now with the new $\varepsilon_{eff}$ and $\mu_{eff}$ and based on the Equations (1) and (2), the AMC width (*W*) is calculated.

Since the width of an AMC patch is larger than the width of an MD unit cells are used as the host for our structure. Unit cell 2 in Figure 3 shows the proposed AMC unit cell. Figure 4b expresses the reflection phase versus frequency of a unit cell at three hypothetical cases along with the proposed MD-AMC unit cell.
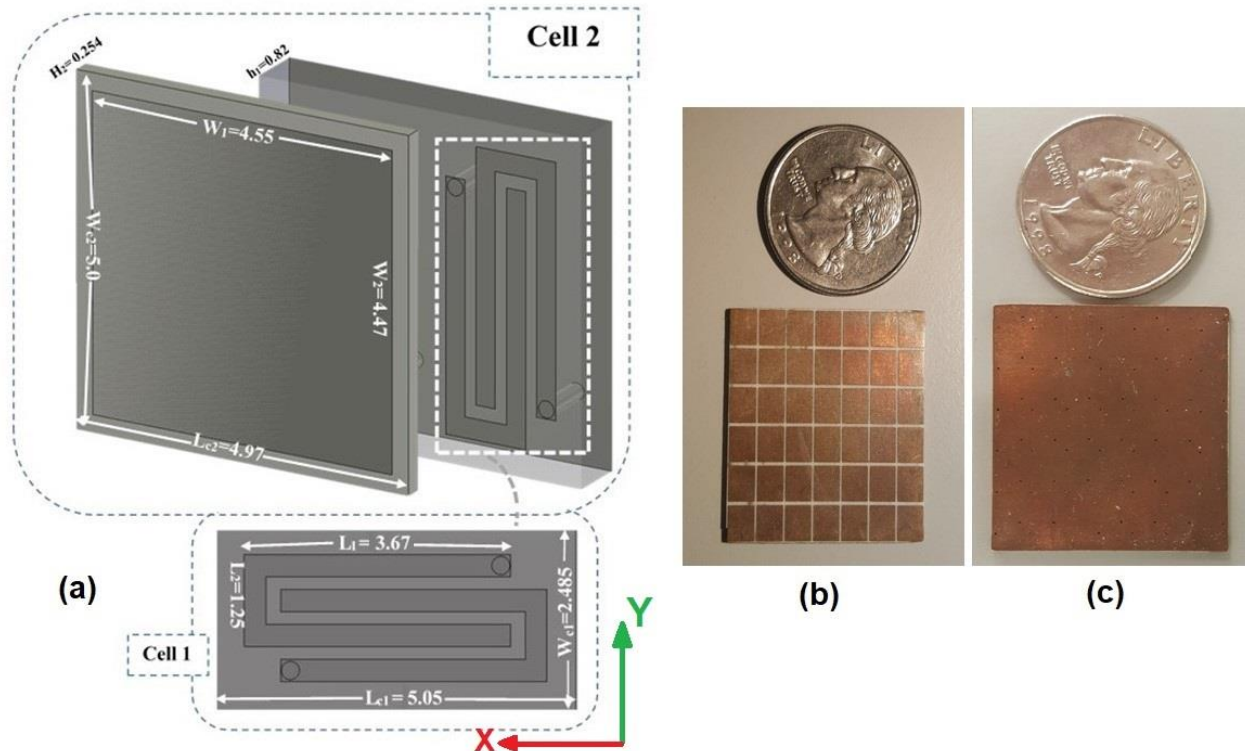


**Figure 3.** (**a**) Magneto Dielectric (MD) unit cell (Cell 1) and MD-AMC unit cell (Cell 2) configuration; The fabricated cells (**b**) Front view, (**c**) Back view.
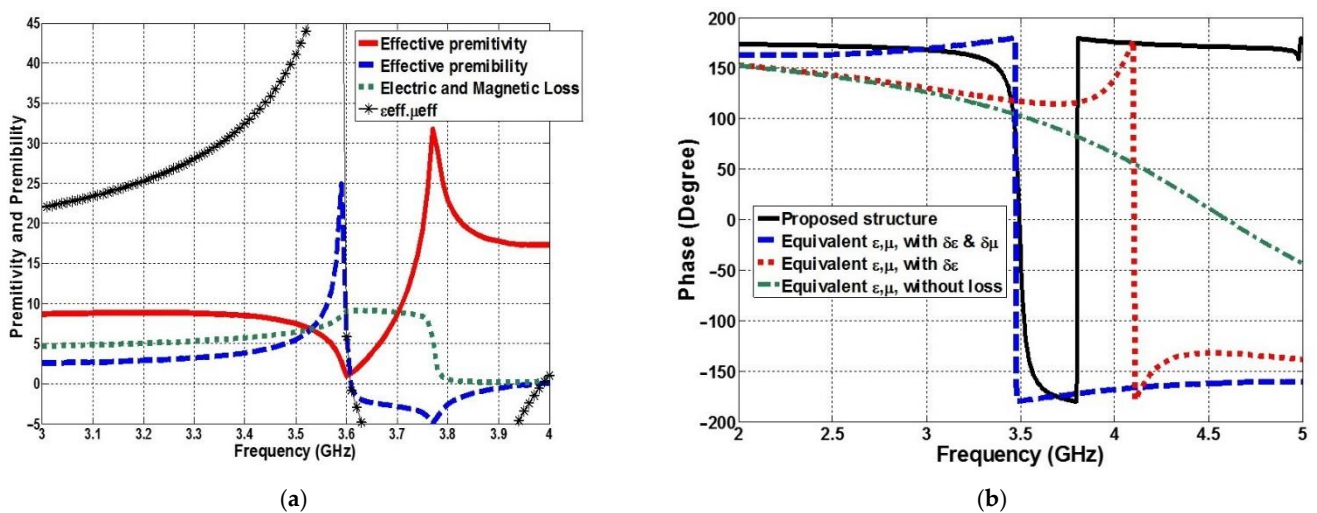


**Figure 4.** (**a**) Effective constitutive parameters of the proposed magneto dielectric substrate. The product of the effective permittivity and permeability is shown by star; (**b**) Reflection phase comparison between proposed substrate structure with other equivalent constitutive parameters.

It is evident from the reflection phase that there is a zero phase at the frequency of 3.5 GHz for the proposed MD-AMC cell. To get a better understanding of the MD-AMC cell, three premises were considered to compare with the proposed structure. It is clear from Figure 4b that for a substrate with the same $\varepsilon$ and $\mu$ of proposed MD and by considering both electrical and magnetic loss, there will be a good matching between reflection phase responses. There is a shift toward higher frequencies when the magnetic loss is not taken into account. In this case, the electrical loss is around six, which is equal to the whole loss extracted from the proposed MD cell. It is interesting to mention that, even for the previous case in which both losses were considered, the electric loss was dominant. As a result, one can conclude that for the proposed MD cell, the electrical loss is dominant. The last hypothetical is when we have used a substrate with the same $\varepsilon$ and $\mu$ of proposed MD, but without any loss. As we can see, the zero reflection phase has moved toward the upper frequency.

The AMC's bandwidth, on the other hand, has greatly increased. When the reflection phase response of the proposed MD-AMC unit cell is compared to that of a substrate with no loss, it can be concluded that the loss will affect the cell efficiency, which in our instance is the cell bandwidth efficiency. It should be noted that by placing the suggested MD cell precisely beneath a patch antenna, the antenna efficiency and gain will be degraded due to the loss. It does, however, make the structure more compact in our design. By comparison of the reflection phase simulation of a unit cell with the substrate of RO4003 and the same height of the proposed MD-AMC cell and structure, it is calculated that more than 70% reduction is achieved by using magneto-dielectric as a host substrate in a conventional AMC.

## 3. Results and Discussion

To validate the performance of the proposed miniaturized AMC, an MD-AMC board is implemented. The CPW antenna is placed 6 mm above the MD-AMC using Rohacell foam, which has a permittivity of $\varepsilon_r = 1.006$. Figure 5 shows the prototype of the whole antenna structure in an anechoic chamber. The distance between CPW and MD-AMC board is optimized based on the best performance regarding the reflection coefficient. The comparison between the simulated and measured reflection coefficient of the proposed antenna is depicted in Figure 6. When compared to a single CPW antenna, the simulated result shows a little shift toward higher frequencies, which is more noticeable in the measured result. The difference between the simulation and measured results is mostly due to the foam material between the antenna and MD-AMC that is not entirely similar to air and, also, there is some glue between the antenna and MD-AMC that affects the result. The comparison between simulation and measurement radiation pattern of the proposed antenna for two principle planes *E*- and *H*-plane is shown in Figure 7, respectively. As it is cleared, the radiation pattern, especially at *E*-field, is similar to a directional antenna which is expected from the MD-AMC board beneath the CPW antenna. The simulation gain result confirms a 3 dB gain enhancement has been achieved compared to the single CPW antenna without an MD-AMC board.
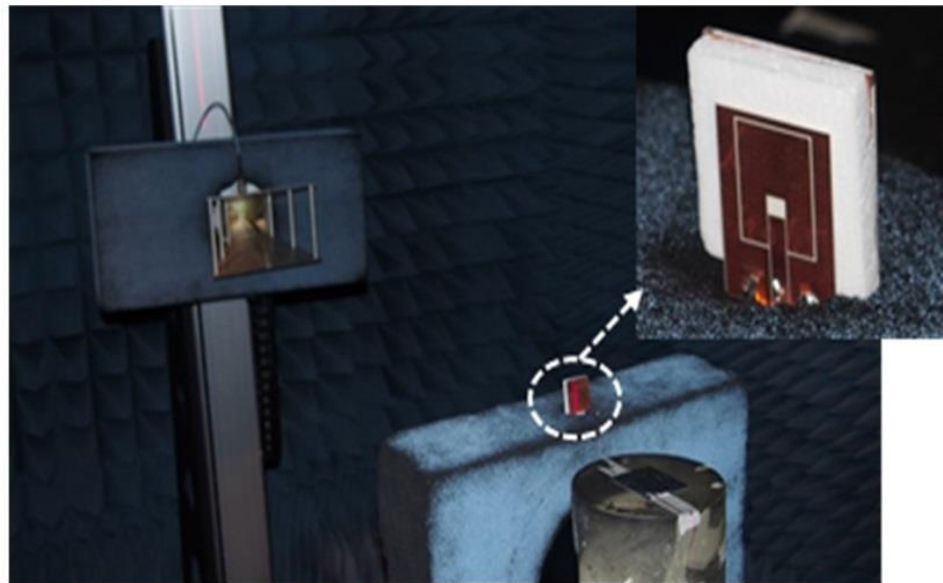
**Figure 5.** Prototype of the fabricated CPW antenna above proposed MD-AMC substrate in an anechoic chamber.
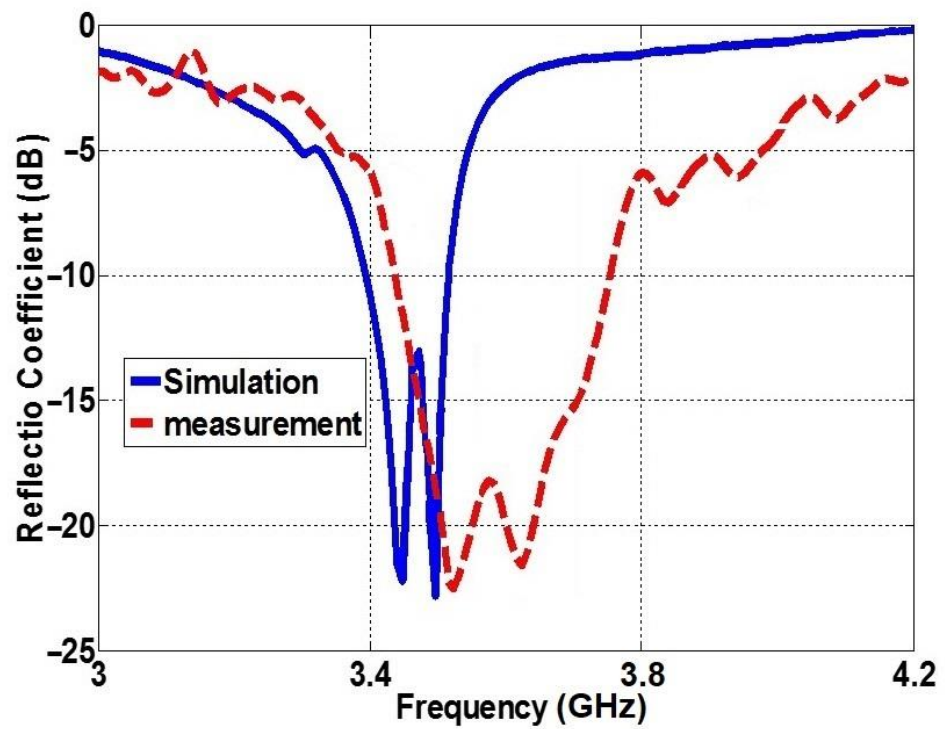


**Figure 6.** Simulation and measurement results of reflected coefficient of the CPW antenna above MD-AMC substrate.
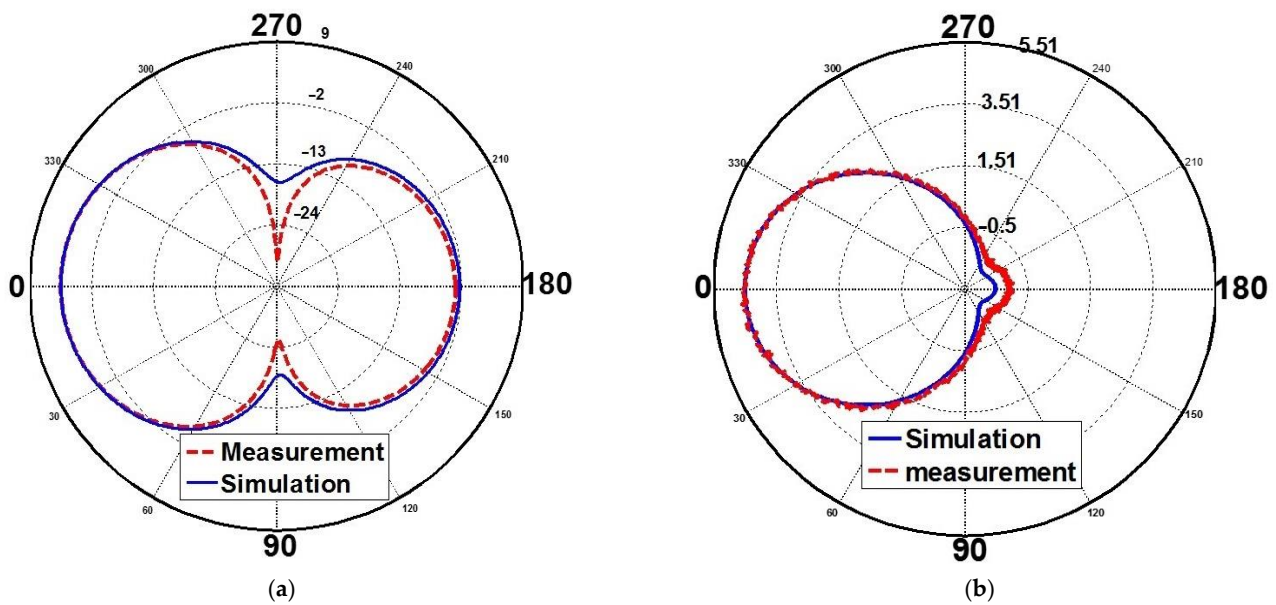
(**a**)

(**b**)

**Figure 7.** Simulation and measurement results of the radiation pattern of the CPW antenna over proposed MD-AMC board for (**a**) E-plane, and (**b**) H-plane.

### 4. SAR Analysis

Specific Absorption Rate (SAR) is a crucial factor to measure the maximum absorption rate by human tissues when exposed to the radiofrequency electromagnetic field. The SAR value according to IEEE Standards Association (IEEE C95.1-2019) and EU standards should not exceed 2 W/kg. From the electric field within the tissues, SAR can be calculated as follow:

$$\text{SAR} = \frac{\sigma |E|^2}{\rho} \tag{6}$$

where $\sigma$ is the conductivity in the tissue in S/m, $E$ is the r oot mean square (RMS) electric field in V/m and $p$ is the mass density of the tissue in kg/m$^3$. To calculate the SAR value of the proposed antenna, a voxel model was utilized. Figure 8, shows the set-up simulation model that is used in the simulation. The original model of the VOXEL is derived from high-resolution CT data segmented into 125 tissue types. The model has a resolution of 3.6 mm$^3$. For a better understanding of the proposed structure effectiveness regarding SAR, two separate simulations were carried out. At the first simulation, the antenna above the MD-AMC board was placed with a distance of 7 mm of the model head, and for the other simulation, the CPW antenna without the MD-AMC board and with the same distance of another simulation was placed near the human head. The SAR distribution for both cases is simulated and shown in Figure 8. The maximum SAR value over the average mass of 10 g tissue is 0.87 W/kg for a single CPW antenna and 0.26 W/kg for the proposed structure. As a result, more than 70% reduction has been achieved with our proposed MD-AMC board. In addition to a geographic SAR distribution in Figure 8, the specific point value of the SAR is depicted in Figure 9a,b for an antenna with MD-AMC board and antenna without MD-AMC board, respectively. For the two latter figures center of the CPW antenna is the origin of the coordinate. As realized, the maximum SAR is for the coordinate of X < 0 and Y = 0. It is expected to have zero SAR for the space between the antenna and the VOXEL model. Consequently, by moving toward negative X-direction, the value of the SAR will decrease.
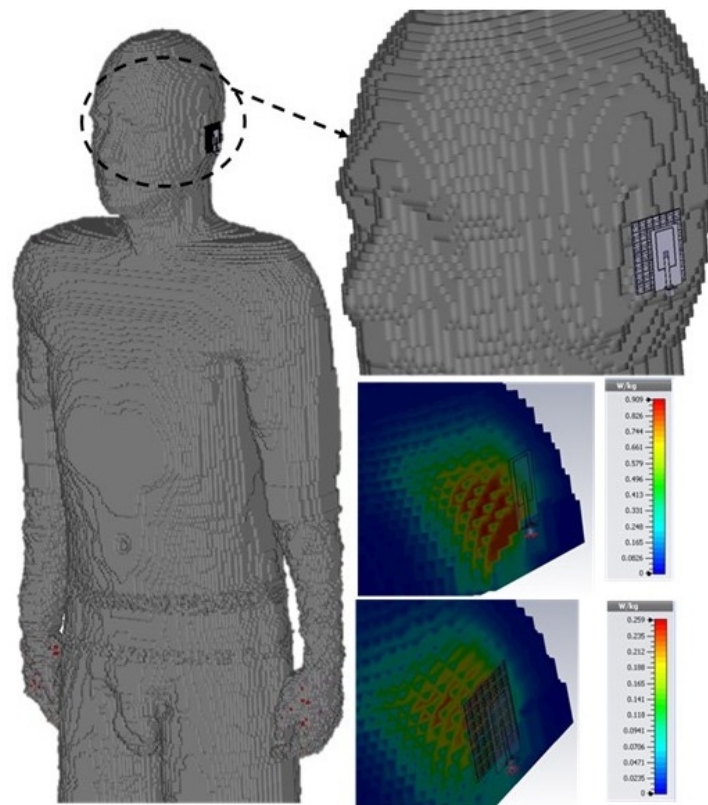
**Figure 8.** 3-D model of SAR calculator set-up that consists of a Voxel model, the antenna close to the head and ear of the body, and simulated SAR results for CPW without MD-AMC board (**top**) and CPW with MD-AMC board (**bottom**).
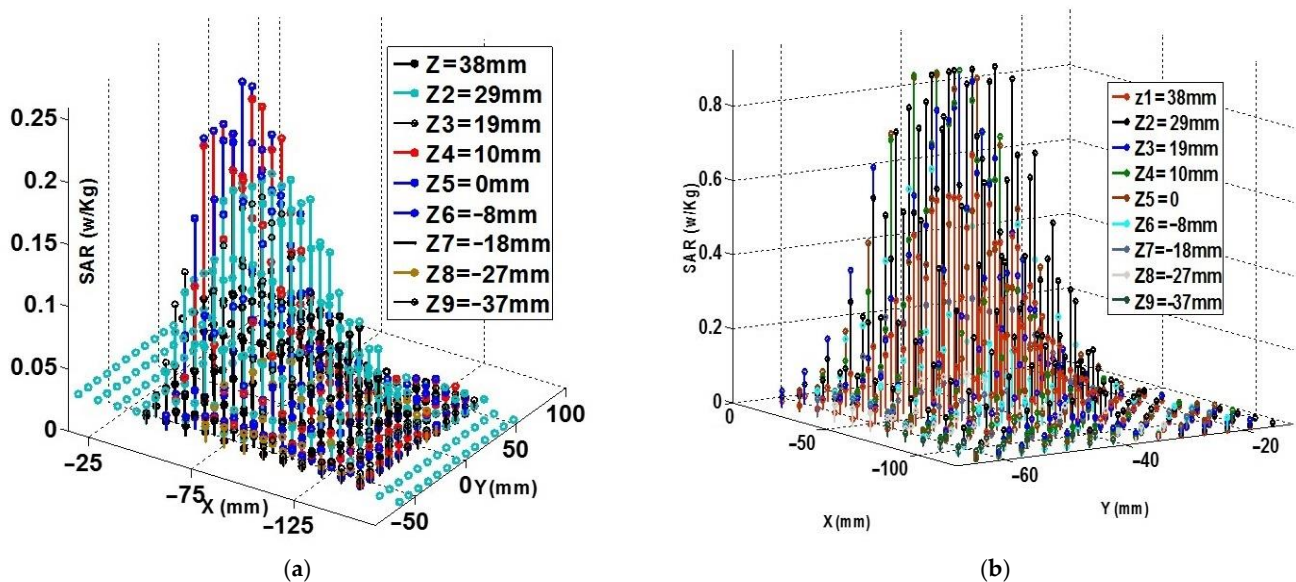


**Figure 9.** Simulation results of SAR distribution (**a**) with MD-AMC board, (**b**) without MD-AMC board.

## 5. Conclusions

Design, simulation, and implementation of a compact artificial magnetic conductor based on a magneto-dielectric substrate has been proposed in this paper to reduce the SAR significantly. To miniaturize a conventional AMC unit-cell, a spiral loop including strip patches and vias are inserted beneath the AMC cell as a host dielectric, which turns a

simple substrate with ($\varepsilon_r = 3.5$, $\mu = 1$) into an artificial substrate with ($\varepsilon_{eff} = 7.48$, $\mu_{eff} = 5.5$). The reflection phase simulation result of the proposed AMC unit cell shows that a 70% reduction has been achieved. To validate the MD-AMC board, a prototype was fabricated and measured with a simple CPW antenna. The measurement result of the reflected coefficient of the antenna shows that it has been followed by simulation one. Finally, a VOXEL model was used to set up a model to calculate the SAR value of the proposed structure. The simulation result shows that more than 70% reduction has been achieved regarding SAR compared to a simple CPW antenna. High efficiency, proper matching, low SAR value and compact structure make this structure a good candidate for biomedical and massive MIMO application.

**Author Contributions:** Conceptualization, R.K.; methodology, R.K. and M.D.A.; software, R.K. and M.D.A.; validation, R.K., M.D.A.; formal analysis, R.K.; investigation, M.D.A.; resources, M.D.A.; data curation, R.K.; writing—original draft preparation, R.K.; writing—review and editing, M.D.A.; visualization, M.D.A.; supervision, S.A. and M.Z.; project administration, R.K.; funding acquisition, S.A. and M.Z. All authors have read and agreed to the published version of the manuscript.

**Institutional Review Board Statement:** Not Applicable.

**Informed Consent Statement:** Not Applicable.

**Data Availability Statement:** Not Applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1.  Ardakani, M.D.; Tabatabaefar, M. A transparent robust quasi-isotropic circularly polarized antenna for Cub-Sat and outdoor wireless. *Eng. Rep.* **2020**, *2*, e12224. [CrossRef]
2.  Faruque, M.R.I.; Islam1, M.T.; Misran, N. Design Analysis of new Metamaterial for EM Absorption Reduction. *Prog. Electromagn. Res.* **2012**, *124*, 119–135. [CrossRef]
3.  Dashti Ardakani, M.; Farahani, M.; Akbari, M.; Tatu, S.O. A Compact Wideband Cubic Dielectric Resonator Antenna for Integrated 60-GHz MHMIC Short-range Transceivers. In Proceedings of the 2020 IEEE International Symposium on Antennas and Propagation and North American Radio Science Meeting, Montreal, QC, Canada, 5–10 July 2020; pp. 71–72.
4.  Karimian, R.; Taravati, S.; Ardakani, M.D.; Ahmadi, S.; Zaghloul, M. Nonreciprocal-Beam Phased-Array Antennas Based on Transistor-Loaded Phase Shifters. *IEEE Trans. Antennas Propag.* **2021**. [CrossRef]
5.  Duncombe, J.U. Infrared Navigation—Part I: An Assessment of Feasibility. *IEEE Trans. Electron. Devices* **1959**, *ED-11*, 34–39.
6.  Kuo, C.-M.; Kuo, C.-W. SAR distribution and temperature increase in the human head for mobile communication. In Proceedings of the Antennas and Propagation Society International Symposium, Columbus, OH, USA, 22–27 June 2003; IEEE: Columbus, OH, USA, 2003; Volume 2, pp. 1025–1028.
7.  Karimian, R.; Pourahmadazar, J.; Nedil, M.; Denidni, T.A. On the design of low SAR CPW antenna with magneto dielectric AMC based ground plane. In Proceedings of the 2016 10th European Conference on Antennas and Propagation (EuCAP), Davos, Switzerland, 10–15 April 2016; pp. 1–5.
8.  Tay, R.Y.S.; Balzano, Q.; Kuster, N. Dipole configurations with strongly improved radiation efficiency for hand-held transceivers. *IEEE Trans. Antennas Propag.* **1998**, *46*, 798–806. [CrossRef]
9.  Karimian, R.; Ardakani, M.D.; Ahmadi, S.; Zaghloul, M. High resolution beam switch antenna based on modified CRLH Butler matrix. *Eng. Rep.* **2021**, *3*, e12287.
10. Ardakani, M.D.; Amiri, R. Mutual coupling reduction of closely spaced MIMO antenna using frequency selective surface based on metamaterials. *Appl. Comput. Electromagn. Soc. (ACES) J.* **2017**, *32*, 1064–1068.
11. Pinto, Y.; Begnaud, X. Mobile phone model with metamaterials to reduce the exposure. *J. Appl. Phys. A* **2016**, *122*, 1–7. [CrossRef]
12. Karimian, R.; Tadayon, H. Compact ultra-wideband antenna with band-notched based on defected ground structure. *J. Eng.* **2014**, *2014*, 30–31. [CrossRef]
13. Shynu, S.V.; Ammann, M.J. A printed CPW-fed slot-loop antenna with narrowband omnidirectional features. *IET Microw. Antennas Propag.* **2009**, *3*, 673–680. [CrossRef]
14. Olmos-Trigo, J.; Sanz-Fernández, C.; Abujetas, D.R.; Lasa-Alonso, J.; de Sousa, N.; García-Etxarri, A.; Sánchez-Gil, J.A.; Molina-Terriza, G.; Sáenz, J.J. Kerker Conditions upon Lossless, Absorption, and Optical Gain Regimes. *Phys. Rev. Lett.* **2020**, *125*, 073205. [CrossRef] [PubMed]
15. Bag, A.; Neugebauer, M.; Wozniak, P.; Leuchs, G.; Banzer, P. Transverse Kerker Scattering for Angstrom Localization of Nanoparticles. *Phys. Rev. Lett.* **2018**, *121*, 193902. [CrossRef]

16. Chen, X.; Grzegorczyk, T.M.; Wu, B.I.; Pacheco, J.; Kong, J.A. Robust method to retrieve the effective constitutive parameters of metamaterials. *Phys. Rev. E Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.* **2004**, *70*, 016608. [CrossRef] [PubMed]

17. Ojaroudi Parchin, N.; Jahanbakhsh Basherlou, H.; Alibakhshikenari, M.; Ojaroudi Parchin, Y.; Al-Yasir, Y.I.; Abd-Alhameed, R.A.; Limiti, E. Mobile-Phone Antenna Array with Diamond-Ring Slot Elements for 5G Massive MIMO Systems. *Electronics* **2019**, *8*, 521. [CrossRef]

18. Ardakani, M.D.; Pourahmadazar, J.; Tatu, S.O. A monopole antenna with notch-frequency function for UWB application. In Proceedings of the 2017 XXXIInd General Assembly and Scientific Symposium of the Internat. Union of Radio Science (URSI GASS), Montreal, QC, Canada, 19–26 August 2017; pp. 1–4.

19. Karimian, R.; Ardakani, M.D.; Koosha, B.; Ahmadi, S.; Zaghloul, M. A Compact Beam Steering Dielectric Resonator Antenna for Wireless Power Transfer. In Proceedings of the 2021 United States National Committee of URSI National Radio Science Meeting (USNC-URSI NRSM), Boulder, CO, USA, 4–9 January 2021; pp. 5–6.

20. Szabo, Z.; Park, G.H.; Hedge, R.; Li, E.P. A Unique Extraction of Metamaterial Parameters Based on Kramers Kronig Relationship. *IEEE Trans. Microw. Theory Tech.* **2010**, *58*, 2646–2653. [CrossRef]

# The Application of Food-Grade Chemical Treatment and Its Effect on the Mechanical Performance Characteristics of Ham Nets

**Md. Al-Amin** [1] , **Charles Freeman** [1,*], **Wes Schilling** [2], **Catherine Black** [1], **Yan Li Campbell** [2], **Wenjie Shao** [2], **Santanu Kundu** [3] **and Anandavalli Varadajan** [3]

[1] Department of Fashion Design and Merchandising, Mississippi State University, Starkville, MI 39762, USA; alamintex20@yahoo.com (M.A.-A.); cblack@humansci.msstate.edu (C.B.)
[2] Department of Food Science, Nutrition, and Health Promotion, Mississippi State University, Starkville, MI 39762, USA; schilling@foodscience.msstate.edu (W.S.); ycampbell@foodscience.msstate.edu (Y.L.C.); ws525@msstate.edu (W.S.)
[3] Department of Chemical Engineering, Mississippi State University, Starkville, MI 39762, USA; santanukundu@che.msstate.edu (S.K.); av234@msstate.edu (A.V.)
**\*** Correspondence: cf617@msstate.edu

**Abstract:** This study involves the use of food-grade chemicals in the integrated pest management of dry-cured ham through the use of 100% polyester weft knitted mesh nets, an idea that was derived from a previously published study in the literature. Tubular mesh nets that are used to contain dry-ageing hams, commonly referred to as ham nets, were treated with a patent-pending food-grade chemical solution (40% Propylene Glycol + 1% Propylene Glycol Alginate + 1% Carrageenan) to control ham mites. Both treated and untreated ham nets were compared for mechanical performance characteristics based on the following standards: abrasion resistance (ASTM D4966), elastic recovery (BS EN 14704-1:2005), breaking strength (ASTM D5034-09), and bursting strength (ASTM D3786). The results indicate that the chemical treatment had minimal to no impact on the mechanical performance characteristics of ham nets. The obtained SEM images also showed no negative effect on the fiber morphology due to the applied chemical solution. The findings support the use of treated ham nets to increase the end-use functionality and provide ham producers an option for integrated pest management without compromising mechanical performance needs.

**Keywords:** dry-cured ham; food-grade chemical; coating; mechanical performance characteristics; chemical treatment; ham nets; pest management; mechanical test; mesh fabric

## 1. Introduction

Dry-cured ham aging involves the use of textiles to hang the cured meat for 3–24 months to develop the desired flavor. During the aging process, one type of mite, *Tyrophagus putrescentiae,* commonly referred to as a ham mite, may grow on the ham's surface [1]. A vital component of a pest management program in the ham industry uses the fumigant Methyl Bromide (MB). Since 1930, MB, as a fumigant, controlled ham mites during aging, but production is no longer permitted globally [2]. However, the dry-cured ham industry exercises an exemption to use the existing stock, and this has significantly increased the MB price. Therefore, it is critical to find a safe, effective, and cheap alternative to MB before depleting the existing stock [3].

A safe and cheap MB alternative is not only functionally significant but also impacts economic and dietary significance. Dry-cured ham is one of the most popular forms of ham among the market's available types [4]. In 2019, the United States consumed approximately 3.3 million tons of hams, which estimates approximately \$2.9 billion and predicted to be \$3.1 billion in 2020 with a growth rate of 2.6% [5]. In Europe, Italy, Spain, and France produce 325,000 tons of dry-cured hams each year [6]. In terms of nutritional

and dietary advantages, dry-cured ham is useful for cardiovascular disorders, high blood pressure, various types of cancer, diabetes, and obesity [7]. The United States Department of Agriculture (USDA) does not allow the sale of homemade country ham to the public due to mites, mold, and fungi on the surface [8]. Therefore, all the dry-cured ham sold in the U.S. is from a commercial or industrial processor. Any delay for an MB alternative before consuming the current MB stock will jeopardize the commercial dry-cured ham industry in the United States.

The application of chemical layers directly on the textile surface adds functionality and value, which enhances the tactile qualities, and aesthetic appearances are considered the most feasible technology to alter textiles' surface performance [9]. Some common examples of surface treatments include water repellency, flame retardant, and lamination [10]. The treatment materials directly influence the performance characteristics of the end product [11]. Apart from the functional advantages, the surface treatments have some disadvantages, such as surface defect and messiness, superficial saturation, large energy consumption, and short-lived adhesion [12]. In terms of polymer deposition on the surface, the polymer and fabric must endure distortion and recover from daily wear, such as twisting and stretching. The polymer coating must retain the surface's structure and shape when exposed to the environment [13]. Various functional additives, such as organic compounds, inorganic particles, and polymers, are used on synthetic textiles to strengthen the mechanical performance characteristics to apply those in the food industry [14]. These treatments improve the textiles' mechanical performance, providing, for example, high mechanical strength, wear, and abrasion resistance, and thus, offer possible methods for varying surface performance. Moreover, the surface treatment also affects the textiles' mechanical performance characteristics [15].

Niekraszewicz et al. (2005) [16] applied an anti-mite modified polypropylene (PP) fiber in the nonwoven textiles to fabricate bedding inserts. The results indicated that the addition of anti-mite agents (0.25% to 2.5% wt) in the PP fiber did not affect the final nonwoven's physical–mechanical properties. The results indicate positive effects on mild allergy conditions and the controlling of the population of dust mites. However, the results of the study report limitations to its efficacy against other species of mites. Rahel et al. (2013) [17] carried out a chitosan surface coating with metal ($Cu^{2+}$, $Ag^+$, $Zn^{2+}$) ions on plasma-treated nonwoven textiles. The results indicated that the chitosan with Ag+ metal ion coating on the ham netting was significantly effective against the ham mites. However, the researchers do not recommend applying chitosan/$Ag^+$ on the food surface due to its high toxicity. The coating effects on the nonwoven textiles' mechanical properties were outside the 2013 study's scope. Abbar et al. (2016) [18] used a coating of propylene glycol, lard, ethoxyquin, or butylated hydroxytoluene on the ham surface. They also used another coating of food-grade ingredient solution that includes PG, PGA, CG on the ham by dipping the ham cubes into the coating solution. The results indicated that the food-grade ingredients were effective in inhibiting the reproduction of ham mites.

Zhang et al. (2017) [3] applied a coating of food-grade ingredients on ham nets by means of dip coating at the laboratory settings. The food-grade ingredient coating includes Xanthan Gum (XG), Propylene Glycol (PG), Propylene Glycol Alginate (PGA), and Carrageenan (CG). Dip coatings effectively inhibited the growth and reproduction of ham mites without affecting the ham's sensory qualities. As in previous studies, the researchers did not examine the coating solutions' effect on the treated ham nets' mechanical performance characteristics. Campbell et al. (2017, 2018) [1,2] used polyester/cotton and cotton nets coated with PG, PGA, and CG on the laboratory scale. The results indicated that the coated nets controlled the infestation of ham mites more effectively than the untreated nets and directly treated ham. However, this study did not conduct a mechanical performance evaluation of the treated versus untreated nets. They also reported that all previous works were not economically feasible in commercial settings.

## 1.1. Research Gap

Prior research on MB alternatives to dry-cured ham nets focuses primarily on the effectiveness of chemical alternatives to control ham mites. Research remains limited in terms of the effect of the applied chemicals on the mechanical performance characteristics of treated ham nets. In the dry-cured ham aging process, a whole green ham weighs around 8 kg–10 kg (17.8 lb–22.22 lb), which exerts pressure on the multi-direction of the nets during aging [2]. The hanging mass on nets imparts forces simultaneously in all directions [19]. The breaking strength test measures the breaking force of the specimen lengthwise or widthwise, which does not conform to the exerted multi-directional pressure [20] of ham to the netting during aging. Dry-cured hams exert heavy stress on the nets for a long time. Therefore, maintaining the treated nets' mechanical performance characteristics compared to untreated nets throughout the aging period is necessary. Campbell et al. (2018) reported that 40% PG, 1% PGA, and 1% CG coating solution, when applied on 100% polyester ham nets, could effectively slow down the ham mite growth. In a follow-up preliminary performance study, Al-Amin et al. (2020) indicated that the application of a coating solution does not affect the mechanical performance characteristics of various fiber ham nets. Abrasion resistance, breaking strength, and elastic recovery testing methods were used to assess the mechanical performance evaluation. However, limitations included failure to evaluate biaxial bursting strength performance, direct net construction comparisons, and exact fiber/yarn size comparisons.

Therefore, based on the preliminary study [21] the following limitations were identified.

No biaxial test, such as the bursting strength test, was performed, which is the most important mechanical test for ham nets.

All mesh nets were not of the same construction, in terms of, for example, weight in grams per square meter (GSM) and stich length, although the fiber contents were of 100% cotton, and thus, a true comparison could not be justified.

The sample size was not consistent across all the performed tests, such as ten for the tensile strength test and six for the abrasion-resistance test, which may have brought inconsistency in the results.

Visual assessment of abrasion resistance with the gray scale for color change did not always reflect the expected outcome for mechanical performance characteristics. Gray scale, for color change, was predominantly used for the colorfastness test.

## 1.2. Purpose of the Study

Therefore, the purpose of this study is to make a true comparison of the effect of food-grade PG solution treatment on the mechanical performance characteristics (abrasion resistance, elastic recovery, breaking strength, and bursting strength) between treated and untreated 100% polyester ham nets based on the identified research gaps and limitations of the preliminary study of Al-Amin et al. (2020).

## 1.3. Hypotheses

Based on the previous studies, and the purposes of those studies, the following hypotheses were formulated.

**$Ha_0$.** *The results will not validate the preliminary findings of no significant differences in the abrasion resistance between treated and untreated ham nets.*

**$Ha_a$.** *The results will validate the preliminary findings of no significant differences in the abrasion resistance between treated and untreated ham nets.*

**$Hb_0$.** *The results will not validate the preliminary findings of no significant differences in the elastic recovery between treated and untreated ham nets.*

**$Hb_a$.** *The results will validate the preliminary findings of no significant differences in the elastic recovery between treated and untreated ham nets.*

**Hc₀.** *The results will not validate the preliminary findings of no significant differences in the breaking strength between treated and untreated ham nets.*

**Hcₐ.** *The results will validate the preliminary findings of no significant differences in the breaking strength between treated and untreated ham nets.*

**Hd₀.** *There will be no significant differences in the bursting strength between treated and untreated ham nets.*

**Hdₐ.** *There will be a significant difference in the bursting strength between treated and untreated ham nets.*

## 2. Materials and Methods

Based on the previous studies and the hypotheses developed, a quantitative research methodology was used to evaluate the comparison of mechanical performance characteristics between a treated knit mesh net and an untreated knit mesh net of the same lot. Earlier research indicated that the treatment of a ham net with a Propylene Glycol (PG), Propylene Glycol Alginate (PGA), and Carrageenan (CG) solution improved the end-use functionality of the ham nets [1]. However, this research (Campbell et al., 2018) did not examine the effect of the treatment on the mechanical performance characteristics of that net. Initial research validates the application of abrasion resistance (ASTM D4966), breaking strength (ASTM D5034-09), elastic recovery (BS EN 14704-1:2005), and bursting strength (ASTM D3786) test methods to evaluate mechanical performance characteristics of treated and untreated ham nets. Apart from the mechanical performance characteristic testing procedures, both treated and untreated ham nets were assessed by visual morphology analysis. Data obtained from the mechanical performance characteristic tests were analyzed by independent sample *t*-test using the Statistical Package for Social Sciences (IBM SPSS, Version 26.0, IBM, New York, NY, USA).

### 2.1. Samples

The untreated circular knit ham nets (Figure 1a) with the technical specifications of (1) 100% polyester, (2) weight of 140 g per square meter (GSM), (3) a stitch length of 63.4 cm per 200 needles, (4) a wale per inch (WPI) of 24.8 ($3.9/cm^2$), (5) a course per inch (CPI) of 39.4 ($6.1/cm^2$), and (6) a diameter of 19.8 cm were provided by Dickson Industries, Inc., Des Moines, IA, USA. Half of the received untreated sample was treated following the treatment procedures outlined by Campbell et al. (2018). Ten specimens from each of the samples were selected for each of the tests in accordance with the recommendation of earlier studies and the respective ASTM and BS test methods.



**Figure 1.** Untreated (**a**) and treated (**b**) ham net sample.

### 2.2. Chemical Solution

According to Campbell et al. (2018), a food-grade chemical solution of 40% PG + 1% PGA + 1% CG was prepared for the surface treatment. Here, PG is the main ingredient to

control the growth of the ham mite. PGA and CG are the binders to keep the PG on the net and prevent evaporation of the PG. The PG was acquired from Hawkins, Inc. (Roseville, MN, USA) with the lot number 384,781; the PGA and CG were provided by TIC Gums, White Marsh, MD, USA. At first, one liter of 400 g PG solution was prepared. Then, 10 g of carrageenan (CG) and 10 g of propylene glycol alginate (PGA) were slowly mixed into the PG solution. After that, all of the solution, containing 420 g of chemicals, was taken into a pot of three-liter volume where 580 g of warm (35 °C–50 °C) water was added slowly by vigorous stirring. The solution was heated up with continuous stirring until the temperature reached at 85 °C. Finally, the solution was transferred to a netting machine (Midwest Metal Craft & Equipment Company, Winsor, MO, USA) to immerse the nets into the solution. The soaked nets were squeezed when passing through a padding roller. The treated nets (Figure 1b) were vacuum packed and kept for conditioning.

### 2.3. Conditioning

Before the mechanical performance characteristic testing, samples were preconditioned according to ASTM D1776/D1776M Standard Practice for Conditioning and Testing Textiles. Both treated and untreated samples were kept in a conditioned laboratory at the temperature of $20 \pm 2$ °C ($68 \pm 4$ °F) with a relative humidity (RH) of $65 \pm 2$%. The treated sample was packed with a (50.8 cm × 71.12 cm) vacuum bag (3 mL standard barrier, nylon/PE clarity pouches, Kansas City, MO, USA) with a dual chamber vacuum packaging machine (Model 2100, Koch equipment LLC, Kansas City, MO, USA) as it contained volatile chemical solution, whereas the untreated samples remained flat on a table surface at 20 °C. However, both the treated and untreated samples received sufficient conditioning before the test, as per the aforementioned ASTM standard.

### 2.4. Mechanical Performance Test

The recommended methodologies from the prior studies were followed to evaluate the mechanical performance characteristics, such as abrasion resistance, elastic recovery, breaking strength, and bursting strength. The apparatuses, specimen selection, procedures, evaluation, and data reporting were followed in accordance with the international standard guidelines set by ASTM International [22–24] and the British Standards Institute.

#### 2.4.1. Abrasion Resistance

A six–head Nu-Martindale abrasion and pilling tester (James Heal, UK) was used to test the abrasion resistance following the ASTM standard method D4966–Standard Test Method for Abrasion Resistance of Textile Fabrics (Martindale Abrasion Tester Method) with 9 kilopascals (kPa) of pressure on each head. The specimens were cut as circular areas with diameters of 38 mm. ASTM recommends testing at least three specimens from each of the samples [25]. Moreover, Nautiyal and Vasugi (2018) [26] evaluated five specimens for their abrasion resistance test of single jersey cotton fabric samples. Hence, as a valid number, ten random specimens from both the treated and untreated samples were evaluated.

According to the ASTM standard D4966, the six-head Nu-Martindale Abrasion and Pilling Tester was calibrated. The circular specimens were cut using a press cutter and inserted on the head of the tester with the support of a circular polyurethane foam of equal area to the specimen in between the specimen and metal insert to keep the specimen flat and even throughout the abrasion cycle. The specimens were abraded until the first thread broke by showing a hole on the abraded surface. After 7000 cycles, the specimens were checked in a regular interval of 500 cycles. The number of cycles abraded to make a hole or break the first yarn was recorded.

#### 2.4.2. Elastic Recovery

A Fabric Extensometer (James Heal, UK) was used to test the elastic recovery following the standard test procedures of BS EN 14704-1:2005. Wardiningsih et al. (2013) [27] evaluated five specimens from each of the samples, for each direction, in their elastic

recovery test of compression garments, which validates that five or more specimens is a valid number for this mechanical test. However, to obtain more accurate results, ten random specimens from each of both the course and wale directions, from the treated and untreated nets, were evaluated according to the following procedures outlined in BS EN 14704-1:2005.

The specimen size was cut as 85 mm × 75 mm. Two reference points were marked on the specimens aligning perpendicular to the two jaws on the tester. The linear distance between the reference points (P) was recorded. One jaw was fixed, and the other was movable. The minimum and initial distance between the jaws was constantly 75 mm (E). After inserting the specimen into the jaws of the extensometer, a 3 kg mount weight was hung to the movable jaws, with the support of a metal wire, to expedite the extension force to the specimen at the time of extending the movable jaw. The distance between the jaws after turning the wheel (L) five times was recorded. The specimen was taken off after at least ten seconds. The distance between the reference points (Q) was recorded. All the measurements were taken in mm. The procedure was undertaken for both directions of the treated and untreated specimens. Calculations were made following the equation outlined in BS EN 14704-1:2005.

$$S = \frac{L - E}{E} \times 100\% \tag{1}$$

$$C = \frac{Q - P}{P} \times 100\% \tag{2}$$

$$D = 100\% - C \tag{3}$$

$$R = \frac{D}{S} \times 100\% \tag{4}$$

Note: S = elongation; C = unrecovered elongation; D = recovered elongation; R = elastic recovery; L = length after 5th cycle; E = initial length; Q = distance between reference marks after 5th cycle; P = initial distance between reference marks.

### 2.4.3. Breaking Strength

A Mark–10 ESM 1500 tensile strength tester was used to test the breaking strength following the ASTM D5034-09 Breaking Strength and Elongation of Textile Fabrics-Grab Test as the standard method for breaking strength, with a load cell of 226.8 kg and a gauge length of 6.35 cm. Specimens were cut with their long dimensions parallel to both the course and wale directions, at a size of 7.62 cm by 5.08 cm (3 inches × 2 inches), from both the treated and untreated samples. The ASTM recommends testing at least three specimens from each of the samples [23]. However, as a valid specimen number, ten random specimens, for each of the directions, from both treated and untreated samples were evaluated.

The cut specimen was placed lengthwise in the jaws of the tensile tester. The bottom jaw was fixed while the upper jaw was movable. Once the upper jaw started moving upward, the specimen also started extending. At a certain point, the applied load broke the specimen. The maximum load (N) at the breaking point was recorded.

### 2.4.4. Bursting Strength

A Truburst[4] (model–1440 Truburst[4], James Heal, UK) pneumatic bursting strength tester was used to test the bursting test following the ASTM D3786 Standard Test Method for Bursting Strength of Textile Fabrics—pneumatic bursting strength tester method. The tester has built-in calibrated test procedures based on the ASTM International and ISO standards. This test requires a maximum air pressure of 1000 kPa (kN/m$^2$)/145 psi/10 bar. The circular rubber diaphragm and the specimen had a diameter of 29.8 cm, whereas the testing area was 7.3 cm$^2$ (diameter 30.5 mm). Nautiyal and Vasugi (2018) evaluated five specimens for the bursting strength test of single jersey cotton fabric using the ASTM D3786 method. Hence, as valid number, 10 random specimens were evaluated from both untreated and treated samples.

The circular specimens were cut at an equal diameter of the rubber diaphragm. At first, the diaphragm was inserted at the center of the test area. The specimen was placed on the diaphragm and the dome was set on the test area. Then, the Truburst[4] was run with an air pressure of 1000 kPa according to the built-in ASTM standard calibrated method, and this continued until the sample burst with the required pneumatic pressure. The maximum bursting pressure (kPa), distension (mm), and time (s) to burst the specimen were recorded.

### 2.4.5. Fiber Morphology Analysis

Scanning Electron Microscope (SEM) images of treated and untreated ham nets were taken using a ZEISS EVO–50 H (Germany) scanning electron microscope to observe the fiber morphology and to investigate the in-depth effect of the chemical treatment on the fiber physiology. All the images were processed at different microscales (20 μm, 50 μm, 100 μm, and 200 μm) and at magnification levels ranging from 95× to 501×.

### 2.5. Statistical Analysis

The Statistical Package for Social Sciences (SPSS) software (version 26.0, IBM, New York, NY, USA) was used to analyze the data to compare the mechanical performance characteristics between the treated and untreated samples. To investigate the mechanical performance characteristics, an independent samples *t*-test, at a 95% confidence interval, was conducted for each of the samples as follows: abrasion resistance, elastic recovery (course and wale direction), breaking strength (course and wale direction), and bursting strength. The independent sample *t*-test is used to compare the means between two sample groups, which is mostly convenient in cases in which there are equal specimen sizes from both sample groups. The equal specimen sizes neutralize possible errors due to unequal standard deviation (SD) [28]. For hypotheses testing, the independent sample *t*-test is applicable for even small sample sizes (N $\leq$ 5) [29]. Therefore, the independent sample *t*-test has been a valid statistical method for the testing of hypotheses when comparing mechanical performance characteristics between untreated and treated ham nets through the evaluation of an equal number of specimens from both groups of samples.

## 3. Results and Discussion

### 3.1. Abrasion Resistance

The average abrasion cycles for both the treated and untreated samples were more than 11,000. Usually, 100% polyester knitted fabrics can withstand up to 10,000 abrasion cycles, given that the mechanical properties of the fabric vary with the type of fibers, yarns, fabric structure, and finish [30]. The results of the independent samples *t*-test were not found to be statistically significant, $t$ (18) = $-0.93$, $p > .05$, resulting in the preliminary findings of no significant difference in abrasion resistance between treated samples (M = 12,800, SD = 3146.43, df = 9) and untreated samples (M = 11,750, SD = 1654.12, df = 9) not being validated. Hence, the null hypothesis was accepted (Table 1). In addition, the significant difference in abrasion resistance between treated and untreated ham nets stated in the preliminary study was not validated. Moreover, the result is consistent with the existing literature on the mechanical properties of 100% polyester fabric under Martindale Abrasion testing.

### 3.2. Elastic Recovery

The independent samples *t*-test was conducted in both the wale and course directions. The independent *t*-test for the wale direction showed a statistically significant difference between the elastic recovery of treated (M = 5.64, SD = .768, df = 9) and untreated (M =7.36, SD = 0.66, df = 9) samples; $t$ (18) = 5.38, $p < .05$. Therefore, the null hypothesis was rejected (Table 1). The results validate the preliminary study of significant effects of food-grade treatment on the elastic recovery of the ham nets. Therefore, the significant difference in elastic recovery in the wale direction between the treated and untreated ham nets stated in the preliminary study was validated.

**Table 1.** Independent samples *t*-test of abrasion resistance, elastic recovery, breaking strength, and bursting strength test.

| Tests | Samples | DF [1] | M [2] | SD [3] | SE [4] | 95% CI [5] | *t* | *p* |
|---|---|---|---|---|---|---|---|---|
| Abrasion Resistance | Treated | 9 | 12,800 | 3146.43 | 994.99 | (10,459 15,050) | | |
| | Untreated | 9 | 11,750 | 1654.12 | 523.08 | (10,566 12,933) | | |
| | Eq. Var. diff [6] | 18 | | 2852.39 | 1124.1 | (−3467 1367) | −0.93 | .37 |
| Elastic Recovery (Wale) | Treated | 9 | 5.64 | 0.768 | 0.24 | (5.09 6.19) | | |
| | Untreated | 9 | 7.36 | 0.66 | 0.21 | (6.89 7.83) | | |
| | Eq. Var. diff | 18 | | 0.77 | 0.32 | (1.05 2.40) | 5.38 | <.05 |
| Elastic Recovery (Course) | Treated | 9 | 4.82 | 0.23 | 0.07 | (4.66 4.99) | | |
| | Untreated | 9 | 5.06 | 0.41 | 0.13 | (4.77 5.35) | | |
| | Eq. Var. diff | 18 | | 0.26 | 0.15 | (−0.075 0.55) | 1.6 | .13 |
| Breaking Strength (Wale) | Treated | 9 | 143.8 | 27.60 | 8.73 | (124 164) | | |
| | Untreated | 9 | 99.5 | 19.95 | 6.31 | (85 114) | | |
| | Eq. Var. diff | 18 | | 28.87 | 10.77 | (−68 −22) | −4.11 | <.05 |
| Breaking Strength (Course) | Treated | 9 | 114.50 | 13.14 | 4.16 | (105 124) | | |
| | Untreated | 9 | 114.50 | 9.00 | 2.85 | (108 121) | | |
| | Eq. Var. diff | 18 | | 15.12 | 5.03 | (−10.58 10.58) | 00 | 1.00 |
| Bursting Strength | Treated | 9 | 120.64 | 3.19 | 1.01 | (118 123) | | |
| | Untreated | 9 | 120.29 | 2.48 | .78 | (119 122) | | |
| | Eq. Var. diff | 18 | | 1.85 | 1.27 | (−3.0 2.33) | −0.277 | .79 |

[1] degree of freedom, [2] mean, [3] standard deviation, [4] standard error, [5] confidence interval, [6] equal Variance difference.

The results obtained for the course direction did not show a statistically significant difference between the elastic recovery of treated (M = 4.28, SD = 0.23, df = 9) and untreated (M = 5.06, SD = 0.41, df = 9) samples; $t (18) = 1.6$, $p > .05$. Therefore, the null hypothesis was accepted (Table 1). Naturally knitted mesh fabrics are elastic due to the structure of loops, which is why the effects of Poisson's Ratio (PR) and Young's Modulus (YM) are predominant in knit fabrics compared to woven fabrics [31]. However, graphical representation of elastic recovery in the wale direction (Figure 2) shows that untreated samples have better elastic recovery values than those of treated samples. The elastic recovery performances are less consistent after the treatment. On the other hand, in the course direction (Figure 3) the untreated samples do not change significantly in terms of their elastic recovery performance characteristics upon receiving food-grade chemical treatment. The values are almost consistent for both types of samples.
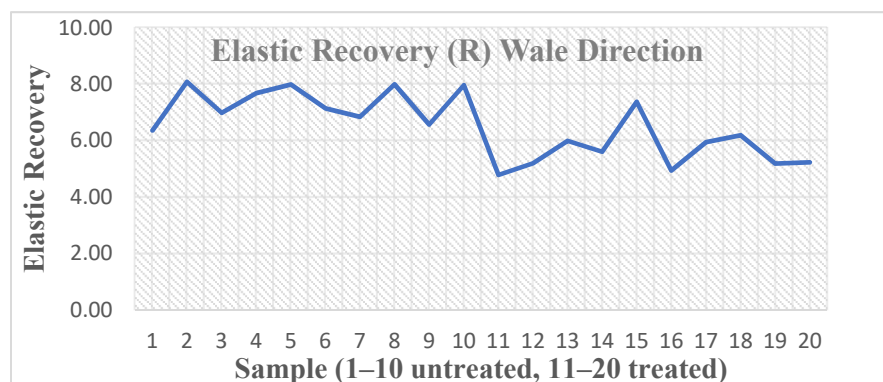


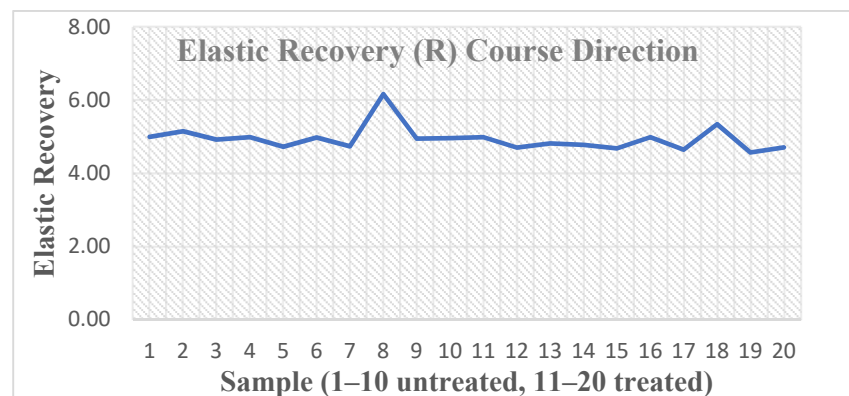**Figure 2.** Elastic Recovery in wale direction.

**Figure 3.** Elastic Recovery in course direction.

The results validate the preliminary study of the significant effects of food-grade treatment on the elastic recovery of the ham nets. Therefore, the significant difference in elastic recovery in the course direction between the treated and untreated ham nets stated in the preliminary study has also been validated. The results from the elastic recovery tests in both directions provide information indicating that the knitted mesh ham net is able to perform in its originally intended state.

*3.3. Breaking Strength*

The breaking strength test in the wale direction shows a statistically significant difference between the treated (M = 143.80, SD = 27.60, df = 9) and untreated (M = 99.50, SD = 19.95, df = 9) samples; t (18) = $-4.11$, $p < .05$. Therefore, the null hypothesis was rejected (Table 1). The results validate the preliminary study of the significant effects of food-grade treatment on the breaking strength of the ham nets. Therefore, the significant difference in breaking strength in the wale direction between the treated and untreated ham nets stated in the preliminary study was validated.

Unlike the wale direction, the results of the independent sample *t*-test for the course direction did not show a statistically significant difference between the breaking strength of treated (M = 114.50, SD = 13.14, df = 9) and untreated (M = 114.50, SD = 9.00, df = 9) samples; $t$ (18) = 0.00, $p > .05$. Therefore, the null hypothesis was accepted (Table 1). The results do not validate the preliminary study of the significant effects of food-grade treatment on the elastic recovery of the ham nets. Therefore, the significant difference in elastic recovery in the course direction between the treated and untreated ham nets stated in the preliminary study was validated.

*3.4. Bursting Strength*

The independent samples *t*-test did not indicate a statistically significant difference between the bursting strength of the treated (M = 120.64, SD = 3.19, df = 9) and untreated (M = 120.29, SD = 2.48, df = 9) samples; t (18) = $-0.277$, $p > .05$, Therefore, the null hypothesis was accepted (Table 1). The results indicate the coating treatment does not significantly affect the bursting strength of the treated ham nets.

*3.5. Morphology Analysis Using SEM Images*

The scanning electron microscope (SEM) images of the untreated sample show a normal geomorphological condition with no visible swelling and relapse of the fibers (Figure 4). However, the visual assessment does not identify any breakage or damage of the inner fibers. On the other hand, the SEM images of the treated sample (Figure 5) show food-grade chemicals adhered to the inner fibers with swelled diameters of the fibers due to the absorption of chemicals. The visual assessment identifies the appearance of swelling in addition to viscous surfaces of the fibers.
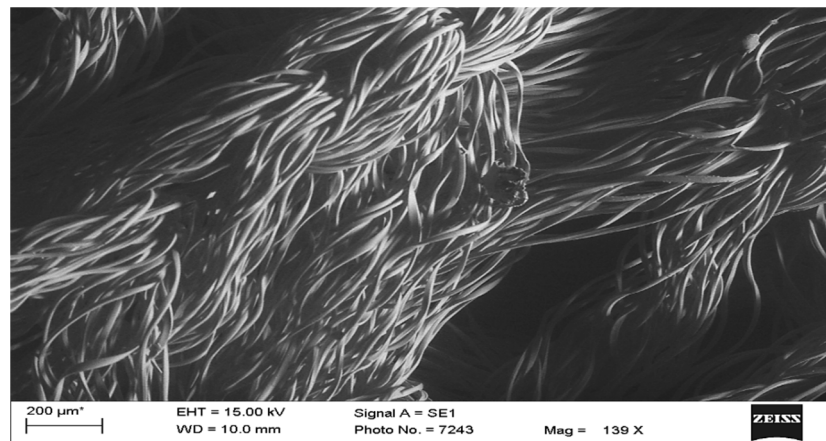
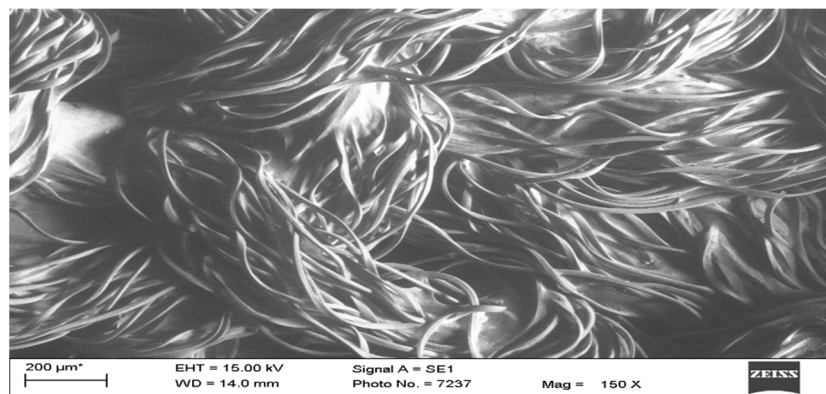**Figure 4.** SEM view of untreated ham net fibers.



**Figure 5.** SEM view of treated ham net fibers.

*3.6. Implications*

Breaking strength plays a vital role during the aging of dry-cured ham. The statistical analysis showed no treatment to have a positive effect on the breaking strength of ham nets, which justifies the effectiveness of the treated ham nets for the dry-cured ham industry.

Similarly, being a multidirectional applied force, bursting strength is the major mechanical performance parameter to evaluate in ham nets during aging. The statistical analysis also indicates that the treatment of ham nets with a PG solution does not affect the bursting strength capabilities and could potentially decrease the risk of net failure during the aging process.

The earlier literature supports the use and efficacy of food-grade chemicals in dry-cured ham aging and the current research indicates no negative implications on the mechanical performance characteristics for treated vs. untreated nets.

**4. Conclusions**

For each of the individual mechanical performance characteristics (abrasion resistance, elastic recovery, breaking strength, and bursting strength), the results indicate the treatment on a circular knit ham net with a food-grade chemical solution of propylene glycol, propylene glycol alginate, and carrageenan has little to no effect on the mechanical performance characteristics of ham nets. For the dry-cured ham producers, this coating treatment on a textile with this solution, as part of an integrated pest management program, will perform just as well as—if not better than—the ham nets that are currently available in the market. Furthermore, this study supports the use of food-grade chemicals and the use of ham nets to increase the end-use functionality and provide ham producers with an option for integrated pest management without compromising in terms of mechanical performance needs.

*Limitations and Future Recommendations*

An additional mechanical performance test, such as a tear test, could have been incorporated to establish the tear strength of the sample ham nets.

The use of ham nets of different fiber contents, such as cotton or a cotton/polyester blend nets, is recommended to investigate the mechanical performance characteristics.

Food-grade PG solution at the polymer extrusion stage can be applied to further investigate the effect of food grade chemical solution on the mechanical performance characteristics of polymer-treated ham nets for a longer time, with a view towards labor saving in the dry-cured ham aging process.

## References

1. Campbell, Y.L.; Zhang, X.; Shao, W.; Williams, J.B.; Kim, T.; Goddard, J.; Abbar, S.; Phillips, T.W.; Schilling, M.W. Use of Nets Treated With Food-Grade Coatings on Dry-Cured Ham to Control Tyrophagus Putrescentiae Infestations without Impacting Sensory Properties. *J. Stored Prod. Res.* **2018**, *76*, 30–36. [CrossRef]
2. Campbell, Y.L.; Zhao, Y.; Zhang, X.; Abbar, S.; Phillips, T.W.; Schilling, M.W. Mite Control and Sensory Evaluations of Dry-Cured Hams with Food-Grade Coatings. *Meat Muscle Biol.* **2017**, *1*. [CrossRef]
3. Zhang, X.; Campbell, Y.L.; Phillips, T.W.; Abbar, S.; Goddard, J.; Schilling, M.W. Application of food-grade ingredients to nets for dry cured hams to control mite infestations. *Meat Muscle Biol.* **2017**, *1*. [CrossRef]
4. Garavaglia, C.; Mariani, P. How Much Do Consumers Value Protected Designation of Origin Certifications? Estimates of Willingness to Pay for PDO Dry-Cured Ham in Italy. *Agribusiness* **2017**, *33*, 403–423. [CrossRef]
5. Processed Meat—United States: Statista Market Forecast. Available online: https://www.statista.com/outlook/40020100/109/ham-bacon/united-states (accessed on 17 August 2021).
6. Swine. Comparison of European & American Systems of Production and Consumption of Dry-Cured Hams. Available online: https://swine.extension.org/comparison-of-european-american-systems-of-production-and-consumption-of-dry-cured-hams/ (accessed on 17 August 2021).
7. Jiménez-Colmenero, F.; Ventanas, J.; Toldrá, F. Nutritional Composition of Dry-Cured Ham and Its Role in a Healthy Diet. *Meat Sci.* **2010**, *84*, 585–593. [CrossRef] [PubMed]
8. Rentfrow, G. How to Make a Country Ham. Available online: https://uknowledge.uky.edu/cgi/viewcontent.cgi?referer=https://scholar.google.com/&httpsredir=1&article=1145&context=anr_reports (accessed on 17 August 2021).
9. Jiang, S.Q.; Newton, E.; Yuen, C.W.M.; Kan, C.W. Chemical Silver Plating on Cotton and Polyester Fabrics and Its Application on Fabric Design. *Text. Res. J.* **2006**, *76*, 57–65. [CrossRef]
10. John, M.J.; Anandjiwala, R.D. Surface modification and preparation techniques for textile materials. In *Surface Modification of Textiles*; Wei, Q., Ed.; Woodhead Publishing: Cambridge, UK, 2009; pp. 1–25.
11. Hall, M.E. Coating of technical textiles. In *Handbook of Technical Textiles*; Horrocks, A.R., Anand, S.C., Eds.; CRC Press: Boca Raton, FL, USA, 2000; pp. 173–186.
12. Nadi, A.; Boukhriss, A.; Bentis, A.; Jabrane, E.; Gmouh, S. Evolution in the Surface Modification of Textiles: A Review. *Text. Prog.* **2018**, *50*, 67–108. [CrossRef]
13. Pei, E.; Shen, J.; Watling, J. Direct 3D Printing of Polymers onto Textiles: Experimental Studies and Applications. *Rapid Prototyp. J.* **2015**, *21*, 556–571. [CrossRef]
14. Aznar, M.; Alfaro, P.; Nerin, C.; Kabir, A.; Furton, K.G. Fabric Phase Sorptive Extraction: An Innovative Sample Preparation Approach Applied to the Analysis of Specific Migration from Food Packaging. *Anal. Chim. Acta* **2016**, *936*, 97–107. [CrossRef] [PubMed]
15. Hasani, H.; Avinc, O.; Khoddami, A. Effects of different production processing stages on mechanical and surface characteristics of polylactic acid and PET fibre fabrics. *Indian J. Fibre Text. Res. (IJFTR)* **2017**, *42*, 31–37.

16. Niekraszewicz, A.; Lebioda, J.; Hoffman, J.; Ruszkowski, K.; Struszczyk, H. Anti-mite modified polypropylene fibres and bedding inserts containing such fibres. *Fibres Text. East. Eur.* **2005**, *6*, 24–27.

17. Rahel, J.; Jonasova, E.; Nesvorna, M.; Klubal, R.; Erban, T.; Hubert, J. The Toxic Effect of Chitosan/Metal-Impregnated Textile to Synanthropic Mites: Chitosan/Metal-Impregnated Textile Kills Mites. *Pest Manag. Sci.* **2013**, *69*, 722–726. [CrossRef] [PubMed]

18. Abbar, S.; Amoah, B.; Schilling, M.W.; Phillips, T.W. Efficacy of Selected Food-Safe Compounds to Prevent Infestation of the Ham Mite, Tyrophagus Putrescentiae (Schrank) (Acarina: Acaridae), on Southern Dry-Cured Hams: PreventingTyrophagus Putrescentiaeinfestation of Hams with Food-Safe Compounds. *Pest Manag. Sci.* **2016**, *72*, 1604–1612. [CrossRef] [PubMed]

19. Wang, X.; Liu, X.; Deakin, C.H. Physical and mechanical testing of textiles. In *Fabric Testing*; Hu, J., Ed.; Woodhead Publishing: Cambridge, England, 2008; pp. 90–124.

20. Özdemir, H.; Mert, E. The Effects of Fabric Structural Parameters on the Tensile, Bursting, and Impact Strengths of Cellular Woven Fabrics. *J. Text. Inst.* **2013**, *104*, 330–338. [CrossRef]

21. Al-Amin, M.; Shao, W.; Varadarajan, A.; Freeman, C.; Campbell, Y.L.; Schilling, W.M.; Kundu, S.; Black, C. Comparison of the Mechanical Properties between Treated and Untreated Ham Nets. *Text. Res. J.* **2020**, *90*, 1237–1244. [CrossRef]

22. D13 Committee. *Test Method for Bursting Strength of Textile Fabrics—Diaphragm Bursting Strength Tester Method*; ASTM International: West Conshohocken, PA, USA, 2018.

23. D13 Committee. *Test Method for Breaking Strength and Elongation of Textile Fabrics (Grab Test)*; ASTM International: West Conshohocken, PA, USA, 2017.

24. D13 Committee. *Test Method for Abrasion Resistance of Textile Fabrics (Martindale Abrasion Tester Method)*; ASTM International: West Conshohocken, PA, USA, 2016.

25. D13 Committee. *Practice for Conditioning and Testing Textiles*; ASTM International: West Conshohocken, PA, USA, 2016.

26. Nautiyal, M.; Vasugi, N. Evaluating the performance attributes of single jersey conventional and organic cotton. *Int. J. Home Sci.* **2018**, *4*, 202–207.

27. Wardiningsih, W.; Troynikov, O.; Molotnikov, A.; Estrin, Y. Influence of Protective Pad Integrated into Sport Compression Garments on Their Pressure Delivery to Athlete's Lower Limbs. *Procedia Eng.* **2013**, *60*, 170–175. [CrossRef]

28. Ross, A.; Willson, V.L. *Basic and Advanced Statistical Tests: Writing Results Sections and Creating Tables and Figures*; Sense Publishers: Rotterdam, The Netherlands, 2017.

29. De Winter, J.C. Using the Student's *t*-test with extremely small sample sizes. *Pract. Assess. Res. Eval.* **2013**, *18*, 10.

30. Uzun, M.; Patel, I. Tribological properties of auxetic and conventional polypropylene weft knitted fabrics. *Arch. Mater. Sci. Eng.* **2010**, *44*, 120–125.

31. Jinyun, Z.; Yi, L.; Lam, J.; Xuyong, C. The Poisson ratio and modulus of elastic knitted fabrics. *Text. Res. J.* **2010**, *80*, 1965–1969. [CrossRef]

# A Critical Look at the Need for Performing Multi-Hazard Probabilistic Risk Assessment for Nuclear Power Plants

Egemen M. Aras and Mihai A. Diaconeasa *

Department of Nuclear Engineering, North Carolina State University, Raleigh, NC 27695, USA; emaras@ncsu.edu
* Correspondence: madiacon@ncsu.edu

**Abstract:** Probabilistic Risk Assessment (PRA) is one of the technologies that is used to inform the design, licensing, operation, and maintenance activities of nuclear power plants (NPPs). A PRA can be performed by considering the single hazard (e.g., earthquake, flood, high wind, landslide) or by considering multi-hazards (e.g., earthquake and tsunami, high wind and internal fire). Single hazard PRA was thought sufficient to cover the analysis of a severe accident until the Fukushima Daiichi NPP accident in 2011. Since then, efforts were made to consider multi-hazards as well; thus, multi-hazard PRAs are starting to be seen as being indispensable for NPPs. In addition to the changing frequency of global and local natural hazards, other reasons to be highlighted are that the number and diversity of NPPs will probably increase. Moreover, advanced reactors are close to becoming a reality by designing them with passive safety systems, smaller, standardized, and even transportable to make them cheaper across the design, licensing construction, and operation stages. Thus, multi-hazards should be addressed in any future full-scope PRA. Although we found a few studies discussing multi-hazards, a general framework for multi-hazard PRA is still missing. In this paper, we argue that the starting point for any multi-hazard PRA general framework should be the Advanced Non-LWR Licensing Basis Event Selection (LBE) Approach and Probabilistic Risk Assessment Standard for Non-Light Water Reactor (non-LWR) Nuclear Power Plants. For Probabilistic Risk Assessment (PRA), history has shown us the path forward before, with Three Mile Accident being seen as one milestone to understand the necessity of PRA. The Fukushima Daiichi NPP Accident is another milestone in the development of PRA, showing the need for performing multi-hazard PRA for the current and future NPPs.

**Keywords:** probabilistic risk assessment; multi-hazards; nuclear safety

## 1. Introduction

Using nuclear technology has many benefits, either for energy or other applications [1] like cancer diagnostics and treatment, non-destructive material testing, thickness, gauge, or level measurements. Nevertheless, just as any other modern technology, it has some disadvantages; for example, depending on the reactor design, nuclear proliferation concerns can be raised [2]. The main requirements of commercial nuclear power are safety, security, and non-proliferation during energy production and the entire fuel cycle. Nuclear safety is necessary to ensure that there is no significant increase in societal health risk compared to other societal risks. More specifically, deterministic and probabilistic safety analyses are performed to assess the likelihood of plant damages and the associated consequences, such as releases of radioactive materials to the environment and acute or latent effects of radiation exposure leading to injuries or deaths [3]. On the other hand, nuclear security ensures nuclear materials and radioactive substances are protected from theft, sabotage, unauthorized access, illegal transfer, or other malicious events [4].

Nuclear safety and security are regulated by national and international agencies, such as the Nuclear Regulatory Commission (NRC) [5] in the United States (US). Regulation of the nuclear industry is based on applying both deterministic and probabilistic methods

to assure the requirements are met. Thermal-hydraulic and reactor physics are examples of deterministic calculations. They are generally based on best estimate plus uncertainty analyses. PRA is one of NRC's endorsed technologies used for risk-informed, performance-based decision-making [6], covering design, construction, operation, and decommissioning stages. PRA generally does quantitative risk estimates for complex, high-risk engineering systems like nuclear power plants, chemical process facilities, waste repositories, and space systems [7]. Specifically, PRA enables the investigation of the probability of accidents and their consequences for nuclear power plants by trying to address the questions: "What can go wrong? How likely is it? What are the consequences?" The answers to these questions form the so-called risk triplet [8].

Although the roots of PRA may extend to the year 1953 through GE-led research called "The Evaluation of Probability of Disaster," WASH-1400 is the first formal PRA for nuclear power plants [6]. More recently, NRC published Regulatory Guides 1.174, "An Approach for Using Probabilistic Risk Assessment in Risk-Informed Decisions on Plant-Specific Changes to the Licensing Basis [9]" and 1.200 "Acceptability of Probabilistic Risk Assessment Results for Risk-Informed Activities [10]" which are indispensable guides to use PRA insights in the nuclear industry. Furthermore, NUREG/CR-2300 "PRA Procedures Guide: A Guide to the Performance of Probabilistic Risk Assessment for Nuclear Power Plants [11]" gives the methods needed for performing PRAs for NPPs.

PRA helps identify the system's weaknesses, both qualitatively and quantitatively. The goal is to support decision-making in three levels for light water reactors (LWRs). Level 1 PRA focuses on the response of systems and operators to the initiating events by evaluating the core damage frequency, level 2 PRA assesses the containment failure by quantifying fission product releases from the containment, and level 3 PRA evaluates the public health consequences [12].

The discussion until now mainly was focused on LWR technology. Presently, interest in Generation IV, small modular reactors (SMRs), and microreactors is higher than ever. Generally speaking, SMRs are new and advanced designs to produce electrical power up to 300 MW [13]. The innovative idea behind the SMRs is the use of pre-fabricated components and systems, which enables a more flexible deployment with the same inherent safety features common to Generation IV reactor designs [14]. Another intended goal for SMRs is to make them economically competitive with a shorter construction period. According to the International Atomic Energy Agency (IAEA) SMR Booklet [13], there are six different categories for SMRs depending on the coolant and neutron spectrum: (1) land-based water-cooled, (2) marine-based water-cooled, (3) high-temperature gas-cooled, (4) fast neutron spectrum, (5) molten salt, and (6) micro-sized SMRs.

This new concept needs a unique perspective in terms of PRA since the three levels approach for LWRs is no longer appropriate since some of the Generation IV reactor designs have no concept of core damage. To address this gap, the Licensing Modernization Project (LMP) [15] aimed to develop a framework to support designers in developing a more reasonable licensing basis and to support the NRC in developing regulations for advanced non-LWR plants. The LMP methodology is technology-inclusive, risk-informed, and performance-based (TI-RIPB). In parallel, a standard for performing a PRA for non-LWRs was developed as well, called "Probabilistic Risk Assessment Standard for Non-Light Water Reactor Nuclear Power Plants" [16] published jointly by the American Nuclear Society (ANS) and the American Society of Mechanical Engineers (ASME) in early 2021.

To sum up, there is no doubt PRA is a vital tool for complete safety analysis of any reactor design at any design stage. Moreover, it has been continuously refined since it was initially developed and used in 1975 with WASH-1400, until it was standardized for advanced non-LWRs in early 2021. Nevertheless, issues such as multi-hazard PRA for nuclear plants still need to be addressed, which we want to cover in this paper.

The rest of the paper will follow this pathway. In Section 2, we intend to introduce single-hazard PRAs. Section 3 will try to address the basics and why we need to consider multi-hazard PRAs. Section 4's focus is the quantification of multi-hazard PRA and

provides two different views of the quantification of multi-hazard PRA. Section 5 introduces the advanced non-LWR Licensing Basis Event Selection (LBE) approach, and Section 6 investigates the latest PRA standard for non-LWR, and especially tries to address definitions related to the multi-hazard PRA.

## 2. Single-Hazard PRA for NPPs

One of the PRA elements is the identification of initiating events. The initiating events are perturbations to the plant during a specific plant operating state that challenges its control and safety systems, whose failure could potentially lead to undesirable consequences, such as radioactive material release [16]. In general, an initiating event may result from human actions or equipment failures from causes internal to the plant (e.g., hardware faults, flood, or fires) or external to the plant (e.g., earthquakes or high winds), or combinations thereof. If such events are brought about by the occurrence of the specific hazard, which is a phenomenon that challenges the safe operation of a facility (e.g., external flood), they are called hazard events. While internal hazard events occur inside the NPP, external hazard events [17] occur outside of the NPP boundaries but have an impact on the NPP site.

In traditional PRA, we use event trees as a step-by-step risk analysis technique to evaluate the progression of system failure events followed by an undesired initiating event. The following event is obtained by asking "What happens next?" to eventually get the consequence. Each of the system's failure events in an event tree is generally obtained by fault tree analysis. A fault tree construction starts with defining the top event, which describes the undesired failure event. We ask, "What caused that?" for the top event and answers will be construction steps for the fault tree. The construction stops either when we reach the basic event or do not have sufficient information to continue. The probabilities for event trees come from fault tree analysis. For single hazard analysis, this process is straightforward since we do not need to model any correlation and interaction of hazards as is needed in multi-hazard analysis. To better understand and quantify the hazard events, classifying the external hazards may be helpful. One may categorize the external events depending on the source of the events as [18]:

- Air-based external events are caused by airspeed, air temperature, air pressure, precipitation, humidity, air contamination, electromagnetic fields, and direct impact from the air.
- Ground-based external events are caused by ground speed, limited ground impact, direct impact from the ground, fire outside the plant, and ground contamination.
- Water-based external events are caused by water speed, water level, water temperature, soil impact, ice impact, solid impurities, water contamination, and direct impact from water.

Other than classifying the events, having a list of potential single external events is crucial. For example, US NRC [19] and International Atomic Energy Agency (IAEA) [17,20,21] have their publications that contain potential single external events.

According to the NUREG-1407 Procedural and Submittal Guidance for the Individual Plant Examination of External Events (IPEEE) for Severe Accident Vulnerabilities [19], the events evaluated for IPEEE are:

- Seismic events
- Internal fires
- High winds and tornadoes
- External floods
- Transportation and nearby facility accidents
- Lightning
- Severe temperature transients, including extreme heat and extreme cold
- Severe weather storms
- External fires
- Extraterrestrial activity
- Volcanic activity

As an example for single-hazard PRA, the Final Safety Evaluation Report Related to Certification of the AP1000 Standard Design (NUREG-1793, Initial Report) [22] may give some insights. The AP1000 PRA analyzed three hazard events, seismic, internal fires, and internal floods. For example, the fire risk analysis was performed for both at-power and shutdown using the available plant-specific design information, fire safety data, and the plant internal events PRA model. The goal was to obtain the core damage frequency (CDF) associated with internal fire. The CDF was about $5.6 \times 10^{-8}$ per year at-power and about $8.8 \times 10^{-8}$ per year during the shutdown. The NRC commented on the results showing the design is capable of withstanding internal fires.

In general, a full-scope PRA study requires the analysis of external flooding, hurricanes, or other external events applicable to the specific site. Thus, the PRA performed for the design certification for any reactor design needs to be updated when site-specific and plant-specific data become available.

### 3. Multi-Hazard PRA for NPPs

Concurrent and successive occurrences of more than one hazard are defined as multi-hazard [23]. In the nuclear industry, multi-hazards are often overlooked in PRA since no general framework is available for such an analysis.

Figure 1 shows the papers published related to the multi-hazard risk assessment between 1983 and 2021, with the given keywords on top of the figures. Figure 1a shows only nuclear-based multi-hazard risk assessment; the figure on the right-hand side (Figure 1b) displays all research based on multi-hazard risk assessment. The interest in the multi-hazard risk assessment tends to increase, especially after the Fukushima NPP accident in March 2011. Moreover, most of the publications on multi-hazards come from the US, leading the published papers with 61 articles. France has 27 articles, Germany, Italy, and Japan have 13 articles each, Canada and the Republic of Korea have 12 articles each, and so on during the same period.

Multi-hazard PRA became a topic after the Fukushima NPP accident in March 2011. According to the World Health Organization (WHO) [24], the Great East Japan Earthquake was a 9.0-magnitude earthquake followed by a tsunami in the eastern coastal. According to the International Nuclear Event Scale (INES) [25], this event, caused by a multi-hazard, led to a level 7 accident at Fukushima Daiichi NPP, the highest level, according to the International Nuclear Event Scale (INES). While the Three Mile Island accident in the US made PRA crucial, the Fukushima Daiichi accident showed the necessity for multi-hazard PRA.

Multi-hazard PRA is a more complex analysis, and it is more challenging to assess its necessity compared to the single hazard analysis [26]. As discussed above, the accidents made PRA crucial for NPP; moreover, climate change and the growing population lead to an increase in the frequency of local, regional, or global hazards. This increase may lead to a higher impact on the critical infrastructures, such as NPPs than anticipated when they were designed.

In comparison to single hazard analysis, multi-hazard PRA requires different analysis methods since every hazard has its own characteristics [27]. Therefore, to better understand and find ways to quantify multi-hazard PRA, a feasible approach could be to categorize the multi-hazards and define standard parameters.

Although there is more than one way to categorize the multi-hazards, in this paper, we prefer to walk through the categorization given in Table 1 [28], which suggests the order of events is also important to classify the events besides the number of events.
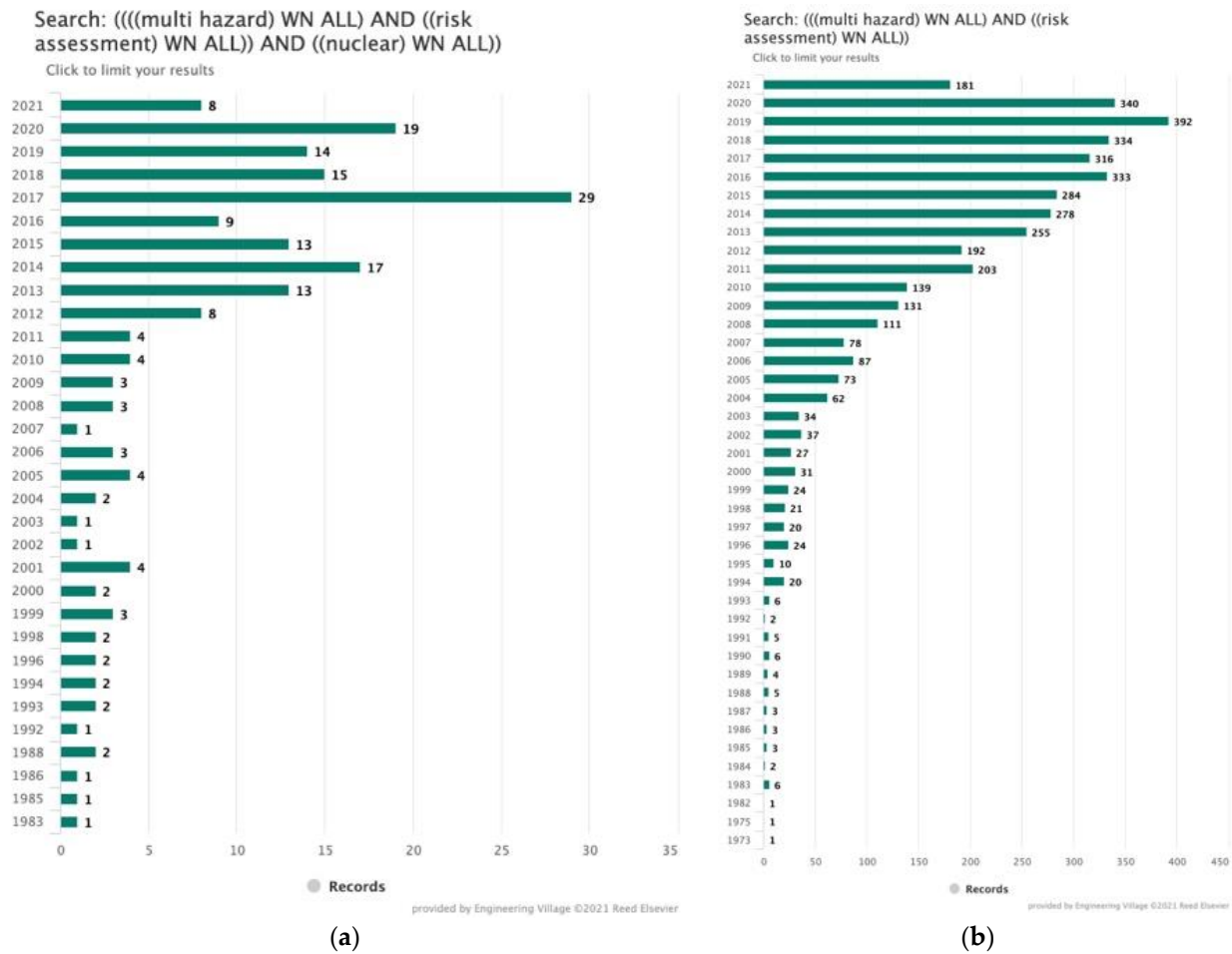
Search: ((((multi hazard) WN ALL) AND ((risk assessment) WN ALL)) AND ((nuclear) WN ALL))

Search: (((multi hazard) WN ALL) AND ((risk assessment) WN ALL))

(**a**)            (**b**)

**Figure 1.** Literature search based on multi-hazard risk assessment on nuclear (**a**) and non-nuclear research (**b**).

**Table 1.** Multi-hazard categorization [28].

| Categories | | Details |
|---|---|---|
| Number of events and hazards | Number of events | Single event: one event and one hazard |
| | | Multi-event: two or more events, including secondary event |
| | Number of hazards | Single hazard: one hazard (it may be caused even by multi-event) |
| | | Multi-hazard: two or more hazards (it may be caused even by a single event) |
| Order of event | Independent event [Independent] | Two or more events that are independent of each other |
| | Simultaneous event [Concurrent] | Two or more events caused by a single source |
| | Sequential event [Successive] | Occurred by secondary event |

Multiple definitions for multi-hazard PRA can be found in the literature; thus, providing descriptive definitions is necessary to understand Table 1:

- *Hazards* are phenomena that challenge the safe operation of a NPP, such as a seismic occurrence or high wind.
- *Hazard event* is an event caused by the occurrence of the specified hazard described in terms of various levels of some characteristic measure of its intensity, such as the peak ground acceleration for seismic hazards or wind speed for high wind hazards.
- *Initiating events* cover natural and human-made perturbations to the plant that can challenge control and safety systems, whose failure can lead to undesired conse-

quences, such as radioactive material release. An initiating event can result from various hazard events internal (e.g., hardware fault, flood, fire) or external to the plant (e.g., earthquakes, high winds).

- *Hazard analysis* is the process of determining an estimate of the expected frequency of exceedance over a specified time interval of various levels of some characteristic measure of the intensity of the hazard, such as water level in a flood.
- *Secondary hazard* is a hazard induced by another hazard, such as a landslide caused by an earthquake.
- *Multi-hazard* is phenomenon in which one hazard occurs concurrently with another hazard, such as seismic and flooding.
- *Multi-hazard (initiating) event* is the occurrence of two or more correlated or uncorrelated events, such as an earthquake of a specific peak group acceleration and high winds of a specific wind speed.

Although some of the definitions are not common and highlight the need for a common language in multi-hazard PRA, the relationship between hazards, which may make analysis trivial or complex, needs to be considered in any approach.

In multi-hazard PRA, both internal and external events need to be addressed. External events are the events that occur outside the NPP. The hazard of the external events may either be the natural environment or man-made. However, multi-hazard PRA is not only related to external events but also related to internal events. For example, a large break loss of cooling (LBLOCA) initiating event can happen during an earthquake [29]. A well-known example is the Fukushima Daiichi NPP accident, an internal accident induced by an external earthquake and tsunami. The lesson learned from this accident is that the PRA for external events, combined events, and external hazards causing internal events needs to be revised. Therefore, there is a need for developing a framework that considers the combination of external events by taking the likelihood of the risk contributors and their effects into account.

In general, hazard events can occur either individually or in a combination of each other. Two combined hazard events occur either simultaneously or within a short duration of time. Moreover, they may also cause an internal event, such as equipment failure. This simple illustration is just like the accident of Fukushima Daiichi NPP. The takeout from this example is that considering the different events as being independent may not always be a reasonable assumption to make. Such simplification ignores the correlation between events and may lead to unintended consequences.

Identification of individual hazards depends on screening analysis, which is established to collect information on plant characteristics concerning internal and external hazards, statistical methods, and experiences concerning the investigation of hazards and their impacts on the plant [30].

Any hazards can be treated in four steps [31]:

1.  Initial data collection can be either site- or plant-specific. Then, the data is the source for screening analysis.
2.  Identification of hazards is the following step upon data gathering. The source for the hazards may be either natural or man-made.
3.  Hazard screening analysis aims to screen out the insignificant item or the items that have insignificant effects.
4.  Detailed hazards analysis analyzes the relevant hazards that affect the structures, systems, or components (SSCs).

The current practice considers two or more hazard events as independent events and evaluates the total frequency as the product of single frequencies. It is straightforward and makes evaluation extremely easy; however, this is not always appropriate.

The studies on multi-hazard PRA in the nuclear industry are not mature yet; however, a couple of efforts are currently ongoing. NARSIS [32] aims to review, analyze and improve the safety assessment methodologies. A practical approach was presented for performing an earthquake-induced landslide PRA for NPP [33]. Another study [34] demonstrated the

results from a survey of multi-hazard PRA that was conducted using a Bayesian network (BN) with Bayesian inference. One comparatively older study [35] developed a systematic methodology to assess and rank the risk from multiple hazards in a community. The final example study [36] describes NRC's Office of Nuclear Regulatory Research's initial efforts to support a portion of the Level 3 project, namely, the multi-hazard Level-2 PRA for LWRs.

## 4. Quantification of Multi-Hazard Risk

A couple of methods and research efforts on multi-hazards are mentioned at the end of the previous section; however, they all focused on specific scenarios rather than a general framework. In addition to these scenarios, there are also some available quantification methods for multi-hazards in the non-nuclear industries. The INFRARISK project developed a reliable stress test on European Critical Infrastructure using integrated tools for decision support [37]. One of the focal points of the INFRARISK was developing a methodology for extreme natural hazards and cascading events. This study quantified the earthquake-induced landslides through a case study. Another effort [38] developed a quantitative approach of multi-hazard risk assessment based on vulnerability surface and joint return period of hazards specifically focused on the risk of crop losses in the Yangtze River Delta region of China. The authors claimed that the methodology can be used in other areas as well. A study [39] used the Choquet integral multiple linear regression model to overcome the problem of nonlinear additivity, which is one of the difficulties for multi-hazard quantification. Additionally, this study considered the effects of magnification on the severity of disasters and the vulnerability of victims in multi-hazard cases.

The main take-away message from the mentioned studies is that the composite individual risk of multi-hazard events is more significant than the simple addition of the risk of each hazard. This conclusion is also the motivation for the work on multi-hazard PRA in the nuclear industry. Current PRA methods do not address the multi-hazard events, so it is crucial to develop a general framework to quantify multi-hazards. This can be done in several steps like modeling the physical phenomena considering the multi-hazards, gathering data, evaluating and generalizing the methodology for different types of reactors and failure modes, and verifying the framework through a peer review panel process. Currently, we are in the phase of modeling physical phenomena considering multi-hazards. This section introduces two recent preliminary efforts on the quantification of multi-hazard risk, specifically in the nuclear industry.

### 4.1. Sampling-Based Multi-Hazard PRA Algorithm

This method starts with an existing methodology for single hazard events and then quantifies multi-hazard events [40]. The current method for a single hazard event quantification is External Event Probabilistic Safety Assessment (EE-PSA) [41], which predicts the possibility of damage to the reactor core due to external hazards. EE-PSA is an integrated process that obtains a single risk value by performing an external analysis, fragility analysis, accident scenario analysis, and risk quantification. In addition, EE-PSA considers the inherent randomness (aleatory uncertainty) and the state of uncertainty (epistemic uncertainty) associated with these analyses and quantifications.

The way to quantify the single hazard EE-PSA is based on the Electrical Power Research Institute (EPRI) separation-of-variable (SOV) method [41]. It is important to note that this method uses the Boolean algebra laws assuming the independence between events to evaluate system fragility, which is valid only if there are no partial dependencies between components. One way to handle partial dependency is using the Bayesian network techniques [34]. Another way is using the sampling-based risk quantification method, also known as Direct Quantification of Fault Tree using the Monte Carlo simulation (DQFM) [42], which considers partial dependencies between components based on the fragility analysis.

Defining fragility analysis is necessary to understand the hazard phenomena better. The fragility analysis is the process used to calculate the conditional probability of failure of a component due to a hazard by considering the aleatory and epistemic uncertainty [40].

Although many approaches are widely used, the safety factor [41] and response factor methods [43] are the most common.

The safety factor method [41] is described in Equation (1) below, where $\Phi[.]$ is the standard Gaussian cumulative distribution of the term in brackets, $a$ denotes the hazard intensity, $A_m$ is the median hazard performance value, and $\beta_c = \sqrt{\beta_r^2 + \beta_u^2}$ is the composite log-standard deviation, including randomness and uncertainty.

$$f_0 = \Phi\left[\frac{\ln \frac{a}{A_m}}{\beta_c}\right] \tag{1}$$

The response factor method [43] is described in Equation (2), where the actual response $R$ and capacity $C$ define the fragility distribution, $R_m$ and $C_m$ are the median values associated with response and capacity, respectively. $\beta_{Rc}$ and $\beta_{Cc}$ represent a log standard deviation concerning the randomness and uncertainty for response and capacity, respectively.

$$f_0 = \Phi\left[\frac{\ln \frac{R_m}{C_m}}{\sqrt{\beta_{Rc}^2 + \beta_{Cc}^2}}\right] \tag{2}$$

The approach for determining fragility uses the response factor method with safety factor inputs in the given study [40]. Using the response factor method provides a good separation between response and capacity. The safety factor needs less input value than the response factor. Moreover, a sampling-based fragility assessment method, called DQFM [42], also has an advantage over the response factor method.

### 4.2. Trustworthiness of Risk Assessment

A novel study [44] focuses on the realism and trustworthiness of risk assessment. The study introduces trustworthiness as a risk assessment metric that shows confidence in the background knowledge, suitability, comprehensiveness, and completeness of PRA. Background knowledge, assumptions, conservatism, and sensitivity analysis fundamentally affect a risk analysis's realism and trustworthiness. Strong background knowledge and the modeling's fidelity are two main attributes of the suggested framework for evaluating trustworthiness. The hierarchical tree for trustworthiness evaluation represents different attributes and a 4-level approach to assessing trustworthiness to provide the reader the means how to evaluate the trustworthiness.

The level of trustworthiness of risk assessment is evaluated using a weighted average of the leaf attributes and using the Equation (3) is where $\omega_i$, which is is the weight of the leaf attribute, measures its relative contribution to the trustworthiness of risk assessment, $n$ is the number of the leaf attributes, and the trustworthiness score, $A_i$. The $i$-th leaf attribute is calculated based on the scoring guidelines, which can be reached in a related work's Appendix.

$$T = \sum_i^n \omega_i A_i \tag{3}$$

The weights are determined based on the Dempster Shafer-Analytical Hierarchy Process (DST-AHP) [45]. After obtaining the trustworthiness, the weighted posterior method integrates the risk index with the trustworthiness of the PRA for a single hazard group. After integration, the risk is expressed in terms of a subjective distribution on the probability that a given consequence will occur. Then, the estimated risk from different hazard groups is aggregated. In other words, risk distributions from different hazard groups are simply added, as shown in Equation (4).

$$Risk_{total} = \sum_i^n (Risk_i/T) \tag{4}$$

$Risk_{total}$ is the total risk considering the level of trustworthiness, $(Risk_i/T)$ is the risk from the hazard group $i$ given in the level of trustworthiness, and $n$ is the number of hazard groups.

The suggested framework can be applicable to risk-informed decision-making. However, even if the Dempster-Shafer is used to capture the uncertainty in the expert elicitation process of the relative weights of the attributes, uncertainty in the scoring was deferred to be addressed in a future study.

*4.3. The Need for a General Multi-Hazard PRA Framework*

As it can be seen in the previous two sub-sections, preliminary efforts have been put into evaluating multi-hazard events for NPPs; however, a general multi-hazard PRA framework still needs to be developed. The general multi-hazard PRA framework should include the necessary quantification methods to estimate the expected exceedance frequency of multi-hazard initiating events, both correlated and uncorrelated. Moreover, their combined effect on the plant control and safety systems should be systematically evaluated to enable the development and quantification of realistic event sequences of safety systems and human responses to such multi-hazard initiating events. The ultimate goal is to design and demonstrate the safety of advanced reactors in a wider range of challenging conditions since ignoring correlated or uncorrelated hazards may lead to unintended consequences.

**5. Advanced Non-LWR Licensing Basis Event Selection (LBE) Approach**

The U.S. NRC provides different guidance to reactor designers and applicants depending on if they are licensing a LWR or a non-LWR: Appendix A to 10 CFR Part 50 lists the general design criteria for LWRs [46] and regulatory guide 1.232 provides the guidance for developing principal design criteria for non-LWRs [47]. The guidance applies for all production and utilization facilities licensed under 10 CFR Part 50 or Part 52.

Moreover, the risk metrics for non-LWRs are different from the traditional risk metrics that are used for licensing LWRs. Core Damage Frequency (CDF), Large Early Release Frequency (LERF), or Conditional Containment Failure Probability (CCFP) are the surrogates for the Quantitative Health Objectives (QHO) for currently operating LWRs [48]. However, for advanced non-LRWs, a set of frequency-consequence criteria called F-C Target is used [15].

There are several different categories of events that need to be included during the licensing process. In LMP, the licensing basis events (LBEs) are the whole set of event sequences considered during the design and licensing phases, which include Anticipated Operational Occurrences (AOOs), Design Basis Events (DBEs), and Beyond Design Basis Events (BDBEs). The definition of design basis accidents (DBAs) is different from the current or common use for LWRs. They have postulated event sequences that are used to set design criteria and performance objectives for the design of safety-related structures, systems, and components. DBAs are derived from DBEs based on the capabilities and reliabilities of safety-related structures, systems, and components needed to mitigate and prevent event sequences, respectively.

The given categories in Figure 2 are based on the 5th and 95th percentiles of the event-sequence frequency of occurrence per plant-year, while the F-C Target (i.e., blue line) is based on both frequencies and consequences, as follows:

- AOOs are expected to occur with a frequency greater than $10^{-2}$/plant-year during the plant's life, either include single or multiple reactor modules. The F-C Target for high-frequency AOOs until $10^{-1}$/plant-year are based on an iso-risk profile defined by annual exposure limits of 10 CFR 20 [49], in other words, 100 mrem/plant-year. The frequencies between $10^{-1}$/plant-year and $10^{-2}$/plant-year, the F-C Target is set at a reference value of 1 rem by considering the Environmental Protection Agency (EPA) Protective Action Guide (PAG) limits [50].
- DBEs are expected to occur between the frequencies $10^{-2}$/plant-year and $10^{-4}$/plant-year, meaning the less frequent events. The F-C Target for this part is 1 rem at

$10^{-2}$/plant-year with 25 rem at $10^{-4}$/plant-year with the dose calculated at the Exclusion Area Boundary (EAB) for the 30 days following the release.

- BDBE frequency is less than $10^{-4}$/plant-year with the upper limit $5 \times 10^{-7}$/plant-year, meaning that rare events can be excluded from the BDBE category. The F-C Target for this case is 25 rem at $10^{-4}$/plant-year with 750 rem at $5 \times 10^{-7}$/plant-year to ensure that the QHO for early health effect does not exceed individual BDBEs.
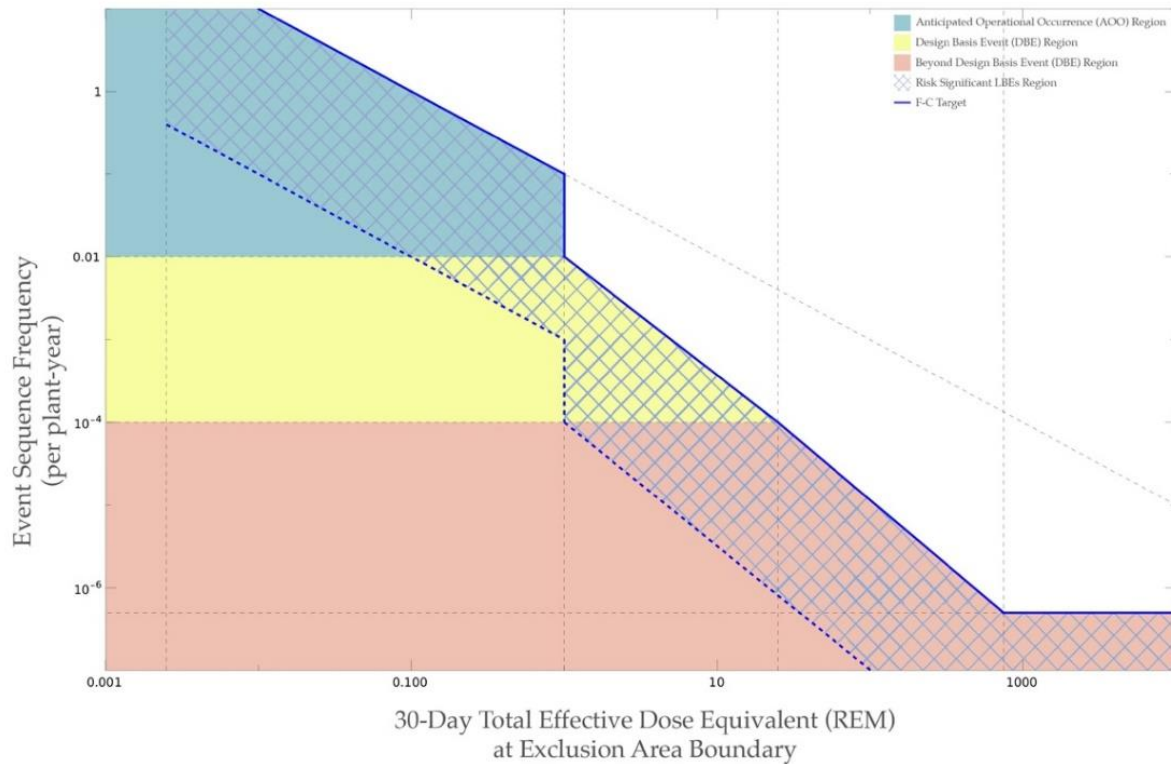


**Figure 2.** Frequency-Consequence Target (reproduced from [51]).

Also, it is essential to note that:

- The frequency-dose evaluation line includes increasing and decreasing risk arrows to evaluate the risk significance of each LBE.
- The frequency-dose anchor points in the figure are used to define the shape of the curve. The lines between anchor points are straight lines on a log-log frequency-dose graph.
- Finally, the event sequences with frequencies less than $5 \times 10^{-7}$/plant-year are kept in the PRA results to confirm no cliff-edge effects.

The current regulatory infrastructure was developed for reactor licensing in the 1960s and 1970s by considering the technology available at the time. On the other hand, the LBE approach addresses the following:

- The plant initiating event and event sequences are associated with the design and site.
- The response of the design and its structures, systems, and components (SSC) to initiating events and event sequence.
- Margins provided by the facility's response regarding prevention and mitigation of radiological releases.
- Adequacy of the defense-in-depth (DID) philosophy.

Moreover, one of the important elements of the LBE approach is having the F-C target. Performing either single or multi-hazard PRA will have a point on the F-C curve. This approach could give an indication on whether performing a multi-hazard PRA for the specific design with specific plant and site conditions is necessary or not.

## 6. Overview of Current Probabilistic Risk Assessment Standard for Non-Light Water Reactor (Non-LWR) Nuclear Power Plants

Probabilistic Risk Assessment Standard for Non-Light Water Reactor Nuclear Power Plants [16] is an American National Standard released on 8 February 2021. The Standard supports risk-informed decisions for advanced non-light water reactor NPPs and describes a method for applying the requirements for specific applications. The standard is applicable beginning from the design phase, and includes licensing, procurement, construction, operation, and maintenance of advanced non-LWR NPPs.

The standard contains 18 PRA elements covering the technical requirements necessary to be addressed in developing a full-scope PRA for advanced non-LWR NPPs. These are Plant Operating State Analysis (POS), Initiating Event Analysis (IE), Event Sequence Analysis (ES), Success Criteria Development (SC), Systems Analysis (SY), Human Reliability Analysis (HR), Data Analysis (DA), Internal Flood PRA (FL), Internal Fire PRA (F), Seismic PRA (S), Hazards Screening Analysis (HS), High Winds PRA (W), External Flooding PRA (XF), Other Hazards PRA (O), Event Sequence Quantification (ESQ), Mechanistic Source Term Analysis (MS), Radiological Consequence Analysis (RC), and Risk Integration (RI).

A critical look at the whole standard is necessary to assess its applicability to multi-hazard PRAs. In this section, we look only at the IE element [52–54] since this element forms the basis for multi-hazard PRA. There are three core steps for IE, that is identifying, grouping, and quantifying.

The first step is the identification of initiating events. The IE is used to identify all initiating events for both modeled plant operating states and the plant pre-operational stage with sources of radioactive material. The point that needs to be highlighted is that the requirement also touches on multi-hazard events. It considers the initiating events caused by a combination of hazards like seismically induced fires and flooding caused by fire sprinkler actuation. After the identification is complete, the initiating events are grouped to make the analysis more manageable. This grouping enables the estimation of the frequency of each modeled event sequence and event sequence family efficiently and realistically. The third step requires the quantification of the annual frequencies for each initiating event group.

In regard to multi-hazard, the standard does not explicitly define multi-hazard events. However, some related definitions in the standard include coexistent, concurrent, primary, and secondary hazards. The explanations for each of them are given below.

- *Coexisting hazard*: a hazard that is secondary to and/or concurrent with another hazard.
- *Concurrent hazard:* a hazard that co-occurs with the occurrence of another hazard resulting from a common cause (e.g., high winds concurrent with storm surge event caused by a hurricane or a moderate wind event concurrent with a significant rainfall event).
- *Primary hazard:* Hazards that are not the consequence of other preceding hazards.
- *Secondary hazard:* It is used in connection with, and in contrast to, a primary hazard. It is an additional hazard effect that is induced by the primary hazard.

Noteworthy to mention is that the coexistent hazards are mentioned many times in the standard. For instance, the effects of coexistent hazards on the fragilities included in the high winds PRA scope should be addressed, if applicable. Another example is the need to identify and address the coexisting hazards for flooding. These examples show that the importance of considering multi-hazard events.

The hazard screening analysis is critical to deciding which multi-hazard events should be included in the detailed analysis. The first objective of hazard screening is identifying the hazards that may affect the NPPs. The first step is vital to capture the multi-hazards. The standard asks not only to identify the site and hazards but also to identify the secondary hazards associated with hazards and hazard groups. Defining screening criteria is the second objective followed by performing previously described quantitative screening criteria. The supporting requirement for this objective also considers that the screened-out hazard or hazard group could not result in worse effects as another hazard with a significantly higher frequency. The defined quantitative screening criteria require addressing the

secondary hazards as well. Finally, even the screening criteria are not well defined at this point, and the multi-hazard events still need to be identified even if, ultimately, they can be screened out.

## 7. Conclusions

Multi-hazard risk quantification has become a necessary ingredient of a full-scope PRA in the nuclear industry since the Fukushima Daiichi NPP accident. Although the frequency of a multi-hazard event is small, even negligible in many cases, the consequences can be sufficient to challenge the regulatory limits.

Currently, 443 NPPs are in operation, and 50 NPPs are under construction [55], meaning at least a 10% increase in nuclear power plants. Assuming that the advanced reactors will be in our daily lives in 10 years, now is the right time to develop a general framework for assessing multi-hazard risks to inform the current design activities of advanced reactors.

A straightforward and verifiable technique applicable for different types of advanced reactors for multi-hazard PRA is inevitable for the next couple of years. As we discuss in this paper, the LMP approach provides the foundation for a licensing basis that can incorporate multi-hazard events. Also, the non-LWR PRA standard prescribes the requirements needed to be addressed for performing a full-scope PRA, although multi-hazards are not fully accounted for across all the PRA elements.

## References

1. Udalova, A.A. Nonpower applications of nuclear technology. In *Nuclear Reactor Technology Development and Utilization*; SKhan, U.-D., Nakhabov, A., Eds.; Woodhead Publishing: Cambridge, UK, 2020; pp. 319–341. [CrossRef]
2. Wu, J.; Ma, Y.; Yu, C.; Zou, C.; Cai, X.; Chen, J. Nuclear non-proliferation review and improving proliferation resistance assessment in the future. *Int. J. Energy Res.* **2021**, *45*, 11399–11422. [CrossRef]
3. Sehgal, B.R. Light water reactor (LWR) safety. *Nucl. Eng. Technol.* **2006**, *38*, 697–732.
4. Security Aspects of Nuclear Facilities | IAEA. Available online: https://www.iaea.org/topics/security-aspects (accessed on 21 February 2021).
5. Nuclear Regulatory Commission. Available online: https://www.nrc.gov/ (accessed on 6 June 2021).
6. U.S. Nuclear Regulatory Commission. *The Reactor Safety Study-The Introduction of Risk Assessment to the Regulation of Nuclear Reactors*; WASH-1400, NUREG/KM-0010; U.S. Nuclear Regulatory Commission: Washington, DC, USA, 2016; p. 60.
7. Apostolakis, G. The concept of probability in safety assessments of technological systems. *Science* **1990**, *250*, 1359–1364. [CrossRef] [PubMed]
8. Kaplan, S.; Garrick, B.J. On the quantitative definition of risk. *Risk Anal.* **1981**, *1*, 11–27. [CrossRef]
9. Gilbertson, A. An approach for using probabilistic risk assessment in risk-informed decisions on plant-specific changes to the licensing basis. In *U.S. Nuclear Regulatory Commission, Regulatory Guide 1.174, Revision 3*; U.S. Nuclear Regulatory Commission: Washington, DC, USA, 2018.
10. An approach for determining the technical adequacy of probabilistic risk assessment results for risk-informed activities. *U.S. Nuclear Regulatory Commission, Regulatory Guide 1.200, Revision 2*; U.S. Nuclear Regulatory Commission: Washington, DC, USA, 2009.

11. PRAProcedures Guide: A Guide to The Performance of Probabilistic Risk Assessments for Nuclear Power Plants (NU-REG/CR-2300) | NRC.gov. Available online: https://www.nrc.gov/reading-rm/doc-collections/nuregs/contract/cr2300/index.html (accessed on 31 August 2021).

12. Electric Power Research Institute. Basics of Nuclear Power Plant Probabilistic Risk Assessment. presented at the Fire PRA Workshop, San Diego, CA, USA and Jacksonville, FL, USA. 2011.

13. Advances in Small Modular Reactor Technology Developments. *A supplement to: IAEA Advanced Reactors Information System (ARIS)*; International Atomic Energy Agency: Vienna, Austria, 2020.

14. Hussein, E.M. Emerging small modular nuclear power reactors: A critical review. *Phys. Open* **2020**, *5*, 100038. [CrossRef]

15. Moe, W.L.; Afzali, A. *Modernization of Technical Requirements for Licensing of Advanced Non-Light Water Reactors Selection and Evaluation of Licensing Basis Events*; INL/EXT-19-55513-Rev000; Idaho National Lab. (INL): Idaho Falls, ID, USA, 2019. [CrossRef]

16. *Probabilistic Risk Assessment Standard for Advanced Non-LWR Nuclear Power Plants*; ASME/ANS RA-S-1.4-2021; ASME: New York, NY, USA, 2021.

17. *Treatment of External Hazards in Probabilistic Safety Assessment for Nuclear Power Plants: A Safety Practice*; International Atomic Energy Agency: Vienna, Austria, 1995. Available online: https://www.iaea.org/publications/5178/treatment-of-external-hazards-in-probabilistic-safety-assessment-for-nuclear-power-plants-a-safety-practice (accessed on 22 July 2021).

18. Knochenhauer, M.; Louko, P. Guidance for external events analysis. In *Probabilistic Safety Assessment and Management*; Spitzer, C., Schmocker, U., Dang, V.N., Eds.; Springer: London, UK, 2004; pp. 1498–1503. [CrossRef]

19. U.S. Nuclear Regulatory Commission. *Procedural and Submittal Guidance for the Individual Plant Examination of External Events (IPEEE) for Severe Accident Vulnerabilities (NUREG-1407)*; U.S. Nuclear Regulatory Commission: Washington, DC, USA, 1991; p. 98.

20. Kimura, C.Y.; Prassinos, P.G. *Evaluation of External Hazards to Nuclear Power Plants in the United States: Other External Events*; U.S. Nuclear Regulatory Commission: Washington, DC, USA, 1989. Available online: http://inis.iaea.org/search/search.aspx?orig_q=RN:20045139 (accessed on 22 July 2021).

21. *External Man-Induced Events in Relation to Nuclear Power Plant Design: A Safety Guide*; International Atomic Energy Agency: Vienna, Austria, 1996. Available online: https://www.iaea.org/publications/5184/external-man-induced-events-in-relation-to-nuclear-power-plant-design-a-safety-guide (accessed on 22 July 2021).

22. Final Safety Evaluation Report Related to Certification of the AP1000 Standard Design (NUREG-1793, Initial Report) | NRC.gov. Available online: https://www.nrc.gov/reading-rm/doc-collections/nuregs/staff/sr1793/initial/index.html#pub-info (accessed on 23 August 2021).

23. Choi, E.; Ha, J.-G.; Hahm, D.; Kim, M.K. A review of multihazard risk assessment: Progress, potential, and challenges in the application to nuclear power plants. *Int. J. Disaster Risk Reduct.* **2021**, *53*, 101933. [CrossRef]

24. Great East Japan Earthquake. Available online: https://www.who.int/westernpacific/emergencies/great-east-japan-earthquake (accessed on 6 January 2021).

25. International Nuclear and Radiological Event Scale (INES) | IAEA. Available online: https://www.iaea.org/resources/databases/international-nuclear-and-radiological-event-scale (accessed on 11 June 2021).

26. Wang, J.; He, Z.; Weng, W. A review of the research into the relations between hazards in multi-hazard risk analysis. *Nat. Hazards* **2020**, *104*, 2003–2026. [CrossRef]

27. Kappes, M.S.; Keiler, M.; von Elverfeldt, K.; Glade, T. Challenges of analyzing multi-hazard risk: A review. *Nat. Hazards* **2012**, *64*, 1925–1958. [CrossRef]

28. Kim, J.H.; Kim, M.K.; Choi, I.-K. Preliminary study on the quantification of component level failure frequency by multi-hazard. In Proceedings of the KNS 2017 Fall Meeting, Charlotte, NC, USA, 4–9 August 2019; p. 6.

29. Yu, Y.; Lv, X.; Niu, F. Large LOCA accident analysis for AP1000 under earthquake. *Ann. Nucl. Energy* **2015**, *77*, 142–147. [CrossRef]

30. Prošek, A.; Wielenberg, A.; Löffler, H.; Raimond, E. Methodology for selecting initiating events and hazards for consideration in an extended PSA. In *Safety and Reliability—Theory and Applications*; Tatras Mountains, Tatranské Matliare, Slovak Republic, Ed.; CRC Press: Boca Raton, FL, USA, 2017; p. 490. [CrossRef]

31. Daniell, J.; Schaefer, A.; Wenzel, F.; Hacker, E. *Review of State-of-the Art for Hazard and Multi-Hazard Characterization*; European Commission: Brussels, Belgium, 2019.

32. Home | NARSIS. Available online: http://www.narsis.eu/ (accessed on 23 July 2021).

33. Kwag, S.; Hahm, D. Development of an earthquake-induced landslide risk assessment approach for nuclear power plants. *Nucl. Eng. Technol.* **2018**, *50*, 1372–1386. [CrossRef]

34. Kwag, S.; Gupta, A. Probabilistic risk assessment framework for structural systems under multiple hazards using Bayesian statistics. *Nucl. Eng. Des.* **2017**, *315*, 20–34. [CrossRef]

35. Li, H.; Apostolakis, G.E.; Gifun, J.; VanSchalkwyk, W.; Leite, S.; Barber, D. Ranking the risks from multiple hazards in a small community. *Risk Anal.* **2009**, *29*, 438–456. [CrossRef] [PubMed]

36. Cooper, S.E.; Xing, J.; Chang, Y.J. What HRA needs to support site-wide, multi-hazard Level 2-PRA. In Proceedings of the ANS PSA 2013 International Meeting on Probabilistic Safety Assessment and Analysis, Columbia, SC, USA, 22–26 September 2013.

37. Choine, M.N.; O'Connor, A.; Gehl, P.; D'Ayala, D.; García-Fernández, M.; Jiménez, M.-J.; Gavin, K.; Van Gelder, P.; Salceda, T.; Power, R. A multi hazard risk assessment methodology accounting for cascading hazard events. In Proceedings of the 12th International Conference on Applications of Statistics and Probability in Civil Engineering, Vancouver, Canada, 12–15 July 2015; p. 8.

38. Ming, X.; Xu, W.; Li, Y.; Du, J.; Liu, B.; Shi, P. Quantitative multi-hazard risk assessment with vulnerability surface and hazard joint return period. *Stoch. Environ. Res. Risk Assess.* **2015**, *29*, 35–44. [CrossRef]

39. He, Z.; Weng, W. A risk assessment method for multi-hazard coupling disasters. *Risk Anal.* **2020**, *41*, 1362–1375. [CrossRef] [PubMed]

40. Kwag, S.; Ha, J.G.; Kim, M.K.; Kim, J.H. Development of efficient external multi-hazard risk quantification methodology for nuclear facilities. *Energies* **2019**, *12*, 3925. [CrossRef]

41. *Seismic Probabilistic Risk Assessment Implementation Guide*; EPRI: Palo Alto, CA, USA, 2013.

42. Watanabe, Y.; Oikawa, T.; Muramatsu, K. Development of the DQFM method to consider the effect of correlation of component failures in seismic PSA of nuclear power plant. *Reliab. Eng. Syst. Saf.* **2003**, *79*, 265–279. [CrossRef]

43. Ebisawa, K.; Abe, K.; Muramatsu, K.; Itoh, M.; Kohno, K.; Tanaka, T. Evaluation of response factors for seismic probabilistic safety assessment of nuclear power plants. *Nucl. Eng. Des.* **1994**, *147*, 197–210. [CrossRef]

44. Bani-Mustafa, T.; Zeng, Z.; Zio, E.; Vasseur, D. A new framework for multi-hazards risk aggregation. *Saf. Sci.* **2020**, *121*, 283–302. [CrossRef]

45. Dezert, J.; Tacnet, J.-M.; Batton-Hubert, M.; Smarandache, F. Multi-criteria decision making based on DSmT-AHP. In Proceedings of the ELIEF 2010: Workshop on the Theory of Belief Functions, Brest, France, 1–2 April 2010; p. 9.

46. Appendix A to Part 50—General Design Criteria for Nuclear Power Plants | NRC.gov. Available online: https://www.nrc.gov/reading-rm/doc-collections/cfr/part050/part050-appa.html (accessed on 30 September 2021).

47. U.S. Nuclear Regulatory Commission. *Guidance for Developing Principal Design Criteria for Non-Light-Water Reactors*; Idaho National Lab. (INL): Idaho Falls, ID, USA, 2015.

48. Oh, K.M.; Ahn, S.K.; Jo, N.C.; Yoon, W.H.; Chung, D.W.; Lee, H.J. Comparison of Frequency-Consequence (F-C) curves or criteria in foreign countries. In Proceedings of the Transactions of the Korean Nuclear Society Autumn Meeting, PyeongChang, Korea, 25–26 October 2007; pp. 593–594.

49. PART 20—Standards FOR Protection against Radiation. Available online: https://www.nrc.gov/reading-rm/doc-collections/cfr/part020/index.html (accessed on 5 August 2021).

50. Protective Action Guides (PAGs) | US EPA. Available online: https://www.epa.gov/radiation/protective-action-guides-pags (accessed on 5 August 2021).

51. Moe, W.L. *NEI 18-04 Risk-Informed Performance-Based Technology Inclusive Guidance for Advanced Reactor Licensing Basis Development*; INL/EXT-19-55375-Rev000; Idaho National Lab. (INL): Idaho Falls, ID, USA, 2019. [CrossRef]

52. Sandia National Laboratories and, U.S. *Nuclear Regulatory Commission. Handbook of Parameter Estimation for Probabilistic Risk Assessment*; NUREG/CR-6823; U.S. Nuclear Regulatory Commission: Washington, DC, USA, 2003.

53. Idaho National Engineering and Environmental Laboratory. *Rates of Initiating Events at U.S. Nuclear Power Plants 1987–1995*; UREG/CR-5750; U.S. Nuclear Regulatory Commission: Washington, DC, USA, 1999. Available online: https://www.nrc.gov/docs/ML0705/ML070580080.pdf (accessed on 24 July 2021).

54. IAEA. *Technical Approach to Probabilistic Safety Assessment for Multiple Reactor Units*; International Atomic Energy Agency: Vienna, Austria, 2019; Available online: https://www.iaea.org/publications/12228/technical-approach-to-probabilistic-safety-assessment-for-multiple-reactor-units (accessed on 22 July 2021).

55. PRIS-Home. Available online: https://pris.iaea.org/PRIS/home.aspx (accessed on 21 February 2021).

# An Investigation into Current Sand Control Testing Practices for Steam Assisted Gravity Drainage Production Wells

**Omar Kotb, Mohammad Haftani * and Alireza Nouri**

Department of Civil and Environmental Engineering, University of Alberta, Edmonton, AB T6G 2E3, Canada; kotb@ualberta.ca (O.K.); anouri@ualberta.ca (A.N.)
* Correspondence: haftani@ualberta.ca

**Abstract:** Sand control screens (SCD) have been widely installed in wells producing bitumen from unconsolidated formations. The screens are typically designed using general rules-of-thumb. The sand retention testing (SRT) technique has gained attention from the industry for the custom design and performance assessment of SCD. However, the success of SRT experimentation highly depends on the accuracy of the experimental design and variables. This work examines the impact of the setup design, sample preparation, near-wellbore stress conditions, fluid flow rates, and brine chemistry on the testing results and, accordingly, screen design. The SRT experiments were carried out using the replicated samples from the McMurray Formation at Long Lake Field. The results were compared with the test results on the original reservoir samples presented in the literature. Subsequently, a parametric study was performed by changing one testing parameter at a test, gradually making the conditions more comparable to the actual wellbore conditions. The results indicate that the fluid flow rate is the most influential parameter on sand production, followed by the packing technique, stress magnitude, and brine salinity level. The paper presents a workflow for the sand control testing procedure for designing the SCD in the steam-assisted gravity drainage (SAGD) operations.

**Keywords:** sand retention test; sand control devices; sand production; sand replication; SAGD

## 1. Introduction

### 1.1. Background

Steam-assisted gravity drainage (SAGD) is a thermal production technique for oil extraction from bituminous reservoirs. A pair of horizontal wells are drilled into the reservoir formation. Steam is injected into the reservoir through the injection well to form a steam chamber, heat the bitumen, and reduce its viscosity through conduction and convection (Figure 1). An emulsified water-in-oil bitumen along with condensed water flow by gravity toward the lower well to be pumped out to the surface [1,2].

The in-situ bitumen reserves in Canada are located in unconsolidated oil sands. These naturally loose reservoirs require sand control devices (SCD) to simultaneously support the well, avoid excessive sand production, and let the fine particles pass through the slots and avoid plugging [2–4]. Different types of standalone screens (SAS) have been employed to complete the SAGD wells. The slotted liner (SL), wire-wrapped screen (WWS), and punched screen (PS) are the main options [5–7]. The slotted liner has been the most common SAS in SAGD operations due to its lower cost, mechanical strength, and reasonable performance in unconsolidated and high-permeable oil sands [4]. The sand control screen's success can be measured by its ability to control sand production and prevent formation permeability impairment [4].
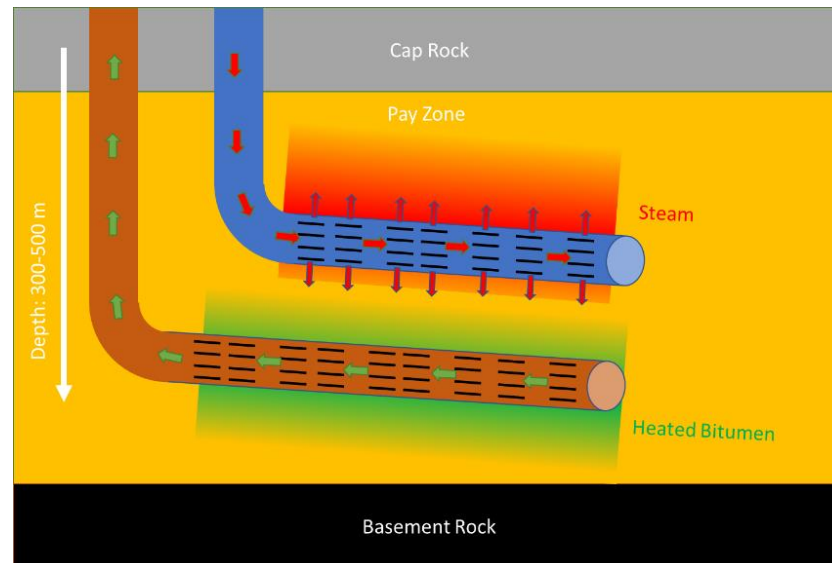
**Figure 1.** Schematic of a pair of horizontal wells in SAGD operations.

*1.2. Existing Testing Setups*

Sand control testing generally aims to evaluate the performance of SCDs under lab conditions emulating in-situ wellbore conditions for screen selection and design purposes. The screens' performance is usually assessed considering sand production and permeability variations.

A few sand control testing setups and procedures are introduced in the literature, attempting to simulate SAGD production wells. The testing devices in the literature can be categorized into two groups: slurry sand retention testing [8–14] and pre-pack sand retention testing [3,4,7,13,15–25]. In the slurry testing, a low concentration slurry of sand and fluids is pumped into an empty cell towards the SCD coupon at a constant flow rate. In contrast, in the pre-pack testing, the sand is already packed in the testing cell before flowing fluids towards the SCD coupon. The pre-pack test can represent SAGD well conditions more realistically as the formation collapses onto the liner at the early stages of production. The formation collapse occurs as bitumen melts when exposed to the steam, destroying the oil sand's intergranular cementation [26].

The first pre-pack setup and procedure for SAGD wells were introduced by Bennion et al. [27] and modified by Romanova et al. [19] and O'Hara [21]. They prepared the sand pack by pouring clean and dry sand particles to fill a cylindrical cell (diameter 6.36 cm, length 40 cm) and then saturating the sand pack with brine. The sand was packed by applying a 500-psi axial load (35.15 kg/m$^2$). Subsequently, flowing fluids (oil, brine, and gas) were injected into the sand pack in 12 stages by incrementally increasing the flow rate. The flow program consisted of four single-phase oil stages, four two-phase stages, three three-phase stages, and a final stage of single-phase oil. The produced solids and pressure drops across the sand pack and near the SCD coupon region were recorded. Bennion et al. [27] employed single-slot coupons underneath the sand pack sample to evaluate the slotted liners' performance.

Devere-Bennett [3] employed the same testing equipment and procedure while scaling down the fluid flow rates used by Bennion et al. [4]. The author argued that the fluid flow rates in past testing were much higher than those expected for SAGD production wells.

Mahmoudi [22] introduced a new testing facility with larger dimensions, accompanied by a modified testing procedure and multi-slot coupons. Commercial sands were used to replicate the oil sands based on the characterization of core samples. The sand was packed layer by layer using the moist tamping method [28] to create a uniform pack. Then, the sample was saturated under 2 psi (0.14 kg/cm$^2$) axial stress to avoid fluidization, and the low stress was maintained during the flow test to simulate the worst-case scenario of sand

production. Brine with a controlled pH and salinity was injected at seven incremental flow rates. Mahmoudi [22] reported the pressure drop across the sand pack and the near-coupon region, the amount of produced fines and sand, and the concentration of fines along the sand pack to develop design criteria for slotted liners.

Anderson [23] developed a large-scale setup to simulate the radial flow regime around the SAGD wells. The work showed a general agreement between the results obtained from the linear and radial flow setups.

Wang et al. [24], Fattahpour et al. [25], and Guo et al. [29] introduced a scaled completion testing (SCT) facility to evaluate the performance of different SCDs at various axial and lateral stress conditions. They used the moist tamping technique and representative brine pH and salinity.

Haftani et al. [30] developed a prepacked full-scale completion testing (FCT) apparatus to simulate the radial-flow regime around the production wells. This facility utilizes a cylindrical-shaped screen to replace a disk-shaped screen coupon employed in other testing devices.

In summary, no standard testing setup and procedure have been established to evaluate the SCD for SAGD wells. It appears that the testing setup and procedure are subjective to the researcher's understanding of the wellbore conditions being simulated.

### 1.3. Essential Factors in SCD Design

Different design criteria have been introduced to select the proper sand control device and design the optimum opening size [8,14,31,32]. Particle size distribution (PSD) of the sand in the reservoir is considered the main factor for selecting the SCD type and aperture size [33,34]. Coberly [8] found that stable bridges would always form when the SCD aperture is below D10, and instability would occur at the sizes above 2 × D10. Rogers [35] and Suman et al. [36] also suggested D10 for WWS and SL aperture sizing. Gillespie et al. [10] recommended that 2 × D50 and 2.5 × D50 can be considered as the maximum aperture width for WWS and PS, respectively. Fermaniuk [31] suggested 2 × D70 as the lower band to let the fine particles freely pass the slot and 3.5 × D50 as the largest slot size to avoid excessive sand production. More recent design criteria incorporate the plugging tendency and sanding level to graphically provide a safe slot window for the slotted liner [24,29,32,37].

High fluid flow rates have been linked to SCD failure due to sand production [38] and near-screen plugging caused by fines migration [13,39]. Excessive sanding [38] and fines migration [40] may occur when the flow velocity exceeds the critical velocity. Therefore, in SRT experiments, injecting the fluids at representative flow rates, reflecting the operational condition, is substantial for design purposes.

Mahmoudi et al. [41] and Haftani et al. [42] reported a significant effect of the brine's pH and salinity on the fines migration from the sensitivity analysis on sand production through several sand retention tests. These observations are consistent with the experimental findings in the relation between fines migration and permeability impairment with the salinity and pH levels of the injected brine [4,40,43]. It has been found that fines migration and permeability reduction are more significant at the lower brine salinity and higher pH levels. Khilar and Fogler [44] introduced a critical salt concentration (CSC) as a certain salinity level below which fine particles detach from the pore walls and migrate inside the porous media. Moreover, it is reported that fines migration is more sensitive to monovalent cations concentration than divalent cations [44]. Therefore, due to the significant effect of the chemical properties of the producing water on the fines migration and permeability variation, it is recommended to conduct the experimental tests emulating the salinity and pH level of the producing water from the wellbores.

Clearly et al. [45] investigated the sand arches' stability at different stress levels and concluded that more stable sand arches could form in higher stress levels. Coskuner and Maini [46] reported that high-stress levels decrease the critical velocity and mobilize the fine particles at a lower flow rate. Wang et al. [24] and Guo et al. [29] performed experimental

studies to evaluate SCD performance and found less sanding and more fines migration at the higher stress magnitudes. Several scholars have applied different axial stress levels in SRT experiments from 2 psi to 500 psi [3,22,47]. The low-stress condition at the screen–oil sands interface has been noted at early SAGD well life when the formation sand collapses over the screen [26].

The sand pack preparation method has also significantly impacted the stress-strain behavior [48–50] and liquefaction resistance of sands [51]. Multiple techniques have been proposed to compact the sands and produce homogenous sand pack samples [28,51–54]. Two packing methods have been employed for SAGD sand control testing in the literature, including dry packing and moist tamping methods. The dry technique involves pouring dry sand into the core holder and applying stress over the sand for compaction [4,19,21]. In the moist tamping method developed by Ladd [28], the sand mixture is packed layer by layer to achieve minimal density and porosity variations [55]. The moist tamping method was employed for sand control testing by Montero et al. [7], Guo et al. [29], and Mahmoudi et al. [56,57].

The literature review reveals that no appropriate protocol has been standardized for sand retention testing to evaluate the SCD's performance. The testing procedures seem to be subjective, leading to uncertainties in the testing outcomes and interpretation. This paper quantifies and compares the relative impact of testing parameters on sand control test results. The paper proposes a workflow for sand control testing that can be used for SRT experimentation of the sand control screens for SAGD production wells.

## 2. Experimental Design

This section describes the experimental setup specifications, testing plan, and testing parameters employed in this research.

### 2.1. Experimental Setup

As shown in Figure 2, the SRT facility employed in this study includes six units of (1) sand pack holder/cell, (2) fluid injection unit, (3) loading frame, (4) data acquisition and monitoring unit, (5) sand and fines production measurement unit, and (6) back-pressure unit.
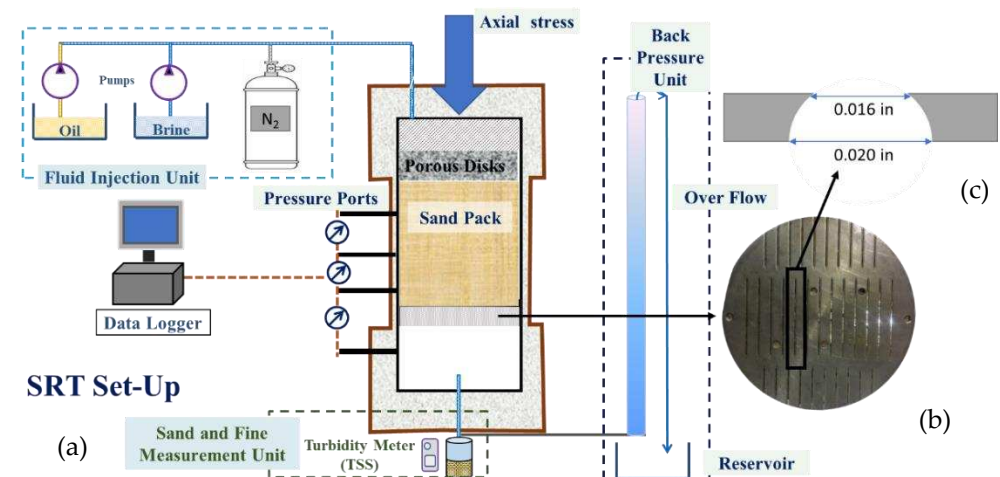


**Figure 2.** (**a**) Schematic of the SRT setup, (**b**) slotted liner coupon, and (**c**) cross-section of a rolled top slot.

The SRT cell comprises a core holder (inner diameter of 17.1 cm and length of 47 cm), top platen, and base plate. The sand is packed inside the core holder made of aluminum with working pressure up to 690 kPa at 20 °C. Three ports are installed on the cell circumference at 5.08, 17.78, and 30.48 cm from the bottom to measure the differential pressure at the bottom, middle, and top intervals of the sand pack. The setup allows using 17.1-cm-diameter SCD coupons, such as the slotted liner (SL) coupon shown in Figure 2. The top

platen transfers the load from the load frame to the sand pack and hydraulically seals the core holder using O-rings. In combination with a porous disk, the top platen provides the conduit for fluid injection, mixing, and uniform distribution into the sand pack. The cell's outlet is connected to a 185-cm-long back-pressure column to generate minor back-pressure (2.5 psi) on the sand pack during the saturation and flow phases. Low flow rates are introduced into the sand pack during the saturation phase through this back-pressure column to avoid the flow channeling.

The fluid injection unit consists of two mechanically actuated triplex diaphragm pumps. Both pumps can inject brine and mineral oil at rates of up to 18 L per hour with an adjustment range of 1:50 and flow rate adjustment of $\pm1\%$. The flow rate output is controlled by adjusting the stroke length and the pump frequency through a variable frequency drive. The triplex design couples with a pulse damper to introduce pulsation-free fluid flow into the cell. Gas is delivered through a high-pressure nitrogen tank connected to a pressure regulator and gas rotameter with a choke to simulate the steam breakthrough in SAGD wells. Stress is applied to the sand pack from the load frame through the top platen. The load frame can apply a maximum force of 8 metric tons.

Three differential transducers with a maximum range of 74 kPa (10.8 psi) and accuracy of $\pm0.022\%$ of the full range are installed to record the pressure difference in certain intervals of the sand pack. The transducers are connected to the National Instruments Data Acquisition System Model USB-6002 to continuously record the differential pressures on a computer using DAQ express software.

The sand and fines production measurement unit consists of a sand trap to collect the produced sand and fines. The sand trap is a flanged cylinder with a blind flange at the bottom to collect the produced sand. A narrow pipe is installed beneath the coupon to take fluid samples for quantifying the concentration of fines inside the produced water using a turbidity meter.

### 2.2. Sand Pack Material Preparation

Commercial sands were used to replicate Long Lake, Alberta's sand prints, as reported in Devere-Bennett [3]. The replica sand was prepared by mixing commercial sands and fines in specific proportions following the technique proposed by Mahmoudi et al. [57]. Figure 3 presents the cumulative PSD curves of the original PSD of Long Lake and the pertinent replica sand mixture. The replicated PSD shows a maximum deviation of 13 µm (0.0005 inches) from the original sand. This deviation is acceptable since it is within the $\pm50$ µm ($\pm0.002$ inches) tolerance in the slotted liner manufacturing.
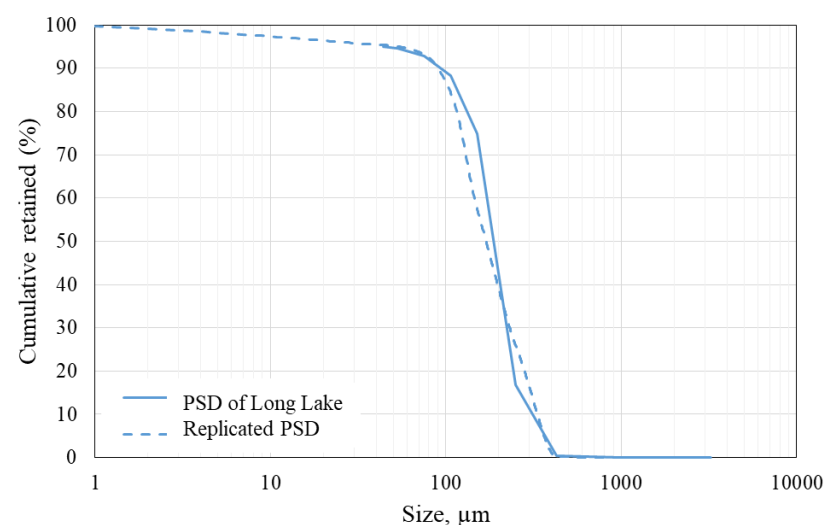


**Figure 3.** Cumulative PSD of Long Lake oil sand print and the relevant replicated sample.

### 2.3. Testing Matrix

The testing matrix in Table 1 is designed to investigate the effect of the experimental setup and testing parameters on evaluating the liner performance for the SAGD operations. Test #1 aims to assess the variations in the testing results due to using a different testing device than the one used by Devere-Bennett [3] while using the same testing procedure. Hence, one of the tests reported by Devere-Bennett [3] is repeated with the setup used in this investigation. This test uses the replica sand with the same PSD as the oil sand used in Devere-Bennett [3].

In Tests #2 through #6, testing parameters are changed, one at a time, from those used in Devere-Bennett [3] to parameters believed to be more representative of SAGD producer conditions. The aim is to assess the effect of testing parameters on sanding and plugging of a rolled top slotted liner with the slot width of 406–508 μm (0.016–0.020 inches), see Figure 2c for the slot cross-section.

**Table 1.** Testing plan.

| Test No. | SCD Coupon | Flow Rates | | | Packing Technique | Water Composition | Stress |
|----------|-----------|------------|--|--|-------------------|-------------------|--------|
| | | Brine | Oil | Nitrogen | | | |
| 1 | WWS 0.006 in/152 μm | Devere-Bennett [3] | | | Dry packing | 1% NaCl | 350 psi 2413 kPa |
| 2 | Slotted Liner RT OFA 2.33%, 0.016 in/ 406 μm | Devere-Bennett [3] rates | | | Dry packing | 1% NaCl | 350 psi 2413 kPa |
| 3 | Slotted Liner RT OFA 2.33%, 0.016 in/406 μm | Devere-Bennett [3] rates | | | Moist tamping | 1% NaCl | 350 psi 2413 kPa |
| 4 | Slotted Liner RT OFA 2.33%, 0.016 in/406 μm | Devere-Bennett [3] rates | | | Moist tamping | 1% NaCl | 60 psi 413 kPa |
| 5 | Slotted Liner RT OFA 2.33%, 0.016 in/406 μm | Representative rates | | | Moist tamping | 1% NaCl | 60 psi 413 kPa |
| 6 | Slotted Liner RT OFA 2.33%, 0.016 in/406 μm | Representative rates | | | Moist tamping | Field representative ion composition | 60 psi 413 kPa |

In Test #1, the testing procedure and parameters such as the sand packing technique, axial stress magnitude, fluid properties, and flow rates were matched with those reported in Devere-Bennett [3]. Test #1 was performed using a WWS coupon. Owing to the smaller size of the testing specimen and screen coupon utilized by Devere-Bennett [3], the original flow rate (Figure 4) was upscaled for the larger screen coupon area used in this work by Equation (1):

$$Q_{Devere-Bennett} = Q_{Devere-Bennett\ orignial} \times \frac{Area_{New\ setup}}{Area_{Devere-Bennett's\ setup}} \tag{1}$$

Test #2 accommodated an SL coupon instead of the WWS in Test #1. Test #3 included all the testing specifications of Test #2, except for the packing procedure. In Test #3, the moist tamping technique [57] was used to pack the sand instead of the dry packing technique used in Tests #1 and 2 and Devere-Bennett's [3] experiments. Rhodes [58] stated that dry packing could cause the segregation of particles across the sample as the finer particles settle at the bottom by the percolation of fine particles or vice versa due to elutriation segregation.

SAGD wells experience a low-stress condition at the early production period [26,29]. Applying excessive stress in the SRT underestimates sand production and overestimates the fines migration [24,29]. However, a minimum magnitude of stress is required during the sand pack saturation to avoid fluidization, which can be estimated using fluidized bed theory [59]. A relatively high magnitude of stress was applied on the sand pack (350 psi or 2413 kPa) in Devere-Bennett's [3] experiment and Tests #1 to 3 in this work. However, Test #4 used a lower stress of 60 psi (413 kPa), which is deemed more representative of the early SAGD wellbore life.
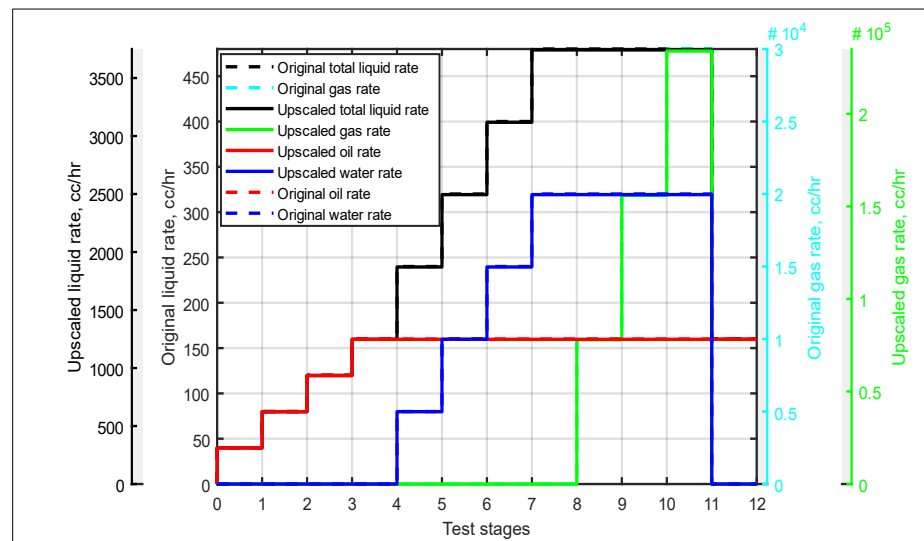
**Figure 4.** Original flow rate employed by Devere-Bennett [3] and upscaled fluid flow rates to the new testing setup (note: the original flow rates presented in the dash lines on the graph are covered by the relevant upscaled flow rates. For instance, the original water rate is covered by the upscaled water rate.).

Typical production well flow rates, well length, and water cut levels were extracted from public regulatory reports for the Long Lake Field [60] to determine representative flow rates of the production wells for experimentation. Subsequently, a range of factors was considered to account for the non-uniformity in fluid flow, slot plugging, and non-contributing liner sections. The non-uniformity of production throughout the wellbore could originate from non-uniform steam chamber growth due to reservoir heterogeneity [61], lack of flow control devices [62], and undulations in the wellbore trajectory [63]. In a case study, Beshry et al. [64] stated that non-uniformity in a SAGD production well could reach up to 50%. Furthermore, the non-contributing sections due to the liner connections of the sand control completion could be approximately 20% of the well length. Plugging of sand control devices as a result of scale deposition [65], corrosion products, and fouling by fines deposition can cause plugging of up to 90% of the slots [66].

Table 2 presents the range of values assigned to each factor affecting the effective flow. For the effective flow rates calculation, the assumption and information presented in Tables 3 and 4 were used as laboratory testing parameters, typical Long Lake well specifications, and field production parameters.

Table 5 presents different fluid flow scenarios to calculate the flow rates for the experimental tests (Table 5). Table 6 shows the parameters for the estimation of the field equivalent stream rates in the lab setup. Steam properties [67] are estimated based on the production wells' downhole pressure and injection wells' temperature conditions in Long Lake Field to simulate the steam breakthrough. As it is difficult to assess the in-situ steam quality in the reservoir, a steam quality of 25% is assumed. The steam properties at downhole conditions are applied to estimate the downhole rates based on the average surface injection rates in Long Lake Field. Subsequently, the influx per unit area is calculated for a typical SAGD with a length of 800 m and liner diameter of 17.1 cm. The influx is then multiplied by the SRT coupon area to determine the test flow rate before applying the effective flow factors.

**Table 2.** Effective flow coefficients for calculating representative rates in the testing.

| Scenarios | Non-Uniform Flow Condition | Plugging Factor | Non-Contributing Liner Sections | Effective Flow Coefficient |
|---|---|---|---|---|
| Favorable condition scenario | 0.8 | 0.5 | 0.8 | 0.32 |
| Non-uniform flow scenario | 0.5 | 0.5 | 0.8 | 0.20 |
| Plugged and non-uniform flow scenario | 0.5 | 0.3 | 0.8 | 0.12 |

**Table 3.** Lab and field parameters.

| Calculation Parameters | | |
|---|---|---|
| Coupon diameter * | 17.1 | cm |
| Coupon area * | 229.7 (0.023) | $cm^2$ ($m^2$) |
| Average steam injection rate ** | 270 | $m^3$/day |
| Injected water density | 1000 | kg/$m^3$ |
| Mass of injection water | 270,000 | kg/day |
| Well length ** | 0.8 | km |
| Assumed steam quality | 0.5 | - |

* Testing setup specifications, ** Data source CNOOC [60].

**Table 4.** Field production information.

| Field Information | | |
|---|---|---|
| Oil rate * | 80 | $m^3$/d |
| Water rate ** | 270 | $m^3$/d |
| Liquid rate | 350 | $m^3$/d |
| Liquid rate | 2201 | bbl/d |
| WOR * | 3.38 | - |
| Length of wells * | 0.80 | Km |
| Oil rate/length | 0.10 | $m^2$/d |
| Water rate/length | 0.34 | $m^2$/d |
| Liquid rate/length | 0.44 | $m^2$/d |
| Oil rate/liner surface area | 0.18 | m/d |
| Water rate/liner surface area | 0.60 | m/d |
| Liquid rate/liner surface area | 0.78 | m/d |
| Liquid rate/liner surface area | 0.46 | bbl/$ft^2$ |

* Data source CNOOC [60], ** Provided by CNOOC for this investigation.

The calculated injection flow rates for representative tests are shown in Figure 5. However, the effect of flow rate on the sanding and plugging levels was investigated by comparing the SRT results conducted at the low flow rates in Test #4 and the customized flow rates for the Long Lake Field in Test #5.

**Table 5.** Scenarios of effective flow represented in the testing.

| Scenarios | Effective Flow | Lab Equivalent Oil Rate, cc/h | Lab Equivalent Water Rate, cc/h | Lab Equivalent Liquid Rate, cc/h |
|---|---|---|---|---|
| Perfect SAGD well condition | 1.0 | 171 | 578 | 749 |
| Favorable condition | 0.32 | 535 | 1807 | 2342 |
| Non-uniform flow | 0.20 | 857 | 2891 | 3747 |
| Plugged and non-uniform flow | 0.12 | 1428 | 4818 | 6246 |

**Table 6.** Steam rate calculations.

| Steam Rate Calculations | | |
|---|---|---|
| Temperature | 210 | °C |
| Pressure | 1600 | kPag |
| Density of steam | 8.42 | kg/m$^3$ |
| Steam viscosity | 0.016 | cp |
| Steam rate | 31,544 | m$^3$/d |
| Length of wells | 0.8 | km |
| Steam rate/length | 39.4 | m$^3$/m |
| Steam rate/surface area | 70.6 | m$^3$/m$^2$ |
| Lab equivalent steam rate | 67,548 | cm$^3$/h |
| Lab equivalent steam rate | 1.1 | L/min |
| Steam quality | 25 | % |
| Good scenario steam rate | 0.9 | L/min |
| Non-uniform scenario steam rate | 1.4 | L/min |
| Plugged and non-uniform scenario steam rate | 2.3 | L/min |
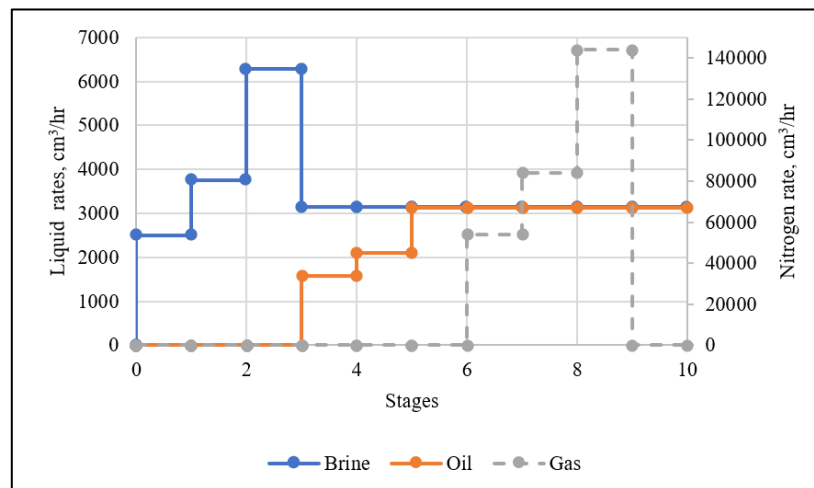


**Figure 5.** Testing stages and associated fluid rates of the testing procedure used herein, used for Tests #5 and #6.

The brine composition of 10,000 ppm employed in Devere-Bennett [3] does not agree with the salinity reported for the produced brine from the Long Lake reservoir, as shown in Table 7. Therefore, the effect of salinity was studied by comparing sanding and plugging under excessively high salinity in Test #5 used by Devere-Bennett [3] with recreated brine composition based on produced brine from the well in Test #6.

**Table 7.** Chemical composition of produced water in Long Lake reservoir and brine used in testing reported by Devere-Bennett [3].

| | Description | Unit | Field Data * | Brine Composition |
|---|---|---|---|---|
| | pH | - | 8.04 | N/A |
| | Sodium (Na) | mg/L | 194 | 3935 |
| Cations | Potassium (K) | mg/L | 3.6 | - |
| | Calcium (Ca) | mg/L | 56.7 | - |
| | Magnesium (Mg) | mg/L | 21.4 | - |
| | Bicarbonate (HCO3) | mg/L | 708 | - |
| | Carbonate (CO3) | mg/L | 0.5 | - |
| Anions | Hydroxide (OH) | mg/L | 0.5 | - |
| | Chloride (Cl) | mg/L | 40.4 | 6065 |
| | Sulfate (SO4) | mg/L | 46.1 | - |

* Provided by CNOOC for this investigation.

### 3. Results and Discussions

Produced sand and pressure gradients are used as the main comparative parameters in the SRT experiments. The pressure gradient, which is the pressure difference over the distance between the pressure ports, allows comparing the sand pack's plugging tendency during the tests. Plugging results in higher pressure drops (dp in Figure 6), and consequently, higher pressure gradients.
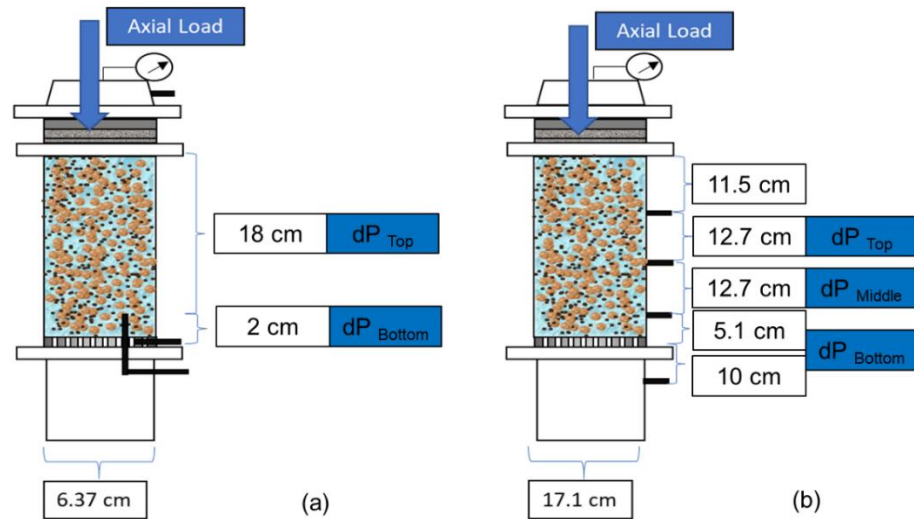


**Figure 6.** Pressure gradient in the context of the testing setups used in; (**a**) Devere-Bennett [3]; and (**b**) the current investigation.

*3.1. Replication Test*

Figure 7 compares the pressure gradients across the cell reported in Devere-Bennett [3] with the replication test (Test #1) on the replicated sand pack. In the first four stages of the replication test, the pressure gradients increase linearly as the flow rate gradually increases. Subsequently, the pressure gradients in the next four stages increase at a more gentle rate than the first four stages. In stages 5 to 8, the pressure drop is affected by the injected water and its associated relative permeability. However, in these stages, the water's lower viscosity compared to the oil overshadows the relative permeability effects. Finally, the pressure gradients steeply increase when gas is introduced, which is attributed to the relative permeability effects. Then, the pressure gradients become stable or show a slight increase as the gas rate increases. The stabilized pressure gradients in the last three stages can be attributed to the liquid expelling from flow channels as the gas injection rate is increased. This occurrence results in a higher gas relative permeability, consequently reducing the effect on pressure gradients.

Figure 7 compares the pressure gradient in Devere-Bennett's [3] experiments with the one obtained from the replication test. At the first look on Figure 7b, it seems that the near-coupon pressure gradient in the replication test (WWS with 152 μm [0.006 inches] sloth size) is not quantitatively comparable with the corresponding data in Devere-Bennett [3]; however, the trends are comparable. This is expected since the pressure ports are located at different distances from the coupon: 2 cm above the coupon in Devere-Bennett [3] and 5 cm above the coupon in the setup used in the current investigations (Figure 6). Some inconsistencies in the pressure gradient data were observed in Devere-Bennett [3] when conducting the test with different SCD coupons, as shown in Figure 8. For instance, the test results on 406 μm (0.016 inches) straight cut slotted liner (SC) show a significant reduction in the pressure gradient after gas injection, which is not consistent with other test results. The test with 304 μm (0.012 inches) WWS also shows a sharp drop in the pressure gradient when brine is introduced and continues to decrease with the increasing water rate.
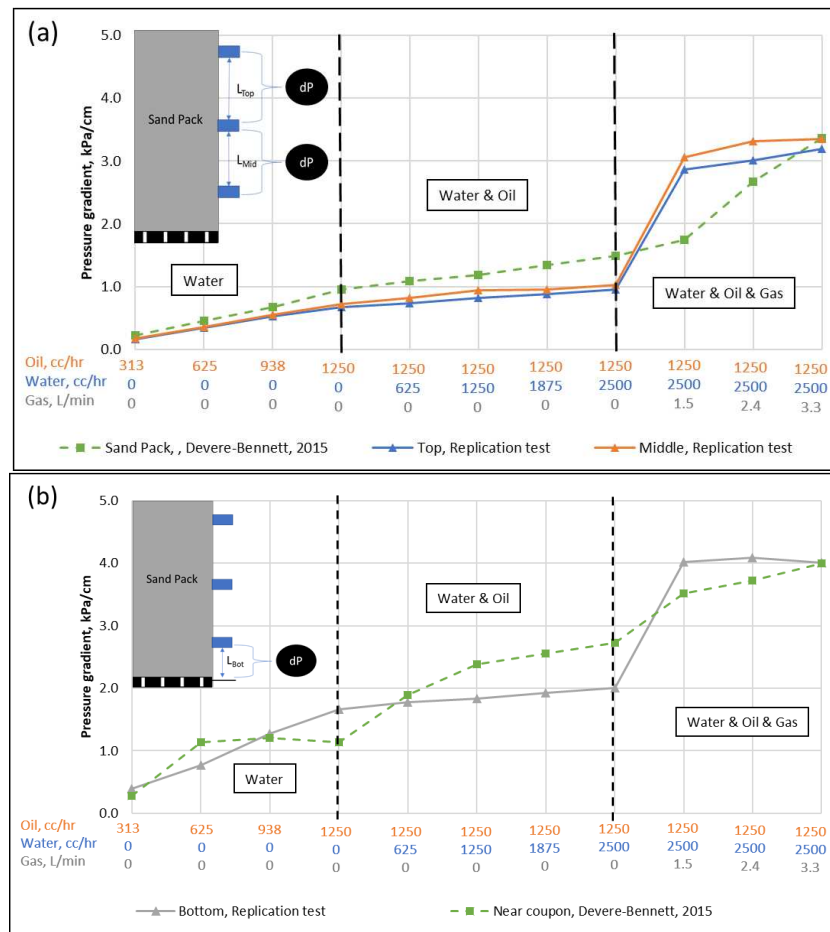
**Figure 7.** Comparison between the pressure gradients across the sand pack from Devere-Bennett [3] with the pressure gradients of the replication test on the replicated sand pack, (**a**) pressure drop in the middle interval, and (**b**) pressure drop near the screen.
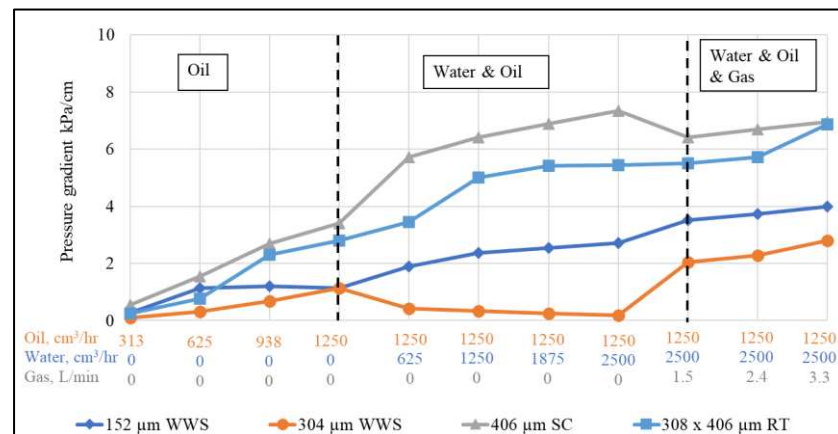


**Figure 8.** Near-coupon pressure gradients of SRT experiments reported in Devere-Bennett [3], where different SCDs were used.

The produced sand, which is normalized by the coupon area, is 0.0083 g/cm² (0.017 lb/ft²) for the replication test, while it was 0.012 g/cm² (0.024 lb/ft²) in Devere-Bennett [3]. The replication test produces 25% less sand than the one in Devere-Bennett [3]. This variation can be attributed to the different procedures in collecting produced sand. Devere-Bennett [3] reported the total produced solids while the produced sands and fines are separated in the replication test. Although sanding in SAGD is usually considered a

problematic phenomenon and is avoided, fines production is desired and should be encouraged to reduce near-wellbore plugging. Hence, it is required to individually measure the produced sand and fines in sand control testing.

### 3.2. Effect of Packing Technique

Two packing methods were used to investigate the impact of packing on SRT results, dry packing (used by [3]) and moist tamping (suggested by [57]). Five core plugs with equal sizes of 5 cm (2 inches) were taken from each sand pack after dry and moist packing to find the porosity distribution along the sand pack (Figure 9). In this Figure, one can notice that the moist tamping shows less porosity variability than the dry packing technique. The porosity of the dry-packed sample shows a standard deviation of 1.05, which is 74% higher than the one for the sand pack prepared by moist packing.
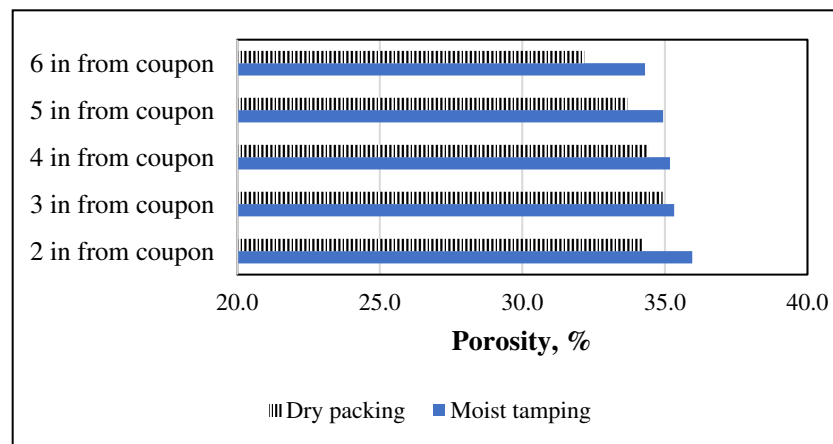


**Figure 9.** Porosity distribution measured across the sand pack using the samples taken from the different intervals.

Figure 10 illustrates that the SRT packed with the moist tamping method has a lower pressure drop, attributed to the higher average porosity of the sand pack. Both tests show a higher pressure gradient (almost two times) at the region near the coupon than other sections inside the sand pack. The higher pressure gradient in the near-coupon area is generated due to flow convergence at the slots [32] and pore plugging due to fines migration at this vicinity.
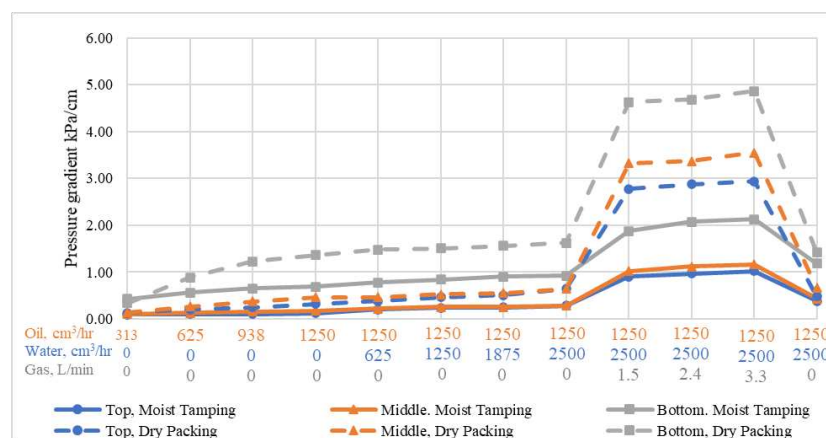


**Figure 10.** Comparison of the pressure gradient across the sand pack in SRT testing packed by moist tamping and dry packing techniques.

The cumulatively produced sand in the moist tamping is 45% less than in the dry packing (Figure 11). The lower produced sand in the moist tamping method can be

attributed to the lower interstitial flow velocity in the near-coupon region due to the higher porosity in the moist tamping method. In the test with the moist tamping method, the smaller interstitial velocity creates weaker drag force on grain particles causing lower sand production in the moist tamping method. Furthermore, as Rhodes [58] showed a higher possibility of segregation in dry packing, the SRT packed by dry technique is more vulnerable to segregation, causing more concentration of smaller-sized particles at the bottom of the cell and consequently more sand production.
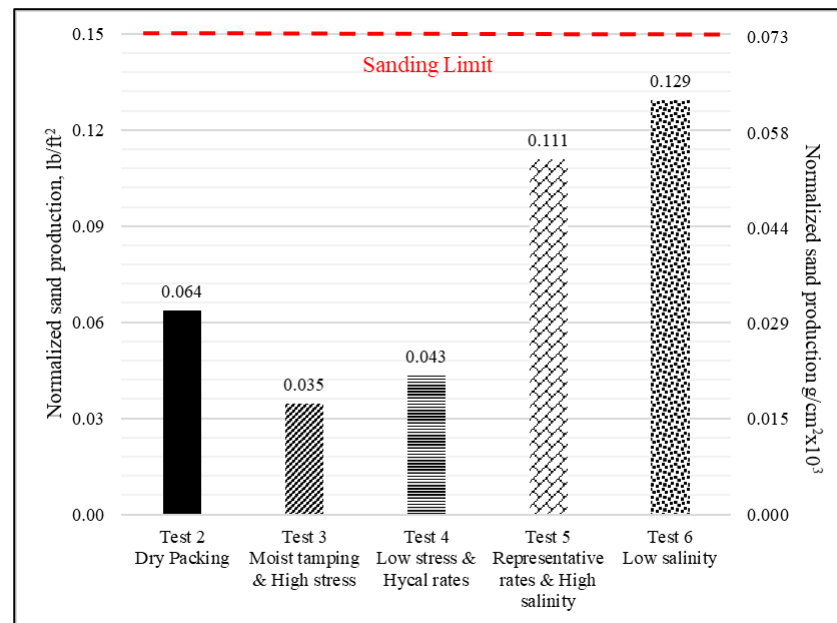


**Figure 11.** Normalized cumulative produced sand in the SRTs with a rolled top SL, slot width was 406–508 μm (0.016–0.020 inch).

*3.3. Effect of Stress Magnitude*

Figure 11 shows that the produced sand in the test with the lower axial stress over the sand pack (413 kPa) is almost 23% higher than that observed in a test with higher axial stress (2413 kPa). Less sand production at a higher stress level agrees with the testing results published in the literature [28,68]. This behavior can be attributed to the higher friction coefficient and the strength of the sample at the higher normal effective stresses [69], which construct more stable sand bridges.

The pressure gradients across the sand pack for two different axial stress levels are presented in Figure 12. In the test with 413 kPa axial stress, the pressure gradient at the near-coupon section is, on average, 23% less compared to the test with 2413 kPa axial stress. Lower stress on the sand pack results in higher porosity and permeability, hence, lower pressure gradients. Further, the sample with higher porosity allows a more effective production of fines, hence, less plugging and lower pressure gradients.

*3.4. Effect of Fluid Flow Rates*

The maximum liquid rates in Devere-Bennett [3], adopted from Bennion et al. [4], were upscaled to the new testing setup diameter and normalized by the OFA for the same aperture-size slotted liner (Figure 13). The results show that the normalized maximum liquid flow rate used in this work is approximately 65% lower than one employed by Bennion [4] (Figure 13). The normalized fluid rate by the OFA in a single-slot coupon results in the maximum liquid flux of 3410 $cm^3/hr/cm^2$, while it is about 1166 $cm^3/hr/cm^2$ when using a larger setup with a multi-slot coupon. Using single-slot coupons and ignoring the slot density with constant flow rates causes excessively high flow velocities for SL

testing compared to WWS. Thus, the testing setup and procedure become biased against SL owing to the large drag forces and flow convergence experienced at a single slot.
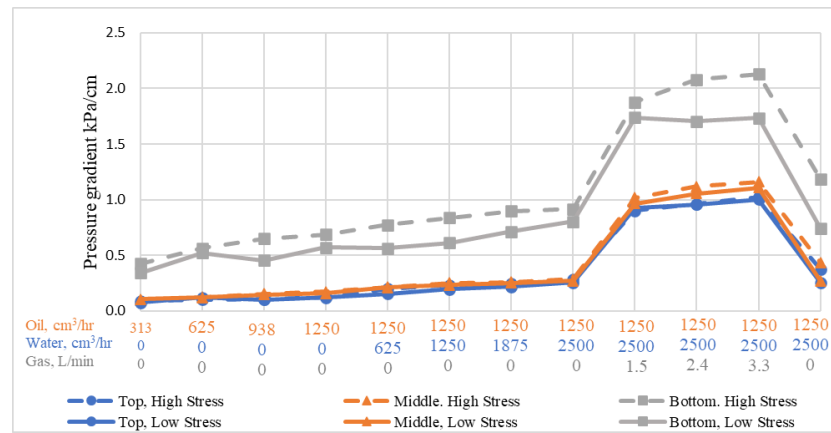


**Figure 12.** Effect of axial stress on the pressure gradient across the sand pack at high-stress (2413 kPa) and low-stress (413 kPa) conditions.
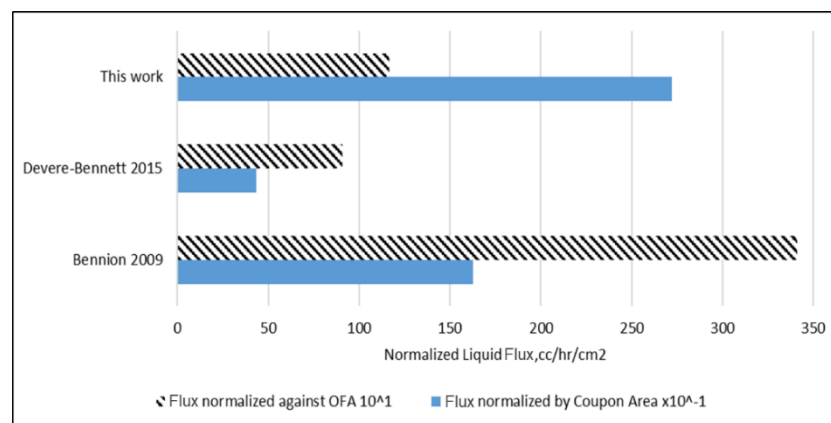


**Figure 13.** Liquid flux employed in the test reported in Devere-Bennett [3] and this study for SL testing.

The water cuts of 50–100% in the testing procedure used in this work are higher than those considered in Devere-Bennett's [3] and Bennion et al. [4] (water cuts of 0–67%). The water cuts applied in this investigation are comparable with the values observed in the SAGD production wells for the Long Lake Field in 2017, where the average water cut was recorded at about 79% [60].

Considering the challenges in quantifying the steam flux during the steam breakthrough events in SAGD wells, it was assumed that all injected steam flows toward the production well during a steam breakthrough. In this regard, the steam properties are determined for the production well at typical thermodynamic conditions using steam tables such as Dahm and Visco [70].

The higher flow rates in Test #4 produced 153% more sand than Test #5 with a low rate, as shown in Figure 11. The higher flow rate generates higher drag forces on sand particles near the slot, enhances the destabilizing forces on sand bridges, and consequently causes more sand production.

As shown in Figure 14, the concentration of the fine in the discharge fluid at stages 1 to 3 of the SRT tests shows an increasing trend correlated with the increasing flow rate. Judging from a significant variation of fines concentration between the first and second stages, one can deduce that the critical flow rate for fines mobilization in these tests is between 2500 and 3570 cm$^3$/hr.
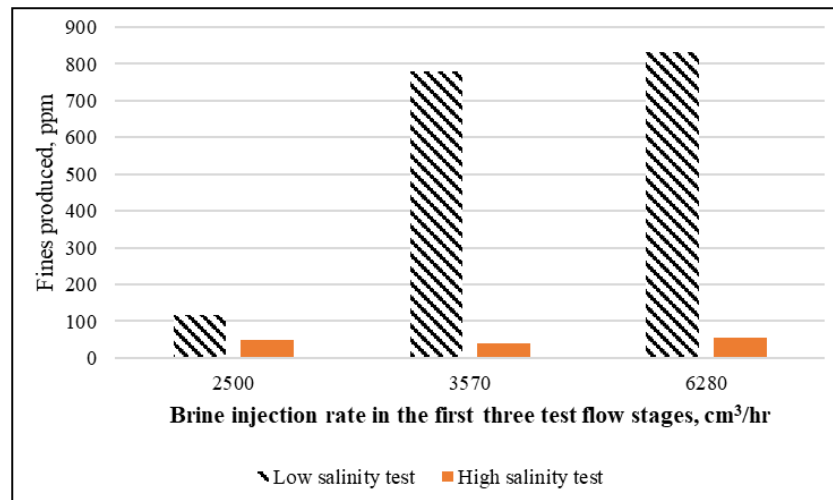
**Figure 14.** Total fines concentration in discharged brine at different testing stages.

*3.5. Effect of Brine Salinity*

The pressure gradient of both high- and low-salinity tests across the sand pack is compared in Figure 15. As expected, a higher fines production is observed in the low salinity injection scenario (Figure 14). The low salinity brine increases the fines particle mobilization compared to the high salinity brine. The mobilized fines become suspended in the flowing fluid and migrate by the fluid. Furthermore, the suspended particles reduce the open flow paths in the porous media and increase the pressure drop in the top and middle intervals, as shown in Figure 15. These sections at the low salinity test exhibit a 45% higher pressure gradient, which can be attributed to the more substantial fines mobilization.
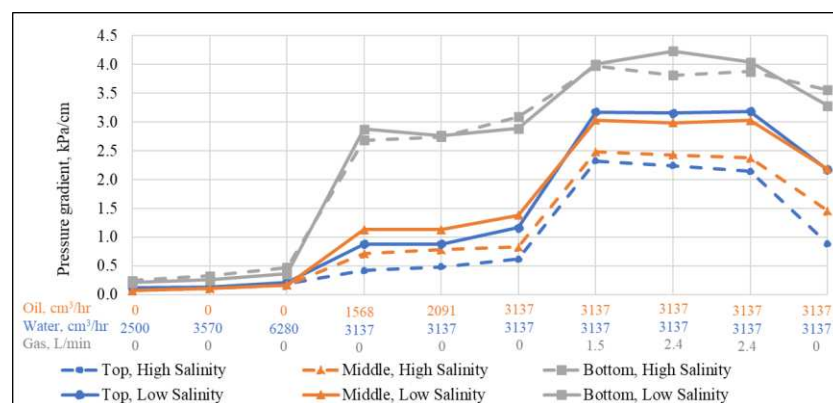


**Figure 15.** Pressure gradient of different sections of the sand pack while injecting brine with different salinities.

The bottom section not only experiences the fines mobilization but also receives some fines from the upper sections. However, the bottom section could discharge the mobilized and receive fines through the screen. Therefore, both tests show similar pressure drops in the near-coupon region.

However, it is believed that if the test is continued for a longer time, the mobilized fines can migrate from the top section to the middle and bottom sections, resulting in a pressure drop reduction in this section. The middle section receives fine particles from the top section and passes them to the bottom section, causing neutral behavior or insignificant pressure drop changes. The bottom section would simultaneously receive and discharge the fine particles. Regarding the pore throat and migratory particle sizes, the fines particles would plug the pore spaces or pass through the screen to the discharge fluid. Therefore, in the bottom section, pressure drop would be increased or reduced.

## 4. Conclusions

This paper investigates the impact of the testing setup and procedure using replicated sands on the sand retention testing results. Sand production, fines production, and pressure gradients within the sand pack and across the coupon are the primary evaluation factors. The replication tests' results support replicating PSD from the commercial sands for sand retention tests.

The results show that the different SRT setups and testing procedures may yield different liner performance results in sanding and plugging. This paper improves the existing SRT setups and testing procedures and introduces a modified SRT apparatus along with an experimental workflow. The packing technique, axial stress magnitude, fluid flow rates, and salinity parameters on test performance were investigated among the influential parameters, and their effects were quantified. The proposed experimental workflow and testing variables' determination procedure enable a more accurate and representative evaluation of sand control devices considering SAGD downhole conditions.

Gradually modifying the testing variables from the ones reported in Devere-Bennett [3] to the new procedure brings several conclusions. (1) The fluid flow rate is the most effective testing parameter on sand production, followed by packing technique, axial stress magnitude, and salinity. (2) The impact of stress is minimal on the pressure drop under the stress conditions applied in this work. (3) Lowering salinity from 10,000 ppm to 350 ppm raises the fines migration and production significantly, causing a higher pressure drop across the sand pack.

## Acronyms

| | |
|---|---|
| CSC | Critical Salt Concentration |
| Ppm | Part Per Million |
| PSD | Particle Size Distribution |
| SAGD | Steam Assisted Gravity Drainage |
| SCD | Sand Control Devices |
| SCT | Scaled Completion Test |
| SL | Slotted Liner |
| SRT | Sand Retention Test |
| UC | Uniformity Coefficient |
| WWS | Wire-Wrapped Screen |

## References

1. Gates, I.D.; Leskiw, C. Impact of steam trap control on performance of steam-assisted gravity drainage. *J. Pet. Sci. Eng.* **2010**, *75*, 215–222. [CrossRef]
2. Shen, J. *Enhanced Oil Recovery Field Case Studies*, 1st ed.; Gulf Professional Publishing: Oxford, UK, 2013; pp. 413–445.
3. Devere-Bennett, N. Using Pre-Pack Sand-Retention Tests (SRT's) to Narrow Down Liner/Screen Sizing in SAGD Wells. In Proceedings of the SPE Thermal Well Integrity and Design Symposium, Banff, AB, Canada, 23 November 2015.
4. Bennion, D.B.; Gupta, S.; Gittins, S.; Hollies, D. Protocols for Slotted Liner Design for Optimum SAGD Operation. *J. Can. Pet. Technol.* **2009**, *48*, 21–26. [CrossRef]

5. Xie, J. Slotted Liner Design Optimization for Sand Control in SAGD Wells. In Proceedings of the SPE Thermal Well Integrity and Design Symposium, Banff, AB, Canada, 23 November 2015.

6. Zhang, Z. An Advanced Sand Control Technology for Heavy Oil Reservoirs. Master's Thesis, University of Calgary, Calgary, AB, Canada, 2017.

7. Montero, J.D.; Chissonde, S.; Kotb, O.; Wang, C.; Roostaei, M.; Nouri, A.; Mahmoudi, M.; Fattahpour, V. A Critical Review of Sand Control Evaluation Testing for SAGD Applications. In Proceedings of the SPE Canada Heavy Oil Technical Conference, Calgary, AB, Canada, 13–14 March 2018.

8. Coberly, C.J. Selection of Screen Openings for Unconsolidated Sands. In Proceedings of the Drilling and Production Practice, New York, NY, USA, 1 January 1937.

9. Underdown, D.R.; Dickerson, R.; Vaughan, W. The Nominal Sand-Control Screen: A Critical Evaluation of Screen Performance. *SPE Drill. Complet.* **2001**, *16*, 252–260. [CrossRef]

10. Gillespie, G.; Deem, C.K.; Malbrel, C. Screen Selection for Sand Control Based on Laboratory Tests. In Proceedings of the SPE Asia Pacific Oil and Gas Conference and Exhibition, Brisbane, Australia, 16 October 2000.

11. Hodge, R.M.; Burton, R.C.; Constien, V.; Skidmore, V. An Evaluation Method for Screen-Only and Gravel-Pack Completions. In Proceedings of the International Symposium and Exhibition on Formation Damage Control, Lafayette, LA, USA, 20 February 2002.

12. Constien, V.G.; Skidmore, V. Standalone Screen Selection Using Performance Master Curves. In Proceedings of the SPE International Symposium and Exhibition on Formation Damage Control, Lafayette, LA, USA, 15 February 2006.

13. Chanpura, R.A.; Hodge, R.M.; Andrews, J.S.; Toffanin, E.P.; Moen, T.; Parlar, M. A Review of Screen Selection for Standalone Applications and a New Methodology. *SPE Drill. Complet.* **2011**, *26*, 84–95. [CrossRef]

14. Ballard, T.J.; Beare, S.P. An Investigation of Sand Retention Testing With a View To Developing Better Guidelines for Screen Selection. In Proceedings of the SPE International Symposium and Exhibition on Formation Damage Control, Lafayette, LA, USA, 15 February 2012.

15. Markestad, P.; Christie, O.; Espedal, A.; Rorvik, O. Selection of Screen Slot Width to Prevent Plugging and Sand Production. In Proceedings of the SPE Formation Damage Control Symposium, Lafayette, LA, USA, 14 February 1996.

16. Ballard, T.; Beare, S. Media Sizing for Premium sand Screens: Dutch Twill Weaves. In Proceedings of the SPE European Formation Damage Conference, The Hague, The Netherlands, 13 May 2003.

17. Ballard, T.; Beare, S. Sand Retention Testing: The More You Do, the Worse It Gets. In Proceedings of the SPE International Symposium and Exhibition on Formation Damage Control, Lafayette, LA, USA, 15 February 2006.

18. Williams, C.F.; Richard, B.M.; Horner, D. A New Sizing Criterion for Conformable and Nonconformable Sand Screens Based on Uniform Pore Structures. In Proceedings of the SPE International Symposium and Exhibition on Formation Damage Control, Lafayette, LA, USA, 15 February 2006.

19. Romanova, U.G.; Gillespie, G.; Sladic, J.; Ma, T.; Solvoll, T.A.; Andrews, J.S. A Comparative Study of Wire Wrapped Screens vs. Slotted Liners for Steam-Assisted Gravity Drainage Operations. In Proceedings of the World Heavy Oil Congress, New Orleans, LA, USA, 5–7 March 2014.

20. Romanova, U.G.; Piwowar, M.; Ma, T. Sand Control for Unconsolidated Heavy Oil Reservoirs: A Laboratory Test Protocol and Recent Field Observations. In Proceedings of the International Symposium of the Society of Core Analysts, St. John's, NL, Canada, 16–21 August 2015.

21. O'Hara, M. Thermal Operations in the McMurray: An Approach To Sand Control. In Proceedings of the SPE Thermal Well Integrity and Design Symposium, Banff, AB, Canada, 23 November 2015.

22. Mahmoudi, M. New Sand Control Design Criteria and Evaluation Testing for Steam-Assisted Gravity Drainage (SAGD) Wellbores. Ph.D. Thesis, University of Alberta, Edmonton, AB, Canada, 2016.

23. Anderson, M. SAGD Sand Control: Large Scale Testing Results. In Proceedings of the SPE Canada Heavy Oil Technical Conference, Calgary, AB, Canada, 15 February 2017.

24. Wang, C.; Pang, Y.; Montero, J.; Haftani, M.; Fattahpour, V.; Mahmoudi, M.; Nouri, A. Impact of Anisotropic Stresses on the Slotted Liners Performance in Steam Assisted Gravity Drainage Process. In Proceedings of the SPE Thermal Well Integrity and Design Symposium, Banff, AB, Canada, 27 November 2018.

25. Fattahpour, V.; Mahmoudi, M.; Wang, C.; Kotb, O.; Roostaei, M.; Nouri, A.; Fermaniuk, B.; Sauve, A.; Sutton, C. Comparative Study on the Performance of Different Stand-Alone Sand Control Screens in Thermal Wells. In Proceedings of the SPE International Conference and Exhibition on Formation Damage Control, Lafayette, LA, USA, 7 February 2018.

26. Fattahpour, V.; Azadbakht, S.; Mahmoudi, M.; Guo, Y.; Nouri, A.; Leitch, M. Effect of Near Wellbore Effective Stress on the Performance of Slotted Liner Completions in SAGD Operations. In Proceedings of the SPE Thermal Well Integrity and Design Symposium, Banff, AB, Canada, 28 November 2016.

27. Bennion, D.B.; Gupta, S.; Gittins, S.; Hollies, D. Protocols for Slotted Liner Design for Optimum SAGD Operation. In Proceedings of the Canadian International Petroleum Conference, Calgary, AB, Canada, 17 June 2008.

28. Selig, E.; Ladd, R.S. Preparing Test Specimens Using Undercompaction. *Geotech. Test. J.* **1978**, *1*, 16. [CrossRef]

29. Guo, Y.; Roostaei, M.; Nouri, A.; Fattahpour, V.; Mahmoudi, M.; Jung, H. Effect of stress build-up around standalone screens on the screen performance in SAGD wells. *J. Pet. Sci. Eng.* **2018**, *171*, 325–339. [CrossRef]

30. Haftani, M.; Kotb, O.; Nguyen, P.H.; Wang, C.; Salimi, M.; Nouri, A. A Novel sand control testing facility to evaluate the impact of radial flow regime on screen performance and its verification. *J. Pet. Sci. Eng.* **2020**, *195*, 107903. [CrossRef]

31. Fermaniuk, B. Sand Control in Steam-Assisted Gravity Drainage (SAGD) Wellbores and Process of Slotted Liner Design and Process. Master's Thesis, University of Calgary, Calgary, AB, Canada, 2013.

32. Mahmoudi, M.; Nejadi, S.; Roostaei, M.; Olsen, J.; Fattahpour, V.; Lange, C.F.; Zhu, D.; Fermaniuk, B.; Nouri, A. Design Optimization of Slotted Liner Completions in Horizontal Wells: An Analytical Skin Factor Model Verified by Computational Fluid Dynamics and Experimental Sand Retention Tests. In Proceedings of the SPE Thermal Well Integrity and Design Symposium, Banff, AB, Canada, 28 November 2017.

33. Tiffin, D.L.; King, G.E.; Larese, R.E.; Britt, L.K. New Criteria for Gravel and Screen Selection for Sand Control. In Proceedings of the SPE International Symposium on Formation Damage Control, Lafayette, LA, USA, 18 February 1998.

34. Price-Smith, C.; Parlar, M.; Bennett, C.; Gilchrist, J.M.; Pitoni, E.; Burton, R.C.; Hodge, R.M.; Troncoso, J.; Ali, S.A.; Dickerson, R. Design Methodology for Selection of Horizontal Open-Hole Sand-Control Completions Supported by field Case Histories. *SPE Drill. Completion* **2003**, *18*, 235–255. [CrossRef]

35. Rogers, E.B. Sand Control in Oil and Gas Wells. *Oil Gas J.* **1971**, 54–60.

36. Suman, G.O.; Ellis, R.C.; Snyder, R.E. *Sand Control Handbook*, 2nd ed.; Gulf Publishing Company, Book Division: Houston, TX, USA, 1983.

37. Wang, C.; Pang, Y.; Mahmoudi, M.; Haftani, M.; Salimi, M.; Fattahpour, V.; Nouri, A. A Set of Graphical Design criteria for Slotted Liners in SteamAssisted Gravity Drainage Production Wells. *J. Pet. Sci. Eng.* **2020**, *185*, 106608. [CrossRef]

38. Stein, N.; Odeh, A.; Jones, L. Estimating Maximum Sand-Free Production Rates From Friable Sands for Different Well Completion Geometries. *J. Pet. Technol.* **1974**, *26*, 1156–1158. [CrossRef]

39. Spronk, E.M.; Doan, L.T.; Matsuno, Y.; Harschnitz, B. SAGD Liner Evaluation and Liner Test Design for JACOS Hangingstone SAGD Development. In Proceedings of the SPE Canada Heavy Oil Technical Conference, Calgary, AB, Canada, 9 June 2015.

40. Khilar, K.C. *Fogler, H.S. Migration of Fines in Porous Media*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 1998; p. 12.

41. Mahmoudi, M.; Fattahpour, V.; Nouri, A.; Leitch, M. An Experimental Investigation of the Effect of pH and Salinity on Sand Control Performance for Heavy Oil Thermal Production. In Proceedings of the SPE Canada Heavy Oil Technical Conference, Calgary, AB, Canada, 7 June 2016.

42. Haftani, M.; Wang, C.; Pallares, J.D.M.; Mahmoudi, M.; Fattahpour, V.; Nouri, A. An Investigation into the Effect of Brine Salinity on Fines Migration in SAGD Operations. In Proceedings of the SPE Western Regional Meeting, San Jose, CA, USA, 22 April 2019.

43. Kia, S.; Fogler, H.; Reed, M. Effect of pH on colloidally induced fines migration. *J. Colloid Interface Sci.* **1987**, *118*, 158–168. [CrossRef]

44. Khilar, K.; Fogler, H. The existence of a critical salt concentration for particle release. *J. Colloid Interface Sci.* **1984**, *101*, 214–224. [CrossRef]

45. Clearly, M.P.; Melvan, J.J.; Kohlhaas, C.A. The Effect of Confining Stress and Fluid Properties on Arch Stability in Unconsolidated Sands. In Proceedings of the SPE Annual Technical Conference and Exhibition, Las Vegas, NV, USA, 23 September 1979.

46. Coşkuner, G.; Maini, B. Effect of net confining pressure on formation damage in unconsolidated heavy oil reservoirs. *J. Pet. Sci. Eng.* **1990**, *4*, 105–117. [CrossRef]

47. Pallares, J.D.M.; Wang, C.; Haftani, M.; Nouri, A. Experimental Correlations for the Performance and Aperture Selection of Wire-Wrapped Screens in Steam-Assisted Gravity Drainage Production Wells. *SPE Prod. Oper.* **2020**, *35*, 537–548. [CrossRef]

48. Oda, M. Initial Fabrics and their Relations to Mechanical Properties of Granular Material. *Soils Found.* **1972**, *12*, 17–36. [CrossRef]

49. Arthur, J.R.F.; Menzies, B.K. Inherent anisotropy in a sand. *Géotechnique* **1972**, *22*, 115–128. [CrossRef]

50. Vaid, Y.P.; Sivathayalan, S. Fundamental Factors Affecting Liquefaction Susceptibility of Sands. *Can. Geotech. J.* **2000**, *37*, 592–606. [CrossRef]

51. Miura, S.; Toki, S. A Sample Preparation Method and its Effect on Static and Cyclic Deformation-Strength Properties of Sand. *Soils Found.* **1982**, *22*, 61–77. [CrossRef]

52. Kuerbis, R.; Vaid, Y. Sand Sample Preparation—The Slurry Deposition Method. *Soils Found.* **1988**, *28*, 107–118. [CrossRef]

53. Suits, L.D.; Sheahan, T.; Ueng, T.-S.; Wang, M.-H.; Chen, M.-H.; Chen, C.-H.; Peng, L.-H. A Large Biaxial Shear Box for Shaking Table Test on Saturated Sand. *Geotech. Test. J.* **2006**, *29*, 1–8. [CrossRef]

54. Huang, Y.-H.; Yang, S.-Q.; Hall, M.R.; Zhang, Y.-C. The Effects of NaCl Concentration and Confining Pressure on Mechanical and Acoustic Behaviors of Brine-Saturated Sandstone. *Energies* **2018**, *11*, 385. [CrossRef]

55. Suits, L.D.; Sheahan, T.C.; Bradshaw, A.S.; Baxter, C.D.P. Sample Preparation of Silts for Liquefaction Testing. *Geotech. Test. J.* **2007**, *30*, 324–332. [CrossRef]

56. Mahmoudi, M.; Fattahpour, V.; Nouri, A.; Rasool, S.; Leitch, M. The Effect of Screen Aperture Size on Fines Production and Migration in SAGD Production Wells. In Proceedings of the 2016 World Heavy Oil Congress, Calgary, AB, Canada, 6–9 September 2016.

57. Mahmoudi, M.; Fattahpour, V.; Nouri, A.; Rasoul, S.; Yao, T.; Baudet, B.A.; Leitch, M.; Soroush, M. Investigation Into the Use of Commercial Sands and Fines to Replicate Oil Sands for Large-Scale Sand Control Testing. In Proceedings of the SPE Thermal Well Integrity and Design Symposium, Banff, AB, Canada, 28 November 2016.

58. Rhodes, M. *Introduction to Particle Technology*; John Wiley & Sons Ltd.: Chichester, UK, 2008.

59. Michaelides, E.; Crowe, C.T.; Schwarzkopf, J.D. *Multiphase Flow Handbook*; CRC Press: Boca Raton, FL, USA, 2016.

60. CNOOC. Long Lake Kinosis Oil Sands Project Annual Performance Presentation. 2018. Available online: https://static.aer.ca/prd/documents/oilsands/insitu-presentations/2018AthabascaNexenLongLakeSAGD9485.pdf (accessed on 10 August 2021).

61. Le Ravalec, M.; Morlot, C.; Marmier, R.; Foulon, D. Heterogeneity Impact on SAGD Process Performance in Mobile Heavy Oil Reservoirs. *Oil Gas Sci. Technol.-Rev. L'ifp* **2009**, *64*, 469–476. [CrossRef]

62. Stone, T.W.; Bailey, W.J. Optimization of Subcool in SAGD Bitumen Processes. In Proceedings of the World Heavy Oil Congress, New Orleans, La, USA, 5–7 March 2014.

63. Sidahmed, A.K. Optimization of Outflow Control Devices Design in Steam-Assisted Gravity Drainage Models with Wellbore Trajectory Excursions. Master's Thesis, University of Alberta, Edmonton, AB, Canada, 2018.

64. Beshry, M.A.; Krawchuk, P.; Brown, G.A.; Brough, B. Predicting the Flow Distribution on Total E&P Canada's Joslyn Project Horizontal SAGD Producing Wells Using Permanently Installed Fiber-Optic Monitoring. In Proceedings of the SPE Annual Technical Conference and Exhibition, San Antonio, TX, USA, 24 September 2006.

65. Brand, S. Results From Acid Stimulation in Lloydminster SAGD Applications. In Proceedings of the SPE International Symposium and Exhibition on Formation Damage Control, Lafayette, LA, USA, 10 February 2010.

66. Romanova, U.G.; Ma, T. An Investigation of the Plugging Mechanisms in a Slotted Liner from the steam-Assisted Gravity Drainage Operations. In Proceedings of the SPE European Formation Damage Conference & Exhibition, Noordwijk, The Netherlands, 5 June 2013.

67. Harvey, A.H. Thermodynamic Properties of Water: Tabulation from the IAPWS Formulation 1995 for the Thermodynamic properties of Ordinary Water Substance for General and Scientific Use, NISTIR 5078. United States Department of Commerce, 1995. Available online: https://nvlpubs.nist.gov/nistpubs/Legacy/IR/nistir5078.pdf (accessed on 10 August 2021).

68. Guo, Y. Effect of Stress Build-Up around SAGD Wellbores on the Slotted Liner Performance. Master's Thesis, University of Alberta, Edmonton, AB, Canada, 2018.

69. Jafarpour, M.; Rahmati, H.; Azadbakht, S.; Nouri, A.; Chan, D.; Vaziri, H. Determination of mobilized strength properties of degrading sandstone. *Soils Found.* **2012**, *52*, 658–667. [CrossRef]

70. Dahm, K.; Visco, D. Fundamentals of Chemical Engineering Thermodynamics. In *Cengage Learning*; Cengage: Boston, MA, USA, 2015.

*Review*

# Simulation and Sensitivity Analysis for Various Geometries and Optimization of Solid Oxide Fuel Cells: A Review

**Mina Tonekabonimoghaddam [1] and Ahmad Shamiri [2],***

[1] Chemical Engineering Department, Faculty of Engineering, University of Malaya,
 Kuala Lumpur 50603, Malaysia; mina.tmoghadam@gmail.com
[2] Reliability and Engineering Department, Dyno Nobel Moranbah, 667 Goonyella Road,
 Moranbah, QLD 4744, Australia
* Correspondence: ahmad.shamiri@incitecpivot.com.au or ahmadshamiri@gmail.com; Tel.: +61-7-4840-3708

**Abstract:** Solid oxide fuel cells (SOFCs) are considered as one of the most promising fuel cell types for application as high efficiency power generators. This work reviews the use of computational fluid dynamics (CFD) to maximise SOFC performance and life, and minimise cost, by considering numerous configurations and designs. A critical analysis of available literature proves that detailed research on the simulation of thermal stress and its damaging impact on the SOFC is still in its early stage of development. Numerical simulation is expected to help optimize the design, operating parameters and fuel cell materials. Therefore, sensitivity analysis of fuel cell parameters using simulation models is analysed to address the issue. Finally, the present status of the SOFC optimization efforts is summarized so that unresolved problems can be identified and solved.

**Keywords:** solid oxide fuel cell; computational fluid dynamics; sensitivity analysis; optimization

## 1. Introduction

State-of-the-art solid oxide fuel cell (SOFC) systems are designed for both small and large-scale stationary power generation systems [1]. This is because the SOFC is simple with high energy efficiency, can work with different hydrocarbon fuels, and can at least partially reform hydrocarbon fuels internally [2–6]. They are generally considered ideal for stationary applications with a high power range, i.e., for several 100 kW to the MW region, such as systems from LG Fuel Cell Systems or Bloom Energy [7], or for power outputs of 1–20 kW, such as the systems from HEXIS, Ceramic Fuel Cells Limited, or Versa Power [8]. Due to the higher power density of SOFC systems compared to other fuel cell types, SOFC systems have also been proposed for portable applications with power ranges of about 20–250 W [8]. Specifically, Delphi Automotive Systems is developing an SOFC that will power auxiliary units in automobiles and tractor-trailers [8,9]. Similarly, Topsoe Fuel cell manufacture SOFCs for stationary combined heat and power (CHP) in steps of 6 kW for residential applications [9,10].

SOFCs have the ability to produce all the required electricity to allow the engine to be smaller and more sufficient due to a high operating temperature. The SOFC can run on the same gasoline or diesel engine and maintains the running of all necessary electrical systems while the engine shuts off when not needed. Rolls-Royce Fuel Cell Systems Ltd. is developing a SOFC gas turbine hybrid system fuelled with natural gas for power generation applications in the range of a megawatt (e.g., Futuregen) [10,11]. Acumentrics, a US company, manufactures SOFC power generators for off-grid applications at a power range of 250 to 1500 W [11,12].

The application of an SOFC is bounded by some limitations, however, which are primarily due to its elevated operating temperature (500–1000 °C), relatively expensive construction materials (in some cases), and sensitivity to thermal stresses within the cells. Major research efforts have been placed on the materials, cell and stack designs, and

improvements in power density along with lowering the operating temperature down to 500–700 °C [13–15]. Most SOFCs are designed to function at elevated temperatures, thereby generating higher cell efficiencies in comparison to other kinds of fuel cells when operating on hydrocarbon fuels [16]. However, many research groups are working to reduce the operating temperature of SOFCs to under 650 °C in order to reduce material degradation, prolong stack-lifetime and decrease stack material costs by enabling the use of common metallic materials [17–19]. Therefore, it is of paramount importance to manage the stack temperature of the SOFC in a uniform manner, particularly for planar SOFCs [20]. To enhance SOFC technology for widespread market penetration, the system should demonstrate extended cell lifetime at reduced cost.

Modelling, simulation, control and optimization of SOFCs have become an important tool to understand, predict and enhance performance, and a means for improving lifetime through dynamic analysis of critical variables and advanced optimization algorithms (Figure 1) [21]. As shown in Figure 1, computational simulations are necessary to have a direct "insight" into the system and to allow the analyses of micro-structure, residual stains and electrochemical effects on macro-structural design, thermo cycling, transport phenomena, different component designs, stack and cell configurations and layouts, fuels, as well as to determine optimum operating conditions. This reduces the number of experimental tests to be performed, and costly and time-consuming physical prototyping methods [21,22]. Moreover, owing to the dynamic model, it is possible to analyse both the time response of fuel cell micro/macrostructure and its full function, thereby allowing the optimization of the system and its control logic. Understanding the fundamental mechanisms is an important step towards SOFC design optimization and performance improvement. In this respect, computational simulations have proved to be a cost-effective method. Many simulation investigations in the literature have been reviewed in this work in order to deliver performance prediction and parameter optimization. Some papers have studied the models of SOFCs for performance optimization [23–30], including the steady- and dynamic-states. Ramadhani et al. [26] have discussed the application of SOFC, in particular, the optimization strategies. They reviewed the decision variable, objective analysis, constraint, method and tools. However, the study did not focus on the sensitivity analysis. A number of recent qualified and comprehensive research and review papers were published with emphasis on different topics of the SOFC which are discussed in the next sections. The rapidly growing number of citations related primary research articles related to optimization of SOFC (Figure 2) is an indication that more comprehensive reviews in this field are necessary in order to draw general conclusions and provide some guided perspectives for future research. As indicated by the number of citations per year, interest is still growing (see Figure 2).

Neither of the previous studies focuses on the comprehensive sensitivity analysis for various geometries of solid oxide fuel cells. Since there has been no extensive review on this subject, the aim of this work is to provide a comprehensive literature review on the simulation, sensitivity analysis for various geometries and optimization of solid oxide fuel cells.

A sensitivity analysis of fuel cell parameters has been performed to capture the variables that most affect system cost and efficiency. Different simulation methods focussing on fuel cell behaviour prediction are discussed and reviewed herein. This work also summarizes the present status of various SOFC optimization efforts so that unresolved problems can be identified and settled.
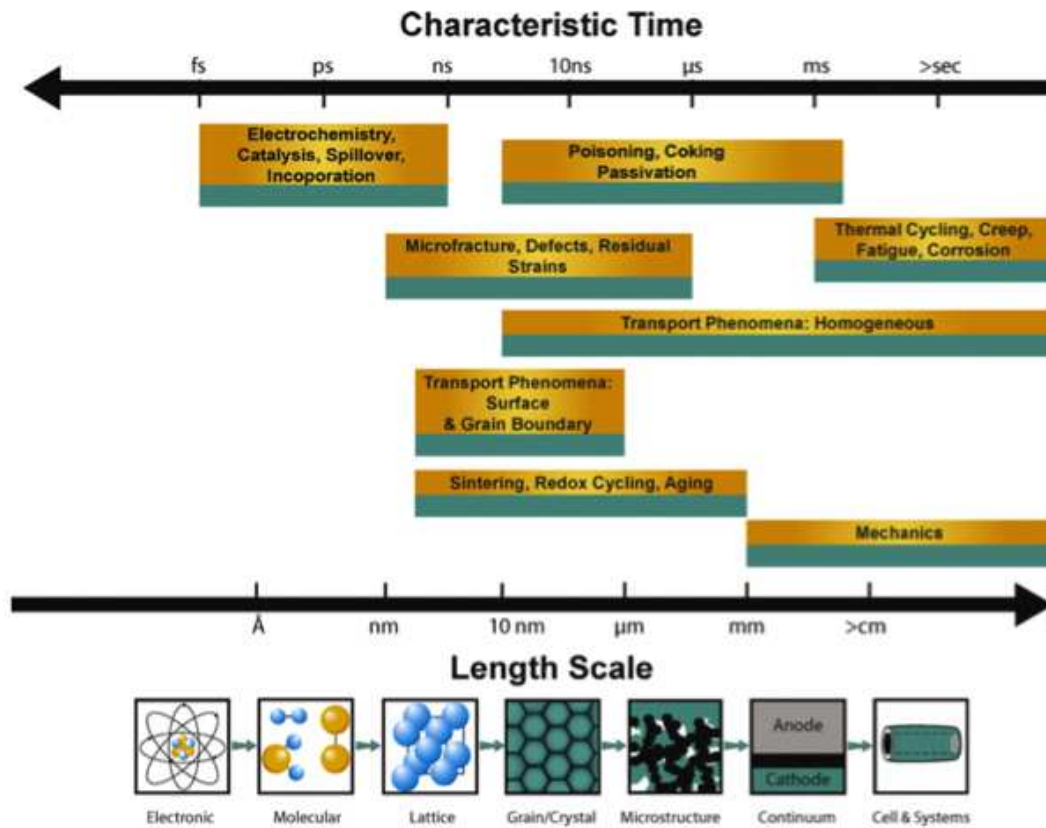
**Figure 1.** Different processes and their corresponding length scales and time frames during the operation of an SOFC (mathematical modelling takes all such processes into account) (figure reprinted from Grew et al. [21] Copyright (2012), with permission from Elsevier).
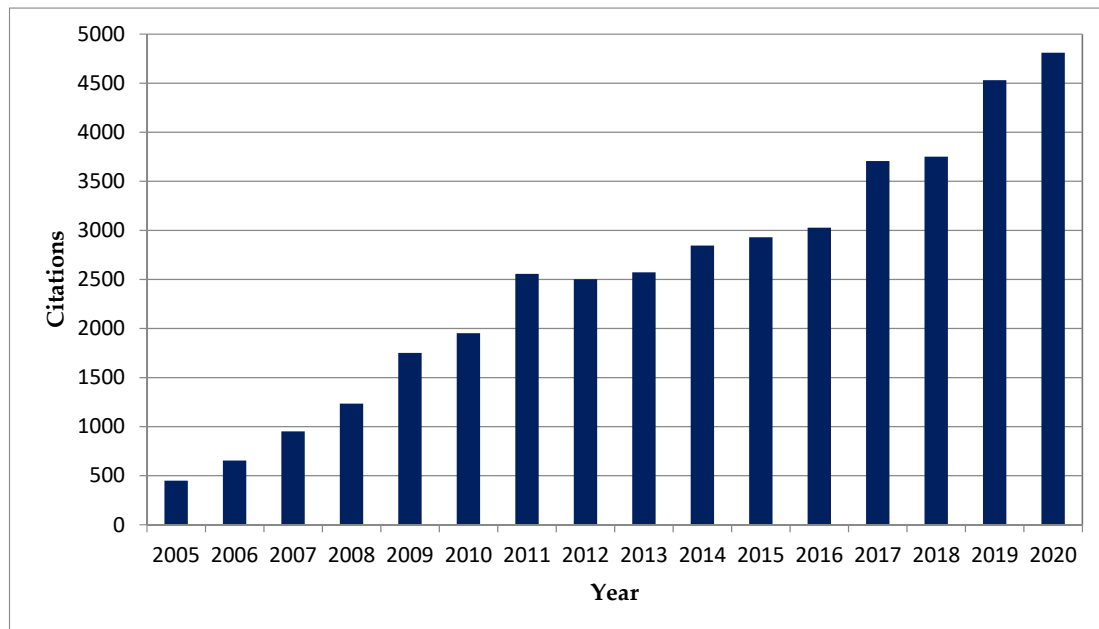


**Figure 2.** Number of citations of articles per year (until December 2020) related to "SOLID OXIDE FUEL CELL OPTIMIZATION", showing the increasing research interest in this topic. Data from Scopus.

## 2. SOFC Configuration Design

The first step of SOFC simulation is geometric design for creating a computerized model of the system that accurately represents its physical dimensions. Furthermore, in

the last decade, there has been numerous research endeavours aimed at increasing the performance of single SOFC systems through configuration design [24,25,31–38].

There are two main classifications for single SOFCs, which are: (i) self-supported, and (ii) externally-supported [31]. As for the self-supported category, the layer that is the thickest functions as support for the structural cell; hence, it could be adopted as electrolyte-supported, anode-supported or cathode-supported (as illustrated in Figure 3). On the other hand, in that of externally-supported, thin layers of single cell are set upon substrate that is porous in nature. In fact, these configurations possess both benefits and drawbacks. For example, configuration based on electrolyte support has a rather solid structure, hence, suggesting more resistance towards oxidation of anodes or reduced cathodes. Nonetheless, due to increased resistance from the electrolyte sheet that is quite thick, the temperature demanded is also high in order to lessen losses of ohmic in electrolyte [32]. In addition, when thin electrolyte is used, the temperatures used may not be high, especially when employing support from anode or cathode. However, it is challenging to develop layers of electrolyte that are dense and those that permeate gas, primarily when the electrolytes become thinner. Compared with the cathode-supported SOFC, there is more attention being given to the anode-supported configuration. Despite that, the low cost of cathode supporting materials may outweigh its disadvantages. Furthermore, when fuel cells are operating with hydrocarbon fuels, a low steam-to-carbon ratio in a relatively thin anode layer prevents carbon deposition on catalyst surfaces. A summary of the characteristics of different self-supported SOFCs is given in Table 1. The simulation studies of SOFCs based on self-supported designs are discussed in detail in Section 4.
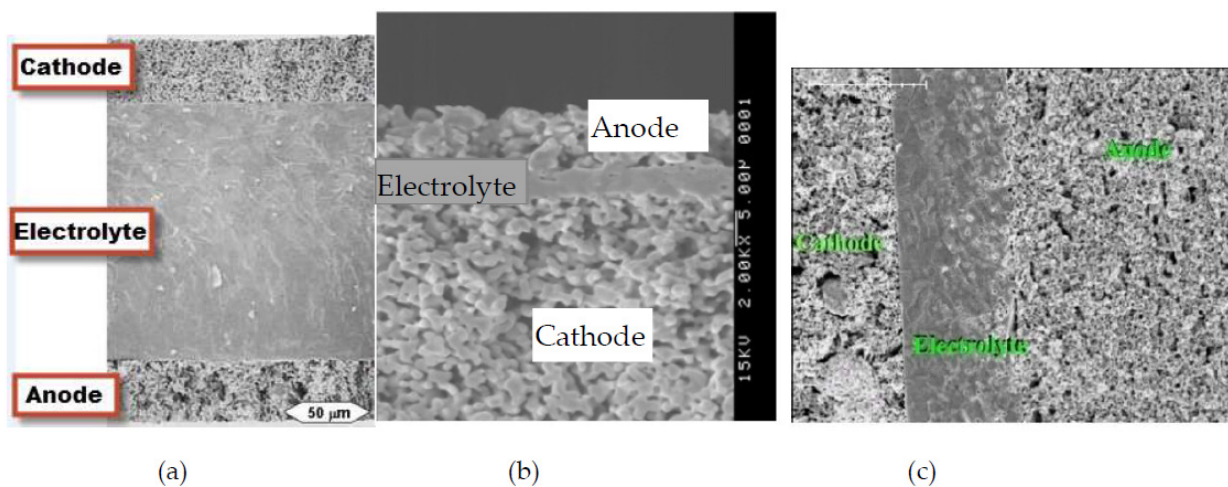


**Figure 3.** SEM micrograph of a single cell (electrolyte: YSZ in center, cathode: porous LSM, anode: porous NiO-YSZ): (**a**) electrolyte-supported, (**b**) cathode-supported, (**c**) anode-supported.

**Table 1.** A summary of the characteristics of different self-supported SOFCs.

| Cell Configurations | Cell Characteristic | Thermo Mechanical Issues: Thermal Expansion Mismatch Stresses and Propensity to Cracking or Delamination |
|---|---|---|
| Electrolyte-supported | High Ohmic Contribution Low Cathodic Concentration Polarization Low Anodic Concentration Polarization | Minimal Tendency for Delamination due to Thermal Expansion Mismatch |
| Cathode-supported | Low ohmic contribution High cathodic concentration polarization Low anodic concentration polarization | Minimal Tendency for Cracking or Delamination due to Thermal Expansion Mismatch |
| Anode-supported | Low ohmic contribution Moderate anodic concentration polarization Low cathodic concentration polarization | Potential for Electrolyte Film Delamination exists as the YSZ Film is in Biaxial Compression |

*2.1. SOFC Design*

There are two principal SOFC configurations: planar and tubular (as shown in Figure 4). In a planar SOFC, the main components are an anode, electrolyte, a cathode and an interconnection between them. The components are all laminated in a plate-type structure. Most fuel cell manufacturers concern themselves with the planar SOFC configuration due to their advantages of low-cost, simplicity, high volume manufacturing as well as high volumetric power densities [33]. Within the electrolyte-supported planar SOFC, the electrolyte generally is made of YSZ with thicknesses of about 100 to 250 μm and an area of $10 \times 10$ cm$^2$ or larger as the supporting part of the cell. The operating temperature of this type of configuration is generally between 850–1000 °C due to the relatively high ohmic resistance of the thick electrolyte. Sealing is a significant factor for a planar SOFC. It inhibits gaseous leakage and reduces cell damage due to the thermal expansion of the cell stack during operation. As a consequence of the cell stack design, which needs a cell support to sustain the air tautness between the air and fuel flow channels, the application of the sealing process tends to be difficult.
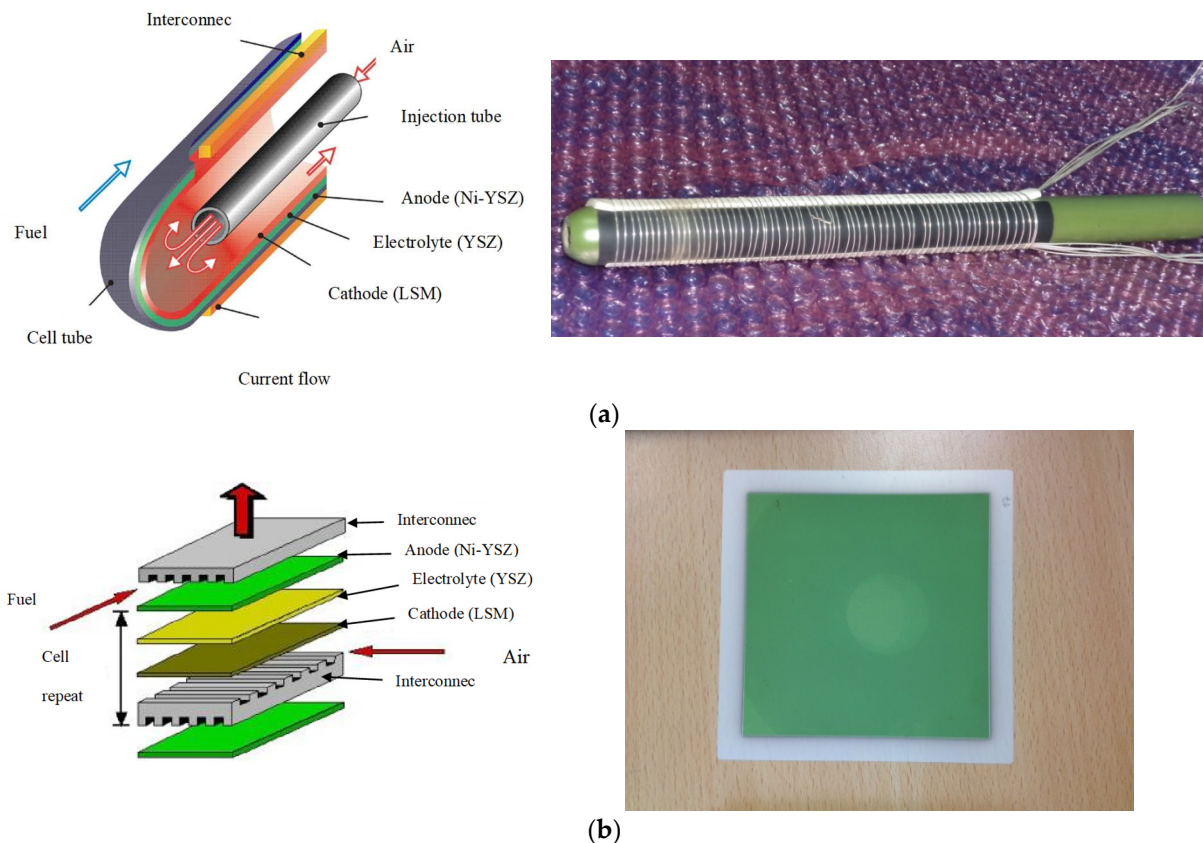


(a)



(b)

**Figure 4.** Schematic of: (**a**) tubular SOFC and (**b**) planar SOFC (figure reprinted from Hajimolana et al. [6] Copyright (2011), with permission from Elsevier).

Planar SOFCs are usually favoured by companies such as Sulzer Hexis. Meanwhile, as for the design of Sulzer Hexis, the supply of fuel is concentrated at the middle of circular cells with 120 mm diameter and supported by electrolyte. After that, the fuel moves towards the cell outer rim to burn, especially fuel that has yet to react with cells. Precisely, the air is supplied, gets heated, and then, moves to the middle part.

The tubular SOFC consists of the same components as a planar SOFC, but with a different structure. A major advantage of the tubular SOFC is that it eliminates the sealing problem, which improves thermal stability, in comparison to the planar configuration [35]. The sealants are required to satisfy all the yardsticks for all the components. They must be

durable in a variety of oxygen partial pressures (air and fuel) while diminishing thermal tensions during high-temperature performance. The trait of the seals should be excellent, since even tiny leaks in such seals may affect the cell potential, causing a decrease in operation. Sealant development is additionally complicated because the optimal sealant depends on the materials of other components resulting in stresses that are induced by a thermal expansion mismatch. Furthermore, this drawback leads to a slow start up capability and less thermal cycling sustainability [34,35].

A number of firms like Westinghouse/Siemens and US-based Acumentrics prefer tubular SOFC to be operated in both manufacturing and army arenas [11,12] for the ease of piling the single cell; hence, this requires low cost for processing [34]. In fact, conventional operations are run on normal fuels, for instance, propane and natural gas, while army related operations employ special fuels like those used for liquid logistics [11,12]. In another instance, a design with standard tubular was employed by TOTO, a firm based in Japan, which was actually initiated with lower industrial cost [35,36], in which 0.5 m length and 16 mm diameter short tubes were selected. Nonetheless, the Japanese firm, Mitsubishi Heavy Industries, has pursued another type of tubular design [6,35,36,38,39]. Besides, the cells that are single are placed in the middle part of the porous support tube, where a series electrical connection is given through interconnector rings made of ceramic, thus increasing the voltage of the tube. In this case, the fuel is pumped into the tube, with air on the external. Moreover, the findings obtained by Singhal [34], Williams et al. [40] and Godfrey et al. [41] portray lower current density in the SOFC tubular, in comparison to that of planar. This occurs due to the in-plane path, where the electrons move from electrodes circumference to interconnecting cells. This in turn increases internal ohmic losses.

Due to the need to increase efficiencies of chemical to electrical energy conversions with less impact upon the environment, several innovations in SOFC technology have been made [34,42]. To date, studies concerning power density escalation have generated a design that purports a flat tube with high power density SOFC (HPD-SOFC), especially to increase power density, but at the same time, maintaining its seal of security [34]. In another case, the Kyocera Corporation developed a design that is supported by anode [41–43]. Moreover, in order to hinder loss of current in the conventional designs of SOFC, the HPD-SOFC has been looked into, in which the design of tubular geometric is redesigned by giving it a flat look and attaching some elements to the electrode that gives flow of current [13]. A flat-tube SOFC has the same components and working principles as that of a tubular SOFC. It is comprised of a cathode, an anode, an electrolyte and interconnections. The design of the geometry and the structure of the stack are the differences between a tubular SOFC and a flat tube SOFC. The flat tube design has a seal-less feature compared to the tubular SOFC. Moreover, a flat tube SOFC can be configured with a thin cathode layer to decrease the concentration polarization, because the internal ohmic resistance can be reduced by adding ribs into the cell stack [13]. In fact, one advantage refers to the limited space that is void among the stacks of cells mainly because of the geometry, hence, making the bundle of cells tighter. As such, the scheme with a tube that is flat should project enhanced progression when compared to the SOFC, although the fuel and air supplies reflect similar approaches. This attribute decreases the resistance of the cell, but heightens the density of cell power.

In another innovation, a joint development between Mitsubishi Heavy Industries and Chubu Electric Power Company manufactured a combination of tubular and planar designs into a so called MOLB-type (Mono-block layer Built) planar SOFC [44]. The cells were manufactured up to a size of $200 \times 200$ mm$^2$. This may not only increase the effective active area but also make the cell design fairly compact. Figure 5 shows the schematic of a monolith and an HPD SOFC.

Rolls-Royce (now LG) has developed an integrated planar SOFC design using a multi-cell MEA concept (Figure 6) [45], which takes advantage of both excellent thermal expansion compliances from tubular design and low-cost component fabrication from the planar configuration. For instance, a new symmetrically bi-electrode supported cell (BSC) was designed by Cable and Sophie by employing an approach known as the freeze tape

casting method [46]. This fresh BSC concept possesses dual exceptional attributes, which are as follows: (i) more stable properties of the mechanical are ensured in the conditions for thermal cycling, and (ii) thick and porous electrodes substitute the flow channels of gas and fuel in order to provide a more compact cell layer. Nonetheless, such tight layer poses an issue, which affects the efficiency of the gas and fuel supply, mainly due to the resistance exerted by the electrodes that are porous.



(**a**)　　　　　　　　　　　　　　　(**b**)
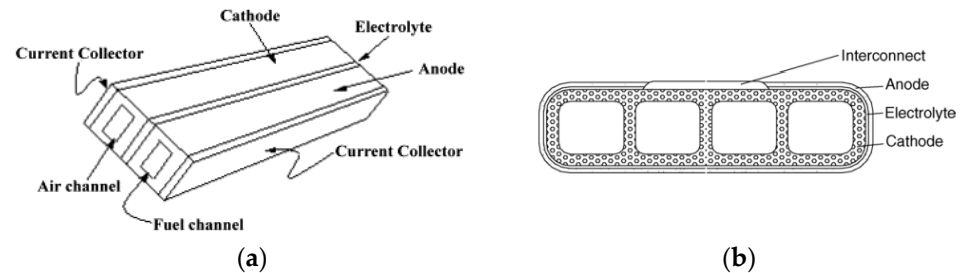
**Figure 5.** Schematic of the: (**a**) monolith type fuel cell modelled and (**b**) HPD-SOFC (figure reprinted from Singhal, S.C [34] Copyright (2000), with permission from Elsevier).
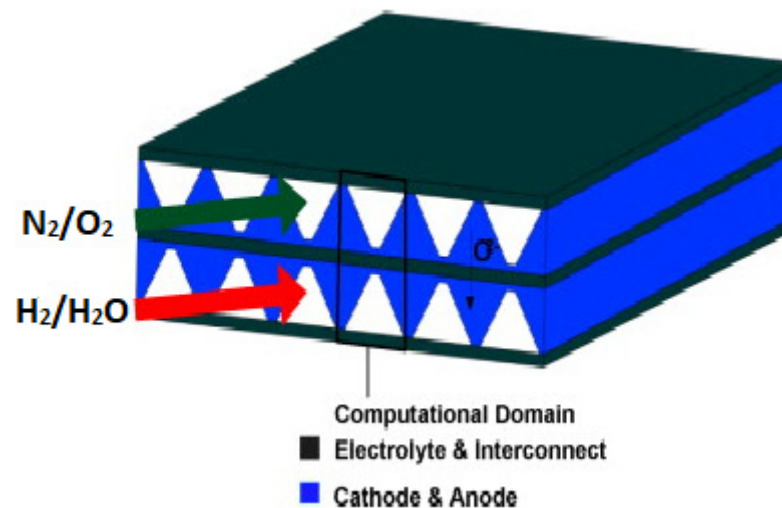


**Figure 6.** Multi-cell MEA concept developed at Rolls-Royce (figure reprinted from Gardner et al. [47] Copyright (2000), with permission from Elsevier).

Meanwhile, a fresh SOFC was looked into by Shi and Xue [13] by applying a simulation with a numerical basis, whereby BSC electrodes that are porous and thick are attached to channels that are micro in size, as illustrated in Figure 7. Thus, it was projected that minute-sized channels would improve the supply of gas and fuel by dismissing the ridiculous pressures of gases, besides taking advantage of the compact and symmetry features of the BSC.

The fuel cell developed by Ceres Power is a unique adaptation of the SOFC technology, which uses a new generation of metal supported ceramic cells. This type of fuel cell invented by the Imperial team is particularly well suited to operation with hydrocarbon fuels such as natural gas [45,46,48]. These properties mean that the fuel cells can run at substantially lower temperatures than conventional designs. Lower temperatures mean that an extremely thin layer of stainless steel can be used in the fuel cells, dramatically reducing material costs, improving resilience and increasing power density, thus making them more attractive for domestic applications such as micro combined heat and power (mCHP).

A very new innovation made at Imperial College London aims to address outstanding technological and economic issues by establishing the feasibility of, and developing, a novel design of SOFCs [46,47,49]. Moreover, fibres that are hollow are used to generate this SOFC, hence, adding to the surface area among the electrodes, which later escalates

power output per unit volume/mass and ensures higher temperatures sealing; and most importantly, these fibres decrease the costs incurred (Figure 8).
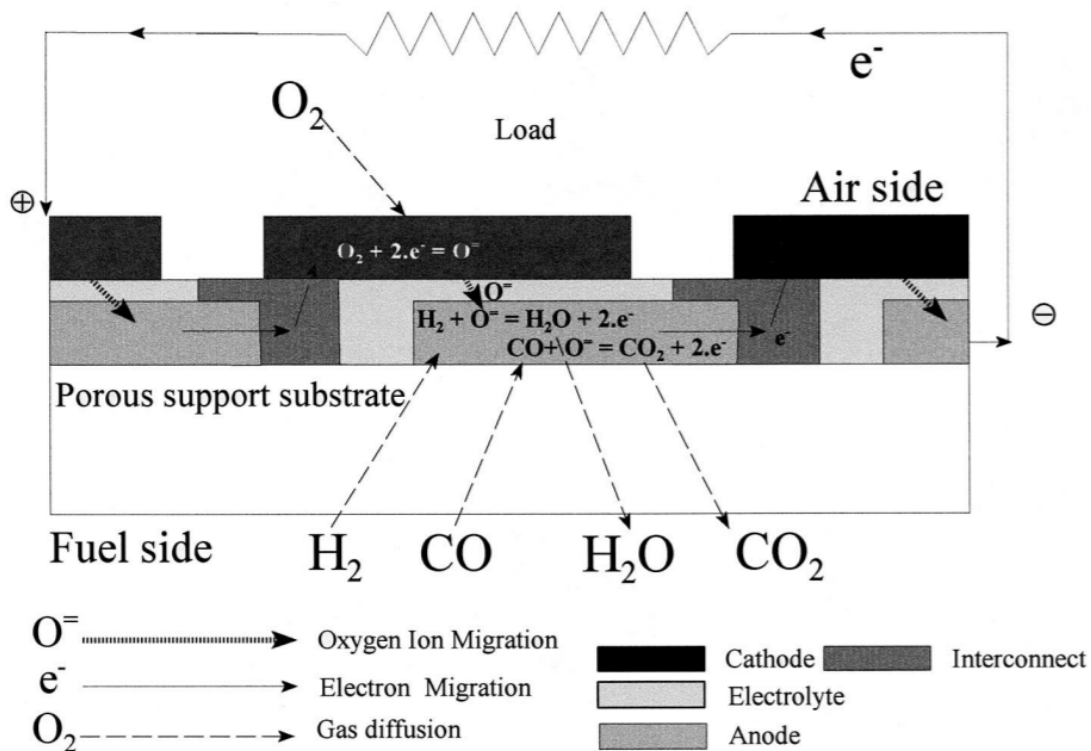


**Figure 7.** Schematic of BSC with micro-channels (figure reprinted from Gardner et al. [47] Copyright (2000), with permission from Elsevier).
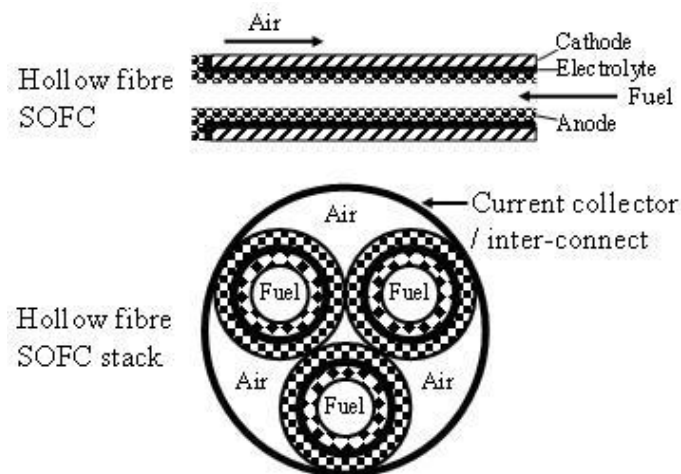


**Figure 8.** Hollow fiber SOFCs designed at Imperial College London (figure reprinted from Gardner et al. [47] Copyright (2000), with permission from Elsevier) [46,47,49].

## 2.2. Micro SOFC

A new concept of micro-SOFC (μ-SOFC) has been looked into, for several reasons: high SOFC power densities, as well as advancement in film and micro technologies [50–53] (Figure 9). In fact, this μ-SOFC requires low power within 1–20 W, similar to the use of a battery in low energy electronic devices like camcorders, industrial scanners, medical devices, battery charger, and laptops. Additionally, this μ-SOFC design offers energy density up to four folds and optimum energy/weight, when compared to batteries that could be charged like Li-ion and Ni metal hydride batteries [54].
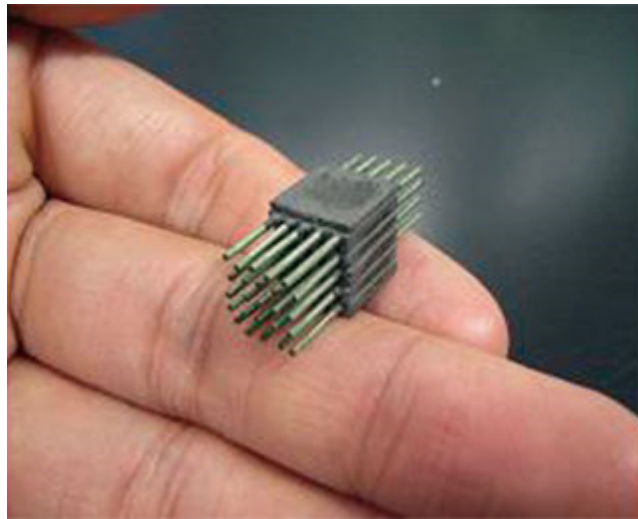
**Figure 9.** Schematic of a micro-tubular SOFC (figure reprinted from Hanna et al. [55] Copyright (2014), with permission from Elsevier).

The thinner SOFC layers reduce the ohmic loss at lower operating temperatures. In addition, thinner layers reduce the thermal mass of the cell [56]. Therefore, a stack of μ-tubular SOFCs is suitable for rapid start-up and shut-down cycles, with only a few minutes compared to hours for conventional designs [57]. μ-tubular SOFCs have been heated from room temperature to 850 °C within 10 min without failure and survived thermal cycling [58].

The small diameter of the μ-tubular SOFCs allows for the 'packing' of a more 'active area' within the same volume, thereby increasing the volumetric power density of a stack. The latter two advantages are especially important for portable power generation applications [57].

Nevertheless, this μ-SOFC projects decreased cell functionality because of the low flow of gas within the micro tube [59], which leads to deprivation of fuel, hence generating a decrease in electromotive force, as well as escalated resistance against anodic polarization. Therefore, some material with porous attribute within the free space could enhance the flow of gas to the cells [59].

Thermal management is a critical issue in the development of μ-SOFCs. In order to control the temperature of a SOFC during single cell testing, a furnace was constructed by Kendall and co-workers [60]. The furnace comprised two refractory bricks that were hollowed out to house electrical elements and a small manifold to hold a micro-tubular SOFC. The heating elements were constructed from Kanthal ribbon, with a resistivity $5.476 \, \Omega \, \text{m}^{-1}$, surrounding the supporting alumina tubing. These heating elements were connected to a power supply fitted with a temperature control device. Simulation approaches can help in the specification of operating conditions and designs that minimize temperature gradients and maximize the use of heat produced in the SOFC stacks. However, few investigations address the issue of thermal stress in μ-SOFCs [58,61–63]. Amongst them is the study conducted by Lockett et al. [58], who investigated the thermal management of a 20-cell μ-SOFC stack using a CFD model.

Different simulation methods have been studied for μ-SOFCs with different designs and operating conditions to enhance their performances and energy efficiencies. Aimed to improve the current efficiency, a new design of a μ-tubular SOFC was proposed [64]. In addition, some hollow fibres with ceramic are used in μ-SOFC to function as supportive element, where the functional layers are laid above. Furthermore, the conditions of thermal-fluid and electrochemical proposed in the design are examined via simulation of CFD, which revealed enhanced performance mainly due to essential decrease in the losses of ohmic, in comparison to micro tubular design supported with anode. In a different study, a parametric analysis was performed to study the effects of temperature, fuel flow rate, fuel composition,

as well as anodic and cathodic pressures on the μ-tubular SOFC performance [65]. Besides, this model of CFD incorporates several features like mass, momentum, species and energy balances, as well as transfer of ionic and electronic charges. In fact, this model was applied in a market-available multi-physics software programme known as COMSOL 3.4 by employing the method of finite element in order to determine the equations of partial differential. Moreover, it was projected that for output current of 0.53 A cm$^{-2}$ at 550 °C; an ionic current density at 0.65 A cm$^{-2}$ would have to be generated by the fuel, in which the variance had been due to leaks that occurred at the internal current. Additionally, temperatures below 500 °C did not project any significant leakage of electronic current.

A CFD simulation based on a multi-dimensional approach, coupling heat transfer, fluid-dynamics and electrochemistry of increasingly complex systems (micro and midi-reactors, constituting 15 and 45 tube assemblies, respectively) was performed to design a μ-tubular SOFC [66]. The system was designed to optimize the geometry of the respective modular assemblies in terms of air and heat release management to acquire the maximum performance in terms of power density whilst keeping the impact of thermal stresses to as low as possible with regards to cell durability. In another effort, a 3D thermal-fluid-dynamic model of a micro-tubular SOFC was presented to analyse both micro- (i.e., a 15 tube assembly) and midi-assemblies (up to 45 tubes) with the aim of optimizing the fuel cell configuration [67]. A summary of the characteristics of different SOFC designs is given in Table 2. Although several innovations to the SOFC in terms of design and configuration have been made to increase the power and performance, still more attempt is required to achieve this goal.

**Table 2.** Summary of different SOFC design characteristics.

| SOFC Design | Characteristics Design | Advantages | Drawbacks | References |
|---|---|---|---|---|
| Planar SOFC | The components are all laminated in a plate-type structure | Low-cost, simplicity in manufacturing and high volume manufacturing with high volumetric power densities. Has higher energy efficiency compared to tubular SOFCs | Difficulties in sealing Internal stresses in cell components due to non-uniform temperature distributions and high manufacturing cost | [33] |
| Tubular SOFC | The components are all laminated in a tube-type structure | Easy to assemble into a cell stack, leading to a lower processing cost | Has a much lower current density than a planar SOFC High ohmic loss | [35] |
| Flat-tube | The tubular geometric design has evolved into a flattened one with ribs in the electrode acting as bridges for current flow | Enhanced power density compared to planar and tubular SOFCs Has a seal-less feature compared to the tubular SOFC Can be configured with a thin cathode layer to decrease the concentration polarization The void space between the cell stacks is reduced | Retains the feature of secure sealing | [34] |
| Bi-electrode supported cell | Uses a freeze tape casting technique | More stable mechanical properties in thermal cycling conditions | This compact design imposes a great challenge on how to effectively feed fuel/gas due to the high diffusion resistance of porous electrodes | [46] |
| μ-SOFCs | | Higher energy densities per volume and specific energy per weight are obtained compared to rechargeable batteries A stack of μ-tubular SOFCs is suitable for rapid start-up and shut-down cycles, with only few minutes compared to hours for conventional designs Low cost Low concentration | Lower cell performance | [50–53] |

### 3. Computational Simulations of SOFC

Hence, comprehending the basic functions is deemed as an essential phase in both optimizing and improving SOFC and its functionality. Although many studies have looked into these SOFC designs [68–70], it is indeed challenging to carry out empirical studies, especially in determining the scenarios of internal transport within the intricacy of internal physical mechanisms. On top of that, SOFC related studies take up a lot of money and time, thus highlighting the importance of computational simulation that saves both cost and time [71].

In addition, optimum designs must be built to ascertain excellent functionality and consistency, as well as lowness in cost. Besides, good simulation models offer ideas about functionality, besides identifying the tools required in devising a design that is functional, which takes into account estimates accuracy, detailed information and varied simulations via computer, especially when building numerical models.

Numerical modelling and simulation must be used very carefully, however, especially when used for prediction purposes [72–75]. The models are typically based on computational and simulation approaches such as FLUENT, CFD-RC, gProm, COMSOL, MATLAB and Star-CD, to evaluate the cell performance at different conditions [71,76–80]. The simulation and computational models are coupled to the 0, 1, 2 and 3 dimensions.

Eichhorn Colombo et al. [81] employed a detailed solid oxide fuel cell (SOFC) model for micro-grid applications to analyze the effect of failure and degradation on system performance. They presented the design and operational constraints on a component- and system level. A degrees of freedom analysis identifies controlled and manipulated system variables which are important for control. They used experimental data to model complex degradation phenomena of the SOFC unit. They assumed that the SOFC unit is consisting of multiple stacks. They studied the failure scenario of the loss of one individual SOFC stack, e.g., due to breakage of sealing or a series of fuel cells. Their simulation results reveal that degradation leads to significant drifts from the design operating point. In addition, failure of individual stacks may bring the still operating power generation unit into a regime where further failures and accelerated degradation is more likely. It is presented that system design, dimensioning, operation and control are strongly linked. Apart from specific quantitative results perhaps the main practical contribution are the collected constraints and the degrees of freedom analysis.

Haoran et al. [82] developed a hybrid model for an on-line analysis of SOFCs at the cell level. Their model combines a multi-physics simulation (MPS) and deep learning, overcoming the complexity of MPS for a model-based control system, and reducing the cost of building a database (compared with the experiments) for the training of a deep neural network. They considered the maximum temperature gradient and heat generation as two target parameters for an efficient operation of SOFCs. Their modelling results reveal that a precise prediction can be achieved from a trained AI algorithm, in which the relative error between the MPS and AI models is less than 1%. Furthermore, an online optimisation is realised using a genetic algorithm, achieving the maximum power density within the limitations of the temperature gradient and operating conditions. Their method can also be used for prediction and optimisation of other non-liner, dynamic systems. Their study shows that the combination of an MPS, a deep neural network and a genetic algorithm provides a promising solution for model-based control systems for precisely and quickly analysing the performances of SOFCs and other non-linear systems. Their deep learning algorithm is trained based only on simulation data under stationary operating conditions. Further training with more experimental data on the algorithm is needed for a wider application under dynamic operating conditions.

Zero-dimensional (0D) methodology, also called box models, simply allows studying processes that can be analysed without taking into account spatial configuration and geometry. Furthermore, the system is investigated using a box and the spatial is averaged for the dimensions. Thus, any difference is dismissed, but the transfer scenario is weighed in to distinguish the output from input variables. Nonetheless, time emerges as the

independent variable for dynamic simulation. Hence, far from being intricate, this model is indeed applicable for SOFC design and does not require numeric. The dynamic variables, flow rates, gaseous compositions, pressures and the basic energy balance to determine the sources of heat to maintain the cell temperature are assumed to have changed between input and output values or in time (dynamic models) [62,83,84]. However, in these models the geometry does not affect the fuel cell performance directly due to the neglect of the spatial variation of variables. Therefore, 0D models can be used in studies where the focus is not on the fuel cell itself but on how the SOFC affects the performance of the entire system. 0D models are usually applied for thermodynamic modelling of SOFCs and numerical analysis of fuel cell-based power systems (such as SOFC/gas turbine systems and combined heat power systems) or μ-SOFCs (see Figure 10). The objective of this kind of analysis could be for optimizing the fuel cell performances and to determine the optimal operating parameters for the chosen application.
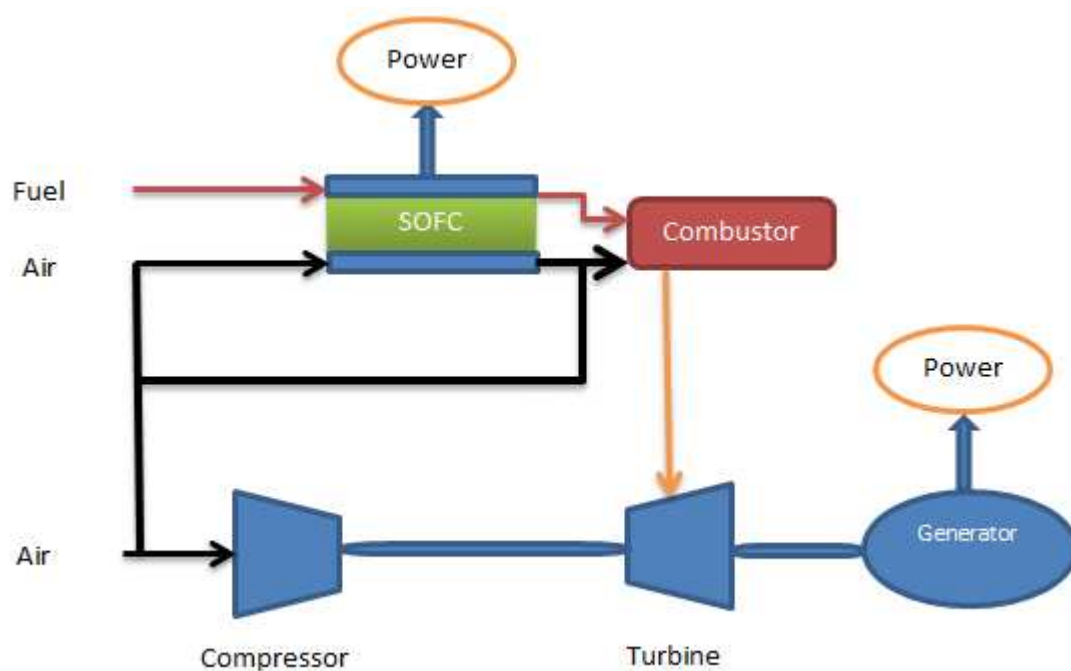


**Figure 10.** Schematic of a hybrid SOFC/GT system for 0D simulation models.

In parallel to the design and optimization of the SOFC/GT system, control and system integration solutions have also been explored in 0D models [85,86]. A simple control relevant model may perform better than its complex 3D counterpart for control design, which on the other hand, may be suitable for design and analysis of the process. A process expressed by a very complex model may be controlled by a regular PID controller. On the other hand, a 0D model may be attributed with plenty of environmental and economic constraints that require multivariate controllers to maintain the optimal performance of the system. Hence, a model-based predictive control strategy, which could handle various operational constraints, has been developed in a 0D model for coordinated power and thermal management [85,86].

In 1D simulation models, the fuel cell is represented by ordinary differential equations and the variation of parameters along the other two directions are negligible. In planar SOFCs, the dimension is usually determined by the gaseous flow direction in the fuel cell. In the 1D, the parameters vary along a direction parallel to the axes of the gas channels for planar SOFCs. Therefore, cross-flow planar SOFCs cannot be simulated by 1D simulation approaches [87].

For tubular SOFCs, the dimension is usually determined by the tube axis which includes the direction of the fuel and oxidant gas flow as in models presented by some

researchers [62,88,89]. Gaseous compositions, flow rates, pressures and temperatures are averaged for each gaseous flow channels. The electrochemical model is considered as a simple prototype.

A rather similar characteristic of current density distribution along the fuel cell is obtained for 0D and 1D models [84]. However, for lower current densities, an average temperature of the fuel cell attained with the 0D model is higher than that computed by means of the 1D model [84]. Besides, numerical findings highlight the significance of temperature for predictions of source and losses, as well as to model properties of materials and transport. Meanwhile, density of current is only required in estimating losses and source. Moreover, information about temperature reveals thermal stresses levels upon fuel cell, hence affecting both the design and the optimizing courses. Therefore, 0D models are not suitable for thermal stress studies.

In two dimensional approaches, two directions of the entire fuel cell operation are considered. In this method it is assumed that the fuel cell's behaviour is the same for each angular section. In case of three-dimensional approaches, mathematical formulae are usually written in the form of partial differential equations. A 3D model is suitable for sensitivity analyses of geometrical parameters of the fuel cell [90,91]. In fact, these researches are rather important in dictating the performance, besides identifying the associated drawbacks to be hindered. Nevertheless, a modelling that is in 3D is definitely costly for fuel cells with solid oxide in layers. Moreover, high power has to be used for implementation of 3D CFD models with layers of SOFC and support elements. Furthermore, time becomes a constraint, hence impeding optimization. Precisely, estimations that are accurate are sought when model optimization is carried out in relation to layers of SOFC, together with reformers or pre-heaters.

## 4. Sensitivity Analysis

In order to improve the efficiency and reduce the cost, sensitivity analysis of fuel cell parameters needs to capture the significant variables. The parameters that affect the efficiency and cost of fuel cells are: (i) fuel cell design parameters, (ii) operating parameters and (iii) fuel cell materials. Numerical simulation is expected to help optimize the design, operating parameters and fuel cell material. Therefore, sensitivity analysis for these parameters using simulation models is conducted to optimize the fuel cell systems operation with higher performance and lower cost to make them commercially viable in the market.

### 4.1. Sensitivity Analysis of Fuel Cell Design Parameters

Optimizing the fuel cell's configuration is a significant means towards enhancing its power density and operation. The optimization of an SOFC's external configuration includes an optimal fuel cell inventory allocation of cross-sectional area and the width of the gas channels so that the SOFC's electrical and pumping powers are balanced accordingly to attain a maximum net value. In addition, resolution of optimally active TPB regions in both electrodes as well as the electrolyte, which is founded on the trade-offs between all the potential losses in the electrodes is also significant for improving the power density.

Most SOFC designs have been reported using numerical modeling and simulation, with a goal to maximize the power density and fuel utilization [13,92–95]. Moreover, the issue of designing a reliable SOFC due to high temperature operation was investigated in details in the literature [2,34,35]. In addition, simulation helps to curtail the non-uniform current density and temperature distributions that directly influence the thermal stresses in a range of SOFC components [22,96–98].

Dang et al. [99] developed and solved a numerical model of gas flow, heat transfer, mass transfer and electrochemical reaction multi-physics field coupling of a planar SOFC. They analyzed distribution of velocity, temperature and concentration inside the SOFC cell. They discussed the influence of cathode inlet flow rate, porosity, rib width and other parameters on the performance of SOFC. Their simulation results showed that within a

certain range, increasing the cathode inlet flow rate can significantly increase the average current density of the cell. Increasing the porosity of the electrode can improve the gas diffusion of the porous electrode, thereby increasing the rate of the electrochemical reaction. Increasing the width of the ribs will result in a significant decrease in cell performance. Therefore, the rib width should be reduced as much as possible within the allowable range to optimize the working performance of the cell.

Parametric investigation of the SOFC flow channel arrangement is the focus of attention for optimal power generation. Significant parameters involve the widths of anodic and cathodic flow conduits, their heights and the width of current collector contacts with the cell. Table 3 summarizes the power density and electrical efficiency with various geometrical configurations of flow channels indicating with types A-F. Decreasing the channel sizes causes a shorter current path and also enables higher heat/mass transport rates [100]. This leads to variations in temperature and concentration distributions, and also affects the power efficiency of an SOFC. The conduits with smaller heights suffer a shorter current path and also result in significant improvement in the mass and heat transfer coefficients. Higher mass transport rates in smaller conduits reduce the concentration polarization in channel flows significantly and enhance the cell's operation. Generally, when the flow channel dimension is reduced from conventional to micro-scale, the temperature/concentration gradient of the SOFC drops, and the heat/mass transfer coefficient is observed to rise accordingly. In addition, minimized sizes are significant to developing low cost and portable SOFCs. Thus, it is expected that concentration loss and operating temperature in smaller SOFCs may be effectively decreased. However, for the smallest cell, the ohmic polarization is higher than those of the larger cells due to a lower operating temperature [101]. Therefore, an optimized fuel cell length should be designed.

**Table 3.** Power density for different geometric arrangements of the flow conduits at 0.6 V [98].

| Type | Configurations (mm) | | | | | | Power Density (mW/cm$^2$) | Electrical Efficiency |
|------|--------------------------|------------------------------|----------------|------------------|--------------------|----------------------|----------------------------|-----------------------|
|      | Anode Channel Height | Cathode Channel Height | Anode Width | Cathode Width | Anode Rib Width | Cathode Rib Width |  |  |
| A | 1 | 1 | 1 | 1 | 1 | 1 | 752.2 | 49.2% |
| B | 0.5 | 1 | 1 | 1 | 1 | 1 | 758.5 (0.83%) | 49.6% |
| C | 1 | 0.5 | 1 | 1 | 1 | 1 | 754.9 (0.36%) | 49.6% |
| D | 1 | 1 | 2 | 1 | 1 | 1 | 755.5 (0.44%) | 49.3% |
| E | 1 | 1 | 1 | 2 | 1 | 1 | 753.1 (0.12%) | 49.7% |
| F | 1 | 1 | 1.4 | 1.4 | 0.6 | 0.6 | 767.4 (2%) | 50.7% |

Alteration in the size of annulus within electrode tubes at high current densities could give significant impact upon tubular SOFC, especially on polarization concentration and activation [102]. Thus, a drop in anodic channel radius for tubular SOFC supported with anode could escalate the decreasing channel pressure primarily due to heightened cell length, as well as a decrease in resistance of ohmic due to the current path that is shorter, but electrode thickness and electrolytes are retained [92]. Generally, difference in tubular SOFC radius brings variation to the electron path circumference length, while alteration in the length of cells modifies the cross-sectional zone for electron movement [103].

For the cell with the minimum rib resistance, SOFC performance enhances moderately with reduced channel dimension [103]. In contrast, cell operation enhances significantly with the decrease in channel dimensions for the case whereby the SOFC possesses higher rib resistances. Nevertheless, a cell that has the least dimension for channel disregards maximum output for the terminal. In fact, there may be two reasons for this occurrence: (i) lower temperature accelerates loss of ohmic at the solid area, hence associated to func-

tional temperature, and (ii) a current path that is shorter and a minute loss of concentration might partially or completely interact with progression of ohmic polarization [104].

A wider rib decreases the enhancement in the temperature rate. Therefore, it may be concluded that such an arrangement could reduce thermal stresses. Figure 11c,d show that for cells having thinner ribs, the oxygen concentrations are more constantly distributed across the whole porous layer (z-direction), which excludes inefficiencies that are usually caused by obstructed "dead zones" below the ribs [104]. When compared with a thicker rib, the oxygen molar fraction next to the surface of the diffusion layer is far higher for the cell with a thinner rib. The decrease in concentration loss for a thinner rib, however, is not significant enough to counteract the enhancement in ohmic loss and thus, the terminal output still reduces. Therefore, it is anticipated that a lower channel width ratio to the rib width may decrease the ohmic losses at the interface and enhance cell performance [104].
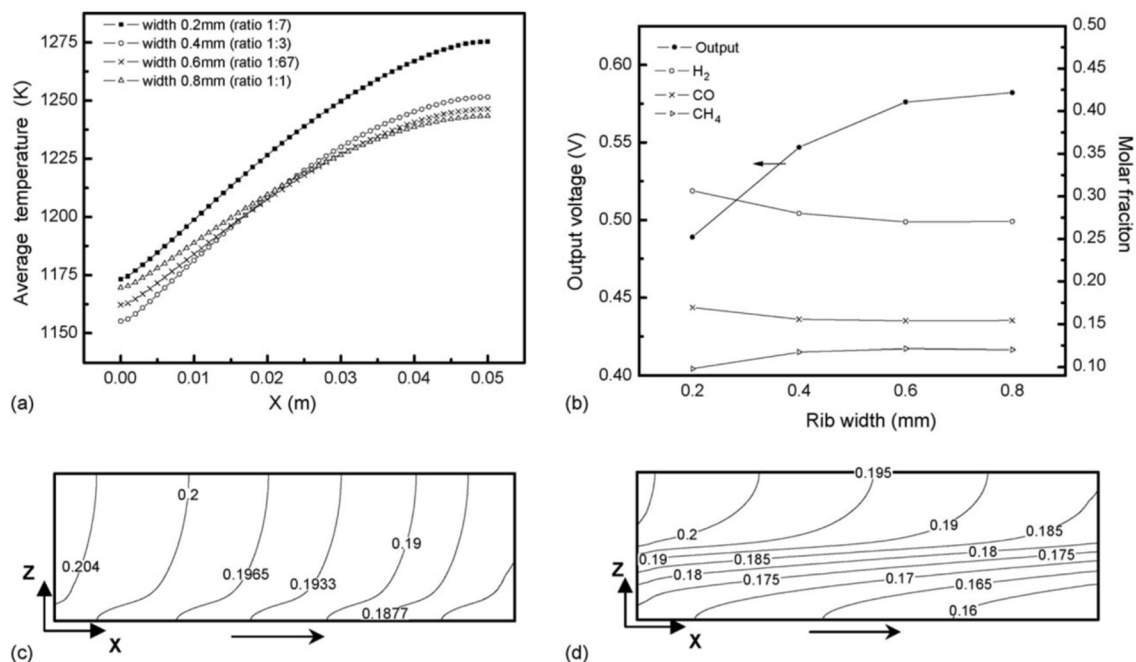


**Figure 11.** Effects of rib width on SOFC performance: (**a**) average top interconnector temperature, (**b**) comparison of performance, (**c**) oxygen at the cathode/electrolyte interface with the thinnest rib (ratio 1:7), (**d**) oxygen at cathode/electrolyte interface with the thickest rib (ratio 1:1) (figure reprinted from Ji et al. [104] Copyright (2006), with permission from Elsevier).

Rib number tends to improve the cell terminal voltage. This enhancement is higher when the rib number rises from a small to a median value. Further enhancement in the rib number, however, reduces the magnitude of the enrichment [105].

The ratio of the support thickness to the conduit diameter is taken as significant design parameter. A reduction in the anode support thickness results in an enhancement in the hydrogen mass transfer rate [105]. Simultaneously, the ohmic polarization tends to decrease as a consequence of the reduction in the electron path. The thin support may reduce the mechanical durability of the cell, however, leading to deformation or cleavage of the cell during manufacture, stacking or when running [105]. For an electrolyte-supported SOFC, a greater ohmic polarization in the electrolyte is a significant issue hence adopting a thin-film electrolyte to considerably decrease the ohmic loss is a valuable consideration.

Although there are many investigations on sensitivity analysis of fuel cell parameter design, still, more research studies are required on optimal design and geometry optimization of SOFCs to enhance the power density and fuel cell volume reduction. Studying novel fuel cell configuration design is necessary for fuel cell cost reduction and power density enhancement.

*4.2. Sensitivity Analysis of Flow Field*

For efficient cell stack design, simulation study of the uniform inlet flow is significant for reducing the temperature changes in the cell stack structure and enhances the cell power density [106–109]. Cross-flow geometries have minimal effect on gas flow mixing. They enhance cell active area, resulting in improved SOFC performance whilst maintaining thermal gradients and current path lengths to the standard planar fuel cell geometry [110].

Counter-flow configuration is the subject of interest as a consequence of its potential for higher efficiency. This is the case because of the SOFC's high efficiency in using the electrochemically generated heat for fuel processing. A design with counter-flow channel generates power density higher when compared to its related co-flow and cross-flow aspects [87]. Nonetheless, the design of layered cells employing channels with counter-flow portrays some integral inadequacies, including irregular flow, as well as intricate designs that demand fuel and air flow inlet and outlet coupling. Therefore, in order to decrease these drawbacks, an intricate configuration based on geometrics is deemed as significant for channels of inlet and outlet meant for fuel and oxidizer.

On top of that, the counter-flow has been discovered to be more efficient when compared to co-flow mainly due to the high current at the entrance area of the anode for its concentration of reactant [111]. In fact, for the same functional setting, higher voltage and bigger power output could be predicted for the design of counter-flow primarily because of low concentration and active polarization [112].

A good design of flow distributors can allocate the fuel and the air uniformly onto the anode and the cathode for achieving consistent diffusion processes through porous electrodes. It has been discovered that by simply employing small guide vanes equally spaced around the feed header of conventionally used rib-channel flow distributors (to effectively improve the degree of flow uniformity in flow distributors), the power density of the single-cell stack may be enhanced in comparison to that where guide vanes are employed under normal experimental conditions [100].

The enhancement of flow uniformity in the distributors is useful to enhance the redox stability of the Ni-based anode, so that a balanced employment of the anodic catalyst may be attained for the continuous cell operation. This is significant in extending the longevity of the cell stack [100]. Thus, the optimization of interconnects or flow distributors is crucial for further enhancement of the SOFC's performance.

*4.3. Sensitivity Analysis of Micro-Structures and Electrodes-Electrolyte Thicknesses for an Efficient Triple Phase Boundary (TPB)*

The essential aspects that accelerate the performance of cells are the provision of effective TPB or electrochemical reaction region, and increased conductivity of ionic and electronic in both ion and electron of the conducting elements [113].

The microstructural features considerably affect transport and reaction phenomena, influencing the electrode performance. For instance, decreasing the particle size substantially improves the number of reaction sites [114], which may be maximized by adapting the electrode composition [113]. The porosity is another significant factor: while a reduction in porosity initially causes an enhancement in electrode efficiency [115], blockage of pores and gas diffusion constrictions tend to hamper the electro-chemical performance to below a critical limit. Figure 12 clearly illustrates that undesirable microstructures such as insufficient porosity reduce cell performance [92].

Thickness is a geometric aspect that has an impact upon the performance of electrode [113] because an increment in thickness improves the amount of sites with reaction. Otherwise, electrons, as well as molecular and oxygen ions, would have to move further in order to arrive at/depart from the sites with reaction. Therefore, electrode thickness could affect both loss of ohmic and over potential concentration, thus influencing the performance of fuel cell [56]. Moreover, the anode thickness in a tubular SOFC could influence the surface area of TPB that involves reactions of both cathode and anode. Besides, an increment in electrolyte thickness alters the resistance of ohmic, which could even halt

the performance [116], as well as affecting the length of current path in cathode and the cathodic active reaction area [116].
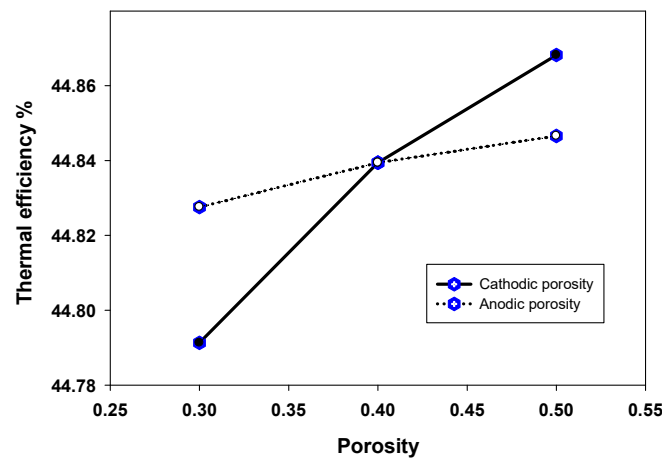


**Figure 12.** Effect of electrode porosity on fuel cell efficiency (figure reprinted from Hajimolana et al. [92] Copyright (2012), with permission from Elsevier).

In general, the over potentials of ohmic have to be hindered due to its significant effect upon the potential loss of cells [117–119]. Specifically, the ionic conductivity that derives from ion conducting elements at reaction region must be enhanced for cell functionality improvement to generate ohmic [120].

Meanwhile, due to charge transfer resistance and active polarization, over potential anodic becomes a factor, whereas the case of over potential of ohmic in anode takes place because of ions and electrons flow resistance at the reactive region and at the anode layer, respectively [62,83]. Besides, the potential of loss in cells often derives from the anode in the SOFC and the four times smaller anodic concentration over potential when at increased current density [120]. Other than that, the effect of cathodic concentration over potential upon total cell potential is insignificant primarily because of the thin-sized cathode, while the over potential of cathodic activation is smaller by two folds, in comparison to the over potential of ohmic [120–122]. Nonetheless, Hussain et al. [120] asserted the significance of cathode and electrolyte thickness, inclusive of those at the backing and reaction regions in SOFC with anode support, as they could affect the loss in the elements.

For instance, over potential of cathode incorporates activation, ohmic and concentration elements [89]. On the other hand, the ion transport resistance in the electrolyte layer generates electrolyte over potential [120]. Meanwhile, as for SOFC with electrolyte support, electrolyte in the form of thin film is required to decrease ohmic loss for huge ohmic polarization appears to be a rather big threat [123–125].

Furthermore, the illustration in Figure 13 shows that certain porosity and particle size do affect the over potential in electrode minimally based on its thickness [126]. Meanwhile, Figure 13a portrays improvement in cathodic over potential for thickness exceeding minimal setting, which is rather little. This improvement derives from the gas transport resistance mainly because one electrode is thicker than the other that is optimal. Hence, as demonstrated in Figure 13b, as porosity decreases, the curves of the slope are enhanced.

According to Figure 13, when the porosity reduces, the minimum value of the cathodic over potential, decreases. Alternatively, for electrodes thicker than the optimal value, the reduction in porosity may cause a sharp rise in the over potential (Figure 13b). These issues clarify how the optimal thickness decreases in general as the porosity and the particle size reduce (highlighted by the dotted lines in Figure 13).
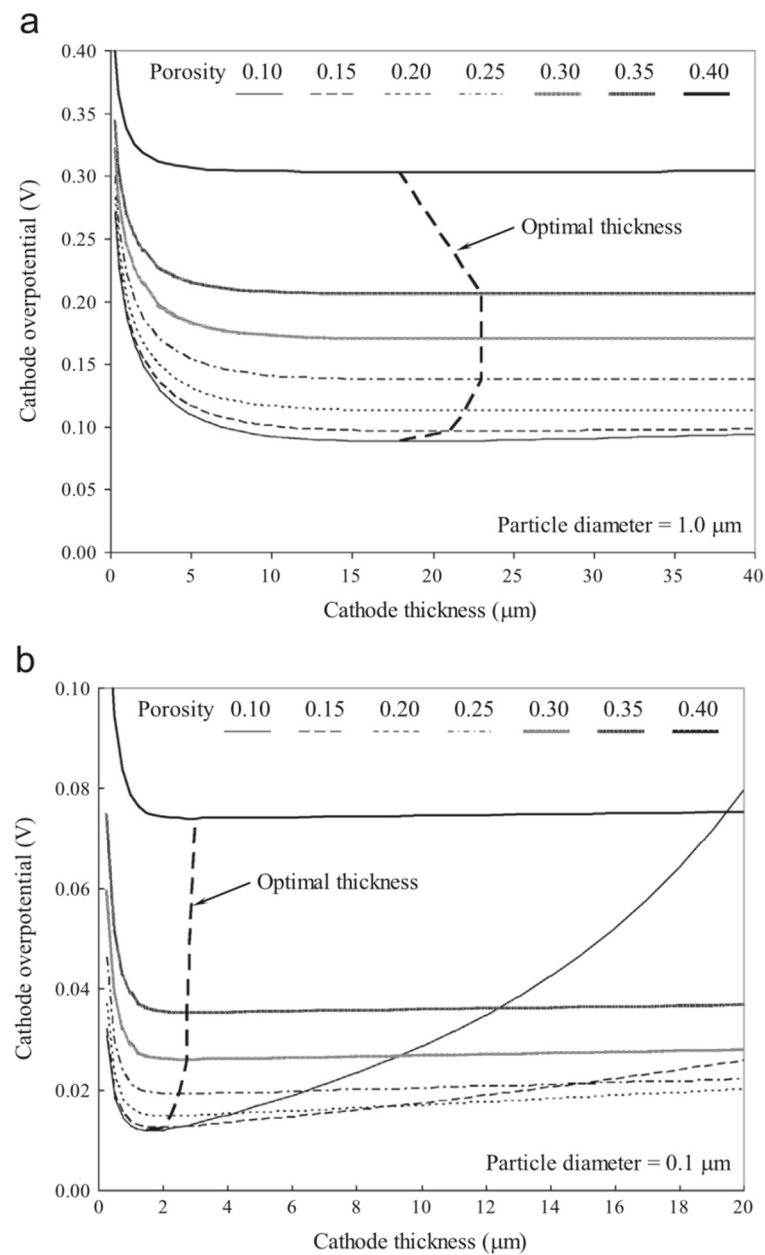
**Figure 13.** Cathodic overpotential as a function of thickness for different porosities: (**a**) particle size equal to 1.0 μm, (**b**) particle size equal to 0.1 μm. In both the simulations, cathodes with equal volume fractions of LSM and YSZ were considered, while the temperature was set to 800 °C (figure reprinted from Bertei et al. [126] Copyright (2013), with permission from Elsevier).

Although considerable progress has been made on the development of simulation and modeling of micro-structure materials and components in recent years, further research is still needed to design and fabricate better fuel cell electrodes and engineer optimum electrode microstructure.

### 4.4. Effect of Fuel Cell Temperature on the Cell Performance

Studying the effect of temperature on fuel cell performance is a key parameter in order to minimize the overall losses, alleviate the thermal stress and improve the SOFC efficiency. Most SOFCs are designed to function at high temperatures, thereby generating higher cell efficiencies in comparison to other kinds of fuel cells [65,127]. The cell current density can be raised by increasing the operating temperature; however, a too high temperature or temperature gradient decreases the cell life time. Therefore, the major challenge is to

optimize the cell durability and to ensure the performance stability of the fuel cell during its operation.

Considering the complexity of the actual energy transport process in SOFCs, simulation is usually a desirable tool to analyze the heat transfer phenomena inside an SOFC and predict its performance. Knowing the temperature distribution within an SOFC is important as it plays a significant role in its operation. Simulation studies focused on thermal distribution and thermal stress were conducted by many investigators to increase the thermal durability and extend the fuel cell life time as well as to determine an optimized operating condition [71,76–80]. Numerical CFD models gave the temperature distribution over the surface of individual cells [89,128], or along a single air channel [103,129]. Since SOFCs operate at high temperatures and the cell scale is tiny, experimental studies on the thermal stresses is difficult. Most advanced research investigations on thermal stresses of SOFCs were performed by establishing computer models and numerical simulations. The issue of designing a reliable SOFC due to high temperature operation was previously detailed in the literature [2,34,35].

There are different methods in the literature to decrease the temperature gradient in fuel cells. Analysis shows that increasing the velocity of the hot air stream, and lowering the Peclet number (by increasing the effective thermal diffusivity of the cell) leads to an optimal design, which minimizes heating time under the constraint of maximum allowable temperature gradients [130]. Significant temperature gradients exist both along the length of the cell and within its cross-section, specifically at the inlet. The high temperatures in SOFCs, further topped with hydrocarbon fuel, generates double reactions, which are: endothermic methane steam reforming reaction and exothermic electrochemical reaction, which happen to occur at various spots in the structure of the PEN [131]. Such huge temperatures, nevertheless, could be brought down with certain percentage recycle of anodic exhaust gases, thus altering the feed composition [127]. Furthermore, the high $H_2$ concentration escalates the reaction of electrochemical, hence leading to a drop in the temperature at the cell inlet [62]. In fact, the high $H_2$, $H_2O$ and CO concentrations could lessen deposition of carbon [127]. Moreover, the high temperatures are usually found in the electrolyte, which is away from channels for flow and unaffected by configuration of SOFC, regardless of electrolyte or anode support [85]. Furthermore, based on estimation of models, temperature for SOFC with electrolyte support is sensitive to over potential heat source, especially at when the gaseous temperature is lower. Nevertheless, an increment in gaseous temperature generates a lower temperature for SOFC with anode support when compared to that with electrolyte support, although functional with similar setting [95].

However, with the decrease in flow channel height, the temperature reduces as well, thus enhancing efficiency of cell mainly because of increased heat/mass transfer coefficient between flow stream and wall of channel, as well as shortened current path [132]. Nevertheless, the bigger cell for fuel that suggests gradual rise in temperature poses risk in relation to thermal stress, which is severe, although it ascertains cell elements' structural integrity. With that, numerous studies have attempted in decreasing the SOFC functional temperature to below 650 °C, especially to avoid degradation of materials, to enhance the lifetime of the layers, as well as to reduce the costs incurred by using common materials that are metallic [17,133–137]. Besides, in order to hinder corrosion that might be severe, especially at higher temperatures, the functional temperature has to decrease, mainly due to the following:

1. Increment in efficiency of thermodynamic conversion can be noted as the temperate for reformed gas (blend of H and CO) is decreased [83].
2. The process of sealing gets intricate as the temperature is lowered [15].
3. Loss of radiation heat is insignificant for a minute system when temperate is decreased. Therefore, management of heat is eased [20].

In fact, another method that could enhance the performance of fuel cell at decreased functional temperature is to lessen the electrolyte thickness [138]. Besides, a fuel cell that

is small could lessen loss of concentration and enhanced performance exerted by the fuel cell [139].

In addition, the variability of SOFC performance depends on an improbable mechanical aspect: the ceramic thin layers are fragile even at moderate stress, which occurs due to several reasons: issues related to fabrication issues or stresses of residual, the variance in TEC (thermal expansion coefficients) layered cell layers, spatial or temporal temperature gradients, oxygen activity gradients, as well as loading of external mechanical. Furthermore, these stresses are affected by the properties of the material, setting of the performance conditions, as well as geometrical design [140]. These residual stresses emerge due to a discrepancy between thermal expansion coefficients and Young's modulus at the layers that adjoin. In fact, such stresses could lead to layers' delamination or cracks in micro-size at the weak layers.

The thermal stresses in the electrolyte are bigger than those of other components for constant temperature distributions. As the temperature rises, the maximum principal stresses in the electrolyte are also enhanced. This occurs as a consequence of the TEC [141].

Based on theory, voltage with reduced functionality offers increased current density and improved density of power, thus hiking both the field and gradient of the temperature. Besides, Table 4 [142] portrays the improvement of thermal stress with decreasing functional voltage meant for anodic porosities at $\varepsilon = 0.2$ and 0.3, respectively (see also Figure 14). Meanwhile, Table 4 presents higher porosity at $\varepsilon = 0.5$, but achieved thermal stress at the maximum level with intermediate voltage at 0.8 V (Figure 14). Other than that, porosity at $\varepsilon = 0.5$ demonstrated an increment in temperature gradient with 0.8 V voltage, hence generating more thermal stress for PEN. Additionally, the literature [143] dictates that reduced elastic modulus for NiO-YSZ anode, in relation to its Ni-YSZ counterpart, has been due to improved anodic porosity in a direct manner. As such, the decrease of stress due to the hike in anodic porosity (as illustrated in Figure 14) could be a result from the low numbers derived from elastic modulus, shear modulus and Poisson's ratio.

**Table 4.** Maximum principal stresses and temperatures at different cell voltages and porosities [129].

| Prosity | Voltage (V) | PEN (MPa) | Gas Distributors (MPa) | PEN $T_{max}$ (°C) | PEN $\Delta T_{max}$ (°C) |
|---------|-------------|-----------|------------------------|--------------------|---------------------------|
| 0.2 | 0.7 | 20.7 | 42.9 | 821.2 | 12.9 |
| 0.2 | 0.8 | 19.6 | 40.5 | 809.9 | 12.8 |
| 0.2 | 0.9 | 18.1 | 37.3 | 808.3 | 12.5 |
| 0.3 | 0.7 | 18.7 | 42.9 | 823.0 | 12.8 |
| 0.3 | 0.8 | 18.4 | 42.1 | 811.3 | 12.6 |
| 0.3 | 0.9 | 16.1 | 36.8 | 806.8 | 11.9 |
| 0.5 | 0.7 | 15.1 | 42.7 | 824.9 | 12 |
| 0.5 | 0.8 | 15.3 | 43.5 | 812.7 | 13 |
| 0.5 | 0.9 | 13.1 | 36.2 | 805.9 | 11.5 |

A critical evaluation of the available bibliography displays that thorough study of the simulation of thermal stresses and the destructive effects of cyclic stresses on the SOFC is still in its early stage of research [144,145]. In most of such reports, the mechanical and thermal physiognomies of the SOFC are taken as constants. According to Nakajo and co-workers [146], however, these characteristics depend mainly on the temperature and the porosity. Furthermore, the temperature changes on the fuel cell surface, so the physiognomies of the materials cannot be considered constant.
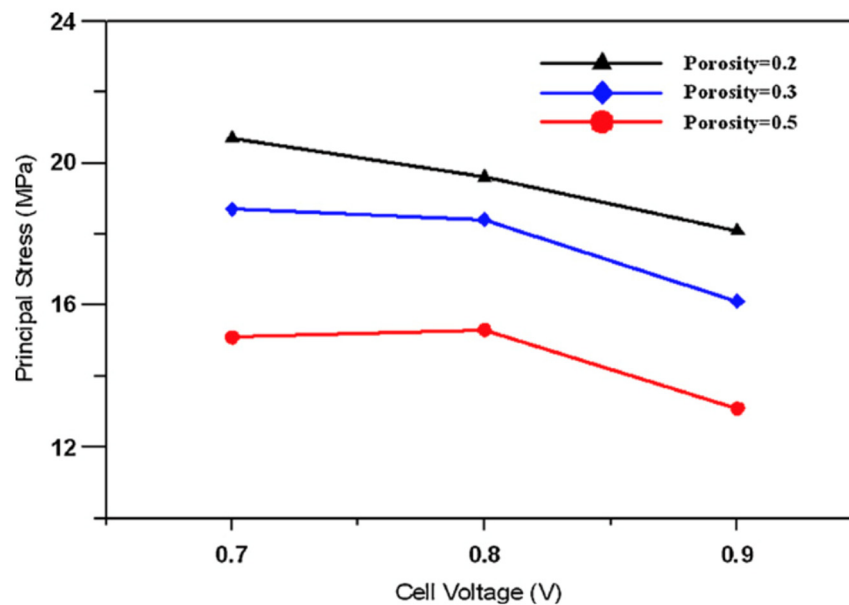
**Figure 14.** Variation of maximum principal stress in PEN with cell voltage for anode porosities of 0.2, 0.3 and 0.5, respectively (figure reprinted from Chiang et al. [142] Copyright (2010), with permission from Elsevier) [143].

## 5. Optimization of SOFC Operational Parameters

Few papers have studied the models of SOFC for the purposes of performance optimization [23,24,90,147], including the steady- and dynamic-states. In order to truly understand the online control strategy for SOFC systems, the optimization of fuel cell operating parameters plays a critical role in improving its performance.

There are models and optimization methods developed by researchers in order to achieve the performance prediction and parameters optimization. The objective of this section is to summarize the present status of such optimization efforts so that unresolved problems can be identified.

Based on the literature, SOFC optimization is classified into: (1) micro-structural, (2) single cell and (3) SOFC hybrid system optimization. In micro-structural optimization, grain size, mean pore diameter, thickness of the cathode/electrolyte/anode functional layer and several other factors are typically considered for optimization [148,149]. For single SOFC optimization strategies, the fuel cell geometry, the design parameters, the cell configuration and the operating conditions were considered [104]. In the single SOFC optimization study performed by Ji et al. [104], a co-flow planar SOFC was studied to figure out the optimum height of the cell. For SOFC hybrid system optimization, different configurations, along with combinations with different equipment and optimum operating conditions were explored [150–152]. A topology of an optimization method was performed to identify characteristics of optimal microstructures of a nano-composite SOFC cathode produced by infiltration of a mixed ionic electron conductor phase into a porous ionic conducting scaffold [153].

In order to optimize fuel cell systems to work in an efficient manner with the least cost and highest power generation, different optimization methods like genetic algorithm [154], clustering Pareto evolutionary algorithm [155], design of experiments [156] and also different methods of multi-objective optimization (MOO) have been investigated. MOO techniques provide the necessary information for a detailed examination of the design and operational trade-offs between conflicting objectives. However, only a few reports address the issues of MOO of SOFC stacks.

A multi-objective evolutionary algorithm (Pareto envelope-based selection algorithm) was used for a combination of photovoltaic panels, batteries and an SOFC using different fuels fed to the latter to determine enough auxiliary power for this hybrid system economi-

cally and ecologically [157]. The most desirable hybrid system was the one that had the least cost, the least emission and the most reliability. However, these objectives were in conflict with each other. It was reported that the most appropriate fuel for the SOFC was the hybrid system involving natural gas due to low annualized cost as well as low $CO_2$ emissions.

Quddus et al. [158] presented multi-objective trade-offs in SOFCs for oxidative coupling of methane. Palazzi et al. [159] performed MOO on an SOFC system. The proposed objectives were efficiency and capital cost per unit power ($/kW). Chakraborty [154] used a genetic programming approach in an SOFC system. Yang et al. [160] presented an improved genetic algorithm to optimize the parameters in the simple SOFC electrochemical model. The results of the optimization parameters showed that the improved genetic algorithm had more precision and stability compared to its simple genetic counterpart.

Nonetheless, in order to enhance one's comprehension pertaining to the performance exerted by the SOFC system, the influence of several significant parameters, as well as the correlations between these parameters in association to the performance of the general system, were looked into by employing empirical optimization [156]. Moreover, the enhanced iterative particle swarm optimization (PSO) algorithm was calculated to determine the SOFC-micro gas turbine hybrid system, especially to seek optimization among the functional parameters with varied loads [161]. In fact, this particular approach is amalgamated with the technique of iteration and the PSO algorithm, thus generating discrete PSO in an iterative manner right until convergence takes place from the control profile to one that is optimal. Other than that, the findings obtained from the MATLAB simulation display that the hybrid model of SOFC/MGT, together with parameters that were optimized, could effectively and efficiently determine power output. In short, the enhanced iterative PSO algorithm could prove to be beneficial when applied for analysis of systems, optimization of designs, as well as real-time control for a hybrid system, in particular, the SOFC/MGT system [161].

A useful method involving thermodynamic optimization has been performed by Bejan [162]. The purpose of thermodynamic optimization is to minimize losses, thereby maximizing power output and efficiency. The two thermodynamic optimizations developed by Bejan are entropy generation minimization and constructal theory. Sciacovelli [163] modeled a single channel of a monolithic cell using CFD (Fluent$^{TM}$) and simulated a stack that also considered a thermal model. The total entropy decreased by 50% using the EGM method which resulted in a power density increase of 10%.

Improvement of fuel cell design results in enhancement of power density. Reducing the anodic thickness first results in an increase of power density due to a decrease of ohmic and gas-phase transport losses. By reducing anodic thickness further, the layer becomes too thin to support the electrochemical conversion. The thickness of the cathode has almost no influence on the system's performance because the cathodic contribution to the power loss of the system is not important. In this regard, Vogler et al. [164] identified an optimum anodic thickness to improve the fuel cell power.

Maximization of power at minimum emission of greenhouse gases is an issue. Therefore, a multi-objective optimization study using non-dominated sorting genetic algorithm is performed to determine the optimal operating conditions and design configurations of the SOFC unit. In this regard, the SOFC power plant is optimized in terms of maximizing the $C_2$ production and minimizing the generation of undesired products such as $CO_x$ or $FCO_x$ [158]. A summary of the objectives for fuel cell optimization is listed in Table 5.

Large-scale power generation benefits from the high efficiency of gas-steam combined cycles. In the lower power range, fuel cells especially SOFCs operating at high temperatures are good candidates to combine with gas turbines to achieve efficiencies exceeding 60%. A configuration of hybrid system technology has been identified using a thermo-economic multi-objective optimization approach [165]. The system is integrated using pinch based methods. Then, a thermo-economic approach is used to compute the integrated system performances, following which, two objectives are optimized: (1) minimization of the

specific cost, and (2) maximization of the efficiency. The optimization results in a design with costs ranging from $2400/kW for a 44% efficient system to $6700/kW for a 70% efficient operation [165]. The PSO method is used for a hybrid power system to determine the optimum power sharing means to enhance the total efficiency of the system [166].

**Table 5.** A summary of objectives for fuel cell optimization.

| Objective | Constraint | Optimization Method | Refs. |
|---|---|---|---|
| Maximize performance of power production from waste biomass | Temperature | Clustering Pareto evolutionary algorithm | [155] |
| Maximize the system efficiency | Operating pressure | Design of experiments | [156] |
| Maximization of system efficiency and the minimization of specific investment cost | Fuel processing temperature Steam to carbon ratio Fuel utilization Oxygen to carbon ratio Air excess ratio | Based on the use of an evolutionary Algorithm: MOO (Multi-Objective Optimizer) | [159] |
| Parameter optimization: Maximize the output voltage and current density to find cathodic shapes that minimize resistance at the base of the cell | Pressure, temperature | Improved genetic algorithm Finite element modeling for topology optimization | [160] |
| Reducing the annual total cost of power, Reducing the heating and cooling generation and the annual $CO_2$ emissions rate | The amount of fuel provided to the SOFC; The operating temperature of the SOFC; the operating pressure of the SOFC | Evolutionary Multi-objective optimizer | [161] |

## 6. Summary and Recommendation

In recent years, considerable progress has been made in SOFC simulation to improve the design and performance of this technology with the aim of reducing its cost. In this work, different designs and configurations of SOFCs (an essential step in simulation to represent the real system as closely as possible) were reviewed. Despite that, such SOFC simulations require much attention to enhance the performance of the system.

The main conclusions of this review are as follows:

- SOFC materials and micro-structures were reviewed due to their significant impact on the simulation reliability as well as system optimization. A literature review has also been performed on simulation studies focused on the effects of microstructure and materials on fuel cell performance. However, most of the studies use only common SOFC materials in their modelling. More simulation studies are required to compare alternative materials in overall modelling to assess their effect on the efficiency and degradation of the cell. Electrochemical modelling that studies the cell voltage behaviour as a function of the microstructure, geometry and material properties is still under development.
- In the literature, different techniques of SOFC simulations are available. In cell and stack level, 0-D, 1-D, 2-D and 3-D techniques may be chosen according to the purpose of the model. Transient modelling is used if any of the heat-up, start-up, shut-down and load changes or a combination of them requires simulation.
- A critical analysis of available literature proved that detailed research on the simulation of thermal stress and damaging impact upon the SOFC is still in its early stage of development. In most of the presented works, the mechanical and thermal characteristics of the SOFC are represented comprehensively [140,144,146]. There is a lack of simulation reports, however, on the radiative heat transfer between the materials used in SOFCs at the high temperatures. In addition, the literature lacks sufficient

information on the energy and exergy analyses of SOFC systems for performance evaluation.

- In order to understand online control and optimization strategies accurately, an effective simulation of the fuel cell system plays a critical role. Therefore, sensitivity analysis of fuel cell parameters using simulation models was reviewed herein. In addition, micro-machined SOFCs are used for portable devices (many electronics and wireless). Hence, optimization and reliability of such devices require further research.
- Combination of an MPS, a deep neural network and a genetic algorithm provides a promising solution for model-based control systems for precisely and quickly analyzing the performances of SOFCs and other non-linear systems. The deep learning algorithm needs to be trained based on experimental data for a wider application under dynamic operating conditions.

## References

1. Liu, M.; Lanzini, A.; Halliop, W.; Cobas, V.; Verkooijen, A.; Aravind, P. Anode recirculation behavior of a solid oxide fuel cell system: A safety analysis and a performance optimization. *Int. J. Hydrogen Energy* **2013**, *38*, 2868–2883. [CrossRef]
2. Aravind, P.; de Jong, W. Evaluation of high temperature gas cleaning options for biomass gasification product gas for solid oxide fuel cells. *Prog. Energy Combust. Sci.* **2012**, *38*, 737–764. [CrossRef]
3. Baldinelli, A.; Barelli, L.; Bidini, G.; Di Michele, A.; Vivani, R. SOFC direct fuelling with high-methane gases: Optimal strategies for fuel dilution and upgrade to avoid quick degradation. *Energy Convers. Manag.* **2016**, *124*, 492–503. [CrossRef]
4. Antonucci, V.; Branchini, L.; Brunaccini, G.; De Pascale, A.; Ferraro, M.; Melino, F.; Orlandini, V.; Sergi, F. Thermal integration of a SOFC power generator and a Na–NiCl$_2$ battery for CHP domestic application. *Appl. Energy* **2017**, *185*, 1256–1267. [CrossRef]
5. Mehr, A.S.; Gandiglio, M.; MosayebNezhad, M.; Lanzini, A.; Mahmoudi, S.M.S.; Yari, M.; Santarelli, M. Solar-assisted integrated biogas solid oxide fuel cell (SOFC) installation in wastewater treatment plant: Energy and economic analysis. *Appl. Energy* **2017**, *191*, 620–638. [CrossRef]
6. Hajimolana, S.A.; Hussain, M.A.; Daud, W.M.A.W.; Soroush, M.; Shamiri, A. Mathematical modeling of solid oxide fuel cells: A review. *Renew. Sustain. Energy Rev.* **2011**, *15*, 1893–1917. [CrossRef]
7. Huang, K.; Singhal, S.C. Cathode-supported tubular solid oxide fuel cell technology: A critical review. *J. Power Sources* **2013**, *237*, 84–97. [CrossRef]
8. Bieberle-Hütter, A.; Beckel, D.; Infortuna, A.; Muecke, U.P.; Rupp, J.L.; Gauckler, L.J.; Rey-Mermet, S.; Muralt, P.; Bieri, N.R.; Hotz, N. A micro-solid oxide fuel cell system as battery replacement. *J. Power Sources* **2008**, *177*, 123–130. [CrossRef]
9. Delphi, Fuel Cells Technology. Available online: http://www.delphi.com/manufacturers/auto/fuelcells (accessed on 11 December 2014).
10. Topsoe Fuel Cell/Products/Fuel Cell. Available online: http://www.topsoefuelcell.com/sitecore/content/Haldor%20Topsoe/Press/News/2007/20070514%20Topsoe%20Fuel%20Cell%20wins%20Funding%20for%20SOFC%20Fuel%20Cell%20Development.aspx/ (accessed on 11 December 2014).
11. Rolls-Royce, Power Systems Company. Available online: http://www.rollsroyce.com/energy/energy_products/fuel_cells/ (accessed on 11 December 2014).
12. Acumentrics: Trusted Power Innovations. Available online: http://www.acumentrics.com/fuel-cell-generators.html (accessed on 11 December 2014).
13. Shi, J.; Xue, X. CFD analysis of a novel symmetrical planar SOFC design with micro-flow channels. *Chem. Eng. J.* **2010**, *163*, 119–125. [CrossRef]
14. Sammes, N.; Du, Y.; Bove, R. Design and fabrication of a 100 W anode supported micro-tubular SOFC stack. *J. Power Sources* **2005**, *145*, 428–434. [CrossRef]
15. Yu, X.; Long, W.; Jin, F.; He, T. Cobalt-free perovskite cathode materials SrFe$_{1-x}$Ti$_x$O$_{3-\delta}$ and performance optimization for intermediate-temperature solid oxide fuel cells. *Electrochim. Acta* **2014**, *123*, 426–434. [CrossRef]

16. Wachsman, E.D.; Marlowe, C.A.; Lee, K.T. Role of solid oxide fuel cells in a balanced energy strategy. *Energy Environ. Sci.* **2012**, *5*, 5498–5509. [CrossRef]

17. Pfeifer, T.; Nousch, L.; Lieftink, D.; Modena, S. System design and process layout for a SOFC micro-CHP unit with reduced operating temperatures. *Int. J. Hydrogen Energy* **2013**, *38*, 431–439. [CrossRef]

18. Ding, J.; Liu, J.; Yin, G. Fabrication and characterization of low-temperature SOFC stack based on GDC electrolyte membrane. *J. Membr. Sci.* **2011**, *371*, 219–225. [CrossRef]

19. Torknik, F.S.; Keyanpour-Rad, M.; Maghsoudipour, A.; Choi, G.M. Effect of microstructure refinement on performance of Ni/Ce$_{0.8}$Gd$_{0.2}$O$_{1.9}$ anodes for low temperature solid oxide fuel cell. *Ceram. Int.* **2014**, *40*, 1341–1350. [CrossRef]

20. Brett, D.J.; Atkinson, A.; Brandon, N.P.; Skinner, S.J. Intermediate temperature solid oxide fuel cells. *Chem. Soc. Rev.* **2008**, *37*, 1568–1578. [CrossRef]

21. Grew, K.N.; Chiu, W.K. A review of modeling and simulation techniques across the length scales for the solid oxide fuel cell. *J. Power Sources* **2012**, *199*, 1–13. [CrossRef]

22. Qu, Z.; Aravind, P.; Dekker, N.; Janssen, A.; Woudstra, N.; Verkooijen, A. Three-dimensional thermo-fluid and electrochemical modeling of anode-supported planar solid oxide fuel cell. *J. Power Sources* **2010**, *195*, 7787–7795. [CrossRef]

23. Sanandaji, B.M.; Vincent, T.L.; Colclasure, A.M.; Kee, R.J. Modeling and control of tubular solid-oxide fuel cell systems: II. Nonlinear model reduction and model predictive control. *J. Power Sources* **2011**, *196*, 208–217. [CrossRef]

24. Bavarian, M.; Soroush, M.; Kevrekidis, I.G.; Benziger, J.B. Mathematical modeling, steady-state and dynamic behavior, and control of fuel cells: A review. *Ind. Eng. Chem. Res.* **2010**, *49*, 7922–7950. [CrossRef]

25. Hajimolana, S.; Hussain, M.; Soroush, M.; Wan Daud, W.A.; Chakrabarti, M.H. Multilinear-model predictive control of a tubular solid oxide fuel cell system. *Ind. Eng. Chem. Res.* **2012**, *52*, 430–441. [CrossRef]

26. Ramadhani, F.; Hussain, M.A.; Mokhlis, H.; Hajimolana, S. Optimization strategies for Solid Oxide Fuel Cell (SOFC) application: A literature survey. *Renew. Sustain. Energy Rev.* **2017**, *76*, 460–484. [CrossRef]

27. Ebrahimi, M.; Moradpoor, I. Combined solid oxide fuel cell, micro-gas turbine and organic Rankine cycle for power generation (SOFC–MGT–ORC). *Energy Convers. Manag.* **2016**, *116*, 120–133. [CrossRef]

28. D'Andrea, G.; Gandiglio, M.; Lanzini, A.; Santarelli, M. Dynamic model with experimental validation of a biogas-fed SOFC plant. *Energy Convers. Manag.* **2017**, *135*, 21–34. [CrossRef]

29. Anyenya, G.A.; Sullivan, N.P.; Braun, R.J. Modeling and simulation of a novel 4.5 kWe multi-stack solid-oxide fuel cell prototype assembly for combined heat and power. *Energy Convers. Manag.* **2017**, *140*, 247–259. [CrossRef]

30. Kang, S.; Ahn, K.-Y. Dynamic modeling of solid oxide fuel cell and engine hybrid system for distributed power generation. *Appl. Energy* **2017**, *195*, 1086–1099. [CrossRef]

31. Ho, T.X.; Kosinski, P.; Hoffmann, A.C.; Vik, A. Modeling of transport, chemical and electrochemical phenomena in a cathode-supported SOFC. *Chem. Eng. Sci.* **2009**, *64*, 3000–3009. [CrossRef]

32. Taher, M.A.; Adjiman, C.; Iora, P.; Chiesa, P.; Brandon, N. Model-Based Evaluation of the Production of Pure Oxygen through SOFC/SOEC Integration. *ECS Trans.* **2011**, *35*, 2997–3006. [CrossRef]

33. Luebbe, H.; Hofmann, H.; Bowen, P.; Aschauer, U.; Schuler, A.; Snijkers, F.; Schindler, H.-J.; Vogt, U.; Lalanne, C. Cathode-supported micro-tubular SOFCs based on Nd$_{1.95}$NiO$_{4+\delta}$: Fabrication and characterisation of dip-coated electrolyte layers. *Solid State Ion.* **2009**, *180*, 805–811. [CrossRef]

34. Singhal, S.C. Advances in solid oxide fuel cell technology. *Solid State Ion.* **2000**, *135*, 305–313. [CrossRef]

35. Singhal, S.C. Solid oxide fuel cells for power generation. *Wiley Interdiscip. Rev. Energy Environ.* **2014**, *3*, 179–194. [CrossRef]

36. Fuji, H. Status of National Project for SOFC Development in Japan. In Proceedings of the Solid Oxide Fuel Cells Meeting, Palm Springs, CA, USA, 18 November 2002.

37. Fardadi, M.; McLarty, D.F.; Jabbari, F. Investigation of thermal control for different SOFC flow geometries. *Appl. Energy* **2016**, *178*, 43–55. [CrossRef]

38. Li, A.; Song, C.; Lin, Z. A multiphysics fully coupled modeling tool for the design and operation analysis of planar solid oxide fuel cell stacks. *Appl. Energy* **2017**, *190*, 1234–1244. [CrossRef]

39. Fujii, H.; Ninomiya, T. Status of National Project for SOFC Development in Japan. In Proceedings of the European Solid Oxide Fuel Cell Forum, Lucerne, Switzerland, 1–5 July 2002; Volume 2, pp. 700–707.

40. Williams, M.C.; Strakey, J.P.; Singhal, S.C. US distributed generation fuel cell program. *J. Power Sources* **2004**, *131*, 79–85. [CrossRef]

41. Godfrey, B.; Föger, K.; Gillespie, R.; Bolden, R.; Badwal, S. Planar solid oxide fuel cells: The Australian experience and outlook. *J. Power Sources* **2000**, *86*, 68–73. [CrossRef]

42. Hwang, J.; Chen, C.O.-K.; Lai, D. Computational analysis of species transport and electrochemical characteristics of a MOLB-type SOFC. *J. Power Sources* **2005**, *140*, 235–242. [CrossRef]

43. Suzuki, M. Residential Energy System Development, 1–3. Available online: http://www.gas.or.jp/en/newsletter/images/22/dl/pdf01.pdf (accessed on 11 December 2014).

44. Murthy, S.; Fedorov, A.G. Radiation heat transfer analysis of the monolith type solid oxide fuel cell. *J. Power Sources* **2003**, *124*, 453–458. [CrossRef]

45. Magistri, L.; Bozzolo, M.; Tarnowski, O.; Agnew, G.; Massardo, A.F. Design and Off-Design Analysis of a MW Hybrid System Based on Rolls-Royce Integrated Planar Solid Oxide Fuel Cells. *J. Eng. Gas Turbines Power* **2007**, *129*, 792–797. [CrossRef]

46. Cable, T.L.; Sofie, S.W. A symmetrical, planar SOFC design for NASA's high specific power density requirements. *J. Power Sources* **2007**, *174*, 221–227. [CrossRef]

47. Gardner, F.J.; Day, M.; Brandon, N.; Pashley, M.; Cassidy, M. SOFC technology development at Rolls-Royce. *J. Power Sources* **2000**, *86*, 122–129. [CrossRef]

48. Ceres Power, Fuel Cell Technology. Available online: http://www3.imperial.ac.uk/engineering/ourresearchhome/impactresearch/cerespower (accessed on 11 December 2014).

49. Imperial Collage, Hallow Fibre Solid Oxide Fuel Cell. Available online: http://www3.imperial.ac.uk/electrochemicalengineering/research/currentprojects/tubularsolidoxidefuelcells (accessed on 11 December 2014).

50. Nikbin, D. Micro SOFCs: Why small is beautiful. *Fuel Cell Rev.* **2006**, *3*, 21–24.

51. Bieberle-Hutter, A.; Beckel, D.; Muecke, U.; Rupp, J.; Infortuna, A.; Gauckler, L. Micro-solid oxide fuel cells as battery replacement. *Mst News* **2005**, *4*, 12. [CrossRef]

52. Huang, H.; Nakamura, M.; Su, P.; Fasching, R.; Saito, Y.; Prinz, F.B. High-performance ultrathin solid oxide fuel cells for low-temperature operation. *J. Electrochem. Soc.* **2007**, *154*, B20–B24. [CrossRef]

53. Baertsch, C.D.; Jensen, K.F.; Hertz, J.L.; Tuller, H.L.; Vengallatore, S.T.; Spearing, S.M.; Schmidt, M.A. Fabrication and structural characterization of self-supporting electrolyte membranes for a micro solid-oxide fuel cell. *J. Mater. Res.* **2004**, *19*, 2604–2615. [CrossRef]

54. Shao, Z.; Haile, S.M.; Ahn, J.; Ronney, P.D.; Zhan, Z.; Barnett, S.A. A thermally self-sustained micro solid-oxide fuel-cell stack with high power density. *Nature* **2005**, *435*, 795–798. [CrossRef] [PubMed]

55. Hanna, J.; Lee, W.Y.; Shi, Y.; Ghoniem, A. Fundamentals of electro-and thermochemistry in the anode of solid-oxide fuel cells with hydrocarbon and syngas fuels. *Prog. Energy Combust. Sci.* **2014**, *40*, 74–111. [CrossRef]

56. Qi, Y.; Huang, B.; Chuang, K.T. Dynamic modeling of solid oxide fuel cell: The effect of diffusion and inherent impedance. *J. Power Sources* **2005**, *150*, 32–47. [CrossRef]

57. Sarkar, P.; Yamarte, L.; Johanson, L. Tubular micro-solid oxide fuel cell for remote power applications. *ECS Trans.* **2007**, *7*, 603–608. [CrossRef]

58. Lockett, M.; Simmons, M.; Kendall, K. CFD to predict temperature profile for scale up of micro-tubular SOFC stacks. *J. Power Sources* **2004**, *131*, 243–246. [CrossRef]

59. Akhtar, N.; Decent, S.; Kendall, K. Numerical modelling of methane-powered micro-tubular, single-chamber solid oxide fuel cell. *J. Power Sources* **2010**, *195*, 7796–7807. [CrossRef]

60. Alyousef, Y.; Kendall, K. Characterization of the electrochemical performance of micro-tubular solid oxide fuel cell (SOFC). *J. Taibah Univ. Sci.* **2009**, *2*, 14–21. [CrossRef]

61. Cui, D.; Liu, L.; Dong, Y.; Cheng, M. Comparison of different current collecting modes of anode supported micro-tubular SOFC through mathematical modeling. *J. Power Sources* **2007**, *174*, 246–254. [CrossRef]

62. Hajimolana, S.A.; Soroush, M. Dynamics and control of a tubular solid-oxide fuel cell. *Ind. Eng. Chem. Res.* **2009**, *48*, 6112–6125. [CrossRef]

63. Akhtar, N. Micro-tubular, single-chamber solid oxide fuel cell (MT–SC-SOFC) stacks: Model development. *Chem. Eng. Res. Des.* **2012**, *90*, 814–824. [CrossRef]

64. Panthi, D.; Tsutsumi, A. Development of a high-performance micro-tubular SOFC based on a hollow fiber support. *ECS Trans.* **2013**, *57*, 789–798. [CrossRef]

65. Serincan, M.F.; Pasaogullari, U.; Sammes, N.M. Effects of operating conditions on the performance of a micro-tubular solid oxide fuel cell (SOFC). *J. Power Sources* **2009**, *192*, 414–422. [CrossRef]

66. Cordiner, S.; Mariani, A.; Mulone, V. An integrated CFD-approch to deisgn micro-tubular Solid Oxide Fuel Cells. In Proceedings of the 2008 Second International Conference on Thermal Issues in Emerging Technologies, Cairo, Egypt, 17–20 December 2008; pp. 85–96.

67. Cordiner, S.; Mariani, A.; Mulone, V. CFD-based design of microtubular solid oxide fuel cells. *J. Heat Transf.* **2010**, *132*, 062801. [CrossRef]

68. Razbani, O.; Wærnhus, I.; Assadi, M. Experimental investigation of temperature distribution over a planar solid oxide fuel cell. *Appl. Energy* **2013**, *105*, 155–160. [CrossRef]

69. Santangelo, P.E.; Tartarini, P. Fuel cell systems and traditional technologies. Part I: Experimental CHP approach. *Appl. Therm. Eng.* **2007**, *27*, 1278–1284. [CrossRef]

70. Santarelli, M.; Leone, P.; Calì, M.; Orsello, G. Experimental evaluation of the sensitivity to fuel utilization and air management on a 100 kW SOFC system. *J. Power Sources* **2007**, *171*, 155–168. [CrossRef]

71. Andersson, M.; Yuan, J.; Sundén, B. SOFC modeling considering electrochemical reactions at the active three phase boundaries. *Int. J. Heat Mass Transf.* **2012**, *55*, 773–788. [CrossRef]

72. Akbari, V.; Borhani, T.N.G.; Shamiri, A.; Aramesh, R.; Hussain, M.A.; Hamid, M.K.A. 2D CFD-PBM simulation of hydrodynamic and particle growth in an industrial gas phase fluidized bed polymerization reactor. *Chem. Eng. Res. Des.* **2015**, *104*, 53–67. [CrossRef]

73. Abbasi, M.R.; Shamiri, A.; Hussain, M.A. Dynamic modeling and Molecular Weight Distribution of ethylene copolymerization in an industrial gas-phase Fluidized-Bed Reactor. *Adv. Powder Technol.* **2016**, *27*, 1526–1538. [CrossRef]

74. Abbasi, M.R.; Shamiri, A.; Hussain, M.A. A review on modeling and control of olefin polymerization in fluidized-bed reactors. *Rev. Chem. Eng.* **2019**, *35*, 311–333. [CrossRef]

75. Akbari, V.; Borhani, T.N.G.; Shamiri, A.; Hamid, M.K.B. A CFD–PBM coupled model of hydrodynamics and mixing/segregation in an industrial gas-phase polymerization reactor. *Chem. Eng. Res. Des.* **2015**, *96*, 103–120. [CrossRef]

76. Al-Dabbagh, A.W.; Lu, L.; Mazza, A. Modelling, simulation and control of a proton exchange membrane fuel cell (PEMFC) power system. *Int. J. Hydrogen Energy* **2010**, *35*, 5061–5069. [CrossRef]

77. Arpino, F.; Massarotti, N. Numerical simulation of mass and energy transport phenomena in solid oxide fuel cells. *Energy* **2009**, *34*, 2033–2041. [CrossRef]

78. Arteaga-Perez, L.E.; Casas, Y.; Peralta, L.M.; Kafarov, V.; Dewulf, J.; Giunta, P. An auto-sustainable solid oxide fuel cell system fueled by bio-ethanol: Process simulation and heat exchanger network synthesis. *Chem. Eng. J.* **2009**, *150*, 242–251. [CrossRef]

79. Autissier, N.; Larrain, D.; Favrat, D. CFD simulation tool for solid oxide fuel cells. *J. Power Sources* **2004**, *131*, 313–319. [CrossRef]

80. Calise, F.; d'Accadia, M.D.; Palombo, A.; Vanoli, L. Simulation and exergy analysis of a hybrid solid oxide fuel cell (SOFC)–gas turbine system. *Energy* **2006**, *31*, 3278–3299. [CrossRef]

81. Colombo, K.W.E.; Kharton, V.V.; Berto, F.; Paltrinieri, N. Mathematical modeling and simulation of hydrogen-fueled solid oxide fuel cell system for micro-grid applications—Effect of failure and degradation on transient performance. *Energy* **2020**, *202*, 117752. [CrossRef]

82. Xu, H.; Ma, J.; Tan, P.; Chen, B.; Wu, Z.; Zhang, Y.; Wang, H.; Xuan, J.; Ni, M. Towards online optimisation of solid oxide fuel cell performance: Combining deep learning with multi-physics simulation. *Energy AI* **2020**, *1*, 100003. [CrossRef]

83. Qi, Y.; Huang, B.; Luo, J. Dynamic modeling of a finite volume of solid oxide fuel cell: The effect of transport dynamics. *Chem. Eng. Sci.* **2006**, *61*, 6057–6076. [CrossRef]

84. Karcz, M. From 0D to 1D modeling of tubular solid oxide fuel cell. *Energy Convers. Manag.* **2009**, *50*, 2307–2315. [CrossRef]

85. Ota, T.; Koyama, M.; Wen, C.-J.; Yamada, K.; Takahashi, H. Object-based modeling of SOFC system: Dynamic behavior of micro-tube SOFC. *J. Power Sources* **2003**, *118*, 430–439. [CrossRef]

86. Bove, R.; Ubertini, S. Modeling solid oxide fuel cell operation: Approaches, techniques and results. *J. Power Sources* **2006**, *159*, 543–559. [CrossRef]

87. Oh, S.-R.; Sun, J.; Dobbs, H.; King, J. Model-based predictive control strategy for a solid oxide fuel cell system integrated with a turbocharger. In Proceedings of the American Control Conference (ACC), Montreal, QC, Canada, 27–29 June 2012; pp. 6596–6601.

88. Achenbach, E. Three-dimensional and time-dependent simulation of a planar solid oxide fuel cell stack. *J. Power Sources* **1994**, *49*, 333–348. [CrossRef]

89. Recknagle, K.P.; Williford, R.E.; Chick, L.A.; Rector, D.R.; Khaleel, M.A. Three-dimensional thermo-fluid electrochemical modeling of planar SOFC stacks. *J. Power Sources* **2003**, *113*, 109–114. [CrossRef]

90. Kapadia, S.; Anderson, W. Sensitivity analysis for solid oxide fuel cells using a three-dimensional numerical model. *J. Power Sources* **2009**, *189*, 1074–1082. [CrossRef]

91. Horr, S.; Mohcene, H.; Bouguettaia, H.; Moussa, H.B. Performance analysis of AS-SOFC fuel cell combining single and sinusoidal flow field: Numerical study. *Renew. Energy Environ. Sustain.* **2021**, *6*, 18. [CrossRef]

92. Hajimolana, S.; Hussain, M.; Daud, W.W.; Chakrabarti, M. Dynamic modelling and sensitivity analysis of a tubular SOFC fuelled with NH 3 as a possible replacement for H 2. *Chem. Eng. Res. Des.* **2012**, *90*, 1871–1882. [CrossRef]

93. Kakac, S.; Pramuanjaroenkij, A.; Zhou, X.Y. A review of numerical modeling of solid oxide fuel cells. *Int. J. Hydrogen Energy* **2007**, *32*, 761–786. [CrossRef]

94. Jiang, W.; Fang, R.; Khan, J.A.; Dougal, R.A. Parameter setting and analysis of a dynamic tubular SOFC model. *J. Power Sources* **2006**, *162*, 316–326. [CrossRef]

95. Calise, F.; d'Accadia, M.D.; Vanoli, L.; von Spakovsky, M.R. Full load synthesis/design optimization of a hybrid SOFC–GT power plant. *Energy* **2007**, *32*, 446–458. [CrossRef]

96. Yakabe, H.; Hishinuma, M.; Uratani, M.; Matsuzaki, Y.; Yasuda, I. Evaluation and modeling of performance of anode-supported solid oxide fuel cell. *J. Power Sources* **2000**, *86*, 423–431. [CrossRef]

97. Weil, K.S.; Koeppel, B.J. Comparative finite element analysis of the stress–strain states in three different bonded solid oxide fuel cell seal designs. *J. Power Sources* **2008**, *180*, 343–353. [CrossRef]

98. Chiang, L.-K.; Liu, H.-C.; Shiu, Y.-H.; Lee, C.-H.; Lee, R.-Y. Thermo-electrochemical and thermal stress analysis for an anode-supported SOFC cell. *Renew. Energy* **2008**, *33*, 2580–2588. [CrossRef]

99. Dang, Z.; Shen, X.; Ma, J.; Jiang, Z.; Xi, G. Multiphysics Coupling Simulation and Parameter Study of Planar Solid Oxide Fuel Cell. *Front. Chem.* **2021**, *8*, 1–11. [CrossRef]

100. Wei, S.-S.; Wang, T.-H.; Wu, J.-S. Numerical modeling of interconnect flow channel design and thermal stress analysis of a planar anode-supported solid oxide fuel cell stack. *Energy* **2014**, *69*, 553–561. [CrossRef]

101. Lin, Y.; Beale, S. Performance predictions in solid oxide fuel cells. *Appl. Math. Model.* **2006**, *30*, 1485–1496. [CrossRef]

102. Bhattacharyya, D.; Rengaswamy, R. Dimensional optimization of a tubular solid oxide fuel cell. *Comput. Chem. Eng.* **2010**, *34*, 1789–1802. [CrossRef]

103. Laurencin, J.; Delette, G.; Lefebvre-Joud, F.; Dupeux, M. A numerical tool to estimate SOFC mechanical degradation: Case of the planar cell configuration. *J. Eur. Ceram. Soc.* **2008**, *28*, 1857–1869. [CrossRef]

104. Ji, Y.; Yuan, K.; Chung, J.; Chen, Y.-C. Effects of transport scale on heat/mass transfer and performance optimization for solid oxide fuel cells. *J. Power Sources* **2006**, *161*, 380–391. [CrossRef]

105. Lu, Y.; Schaefer, L. Numerical study of a flat-tube high power density solid oxide fuel cell: Part II: Cell performance and stack optimization. *J. Power Sources* **2006**, *153*, 68–75. [CrossRef]

106. Recknagle, K.P.; Ryan, E.M.; Koeppel, B.J.; Mahoney, L.A.; Khaleel, M.A. Modeling of electrochemistry and steam–methane reforming performance for simulating pressurized solid oxide fuel cell stacks. *J. Power Sources* **2010**, *195*, 6637–6644. [CrossRef]

107. Huang, C.; Shy, S.; Lee, C. On flow uniformity in various interconnects and its influence to cell performance of planar SOFC. *J. Power Sources* **2008**, *183*, 205–213. [CrossRef]

108. Huang, C.; Shy, S.; Li, H.; Lee, C. The impact of flow distributors on the performance of planar solid oxide fuel cell. *J. Power Sources* **2010**, *195*, 6280–6286. [CrossRef]

109. Yuan, P.; Liu, S.-F. Numerical analysis of temperature and current density distribution of a planar solid oxide fuel cell unit with nonuniform inlet flow. *Numer. Heat Transf. Part A Appl.* **2007**, *51*, 941–957. [CrossRef]

110. Christman, K.L.; Jensen, M.K. Solid Oxide Fuel Cell Performance With Cross-Flow Roughness. *J. Fuel Cell Sci. Technol.* **2010**, *8*, 024501–024505. [CrossRef]

111. Mehrpooya, M.; Akbarpour, S.; Vatani, A.; Rosen, M.A. Modeling and optimum design of hybrid solid oxide fuel cell-gas turbine power plants. *Int. J. Hydrogen Energy* **2014**, *39*, 21196–21214. [CrossRef]

112. Fardadi, M.; McLarty, D.F.; Brouwer, J.; Jabbari, F. Enhanced performance of counter flow SOFC with partial internal reformation. *Int. J. Hydrogen Energy* **2014**, *39*, 19753–19766. [CrossRef]

113. Costamagna, P.; Selimovic, A.; Del Borghi, M.; Agnew, G. Electrochemical model of the integrated planar solid oxide fuel cell (IP-SOFC). *Chem. Eng. J.* **2004**, *102*, 61–69. [CrossRef]

114. Cai, Q.; Adjiman, C.S.; Brandon, N.P. Modelling the 3D microstructure and performance of solid oxide fuel cell electrodes: Computational parameters. *Electrochim. Acta* **2011**, *56*, 5804–5814. [CrossRef]

115. Haanappel, V.; Mertens, J.; Rutenbeck, D.; Tropartz, C.; Herzhof, W.; Sebold, D.; Tietz, F. Optimisation of processing and microstructural parameters of LSM cathodes to improve the electrochemical performance of anode-supported SOFCs. *J. Power Sources* **2005**, *141*, 216–226. [CrossRef]

116. Patcharavorachot, Y.; Arpornwichanop, A.; Chuachuensuk, A. Electrochemical study of a planar solid oxide fuel cell: Role of support structures. *J. Power Sources* **2008**, *177*, 254–261. [CrossRef]

117. Andersson, M.; Yuan, J.; Sundén, B. SOFC modeling considering hydrogen and carbon monoxide as electrochemical reactants. *J. Power Sources* **2013**, *232*, 42–54. [CrossRef]

118. Ferguson, J.; Fiard, J.; Herbin, R. Three-dimensional numerical simulation for various geometries of solid oxide fuel cells. *J. Power Sources* **1996**, *58*, 109–122. [CrossRef]

119. Jia, J.; Abudula, A.; Wei, L.; Jiang, R.; Shen, S. A mathematical model of a tubular solid oxide fuel cell with specified combustion zone. *J. Power Sources* **2007**, *171*, 696–705. [CrossRef]

120. Hussain, M.; Li, X.; Dincer, I. Mathematical modeling of planar solid oxide fuel cells. *J. Power Sources* **2006**, *161*, 1012–1022. [CrossRef]

121. Tanaka, T.; Inui, Y.; Urata, A.; Kanno, T. Three dimensional analysis of planar solid oxide fuel cell stack considering radiation. *Energy Convers. Manag.* **2007**, *48*, 1491–1498. [CrossRef]

122. Chaisantikulwat, A.; Diaz-Goano, C.; Meadows, E. Dynamic modelling and control of planar anode-supported solid oxide fuel cell. *Comput. Chem. Eng.* **2008**, *32*, 2365–2381. [CrossRef]

123. Chen, X.; Wu, N.; Smith, L.; Ignatiev, A. Thin-film heterostructure solid oxide fuel cells. *Appl. Phys. Lett.* **2004**, *84*, 2700–2702. [CrossRef]

124. Gorman, B.P.; Anderson, H.U. Processing of composite thin film solid oxide fuel cell structures. *J. Am. Ceram. Soc.* **2005**, *88*, 1747–1753. [CrossRef]

125. Tang, Y.; Stanley, K.; Wu, J.; Ghosh, D.; Zhang, J. Design consideration of micro thin film solid-oxide fuel cells. *J. Micromech. Microeng.* **2005**, *15*, S185. [CrossRef]

126. Bertei, A.; Nucci, B.; Nicolella, C. Microstructural modeling for prediction of transport properties and electrochemical performance in SOFC composite electrodes. *Chem. Eng. Sci.* **2013**, *101*, 175–190. [CrossRef]

127. Nikooyeh, K.; Jeje, A.A.; Hill, J.M. 3D modeling of anode-supported planar SOFC with internal reforming of methane. *J. Power Sources* **2007**, *171*, 601–609. [CrossRef]

128. Haberman, B.; Young, J. Diffusion and chemical reaction in the porous structures of solid oxide fuel cells. *J. Fuel Cell Sci. Technol.* **2006**, *3*, 312–321. [CrossRef]

129. Iora, P.; Aguiar, P.; Adjiman, C.; Brandon, N. Comparison of two IT DIR-SOFC models: Impact of variable thermodynamic, physical, and flow properties. Steady-state and dynamic analysis. *Chem. Eng. Sci.* **2005**, *60*, 2963–2975. [CrossRef]

130. Damm, D.L.; Fedorov, A.G. Reduced-order transient thermal modeling for SOFC heating and cooling. *J. Power Sources* **2006**, *159*, 956–967. [CrossRef]

131. Li, J.; Kang, Y.-w.; Cao, G.-y.; Zhu, X.-J.; Tu, H.-Y.; Li, J. Numerical simulation of a direct internal reforming solid oxide fuel cell using computational fluid dynamics method. *J. Zhejiang Univ. Sci. A* **2008**, *9*, 961–969. [CrossRef]

132. Yuan, J.; Rokni, M.; Sundén, B. Three-dimensional computational analysis of gas and heat transport phenomena in ducts relevant for anode-supported solid oxide fuel cells. *Int. J. Heat Mass Transf.* **2003**, *46*, 809–821. [CrossRef]

133. Dusastre, V.; Kilner, J. Optimisation of composite cathodes for intermediate temperature SOFC applications. *Solid State Ion.* **1999**, *126*, 163–174. [CrossRef]

134. Bailly, N.; Georges, S.; Djurado, E. Elaboration and electrical characterization of electrosprayed YSZ thin films for intermediate temperature-solid oxide fuel cells (IT-SOFC). *Solid State Ion.* **2012**, *222*, 1–7. [CrossRef]

135. Jing, Y.; Patakangas, J.; Lund, P.D.; Zhu, B. An improved synthesis method of ceria-carbonate based composite electrolytes for low-temperature SOFC fuel cells. *Int. J. Hydrogen Energy* **2013**, *38*, 16532–16538. [CrossRef]

136. Khaerudini, D.S.; Guan, G.; Zhang, P.; Hao, X.; Kasai, Y.; Kusakabe, K.; Abudula, A. Structural and conductivity characteristics of $Bi_4Mg_xV_{2-x}O_{11-\delta}$ ($0 \leq x \leq 0.3$) as solid electrolyte for intermediate temperature SOFC application. *J. Alloys Compd.* **2014**, *589*, 29–36. [CrossRef]

137. Yang, C.; Ren, C.; Yu, L.; Jin, C. High performance intermediate temperature micro-tubular SOFCs with $Ba_{0.9}Co_{0.7}Fe_{0.2}Nb_{0.1}O_{3-\delta}$ as cathode. *Int. J. Hydrogen Energy* **2013**, *38*, 15348–15353. [CrossRef]

138. Xia, C.; Liu, M. Low-temperature SOFCs based on $Gd_{0.1}Ce_{0.9}O_{1.95}$ fabricated by dry pressing. *Solid State Ion.* **2001**, *144*, 249–255. [CrossRef]

139. Amiri, S.; Hayes, R.; Nandakumar, K.; Sarkar, P. Modelling heat transfer for a tubular micro-solid oxide fuel cell with experimental validation. *J. Power Sources* **2013**, *233*, 190–201. [CrossRef]

140. Cui, D.; Cheng, M. Thermal stress modeling of anode supported micro-tubular solid oxide fuel cell. *J. Power Sources* **2009**, *192*, 400–407. [CrossRef]

141. Fan, P.; Li, G.; Zeng, Y.; Zhang, X. Numerical study on thermal stresses of a planar solid oxide fuel cell. *Int. J. Therm. Sci.* **2014**, *77*, 1–10. [CrossRef]

142. Chiang, L.-K.; Liu, H.-C.; Shiu, Y.-H.; Lee, C.-H.; Lee, R.-Y. Thermal stress and thermo-electrochemical analysis of a planar anode-supported solid oxide fuel cell: Effects of anode porosity. *J. Power Sources* **2010**, *195*, 1895–1904. [CrossRef]

143. Jiang, L.; Wei, T.; Zeng, R.; Zhang, W.-X.; Huang, Y.-H. Thermal and electrochemical properties of $PrBa_{0.5}Sr_{0.5}Co_{2-x}Fe_xO_{5+\delta}$ ($x = 0.5, 1.0, 1.5$) cathode materials for solid-oxide fuel cells. *J. Power Sources* **2013**, *232*, 279–285. [CrossRef]

144. Lin, C.-K.; Chen, T.-T.; Chyou, Y.-P.; Chiang, L.-K. Thermal stress analysis of a planar SOFC stack. *J. Power Sources* **2007**, *164*, 238–251. [CrossRef]

145. Bujalski, W.; Paragreen, J.; Reade, G.; Pyke, S.; Kendall, K. Cycling studies of solid oxide fuel cells. *J. Power Sources* **2006**, *157*, 745–749. [CrossRef]

146. Nakajo, A.; Stiller, C.; Härkegård, G.; Bolland, O. Modeling of thermal stresses and probability of survival of tubular SOFC. *J. Power Sources* **2006**, *158*, 287–294. [CrossRef]

147. Sarmah, P.; Gogoi, T.K. Performance comparison of SOFC integrated combined power systems with three different bottoming steam turbine cycles. *Energy Convers. Manag.* **2017**, *132*, 91–101. [CrossRef]

148. Deseure, J.; Bultel, Y.; Dessemond, L.; Siebert, E. Theoretical optimisation of a SOFC composite cathode. *Electrochim. Acta* **2005**, *50*, 2037–2046. [CrossRef]

149. Costamagna, P.; Costa, P.; Arato, E. Some more considerations on the optimization of cermet solid oxide fuel cell electrodes. *Electrochim. Acta* **1998**, *43*, 967–972. [CrossRef]

150. Calise, F.; d'Accadia, M.D.; Vanoli, L.; Von Spakovsky, M. Single-level optimization of a hybrid SOFC–GT power plant. *J. Power Sources* **2006**, *159*, 1169–1185. [CrossRef]

151. Chachuat, B.; Mitsos, A.; Barton, P.I. Optimal design and steady-state operation of micro power generation employing fuel cells. *Chem. Eng. Sci.* **2005**, *60*, 4535–4556. [CrossRef]

152. Baratto, F.; Diwekar, U.M. Multi-objective trade-offs for fuel cell-based auxiliary power units: Case study of South California Air Basin. *J. Power Sources* **2005**, *139*, 197–204. [CrossRef]

153. Song, X.; Diaz, A.; Benard, A.; Nicholas, J. A 2D model for shape optimization of solid oxide fuel cell cathodes. *Struct. Multidiscip. Optim.* **2013**, *47*, 453–464. [CrossRef]

154. Chakraborty, U.K. Static and dynamic modeling of solid oxide fuel cell using genetic programming. *Energy* **2009**, *34*, 740–751. [CrossRef]

155. Facchinetti, E.; Gassner, M.; D'Amelio, M.; Marechal, F.; Favrat, D. Process integration and optimization of a solid oxide fuel cell–Gas turbine hybrid cycle fueled with hydrothermally gasified waste biomass. *Energy* **2012**, *41*, 408–419. [CrossRef]

156. Yi, Y.; Rao, A.D.; Brouwer, J.; Samuelsen, G.S. Analysis and optimization of a solid oxide fuel cell and intercooled gas turbine (SOFC–ICGT) hybrid cycle. *J. Power Sources* **2004**, *132*, 77–85. [CrossRef]

157. Sadeghi, S.; Ameri, M. Study the Combination of Photovoltaic Panels With Different Auxiliary Systems in Grid-Connected Condition. *J. Sol. Energy Eng.* **2014**, *136*, 041008. [CrossRef]

158. Quddus, M.R.; Zhang, Y.; Ray, A.K. Multi-objective optimization in solid oxide fuel cell for oxidative coupling of methane. *Chem. Eng. J.* **2010**, *165*, 639–648. [CrossRef]

159. Palazzi, F.; Autissier, N.; Marechal, F.M.; Favrat, D. A methodology for thermo-economic modeling and optimization of solid oxide fuel cell systems. *Appl. Therm. Eng.* **2007**, *27*, 2703–2712. [CrossRef]

160. Yang, J.; Li, X.; Jiang, J.H.; Jian, L.; Zhao, L.; Jiang, J.G.; Wu, X.G.; Xu, L.H. Parameter optimization for tubular solid oxide fuel cell stack based on the dynamic model and an improved genetic algorithm. *Int. J. Hydrogen Energy* **2011**, *36*, 6160–6174. [CrossRef]

161. Wu, X.J.; Zhu, X.J. Optimization of a solid oxide fuel cell and micro gas turbine hybrid system. *Int. J. Energy Res.* **2013**, *37*, 242–249. [CrossRef]

162. Bejan, A. *Advanced Engineering Thermodynamics*, 2nd ed.; Wiley: New York, NY, USA, 1997.

163. Sciacovelli, A. Thermodynamic optimization of a monolithic-type solid oxide fuel cell. *Int. J. Thermodyn.* **2010**, *13*, 95–103.

164. Vogler, M.; Horiuchi, M.; Bessler, W.G. Modeling, simulation and optimization of a no-chamber solid oxide fuel cell operated with a flat-flame burner. *J. Power Sources* **2010**, *195*, 7067–7077. [CrossRef]

165. Autissier, N.; Palazzi, F.; Maréchal, F.; Van Herle, J.; Favrat, D. Thermo-economic optimization of a solid oxide fuel cell, gas turbine hybrid system. *J. Fuel Cell Sci. Technol.* **2007**, *4*, 123–129. [CrossRef]

166. Baharanchi, A.A. Multidisciplinary Modeling, Control, and Optimization of a Solid Oxide Fuel Cell/Gas Turbine Hybrid Power System. Master's Thesis, University of Miami, Coral Gables, FL, USA, May 2009.

# Effect of Martensitic Transformation and Grain Misorientation on Surface Roughening Behavior of Stainless Steel Thin Foils

Abdul Aziz [1,*], Ming Yang [1], Tetsuhide Shimizu [1] and Tsuyoshi Furushima [2]

1 Advance Materials Engineering Laboratory (AMPEL), Tokyo Metropolitan University, 6-6 Asahigaoka, Hino City, Tokyo 191-0065, Japan; yang@tmu.ac.jp (M.Y.); simizu-tetuhide@tmu.ac.jp (T.S.)
2 Institute of Industrial Science, The University of Tokyo, 4-6-1, Meguro City, Tokyo 153-8505, Japan; tsuyoful@iis.u-tokyo.ac.jp
* Correspondence: abdul.azizayahnajib@gmail.com

**Abstract:** The surface roughening (Ra), martensitic phase transformation (MPT), and grain misorientation (GMO) behavior of stainless steel 304 and 316 in various grain sizes (Dg) were studied experimentally, including five cycles of sequential uniaxial tensile stress testing and Scanning Electron Microscope-Electron Back Scattered Diffraction (SEM-EBSD) investigation. The MPT and GMO characteristics were sequentially investigated using tensile testing and SEM-EBSD analysis. The correlation between MPT, GMO, martensitic volume fraction (Mf), and Ra behavior were investigated. The experimental results showed that increasing the total strain from 5.0% to 25.0% increased the MPT, GMO, and Mf, which were transformed from the metastable austenitic phase in stainless steel (SUS) 304. The increasing total strain increased Ra for all kinds of Dg. Furthermore, SUS 304 and SUS 316 were used to compare the roughening mechanism. The MPT was very high and spread uniformly in fine grain of SUS 304 thin foil, but the MPT was low and not uniform in coarse grain of SUS 304 thin foil. There was no MPT in SUS 316 thin foil, both in coarse and fine grain. The GMO in fine grains, both in SUS 304 and SUS 316 thin foils, spread uniformly. The GMO in coarse grains, both in SUS 304 and SUS 316 thin foils, did not spread uniformly. Surface roughness increased higher in coarse grain than fine grain for both of SUS 304 and SUS 316 thin foil. SUS 304 increased higher than SUS 316 thin foil. The effect of inhomogeneous deformation due to the MPT is a more important factor than GMO in coarse grain.

**Keywords:** martensitic phase transformation (MPT); grain misorientation (GMO); surface roughening (Ra); grain size (Dg)

## 1. Introduction

Micro metal forming, particularly for austenitic stainless steel (ASS), is a promising approach in the biomedical, electronic, chemical, electrical power, food, and nuclear industries [1], ASS has excellent corrosion resistance and processability [1,2]. In addition, the high demand for microparts has received significant attention in recent decades [2,3]. However, microforming technology has a number of challenges, such as the limitation of material applications and the high cost of mass production [1–3]. When ASS is subjected to plastic deformation, martensitic-induced transformation occurs in the ASS [1]. The transformed martensitic volume fraction increases with the increase in plastic deformation [1–3]. Martensitic transformation reduces the toughness but increases the strength of ASS [4,5]. When subjected to plastic deformation, austenitic stainless steel, as a metastable phase, undergoes a transformation from Face Centered Cubic (FCC) austenite to Body Centered Tetragonal (BCT) martensite at room temperature [1–4]. The martensitic transformation enhances the strength of thin metal foils and results in elongation because an increasing work hardening rate can delay the occurrence of plastic instability in thin metal foils [1–5]. Factors that may result in martensitic transformations are chemical composition, strain path, grain size, strain level, and strain rate [6]. Xue et al. [1] found that the

martensitic-transformed volume fraction can be controlled by controlling the stainless strip steel deformation. Martensitic transformation on the surface is larger than that in the interior with the same strain [1–4]. When ASS is subjected to plastic deformation, in addition to martensitic phase transformation (MPT), dislocation interaction and twinning also occur [1,5]. Twinning is greater on the surface than in the interior of ASS [1,6]. Increasing martensitic transformation results in the increasing stacking fault energy (SFE) of martensite; hence, the martensitic transformation occurs more readily on the surface than in the interior [1]. Furthermore, surface roughening also occurs more readily on the surface than in the interior of thin metal foils [1–3]. It is necessary to investigate the dependency between surface roughening behavior and MPT in ASS thin metal foils. Peng et al. [7] concluded that MPT is caused by an increasing strain rate, which results in an increase in the local temperature. A large increase in the strain rate suppresses MPT [7,8]. The Olson and Cohen model, called the OC model, is fundamental to the description of the kinetics of strain induced in martensite [1,7,9,10]. The shear band intersection, as a dominant nucleation site, is considered [7,11]. The transformation curve is constructed only as a function of the plastic strain and a constant environmental temperature in the OC model [7,10,12]. Tomita et al. [8] found that the number of shear band intersections, as the location of the nucleation site, increases as the strain rate increases, while the probability that a shear band intersection forms an embryo decreases. This conclusion was considered only for a constant temperature, and the increasing temperature caused by self-heating in the tensile test was ignored. Zandrahimi et al. [9], concluded that transformation of austenite to martensite (MPT) in AISI 304 resulted in surface hardening, which led to deterioration of the wear resistance. It is necessary to investigate the surface roughness caused by MPT because surface roughening is caused by grain deformation on the surface and affects thin metal foil properties, particularly on the surface. Zihao et al. [10] concluded that uniform elongation (UEL), ultimate tensile strength (UTS), and martensitic volume fraction decreased because of increasing strain rate and temperature. Jeom et al. [11] concluded that, in duplex stainless steel, strain-induced martensitic transformation (SIMT) occurred after plastic deformation. The transformation from austenite to $\varepsilon$ martensite occurred at a low strain level, and transformation from austenite to $\alpha'$ martensite occurred at a high strain level [10,11].

In addition, in miniaturization and metal forming processes, with decreasing sheet thickness, the ratio of surface roughness to thickness increases in metal foils, and this is known as a non-uniform thickness [12–14]. The surface roughening phenomena of the sheet material have a strong effect on necking and fracture behavior of materials [12,13]. The inhomogeneous deformation of each grain located near the free surface will cause surface roughening phenomena on the free surface of polycrystalline metals [13,14]. Thus, surface roughening is highly important in the field of microforming technology using sheet or thin metal foils [12–14]. Surface roughening results in the size effect of thin metal foils [15]. Furushima et al. [3] concluded that fracture behavior at the microscale is the major factor causing surface roughening. The failure behavior and material flow of thin sheet metals are influenced by the size effect, which is mainly caused by the low number of grains in the straining zone [16]. Meng et al. [17] found that, when the surface non-uniformity increases to the same magnitude as the sample thickness, the effect of the free surface roughening on fracture strain and flow behavior is significant. Stoudt et al. [18] found that a linear dependency exists between grain size and roughening rate, as well as demonstrated that the roughening rate was dependent on Mg-Al alloy grain size. Furushima et al. [19] found that pure titanium and pure copper dramatically decrease the fracture strain for thicknesses from 0.3 to 0.1 mm. Fracture strain is significantly affected by material thickness [18,19]. The size effect is challenging for micro metal forming of ultra-thin metal foils [17–19]. When plastic strain increases, the ratio of the surface roughness to the thickness for each material linearly increases [16–19]. Rabee et al. [20] found that clarification is needed in respect of the relationship between deformation-induced surface roughness and the local microstructure. Furushima et al. [21] concluded that fractures were caused by free surface roughening

because dimpling did not occur for pure copper thin foil in the range of thicknesses from 0.05 to 0.1 mm. Surface roughness increased under uniaxial deformation because of decreased thickness in the same area [19–21]. This indicates that, when the number of grains is decreased, the surface roughness increases, as shown by uniaxial tensile testing with the same strain level. The fracture strain was low and surface roughness significantly increased in a uniaxial tensile test with the same strain level due to the five-fold decrease in the quantity of grain size (Dg) [20,21]. Lei Zhang et al. [22] found that investigation is limited regarding the quantitative description of the evolution of the surface roughness of FCC polycrystalline metals. Research and investigation of surface roughness, other than that of FCC structure polycrystalline metals, is, therefore, needed [18–22]. Kengo Yoshida [23] found that surface roughness is mainly governed by the Dg, and investigation is needed in respect of surface roughness behavior with Dg below 10 μm. Shimizu et al. [24] concluded that the mutuality of grains affects the surface roughening behavior. Different individual deformations of grains affect the surface roughening behavior [16–24]. It is necessary to investigate surface roughness behavior with different Dg, which may result in different mechanisms of surface roughness behavior. Linfa Peng et al. [25] found that the individual grains, and particularly the surface grains, are less restricted due to the decrease in grain boundary density caused by the increasing Dg. As the grain orientation and the structures of individual grains are randomly distributed, the inhomogeneity of grains is more significant, which leads to a significant increase in surface roughness [24,25]. Thus, investigation and understanding are needed in respect of surface roughening behavior with various Dg.

Furushima et al. [26] found that plastic deformation preferentially occurs in weak grains, which have lower flow stress in pure copper (C 10220-O) ultra-thin metal foil with a thickness of 0.05 mm. The roughening phenomena on the free surface of polycrystalline metals are affected by the inhomogeneous deformation in each grain located near the free surface [14,26]. P. Groche et al. [27] found that, due to the Hall–Petch theory, the flattening of the surface asperities is hindered by the increase in yield stress. Cheng et al. [28] found that the mechanical properties are affected by free surface roughening in thin metal foils, rather than the growth of the voids in the side materials, in thin sheet metal or thin metal foils. Thus, research and study of surface roughening is needed to understand the mechanisms of surface roughening, particularly for small Dg. Aziz et al. [29] found that surface roughening increased proportionally without annealing in coarse grains of SUS 304. The increase in surface roughening was not proportional without annealing in fine grains in either SUS 304 or SUS 316 thin metal foils. Furushima et al. [30] found that the grain strength variation is a representation parameter that determines material inhomogeneity. The decrease in the ratio of thickness to grain size, t/d, leads to inhomogeneous deformation, which affects the surface roughness behavior [26–30]. As a result, the increase in Dg in the same thickness leads to inhomogeneity of the grain strength [26–31].

For stainless steel (SUS) 304, the grain has high volume fraction of MPT induced by plastic deformation, and the grain strength increases, and grain deformation becomes more difficult, compared to grains with lower MPT. Furthermore, the effect of MPT on the surface roughness could depend on the microstructure, which means that the effect of MPT depends on the grain size. However, the effect of the MPT for different grain size on the surface roughness is unclear yet. The aim of this study was to clarify how the MPT and GMO result in inhomogeneous grain strength which is shown by the surface roughening behavior of thin metal foils with various Dg. In this study, ASS SUS 304 and SUS 316 thin metal foils were subjected to five cycles of sequential uniaxial tensile stress state testing to measure surface roughening behavior, and the MPT, GMO, and grain deformation mechanisms were analyzed using SEM-EBSD.

## 2. Materials and Research Method

### 2.1. Materials

According to Tables 1 and 2, SUS 304 thin metal foil has a higher quantity of carbon and chromium. As a result, the MPT occur more easily in SUS 304 thin metal foil, and a low volume of carbide compound may occur in the grain matrix. The other element, such as silicon, mangan, phosphor, and sulfur, have the same quantity that gives the same effect in mechanical properties. MPT increase the strength of thin metal foils that occur in the grain matrix. Plastic deformation promotes dislocation motion. MPT occurs because of the dislocation motion during plastic deformation, which changes the crystal structure from face center cubic (FCC) to body center tetragonal (BCT). The amount of nickel in SUS 316 is almost 4–5% higher than in SUS 304 thin metal foils. This indicates that austenite in SUS 316 thin metal foils is more stable than austenite in SUS 304 thin metal foils. However, the quantity of the chromium element as a ferrite former in SUS 304 is higher than that in SUS 316 thin metal foils. This means chromium carbide is more easily formed in the grain matrix or at the boundary, as well as increases the strength of thin metal foils when they are subjected to plastic deformation. The greater the quantity of nickel in stainless steel, the more difficult the transformation from austenite to martensite. This is because nickel is an austenitic stabilizer, and higher energy is required for the transformation of the austenitic phase into martensite [29,31–33].

**Table 1.** Chemical composition thin metal foils of SUS 304 (wt%).

|  | C | Si | Mn | P | S | Ni | Cr |
|---|---|---|---|---|---|---|---|
| Min | - | - | - | - | - | 8.00 | 18.00 |
| Max | 0.08 | 1.00 | 2.00 | 0.045 | 0.030 | 10.50 | 20.00 |

**Table 2.** Chemical composition thin metal foils of SUS 316 (wt%).

|  | C | Si | Mn | P | S | Ni | Cr | Mo |
|---|---|---|---|---|---|---|---|---|
| Min | - | - | - | - | - | 12.00 | 16.00 | 2.00 |
| Max | 0.030 | 1.00 | 2.00 | 0.045 | 0.030 | 15.00 | 18.00 | 3.00 |

Figure 1 shows typical microstructures of various grain size (Dg) after annealing at 400 °C for one hour. The reason of annealing is to remove residual stress after rolling. The grain size (Dg) did not change before and after annealing treatment. Annealing subjected to various Dg ranging from fine grain to coarse grain. According to the Hall–Petch theory and equation, it is well known that grains size affects the mechanical properties of thin foil materials. Thin metal foils having fine grains have higher tensile and yield strength than thin metal foils having coarse grains. SUS 304 and SUS 316 thin metal foils exhibit different surface roughness values in a uniaxial tensile test applied with the same strain level in both thin metal foils. In previous research, a uniaxial tensile test was applied to thin metal foils without annealing. The strength of thin foil SUS 304 and SUS 316 were similar after annealing at 400 °C compared to as received materials [29,34].

a. Dg 1.5 μm SUS 316 as received

b. Dg 1.5 μm SUS 304 as received

c. Dg 3.0 μm SUS 316 as received

d. Dg 3.0 μm SUS 304 as received

e. Dg 9.0 μm SUS 316 as received

f. Dbg 9.0 μm SUS 304 as received
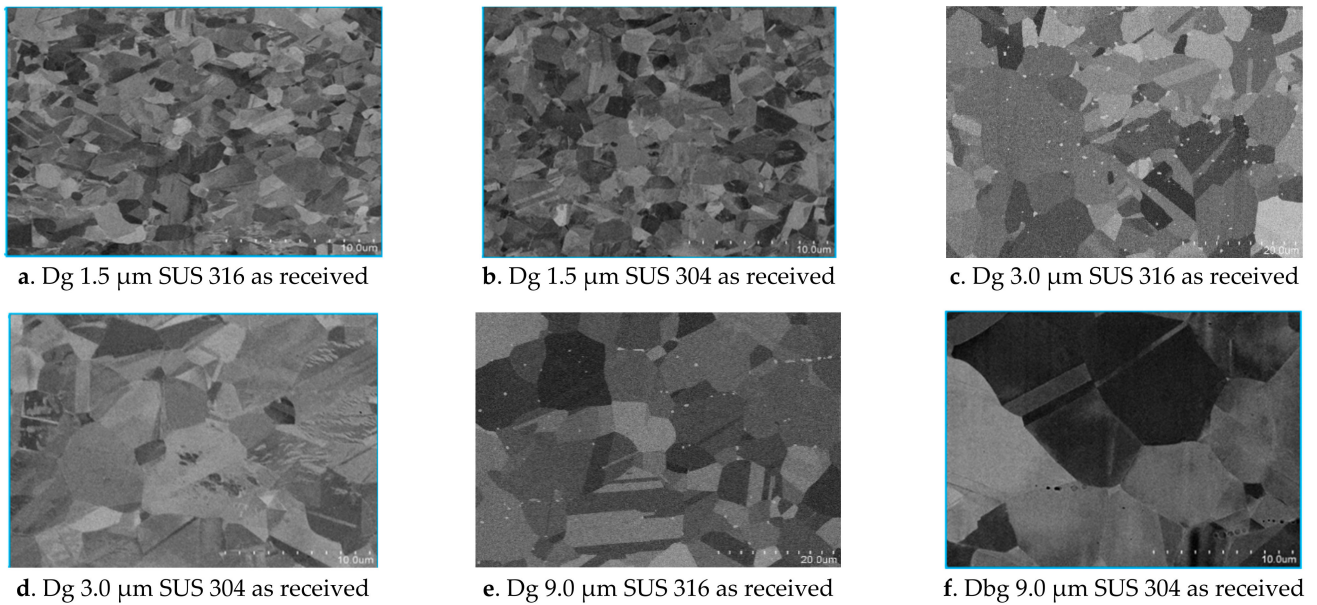
**Figure 1.** SUS 304 (**b**,**d**,**f**) and 316 (**a**,**c**,**e**) thin metal foils with different grain sizes (Dg) as received.

Figure 2a shows thin metal foils of commercially available SUS 304 and 316 with a width of 4.0 mm, thickness of 0.1 mm, and gauge length of 20 mm, using the DIN 50125 standard. The sample was made as a dog bone type, as shown in Figure 2a. The grain size measurement technique is shown in Figure 2b.



(**a**)



(**b**)
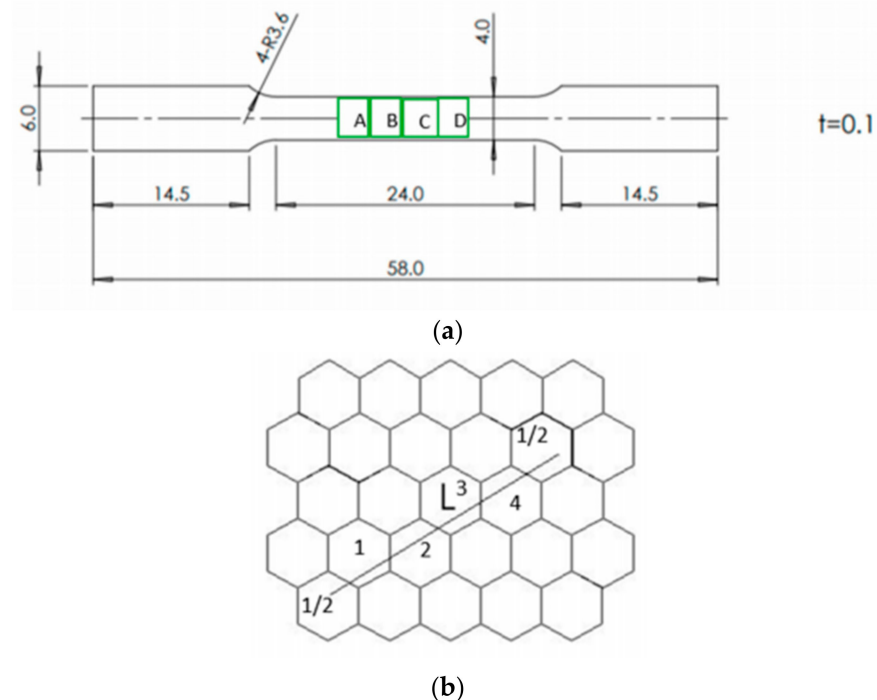
**Figure 2.** (**a**). Technical drawing specimen of tensile test. (**b**). The crystal structure photo micrograph analogy.

## 2.2. Method

Before each sample was subjected to the uniaxial tensile stress test, cleaning was performed using ethanol combined with ultrasonic vibration for 30 min to increase the cleanliness of the surface. The sample was subjected to five cycles of tensile stress with constant strain. After the sample was subjected to uniaxial tensile stress, surface roughness

was measured using a confocal laser microscope (OLS-5000, produced by Olympus, Co., Tokyo, Japan. The samples were subjected to five cycles of uniaxial tensile stress using a commercial Tensile Machine.

The uniaxial tensile test was performed sequentially for five cycles with a constant strain level on SUS 304 and SUS 316 stainless steel thin metal foils with various Dg. Then, the surface roughening behavior was measured sequentially until stage five. Thin metal foils of SUS 304 and SUS 316 were sequentially subjected to the uniaxial tensile test over the yield point using a constant strain level. The thin metal foils of SUS 304 and SUS 316 had a 20 mm gauge length, and the thickness of the thin foil was 0.1 mm. The width of thin foil was 4.0 mm. The fillet radius was 3.6 mm. This experiment used an Autograph AGX-50 KNVD Shimadzu commercial tensile test machine, produced by Shimadzu Corporation, Japan. This machine was used to ensure experimental consistency in the tensile test. During the tensile test, the strain rate was $1.6 \times 10^{-3}$ m s$^{-1}$. After the tensile test, the surface roughness was observed and measured for different materials, different Dg, and constant thickness. The tensile test and surface roughness investigation were repeated until stage five. As the contact extensometer could not be pasted onto the metal foils, the elongation was measured optically with a video non-contact extensometer (DVE-201, Shimadzu corp). The tensile test was sequentially applied to thin metal foils of SUS 304 and SUS 316. At each step, the tensile test was halted, and the specimen was taken from the chuck of the universal tensile test machine. Then, the surface roughness behavior was measured using a confocal laser microscope (OLS-5000, produced by Olympus Co., Japan). Surface roughness was observed at the center point of the A, B, C, and D area in the rolling direction of each step, as shown in Figure 2a. The area of the surface roughness measurement had a length of 0.7 mm, and the surface roughness was measured at the same position until stage five using the same strain level.

A SEM SU-70, produced by Hitachi High Technology corp, Japan, was used in normal mode to investigate the microstructure behavior after the tensile test. During the microstructure investigation, the SEM SU-70 acceleration voltage was 5 kV, the working distance was 10 mm, and the emission current was 16 µA. Phase transformation was investigated using an EBSD Digi View (EDAX) in field-free mode, with an emission current of 16 µA, acceleration voltage of 15 KV, and working distance of 20 mm. The pixel binning was $8 \times 8$, and the step (resolution) of the EBSD machine was 0.1 µm.

A JIS G0551 standard was used for measuring Dg in both SUS 304 and SUS 316 thin metal foils. As shown in Figure 2b, Dg was measured using the draw line (length is L) on a photomicrograph. Then, from the line, the number of crystals could be counted, including $\frac{1}{2}$ at the end of the line. The crystal structure photomicrograph was an analogy, as shown in Figure 2b. The Dg can be calculated using the following formula. The average Dg was calculated as $= \frac{4 + \frac{1}{2} + \frac{1}{2}}{L}$.

## 3. Experimental Results of Tensile Test

Figure 3 shows the stress–strain curves of SUS 304 and SUS 316 thin metal foils with various Dg. In SUS 304 thin metal foils, fracture strain is influential factor in this study. In contrast, tensile strength decreases with increasing Dg in SUS 304 thin metal foils. The strength and ductility of SUS 304 are higher than those of SUS 316 thin metal foils, as shown in Figure 3. According to the Hall–Petch theory, materials with fine grains exhibit higher strength than materials with coarse grains. The strength significantly improved when the grain size decreased from coarse to fine [34]. The correlations between surface roughness behavior with strain level without annealing have been revealed in previous study, showing that surface roughness increases proportional in coarse grain and does not increase proportional in fine grain during uniaxial tensile test [29].
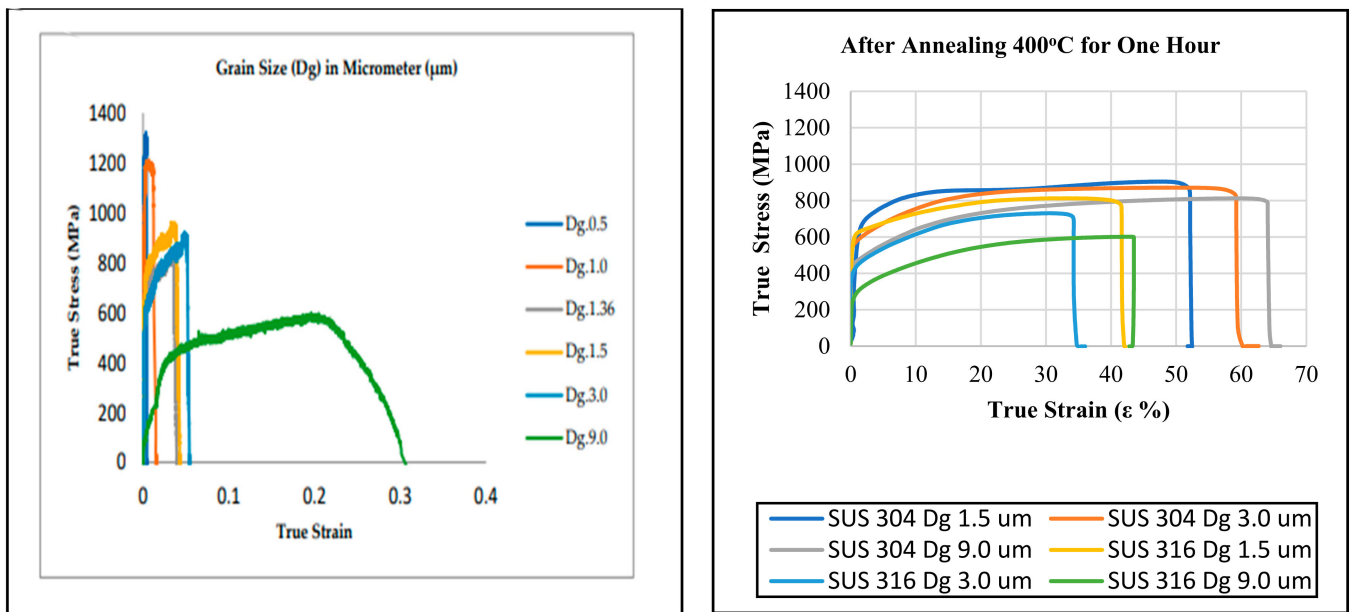
(**a**). Stress–strain curve before annealing (30)

(**b**). Stress–strain curve after annealing

**Figure 3.** Stress–strain curve of SUS 304 and SUS 316 thin foil before (**a**) and after annealing (**b**).

Figure 4 indicates the relation between the surface roughness and true strain at a low strain level of 5% (1.0% at one stage of the tensile test) and at a high strain level of 25% (5.0% at one stage). The samples of both SUS 304 and SUS 316 thin metal foils were subjected to five stages of a tensile test at low strain and high strain levels. Thus, the accumulation of the strain level for low strain was 5.0%, and the accumulation of the strain level for high strain was 25.0%.
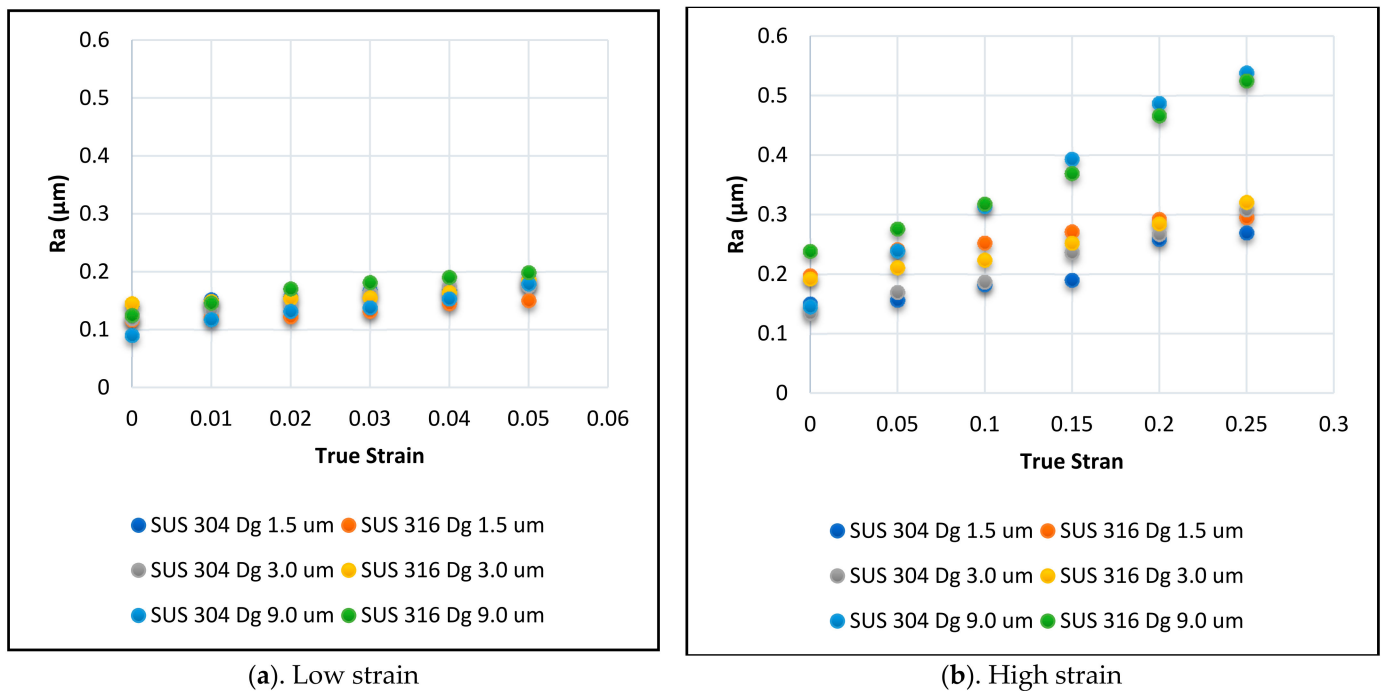


(**a**). Low strain

(**b**). High strain

**Figure 4.** Surface roughening behavior in SUS 304 and SUS 316, low strain (**a**) and high strain (**b**).

According to Figure 3a,b, Dg 1.3 μm was SUS 316 thin foil without annealing, and Dg 0.5 μm, Dg 1.0 μm, Dg 1.5 μm, Dg 3.0 μm, Dg 9.0 μm were SUS 304 thin foil without

annealing [29]. The strength of thin foils between SUS 304 and SUS 316 as received and after annealing are similar, as shown in Figure 3a,b.

The correlation between surface roughness and true strain is shown in Figure 4a,b, in both low and high strain level. Surface roughness (Ra) increased proportional both in low and high strain level. Based on Figure 4a, the surface roughness increased proportionally in the low strain level in both the fine and coarse grains. The tendency of surface roughness behavior is similar in fine and coarse grains at the low strain level. Based on Figure 4B, the surface roughness increased proportionally at the high strain level in both fine and coarse grains. The increase in surface roughness was higher for coarse grain than fine grain, in both SUS 316 and SUS 304 thin metal foils.

Figure 5 indicates the relation between the increase in surface roughness (ΔRa) and the true strain instead of Ra. ΔRa means the difference between the initial surface roughness and the surface roughness after the five steps of the tensile test. At a high strain level, ΔRa was proportional to true strain in both fine and coarse grains, and ΔRa increased to a higher value in coarse grains than in fine grains. The greater increase in Ra in coarse grains under high strain conditions indicates that the mechanisms of fine and coarse grains are different at a high strain level.



(**a**). Low strain            (**b**). High strain
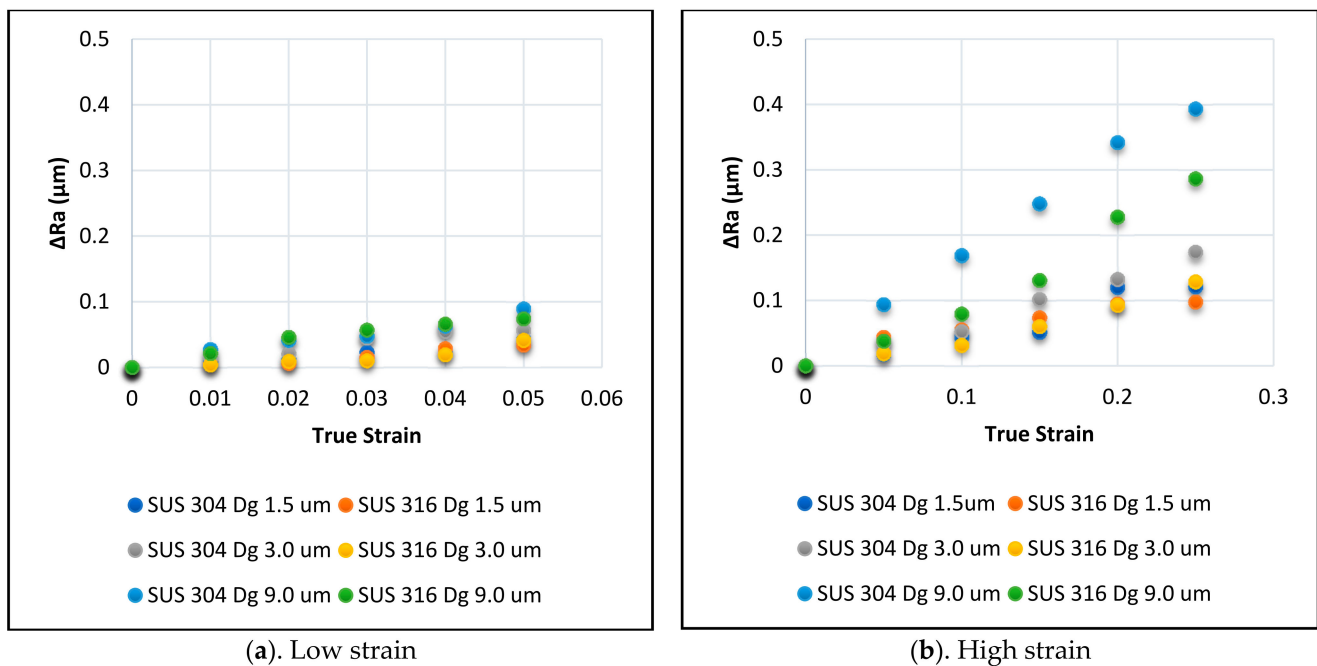
**Figure 5.** The increase in surface roughening behavior in SUS 304 and SUS 316, low strain (**a**) and high strain (**b**).

Figure 6 indicates the ΔRa values for SUS 304 and SUS 316 thin metal foils at a high strain level with various grain sizes. The ΔRa increased proportional to Dg, but the gradient depends on the material.

Since the grain strength of fine grains is higher than that of coarse grains, the grain strength may have an effect on surface roughness increasing. The higher the strength in fine grains, the more difficult it is to deform them. According to the Hall–Petch effect, the higher strength in fine grains than coarse grains is also indicated by the higher tensile strength in fine grains compared to coarse grains, as shown in Figure 3.
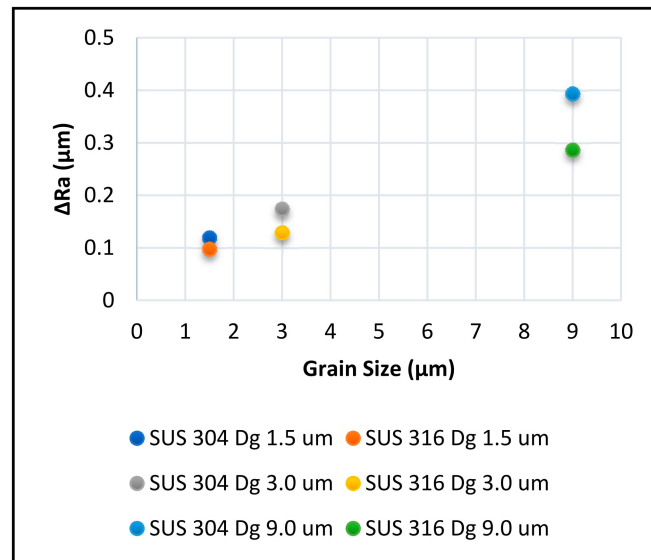
**Figure 6.** Increasing Ra at high strain level.

The surface roughness also depends on the inhomogeneity of the grains; higher inhomogeneity of grains will result in higher surface roughness after plastic deformation with the same strain level [26].

The $\Delta Ra$ of SUS 304 coarse grains is higher than the $\Delta Ra$ of SUS 316 coarse grains because the grain strength in SUS 304 is higher than that of SUS 316, as indicated by the higher tensile strength in SUS 304 than SUS 316 thin metal foil fine grains, as shown in Figure 3.

Based on previous studies [30], $\Delta R_a$ increases linearly to product of grain size ($Dg$) and strain levels or true strain ($\varepsilon$), as shown in Equation (1),

$$R_a = C \cdot D_g \cdot \varepsilon + R_0, \tag{1}$$

where $R_o$ is initial surface roughness, and $C$ is the material constant. Since $\Delta R_a = R_a - R_0$, we could write the equation as:

$$\Delta R_a = C \cdot D_g \cdot \varepsilon. \tag{2}$$

The equation means that coefficient $C$ only depends on the material but not on $Dg$ and $\varepsilon$. By substituting the results shown in Figure 5 into Equation (2), the correlation between product of grain size ($Dg$), true strain ($\varepsilon$), and $\Delta R_a$ can be plotted as shown in Figure 7. The $\Delta R_a$ increase proportional with the increasing $D_g \cdot \varepsilon$. However, there are differences in the gradients. The difference of the gradients may depend on the homogeneous and inhomogeneous grain strength of thin metal foils due to the grain sizes. Different mechanisms may also result in different tendencies in surface roughness behavior in SUS 304 and SUS 316 s.

(**a**). Low strain                       (**b**). High strain
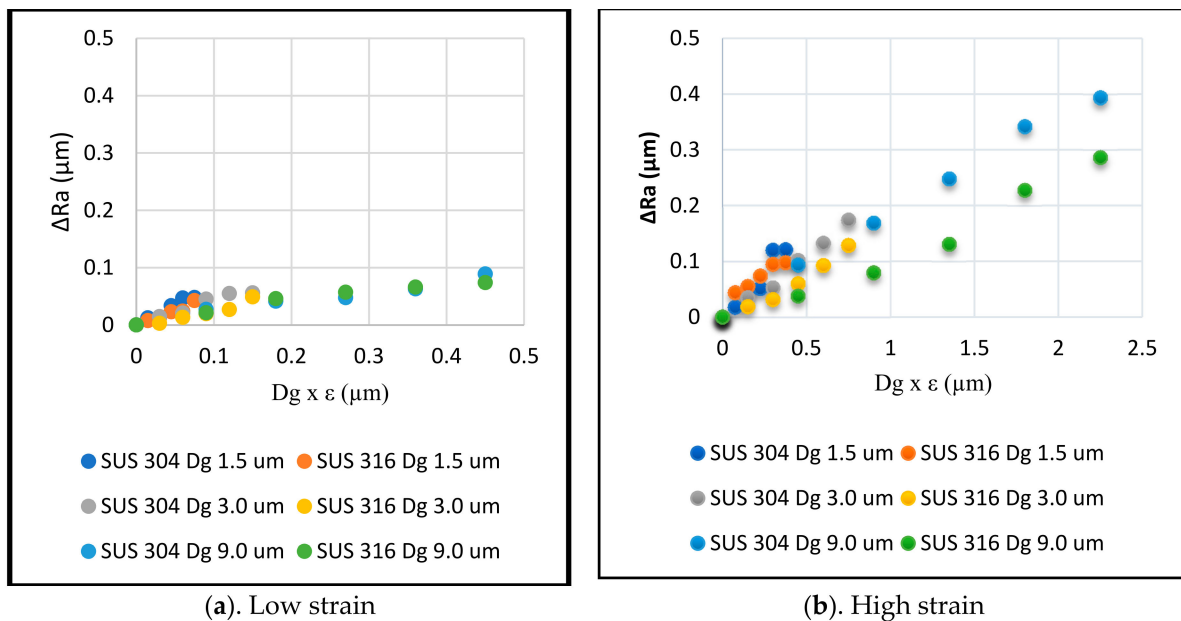
**Figure 7.** The increase in surface roughening due to grain size and true strain of SUS 304 and SUS 316, low strain (**a**) and high strain (**b**).

For investigation of the different in deformation mechanisms, microstructures of the material after the tensile test was analyzed by using SEM-EBSD, discussed in next section.

## 4. Results of SEM-EBSD and Discussion

In this research, SEM-EBSD was used to obtain a phase map, IPF map, and KAM map before and after the sequential uniaxial tensile test. The Phase map consists of austenite and martensite phases in the SEM-EBSD result. The inverse pole figure (IPF) map derived by EBSD reflects the locally discovered orientation. From the IPF map, we could obtain grain deformation behavior with the changing of a color in a grain. The kernel average misorientation (KAM) map is a map that shows grains orientation by changing collective color in the EBSD result. The KAM map comes from the calculation and averaged grain misorientation in the center with the surroundings.

The microstructures, consisting of MPT derived from phase map and GMO derived from KAM map, were obtained using SEM-EBSD analysis. The MPT and GMO were investigated in fine and coarse grains of SUS 304 and SUS 316 thin metal foils at both low and high strain levels and related to surface roughness behavior. In Figure 8(A1–A4), the phase map is shown. The red color is the gamma ($\gamma$) or austenite phase, and the green color is martensite phase transformation (MPT), also known as alpha ($\alpha$) iron; Figure 8(B1–B4) show the IPF normal direction (ND) map. Grain deformation is more severe in coarse grains, as shown in Figure 8(B3,B4), than fine grains, as shown in Figure 8(B1,B2), in both SUS 304 and SUS 316 thin metal foils. There was no MPT in SUS 316 fine grain and coarse grain, as shown in Figure 8(A1,A4), even after 25% total strain level. The grain deformation in coarse grain of SUS 316 was more severe, as shown in Figure 8(B4), compared to fine grain, as shown in Figure 8(B1). The MPT was very high and spread uniformly in fine grain of SUS 304, as shown in Figure 8(A2). The MPT was low and not uniformly spread in coarse grain, as shown in Figure 8(A3).
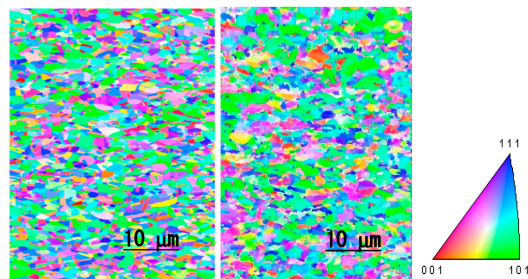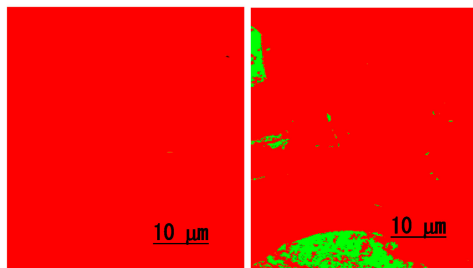
A.1.  EBSD phase mapping in Dg 1.5 µm SUS 316

B.1. EBSD IPF mapping in Dg 1.5 µm SUS 316

A2. EBSD phase mapping in DG 1.5 µm SUS 304

B2. EBSD IPF mapping in DG 1.5 µm SUS 304

A3. EBSD phase mapping in DG 9.0 µm SUS 304

B3. EBSD IPF mapping in DG 9.0 µm SUS 304

A4. EBSD phase mapping in DG 9.0 µm SUS 316

B4. EBSD IPF mapping in DG 9.0 µm SUS 316

**Figure 8.** SEM-EBSD mapping for SUS 304 and SUS 316 thin metal foils.

During the tensile test with the same strain level, the slip band intersection was higher in fine-grain SUS 304 thin metal foils than the coarse grains of SUS 304 thin metal foils [8–10]. As a result, MPT in fine grains was greater than that in coarse grains of SUS 304 thin metal foils. The slip band intersection is the location of the martensitic embryo and nucleation [10,11,35]. Fine grains have a higher probability of the slip band intersection than coarse grains [10], thus, MPT is greater in fine grains than coarse grains with the same strain level. Therefore, fine grains have greater grain strength than coarse grains, and fine grains have lower inhomogeneous grain strength than coarse grains. As a result, coarse grains with lower MPT become inhomogeneous grains, and fine grains that consist of higher MPT become homogeneous grains, which indicates the higher surface roughening behavior in coarse grains than in fine grains of SUS 304 thin metal foils at the same strain level.

As shown in Figure 9(A1–A4), it was known that the KAM map consists of three kinds of color, the blue one was 0° misorientation, the green one was 2° misorientation, and the red one was 5° misorientation [29]. The effect on the grain strength of 5° misorientation was higher than 2° grain misorientation, and the effect on the grain strength of 2° misorientation was higher than 0° grain misorientation [24,27,29]. There was no effect on 0° grain misorientation in grain strength [24,29]. According to the grain misorientation, the grain strength of SUS 316 thin foil was high, and the grain strength of SUS 304 fine grain was very high. Figure 9(A1,A2) show that the amount of 5° grain misorientation was similar between SUS 304 and SUS 316 fine grain thin foils, but 2° green misorientation in SUS 316 thin metal foils was lower than the green misorientation in SUS 304 thin metal foils. This means that a grain of SUS 316 is weaker than a grain of SUS 304.



Dg 1.5 µm. SUS316
As received

Dg 1.5 µm SUS316.
25.0% Strain

A1. EBSD KAM in DG 1.5 µm SUS 316

Dg 1.5 µmSUS 304
As received

Dg 1.5 µm SUS304
25.0% strain

A2. EBSD KAM in DG 1.5 µm SUS 304

Dg 9.0 µm SUS 316
As received

Dg 9.0 µm SUS316
25.0% Strain

A3. EBSD KAM in DG 9.0 µm SUS 316

Dg 9.0 µm SUS304
As received

Dg 9.0 µm SUS304
25.0% Strain

A4. EBSD KAM in DG 9.0 µm SUS 304

● = 5° grain misorientation
● = 2° grain misorientation
● = 0° grain misorientation

**Figure 9.** KAM mapping for SUS 304 and SUS 316 thin metal foils.

The type of grain deformation mechanism, in both SUS 304 and SUS 316 fine grain thin metal foils, was intergrain deformation [24,27]. Intergrain deformation results in a small change in the surface roughness in fine grain SUS 304 and SUS 316 thin metal foils because the intergrain deformation affects the low inclination of a grain from the normal direction that affects more homogeneous grain strength values [22,24,25,36]. The surface roughness in fine-grain SUS 304 thin metal foils was almost the same as that of fine grain SUS 316 thin metal foils because the grain inclinations were similar after plastic deformation, even though the grain strength of SUS 316 thin metal foil is weaker than that of SUS 304 thin metal foil.

## 5. Conclusions

First, surface roughness increased proportionally with true strain until a high strain level (25%) in SUS 304 and SUS 316 thin metal foils with both coarse and fine grains. The increase in surface roughness for coarse grain was greater in SUS 304 than in SUS 316 thin metal foil because the deformation in SUS 304 thin metal foils are more inhomogeneous than SUS 316 thin metal foils due to the MPT.

Second, the effect of MPT on the inhomogeneous grain deformation in coarse grain was larger than GMO, as shown by the higher surface roughness in SUS 304 coarse grain thin foil, compared to SUS 316 thin foil coarse grain, at the same strain level.

Third, MPT is the important factor that enhances the surface roughness in coarse grain of SUS 304. On the other hand, MPT has little effect in increasing the surface roughness of fine grain SUS 304 thin metal foil due to different deformation mechanisms.

Fourth, the coefficient for relation of $\Delta R_a$ and $D_g \cdot \varepsilon$ does not depend only on material but also on microstructure-dependent deformation.

## References

1. Xue, Z.-Y.; Zhou, S.; Wei, X.-C. Influence of pre-transformed martensite on work-hardening behavior of SUS 304 metastable austenitic stainless steel. *J. Iron Steel Res. Int.* **2010**, *17*, 51–55. [CrossRef]
2. Engel, U.; Eckstein, R. Microforming—From basic research to its realization. *J. Mater. Process. Technol.* **2002**, *125–126*, 35–44. [CrossRef]
3. Furushima, T.; Hitomi, T.; Yutaro, H. Fracture and surface roughening behaviors in micro metal forming. In Proceedings of the 17th International Confrence on Metal Forming, Metal Forming 2018, Toyohashi, Japan, 16–19 September 2018; Procedia Manufacturing. Volume 15, pp. 1481–1486.
4. Milad, M.; Zreiba, N.; Elhalouani, F. The effect of Cold Work on Structure and Properties of AISI 304 Stainless Steel. *J. Mater. Process. Technol.* **2008**, *203*, 80. [CrossRef]
5. Jha Abhay, K.; Sivakumar, D.; Sreekumar, K. Role of Transformed Martensite in the Cracking of Stainles Steel Plumbing Lines. *Eng. Fail. Anal.* **2008**, *15*, 1042.
6. Huang, J.X.; Ye, X.N.; Xu, Z.J. Corrosion Resistance of High Performance Weathering Steel for Bridge Building Applications. *Iron Steel Res. Int.* **2012**, *19*, 59–63. [CrossRef]

7. Peng, F.; Dong, X.H.; Liu, K.; Xie, H.Y. Effects of strain rate and plastic work on martensitic transformation kinetics of austenitic stainless steel 304. *J. Iron Steel Res. Int.* **2015**, *22*, 931–936. [CrossRef]

8. Tomita, Y.; Iwamoto, T. Constitutive modeling of TRIP steel and its application to the improvement of mechanical properties. *Int. J. Mech. Sci.* **1995**, *37*, 1295–1305. [CrossRef]

9. Zandrahimi, M.; Bateni, M.R.; Poladi, A.; Szpunar, J.A. The formation of martensite during wear of AISI 304 stainles steel. *Wear* **2007**, *263*, 674–678. [CrossRef]

10. Zihao, Q.; Yong, X. Role of strain—Induced martensitic phase transformation in mechanical response of 304L steel at different strain—Rates and temperatures. *J. Mate. Process. Tech.* **2020**, *280*, 116613.

11. Jeom, Y.C.; Jung, H.J.; Si, W.H.; Kyung, T.P. Strain induced martensitic transformation of Fe-20Cr-5Mn-0,2Ni duplex stainless steel during cold rolling: Effects of nitrogen addition. *Mater. Sci. Eng. A* **2011**, *528*, 6012–6019.

12. Yamaguchi, K.; Mellor, P.B. Thickness and Grain Size Dependence of Limit Strains in Sheet Metal Stretching. *Int. J. Mech. Sci.* **1976**, *18*, 85–90. [CrossRef]

13. Osakada, K.; Oyane, M. On the Roughening Phenomena of Free Surface in Deformation Process. *Trans. Jpn. Soc. Mech. Eng.* **1970**, *36*, 1017–1022. [CrossRef]

14. Fukuda, M.; Yamaguchi, K.; Takakura, N.; Sakano, Y. Roughening Phenomenon on Free Surface of Products in Sheet Metal Forming. *J. Jpn. Soc. Technol. Plast.* **1974**, *15*, 994–1002.

15. Yamaguchi, K.; Takakura, N.; Fukuda, M. FEM simulation of surface roughening in FCC metals Using Direct Numerical Simulation. *Acta Mater.* **2004**, *52*, 5791–5804.

16. Vollertsen, F.; Schulze Nichoff, H.; Hu, Z. State of the art in micro forming. *Int. J. Mach. Tools Manuf.* **2006**, *46*, 1172–1179. [CrossRef]

17. Meng, B.; Fu, M.W. Size effect on deformation behavior and ductile fracture in microforming of pure copper sheets considering free surface roughening. *Mater. Des.* **2015**, *83*, 400–412. [CrossRef]

18. Stoudt, M.R.; Hubbard, J.B.; Leigh, S.D. On the relationship between deformation-induced surface roughness and plastic strain in AA5052. *Met. Mater. Trans. A* **2011**, *42*, 2668–2679. [CrossRef]

19. Tsuyoshi, F.; Hitomi, T.; Manabe, K.-I.; Yang, M.; Sergei, A. Influence of Free Surface Roughening on Ductile Fracture Behavior Under Uni-axial Tensile State For Metal Foils. In Proceedings of the 13th International Conference on Fracture, Beijing, China, 16–21 June 2013.

20. Raabe, D.; Schahtleber, M.; Weiland, H.; Scheele, G.; Zhao, Z. Grain-scale micromechanics of polycrystal surfaces during plastic straining. *Acta Mater.* **2003**, *51*, 1539–1560. [CrossRef]

21. Furushima, T.; Tsunezaki, H.; Manabe, K.; Alexandrov, S. Ductile fracture and free surface roughening behaviors of pure copper foils for micro/meso-scale forming. *Int. J. Mach. Tools Manuf.* **2014**, *76*, 34–48. [CrossRef]

22. Zhang, L.; Xu, W.; Cheng, L.; Ma, X.; Long, J. Quantitative Analysis of Surface Roughness Evolution in FCC Polycrystalline Metal During Uniaxial Tension. *Comput. Mater. Sci.* **2017**, *132*, 19–29. [CrossRef]

23. Kengo, Y. Effect of Grain Scale Heterogeneity on Surface Roughness and Sheet Metal Necking. *Int. J. Mech. Sci.* **2014**, *83*, 48–56.

24. Ichiro, S.; Takao, O.; Takeji, A.; Hideaki, T. Surface Roughening and Deformation of Grains during Uniaxial Tension of Polycrystalline Iron. *JSME Int. J. Ser. A* **2001**, *44*, 499–506.

25. Peng, L.; Xu, Z.; Gao, Z.; Fu, M. A Constitutive Model for Metal Plastic Deformation At Micro/Meso Scale With Consideration of Grain Orientation And Its Evolution. *Int. J. Mech. Sci.* **2018**, *138–139*, 74–85. [CrossRef]

26. Tsuyoshi, F.; Hitomi, T.; Tomoko, N.; Manabe, K.-I.; Alexandrov, S. Prediction of Surface Roughening and Necking Behavior for Metal Foils by Inhomogeneous FE material Modelling. *Key Eng. Mater.* **2013**, *554–557*, 169–173.

27. Groche, P.; Schafer, R.; Justinger, H.; Ludwig, M. On the correlation between crystallographic grain size and surface evolution in metal forming process. *Int. J. Mech. Sci.* **2010**, *52*, 523–530. [CrossRef]

28. Cheng, C.; Wan, M.; Meng, B. Size Effect on the forming limit of sheet metal in micro-scaled plastic deformation considering free surface roughening. *Procedia Eng.* **2017**, *207*, 1010–1015. [CrossRef]

29. Abdul, A.; Yang, M. Effect of Martensitic Transformation on The Surface Roughening of SUS 304 and 316 thin metal foils. *Eng* **2020**, *1*, 167–182.

30. Tsuyoshi, F.; Tomoko, N.; Kanta, S. A new theoretical model of material inhomogeneity for prediction of surface roughening in micro metal forming. *CIRP Ann.-Manuf. Technol.* **2019**, *68*, 257–260.

31. Stoudt, M.R.; Ricker, R.E. The relationship between grain size and the surface roughening behavior of Al-Mg Alloys. *Metall. Mater. Trnsactions A* **2002**, *33*, 2883–2889. [CrossRef]

32. Tsuyoshi, F.; Tetsuro, M.; Manabe, K.-I.; Alexandrov, S. Prediction of Free Surface Roughening by 2D and 3D model considering Material Inhomogeneity. *J. Solid Mech. Mater. Eng.* **2011**, *5*, 978–990.

33. Tetsuhide, S.; Masahiro, O.; Ming, Y.; Manabe, K.-I. Plastic anisotropy of ultra thin rolled phosphor bronze foils and its thickness strain evolution in micro-deep drawing. *Mater. Des.* **2014**, *56*, 604–612.

34. Su, Y.; Song, R.; Wang, T.; Cai, H.; Wen, J.; Guo, K. Grain size refinement and effect on tensile properties of a novel low-cost stainless steel. *Mater. Lett.* **2020**, *260*, 126919. [CrossRef]

35. Olson, G.B.; Cohen, M. The formability of austenitic stainless steels. *Met. Trans.* **1975**, *6*, 791–795. [CrossRef]

36. Xu, J.; Guo, B.; Shan, D.; Li, M.; Wang, Z. Specimen Dimension, Grain Size Effect and Deformation Behavior in Micro Tensile of SUS 304 Stainless Steel Foil. *Mater. Trans.* **2013**, *54*, 984–989. [CrossRef]

# Ultrasonic Effects on Foam Formation of Fruit Juices during Bottling

**Julian Thünnesen** [1,*]**, Bernhard Gatternig** [1,2,3] **and Antonio Delgado** [1,2]

1   Institute of Fluid Mechanics, Friedrich-Alexander University Erlangen-Nuremberg, Cauerstr. 4,
    91058 Erlangen, Germany; bernhard.gatternig@fau.de (B.G.); antonio.delgado@fau.de (A.D.)
2   German Engineering Research and Development Center LSTME Busan, Busan 46742, Korea
3   Hochschule Weihenstephan·Triesdorf, Verfahrenstechnik und Kreislaufwirtschaft, Markgrafenstr. 16,
    91746 Weidenbach, Germany
*   Correspondence: julian.thuennesen@fau.de

**Abstract:** Non-carbonated fruit juices often tend to foam over during bottling. The resulting foam height corresponds to the equilibrium of foam formation and decay. Therefore, the foam unexpectedly occupies more space in the bottle and carries parts of the juice out of the bottle, resulting in product loss under filled containers and hygienic problems in the plant. Chemical antifoams are likewise undesirable in most cases. Recent ultrasonic defoamers are effective but only capable outside the container and after the filling. In this article, a lateral ultrasonication through the bottle wall with frequencies between 42 and 168 kHz is used in-line for non-invasive foam prevention during filling. Foam formation during hot bottling of orange juice, apple juice, and currant nectar at 70 °C happens at flow rates between 124–148 mL/s. The comparably high frequencies have a particular influence on the fresh foams, where a large fraction of small resonant bubbles is still present. Foam volume reductions of up to 50% are reached in these experiments. A low power of 15 W was sufficient for changing the rise of entrained bubbles and minimizing the foam development from the start. The half-life of the remaining foam could be reduced by up to 45% from the reference case. The main observed effects were a changed rise of entrained bubbles and an increased drainage.

**Keywords:** ultrasound application; glass bottles; fruit juice; foam drainage; foaming behavior

## 1. Introduction

In the food and beverage industry, products tend to foam up unintentionally during processing due to their chemical composition. This phenomenon can be observed in the bottling not only of carbonated beverages but also of non-carbonated fruit juices. Similar to everyday pouring into a jar, in the worst case, the foam increases in volume disproportionately to the actual liquid poured in, takes up the planned space in the vessel, and threatens to overflow. Particularly in the process of hot filling already-pasteurized beverages into glass bottles, the containers are sometimes considerably underfilled, as the headspace fills with foam before the required legal minimum volume has been reached. Below this volume, a bottle of this type may not be sold, which means that the economic loss corresponds with a loss in sales price. The result is lower line efficiency by up to 5%, underfilling of the containers, higher production losses, and greater cleaning efforts for the line. Therefore, it is already of interest to medium-sized companies to increase plant utilization by even 1% through foam prevention and thus save up to €50,000 per year [1].

In practice, it is only possible to passively reduce foam formation by reducing the filling speed, since a decreased velocity of the free jet carries less gas from the headspace into the liquid. The entrained bubbles rise to the surface and agglomerate to form the foam. The resulting foam height is obtained from the net foam balance, which is the time ratio of foam formation to foam decay. Especially in the short time window of a few seconds for bottle filling, foam decay is dominated by drainage, where films and plateau edges

lose liquid and decay below a critical thickness [2–4]. Deposited particles, e.g., from pulp, and capillary forces in the plateau areas inhibit drainage; hence, foam persists longer than the time window permits. Naturally fluctuating material properties of the juices, such as the content of proteins, polyphenols, sugars, and soluble solids, etc., constantly change the formation and half-life of the foam, which is why it is constantly necessary to manually adjust the filling speed to avoid over-foaming despite optimized filling tube geometries [5–10].

A method of active and adaptive foam destruction during filling is currently not found in literature or in practice. A look at other industrial fields reveals chemical, thermal, and mechanical defoaming methods, but most of these are not applicable to bottling. Chemical methods, such as the use of defoamers, are limited due to strict national food laws and face the disadvantage of more expensive and environmentally hazardous consequences [11,12]. Therefore, mechanical foam destruction based on ultrasonic waves is of particular interest.

Ultrasound is noninvasive and propagates as a mechanical wave in gases, liquids, and solids, creating localized pressure differences. It is assumed that ultrasonic waves penetrate the lamellae, atomize the liquid from the lamellae, or hit the resonant frequencies of the bubbles, creating surface waves along the lamellae and enhancing drainage [13–17]. Current knowledge on ultrasound-based foam destruction is largely based on laboratory experiments with airborne ultrasound. Ultrasonic foam destruction was initially studied using sirens and Hartman whistles [18–20]. The frequency range from audible 0.7 to 29 kHz showed effective defoaming only at sound power levels of at least 145–148 dB [21]. Additional work on foam control in fermentation vessels showed that defoaming ability increased with higher frequencies from 26 to 34 kHz and rapidly disintegrated foam at high sound intensities of 10 W/cm$^2$ or 120 dB [20]. More recent work has investigated airborne sonication of the foams with piezo-acoustic actuators [14–16,22]. Most of the works used a sonotrode with a frequency of 20 kHz and considered the degree of defoaming dependent on the foam and power applied. It was found that ultrasound is more effective with decreasing viscosity and is just suitable for aqueous foams with average bubble diameters of 0.5–5 mm. However, higher foam formation rates and water contents in the case of foams with SDS required stronger amplitudes, since the liquid migrates only to the neighboring plateau areas when the top foam layer is destroyed and the foam thus loses little liquid hold-up [22,23]. With increasing liquid hold-up and lamella thickness, the reflection coefficient at the foam interface increase. The acoustic power to be applied is so high that Dedhia et al. [22] considered pulsed sonication in terms of economy.

However, for an inline application during filling, airborne high-power ultrasound has several significant drawbacks. For very wet foams, defoaming should not occur by cavitation formation and atomization of the liquid in the lamella. On the one hand, the generation of transient cavitation for a decent foam decay requires a sound intensity of about 100 W/cm$^2$, which is equivalent to a burnout of an heating wire for boiling water. This negatively influences the energy efficiency of the plant, and thermal product degradation could occur. On the other hand, the product is partially atomized, and an aerosol is produced that contaminates the filling system and thereby causes hygienic problems [24]. At the same time, these methods require ultrasonic actuators with relatively large diameters compared to common bottle dimensions. As a result, due to the limited accessibility to the foam and the small installation space, the sonotrode can only remove the excess foam after filling [14,15].

Apart from airborne insonication, Winterburn and Martin [25] demonstrated that sonication via the liquid also accelerates drainage at 40 kHz and at already 4 W total electrical power, as the impedance between air and water is improved. Although the pressure amplitude was too low to rupture the lamellae in the experiments, as is assumed for airborne ultrasound [14], the authors detected stronger effects at 40 kHz than at 28 kHz. However, a fundamental understanding of the mechanisms of foam destruction is lacking.

This leads to the research question of an innovative, energy-efficient, and product-friendly alternative in the form of a resonance-based method of foam destruction for

beverage filling using ultrasound. By using resonant effects, the acoustic power and thus the energy consumption can be significantly reduced. High temperatures and peak pressures from cavitation events no longer occur; thus, product degradation and aerosol formation may not occur. The lower power requirements allow for a coupling of the ultrasound transducer through the bottle wall and liquid phase, giving plant designers more flexibility for system integration.

The primary resonance excitation in foam promotes the generation of surface waves and lowers the viscosity of the liquid in the lamellae. Drainage increases, causing the dynamic foam height to be lower. In published literature, ultrasound has been used in food foams only to determine bubble size distribution and gas content, thus providing information on the frequency-specific resonance behaviors of bubbles [26–28]. In all work, bubbles are assumed to absorb sound in a frequency-specific manner and oscillate in the process. However, it is not yet fully understood how the resonance effects behave within the foam regarding to enhance defoaming. On the one hand, an increasing gas content in the foam lowers the resonant frequency of a bubble [28], while the presence of neighboring resonant bubbles at a distance of half a wavelength would increase the frequency [29]. In the latter case, bubbles with a radius of about 100 μm would resonate at 86 kHz instead of 31 kHz with a gas content of 26%.

The article addresses the following research objectives: (i) discovering which ultrasonic frequencies can minimize the resulting foam height during filling due to resonance effects. Frequencies that are still high enough to keep the potential of transient cavitation low (>40 kHz) and low enough that their wavelengths roughly correspond to the theoretical resonance lengths of the smallest lamellae or bubbles (<200 kHz) are considered. The bubble fractions, which tend to be smaller, are usually located at the bottom of the foam, which is why objective (ii) is whether sonication of the liquid through the wall of the bottle from the side or bottom can achieve the desired effect. This promises no aerosol formation and a lower transmission loss than with the air gap between the transducer and the foam. Finally, for objective (iii), the required ultrasound power and duration of sonication are investigated to deduce the energy requirements of the process. The results are considered in terms of industrial feasibility.

## 2. Materials and Methods

### 2.1. Experimental Set-Up

The aim of the tests was to investigate the influence of different frequencies on foam formation during atmospheric filling under realistic conditions. The test conditions were based on the standard filling specifications for fruit juices in refillable glass bottles, in which a maximum filling time of 10 s is specified. The required volume flow was achieved via the height difference between the reservoir tank (100 × 100 × 250 mm) and the filling tube opening of 930 mm (Figure 1).

Before each filling, 1.2 L of the juice, previously tempered to 70 °C, was taken from a thermobath (3 L) to a reservoir tank. This quantity of juice allowed for the avoidance of additional air intake into the pipe towards the end of the filling.

A magnetic valve and a flow meter below the reservoir tank controlled the filling interval and recorded the flow rate, respectively. The juice entered the bottle via a height-adjustable filling nozzle (length: 250 mm, inner diameter 8 mm). The adjusted height between the nozzle outlet and the bottom of the bottle was 250 mm for the trials with apple juice and currant nectar. For orange juice, the filling nozzle was set at a height of 50 mm. This allowed the foam heights within replicates to be kept as low as possible.

A CCD camera (acA2500-60 uc, Basler GmbH, Ahrensburg, Germany) recorded foam formation at 1 fps during filling. Control of the valve and data recording were done centrally via a mini-PC (Raspberry Pi 4 Model B). The sampling period of all sensors was synchronized to the camera's recording rate and was 1.03 s between samples.
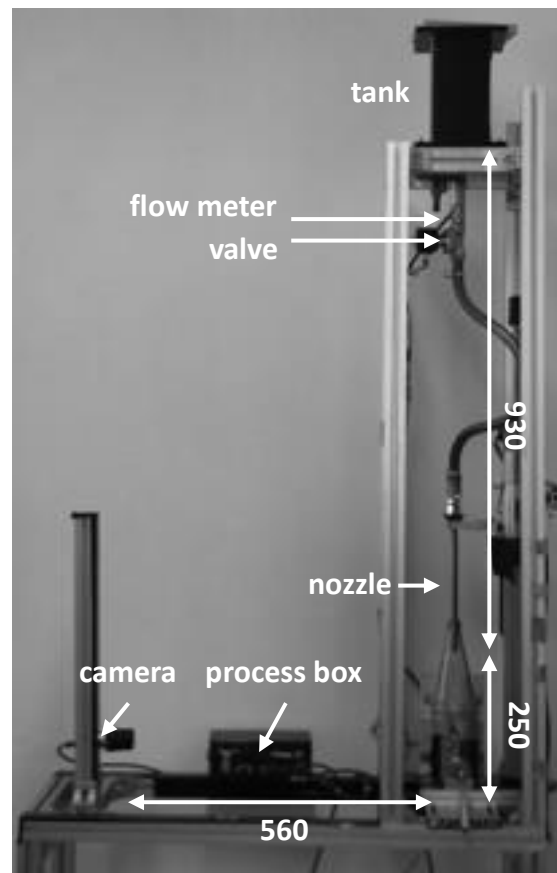
**Figure 1.** Filling setup (here: without heating bath) with given distances [mm].

The individual filling cycles were divided into fixed time periods: initialization (1 s), filling (10 s), and decay (9 s) phases. During the filling phase, the volume flows were kept constant for each of the juices (Table 1). After the filling phase, the decay phase additionally served to assess the pure foam decay during sonication.

**Table 1.** Measured flowrate of currant nectar, apple juice, and orange juice at 70 °C.

| Flowrate (L/s) | Black Currant | Apple | Orange |
|---|---|---|---|
| Mean (n = 630) | 0.1277 | 0.1361 | 0.1480 |
| Std (n = 630) | 0.0015 | 0.0059 | 0.0045 |

The vessel used was a 300 mm high commercial glass bottle (model: 1 L VdF bottle) consisting of a 149 mm high cylinder and a cone above it. After measuring three different bottles with a caliper, the inner radius *r* can be described as a function of the height *h*:

$$r(h)[\text{mm}] = \begin{cases} 43.75, & h < 149 \text{ mm} \\ 0.207 * (149.5 - h) + 43.75, & h \geq 149 \text{ mm} \end{cases} \quad (1)$$

A clamp fixed the bottle to the ultrasonic actuator on the side, which rested against the cylindrical wall of the bottle at a height of 30–70 mm (Figure 2a).

(**a**)                                             (**b**)
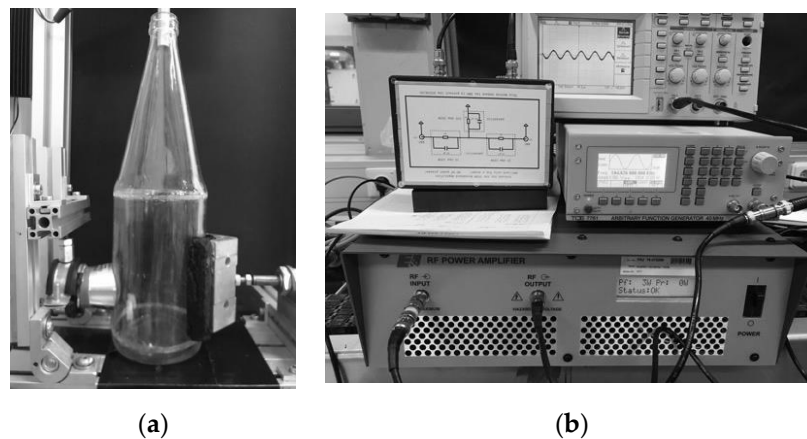
**Figure 2.** Fixing device for the bottle to the modified transducer (**a**) and amplifier system (**b**).

The ultrasonic transducer was a modified 40 kHz Langevin transducer (Hesentec, Rank E), which had a front face adapted to the curvature of the bottle wall (radius r = 44 mm). This modification changed the resonant frequencies of the transducer to 42, 56, 85, 101, and 168 kHz. A arbitrary function generator (TOE7761, Toellner, Herdecke, Germany) produced a sinusoidal signal of the corresponding frequency, which was amplified by 51 dBV through a voltage amplifier (1040 L, Electronics and Innovation, LTD, Rochester, NY, USA). A parasitic resistor box was connected between the amplifier and the transducer to protect the amplifier from changing impedances of the transducer (Figure 2b).

Power outputs at 10, 15, and 20 $W_{el}$ were investigated. The maximum generated pressure amplitudes inside a filled bottle are shown in Table 2 according to hydrophone measurements.

**Table 2.** Maximum pressure amplitudes of output frequencies.

| $W_{el}$ | $p$ ($\times 10^5$ Pa) | | | | |
|---|---|---|---|---|---|
| | 42 kHz | 56 kHz | 85 kHz | 101 kHz | 168 kHz |
| 10 | 1.033 | 0.078 | 2.376 | 0.761 | 0.666 |
| 15 | 2.284 | 0.1 | 3.036 | 1.332 | 0.772 |
| 20 | 2.773 | 0.119 | 3.63 | 1.543 | 0.772 |

*2.2. Juices*

Orange juice (100% fruit content with pulp, Erwin Dietz GmbH, Osterburken, Germany), filtered apple juice (100% fruit content, Erwin Dietz GmbH, Osterburken, Germany), and currant nectar (25% fruit content, Granini, Nieder-Olm, Germany) tempered to 70 °C were used.

Their density, viscosity, and surface tension were triply determined using a hydrometer, a capillary viscometer, and a contact angle meter (Wilhelmy plate method, K100, Krüss, Hamburg, Germany) at 70 °C as shown in Table 3.

**Table 3.** Density, viscosity, and surface tension of apple juice, orange juice, and currant nectar at 70 °C.

| | Density (kg/m$^3$) | Viscosity (mPa $\times$ s) | Surface Tension (mN/m) |
|---|---|---|---|
| Apple juice | 1014.82 | 0.498 | 28.06 |
| Orange juice | 1011.50 | 2.344 | 26.85 |
| Black currant nectar | 1018.50 | 0.450 | 38.47 |

*2.3. Statistical Methods*

A total of 108 fillings were considered with six replicates per juice, power level and frequency. Results are given as means with standard deviations. For the statistical

evaluation with a two-factorial ANOVA and the calculation of the Spearman correlation coefficient in Excel 2016, the foam heights at time $t$ = 8 s were used. At this time, the foam is still in the cylindrical region of the bottles and is not yet affected by the increasing influences during bottle tapering, which would shift the normal distribution to a right-handed one. A two-factor ANOVA (Excel 2016) revealed that trials differed between juices ($p$-value = $2.51 \times 10^{-42}$), sonication significantly decreased foam ($p$-value = $1.86 \times 10^{-11}$), and interactions between frequencies and juices also occurred ($p$-value = $1.035 \times 10^{-16}$).

### 2.4. Measurement of Foam Heights

The temporal development of foam was recorded via the camera's image series. Only the pixel lengths in vertical direction were considered. To this end, constant heights of the bottle opening $y_{bottle,\ top}$ and bottom $y_{bottle,\ bottom}$ were measured once for each run, and the heights of the top and bottom edges of the foam were determined ($y_{foam,\ top}$ and $y_{foam,\ bottom}$) for the individual images. The origin of the coordinate system in the image is in the upper left corner, thus accounting for the y-values increase towards the bottom. The process was performed using a semi-automated script in OpenCV, i.e., foam edges were manually selected and the according y-values were calculated and saved automatically. The resulting foam height $h_f$ and liquid height $h_l$ were calculated by taking the difference of the respective y-values and including the conversion ratio $R$ of the image.

$$R \left[ \frac{Px}{mm} \right] = \frac{y_{bottle,\ bottom} - y_{bottle,\ top}}{h_{bottle}} \tag{2}$$

$$h_l\ [mm] = \frac{y_{bottle,\ bottom} - y_{foam,\ bottom}}{R} \tag{3}$$

$$h_f\ [mm] = \frac{y_{foam,\ bottom} - y_{foam,\ top}}{R} \tag{4}$$

Finally, the foam volume was calculated from the volume of a truncated cone with the radii according to equation 1 and the foam height. In the region between the cylinder and cone at $h_l$ < 149 mm and $h_l + h_f$ > 149 mm, the foam volume was the sum of a truncated cone of a height $h_1 = (h_l + h_f - 149$ mm$)$ and a cylinder of $h_2 = (149$ mm $- h_f)$.

### 2.5. Uncertainty Analysis of Measurements

The heights of the bottle, foam, and liquid were dependent on the accuracy of the computer display. At the same time, the edges of the foam in the image were not always horizontal, resulting in measurement uncertainty when measuring the average height.

The single-sample analysis by Moffat calculated the uncertainties separately. Measurements of two filling tests with 20 images each provided the data. Each bottle edge was measured in the respective image. For the changing foam and filling heights, the y-values were determined tenfold in each image. The uncertainties were caused by the repetition and measurement errors by the author and are indicated by the standard deviation σ:

$$\sigma = \sqrt{\frac{1}{(N-1)} \sum_{j\,=\,1}^{N} \left( \bar{y} - y_j \right)^2} \tag{5}$$

The standard deviations are shown in Table 4. The relative errors of the foam edges varied depending on the fill level, which is why they are not shown in the table. Instead, the relative errors of the foam height based on them are given later.

**Table 4.** Uncertainties $dX_i$ and the relative uncertainties $\frac{\sigma}{X_i}$ of individual measured variables $X_i$, given in the respective units, n = 40.

| $\delta X_i$ | $\delta X_i = \sigma$ | $\frac{\sigma}{X_i} \times 10^2$ [%] |
|---|---|---|
| $y_{bottle,\ bottom}$ [Px] | 3.16 | 0.21 |
| $y_{bottle,\ top}$ [Px] | 1.71 | 2.12 |
| $y_{foam,\ bottom}$ [Px] | 2.76 | |
| $y_{foam,\ top}$ [Px] | 5.45 | |
| *bottle height* [mm] | 0.50 | 0.2 |

The uncertainties contributed by multiple input variables were calculated by a combination of the root-sum-square method:

$$\delta R = \sqrt{\sum_{i=1} \left( \frac{\delta R}{\delta X_i} \sigma \right)^2} \tag{6}$$

where the measurement uncertainty $\delta R$, is characterized by the individual uncertainties of the measured variables $\delta X_i = \sigma$. To calculate the uncertainties of the conversion ratio, the foam and the liquid height, the Equation (2) were, respectively, put into (3) and (4) and partially derived according to the individual variables $X_i$.

For the conversion coefficient $R$, the measurement uncertainty was $\delta R$ = 0.014 Px/mm and a relative error of 0.31%. Table 5 shows the absolute measurement uncertainties of the liquid height $\delta h_l$ and the foam height $\delta h_f$, as well as their relative deviations for representative heights, respectively.

**Table 5.** Uncertainties $dX_i$ and the relative uncertainties $\frac{\sigma}{X_i}$ at given values of liquid height $h_l$ and foam height $h_f$.

| Liquid Height $h_l$ | | | Foam Height $h_f$ | | |
|---|---|---|---|---|---|
| $X_i$ (mm) | $dX_i$ (mm) | % error | $X_i$ (mm) | $dX_i$ (mm) | $\frac{\sigma}{X_i} \times 10^2$ (%) |
| 99.93 | 1.39 | 1.24 | 5.40 | 1.36 | 25.20 |
| 88.87 | 1.38 | 1.37 | 9.93 | 1.36 | 13.70 |
| 77.82 | 1.38 | 1.54 | 18.13 | 1.36 | 7.50 |
| 66.77 | 1.37 | 1.75 | | | |
| 55.72 | 1.37 | 2.03 | | | |
| 44.67 | 1.37 | 2.43 | | | |
| 33.62 | 1.36 | 3.03 | | | |
| 22.57 | 1.36 | 4.02 | | | |
| 11.51 | 1.36 | 5.98 | | | |
| 0.46 | 1.36 | 11.71 | | | |

Deviations < $\pm$1.37 mm around the mean foam heights shown below are due to measurement uncertainty. Above the value, the deviations are attributable to the system.

## 3. Results and Discussion

### 3.1. Evolution of Foam Formation

During the filling phase, three phases of foam formation were evident in all reference tests of the juices (see Figure 3). Initially, the filling jet hit the bottom of the bottle, causing the juice to rise radially along the bottle wall and trap air as it falls back to the center. The flow is chaotic in the first second, which means that the liquid level is not horizontal. Almost two seconds later at a filling height of 45 mm, the entrapped bubbles rise to the surface and form the first layer of foam. In the following phase, continuous foam formation occurs in the cylindrical part of the bottle due to the introduced of air by the free jet. The foam formation happens almost constantly during this phase. The free jet deforms

the gas–liquid interface and causes the formation of gas pockets. If the destabilizing inertial forces caused by the free jet outweigh the stabilizing surface forces, the interface is strongly deformed and bubble entrapment occurs [17]. According to the literature, the gas entrapment shows a nonlinear dependence on the liquid volume flow rate [17]. At the same time, foam development depends on the ascent time of the bubbles rising to the surface. This time either depends on the diameter of the gas bubbles, the viscosity and density of the juice, or the ratio of bubble ascent speed to filling speed. In the case of orange juice, no additional gas entrainment occurred at a filling level of 50 mm due to the already immersed filling nozzle. As such, the foam volume remained constant with some delay due to rise of the bubbles from the 70 mm filling level.
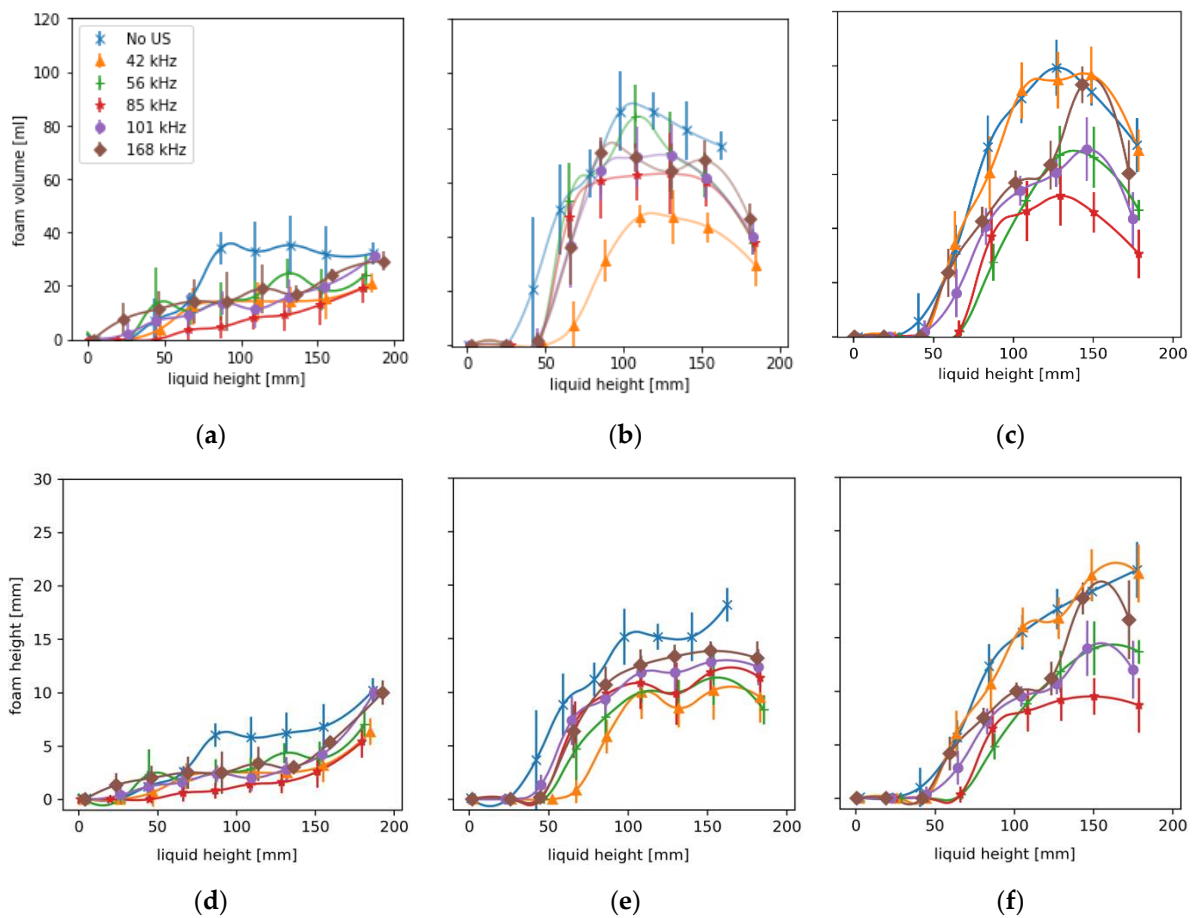


**Figure 3.** Effect of frequency dependent, permanent ultrasound insonication on the evolution of foam volume (**a**–**c**) and foam height (**d**–**f**) for orange juice (**a**,**d**), apple juice (**b**,**e**), and currant nectar (**c**,**f**), each at 70 °C, n = 6.

In the final phase, the foam enters the cone at a height of 149 mm where the foam volume remains constant in the case of orange juice and decreases in the case of apple juice and currant nectar. Due to the cross-sectional taper of the cone, shear forces along the wall crushed larger bubbles in the foam. This shear causes a reduction in volume but not to the extent that it causes the foam heights to remain constant while the cross-sectional taper is reduced (compare Figure 3).

The effectiveness of forced foam decay was very consistent at power levels 10, 15, and 20 W, while it showed a larger dependency on frequency. Therefore, the presented results of 15 W also represents the other power levels. As seen in Figure 3, sonication immediately causes less foaming at a liquid level between 45 and 70 mm. The delayed and lower foaming that occurred initially was independent of the frequency, height of the filling nozzle, and juice, but this differs somewhat in the later phases. With a low nozzle attachment for orange juice, the frequencies 42, 56, and 85 kHz are most effective.

Specifically, sonication at 42 and 85 kHz delays formation until filling is concluded and causes a lower foam height. The higher frequencies, 101 and 168 kHz, only allow more foam formation from a filling height of 130 mm compared with the reference tests.

In the case of a high nozzle attachment for apple juice and currant nectar, sonication also caused a delayed foam formation. Especially in apple juice, the frequencies 42, 85, 101, and 168 kHz significantly reduced foam formation at 80 mm filling level when the transducer fully submersed in the liquid. Furthermore, ultrasound at frequencies 56, 101, and 168 kHz reduced foam formation also in currant nectar. The frequency of 85 kHz is particularly effective here, where the first foam just formed at 70 mm. Meanwhile, 42 kHz provoked a higher foam volume.

As aforementioned, foam formation primarily happened through gas entrainment of the free jet, whose generated bubble size distribution depends on the viscosity, density, and surface tension of the respective juice. Accordingly, the foam characteristics are dependent on the respective juices. Figure 4 shows the resulting foam heights after filling against the applied sonic frequency and density of the respective juice. The main influence of the foam's height happened along the density. This is obvious in the case of orange juice and a lower filling nozzle. However, it is interesting to note that, compared with the respective reference tests, sonication reduces the foam height by about 50% on average at frequencies between 42 and 101 kHz. This reduction becomes less at 169 kHz.
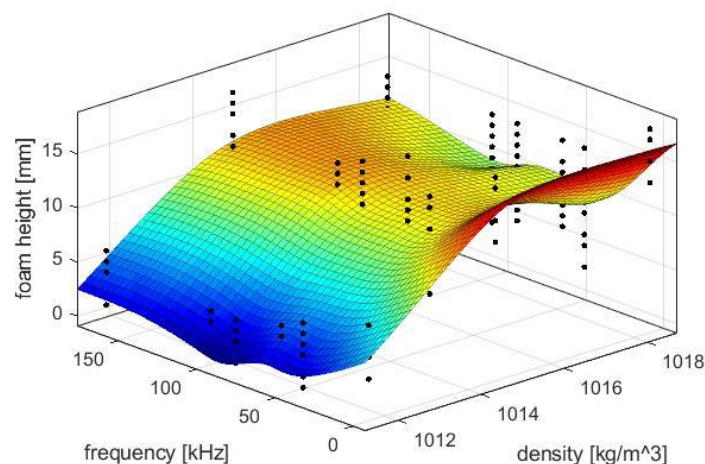


**Figure 4.** Resulting foam heights after filling at *t* = 8 s plotted against the density of the juice and the frequency used.

Figure 4 accordingly shows that the absolute foam height at the end of the respective filling at *t* = 8 s differs between the respective juices but is consistently reduced by the respective frequencies. Especially the frequencies 85 and 101 kHz reduce foam within a respective juice.

Figure 5 shows that the foam formation of apple juice and currant nectar is less inhibited by a delayed sonication starting at 50 mm. However, a strikingly lower foam formation is observed from a height of 60–70 mm. Especially for the frequencies 101 and 168 kHz, a later insonification was more effective for foam decay. Meanwhile, more foam developed throughout at the frequencies 42 and 85 kHz. This is most likely because some bubbles already reached the surface and formed foam before insonication started.
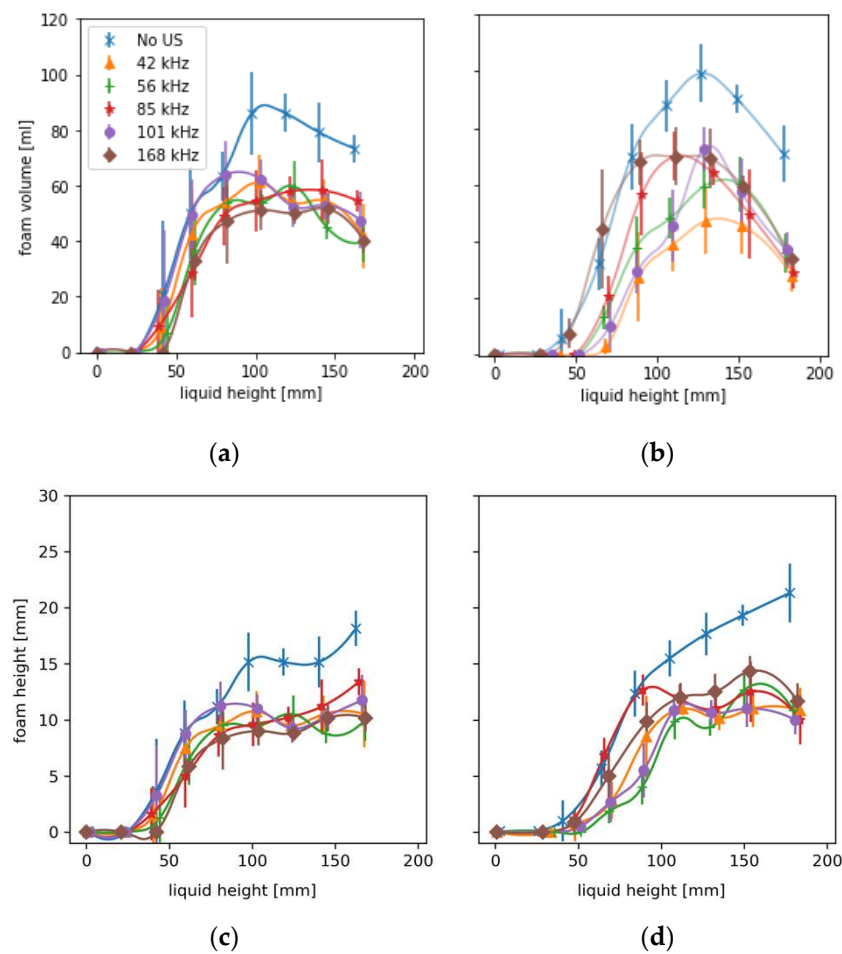
**Figure 5.** Effect of frequency depended ultrasound insonication with 2 s delay on the evolution of foam volume (**a**,**b**) and foam height (**c**,**d**) for apple juice (**a**,**c**) and currant nectar (**b**,**d**), each at 70 °C, n = 6.

## 3.2. Spatial and Temporal Effects

In addition to the respective heights in Figures 3 and 5, the captured images showed qualitative information on bubble rise and foam structure. The spatial change in bubble rise was already obvious at the beginning of filling at a liquid height of 30 mm, in which characteristic bubble-free areas are formed directly in front of the transducer and on the opposite side (Figure 6). Here, ultrasound caused pressure nodes and antinodes, which direct the bubbles through corridors to the surface. At the same time, the bubbles experience Bjerknes forces that first hold the bubbles in pressure antinodes and then cause two oscillating bubbles to coalesce [30,31]. This increases the ascent time directly due to the ultrasonic influence and indirectly due to the larger bubble diameter. As a result, the supply of new bubbles to the foam is interrupted, and the ratio of formation effects to decay effects decreases. This effect was evident again at a filling height of about 70 mm in the juices, where the transducer was lower than the liquid level and foam formation decreased (see Figures 3 and 5).
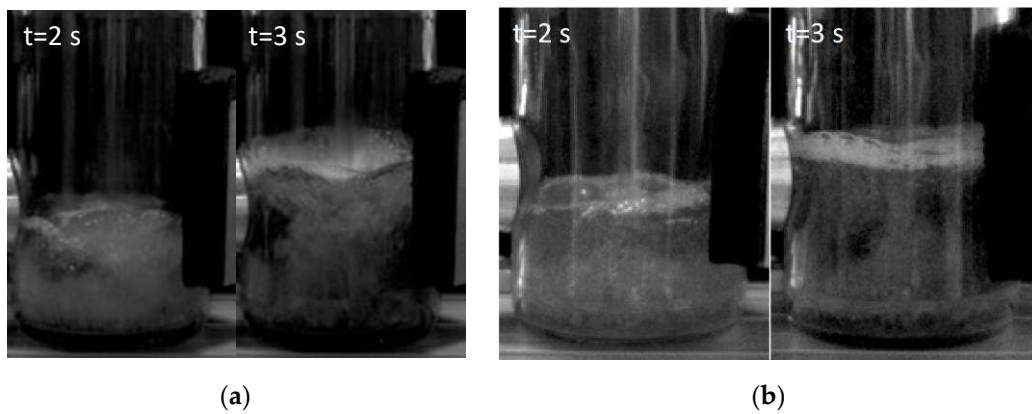
**Figure 6.** Bubble-free regions during filling time events $t = 2$ s and $t = 3$ s for apple juice at 45 kHz, 15 W (**a**) and unchanged bubble distribution without ultrasound (**b**).

The more effective the foam reduction by the ultrasound was, the drier and coarser the foams were by the end of filling. This was particularly evident in black currant nectar and apple juice, where defoaming was most effective (Figure 7). In return, the foams were denser when the ultrasonic effect was absent. Especially in apple juice, the coalescence of the bubbles happened in the end of the filling (Figure 7g–l). This also suggests the explanation that the larger rising bubbles coarsen the foam, thus minimizing its half-life. As a result, drier foams with lower foam heights are present towards the end of the filling process. These observations and the changed bubble rise in Figure 6 show that ultrasound already has an influence on the resulting foam during bubble ascent. However, low frequencies between 20 and 40 kHz bear the risk that, below a certain sound power, the pressure areas are not sufficiently developed in the liquid, whereby shear forces also caused by the sound waves are dominant. In this case, bubbles are dispersed, resulting in undesirable fine foams, as in the case of sonication of currant nectar at 42 kHz [31].
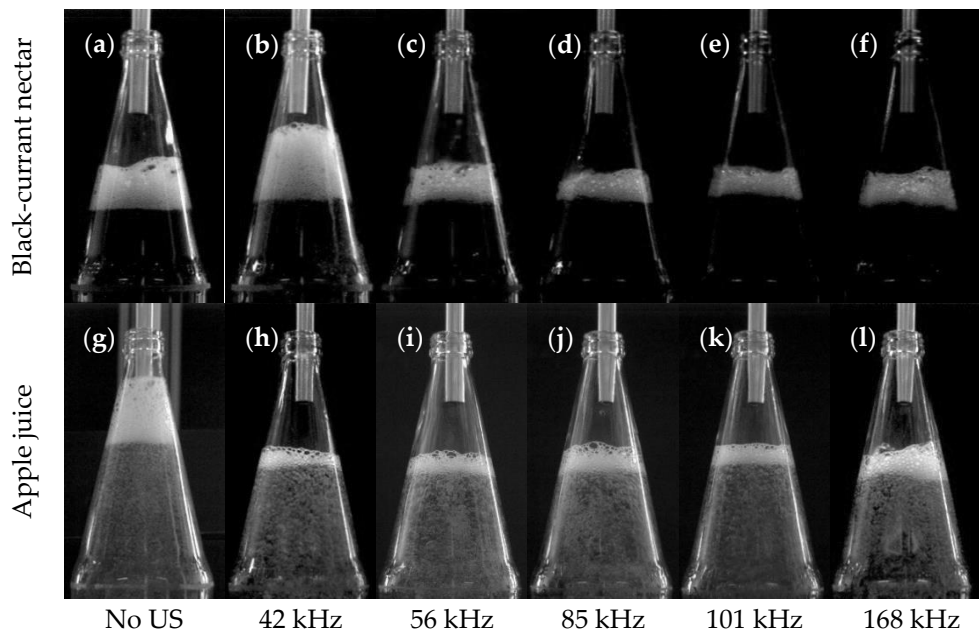


**Figure 7.** Insonification causes coarser, reduced foams of black currant nectar (**a**–**f**) and apple juice (**g**–**l**) and larger ascending bubbles in apple juice at the end of filling $t = 10$ s (**h**–**l**).

### 3.3. Reduced Half-Lives Due to Ultrasound

After filling, sonication continued for another 9 s at the respective frequency and power to enhance foam decay. At that point, no more gas bubbles ascended to the surface and no foam formation happened, resulting in pure foam decay.

The foam volume decayed exponentially during the observation time of 9 s according to the natural decay $V(\tau) = V_0 e^{-\lambda \tau}$, where $V_0$ is the respective foam volume directly after the end of filling, $\lambda$ is the decay constant, and $\tau$ is the time in seconds after filling. The half-lives $\tau_{1/2} = \ln(2)/\lambda$ derived from this are shown normalized to half-lives of the reference foam of the respective juice in Figure 8. Here it became clear that sonication reduced the half-lives of the foam volume to between 63% (101 kHz) and 45% (85 kHz) in apple juice. The foams decayed completely except for a few individual foam bubbles. In the case of currant nectar on the other hand, the half-lives reduced to between 80% (101 kHz) and 57% (85 kHz), as the foam dried out quickly but larger bubbles remained. Nevertheless, the two juices showed similar frequency-dependent decay effects between 56 and 168 kHz.
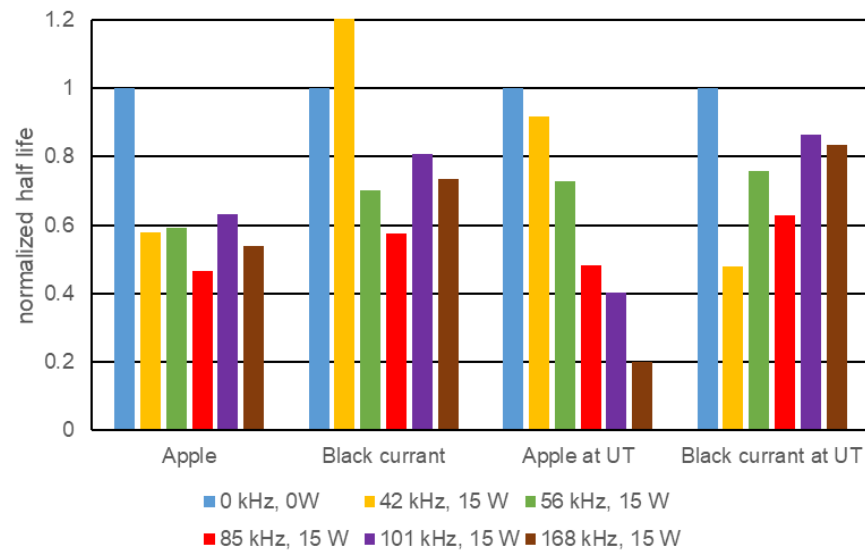


**Figure 8.** Normalized half-lives of the foams for all juices and frequencies related to reference half-lives (Apple: 5.5 s, black currant: 5.66 s, apple at UT: 8.77 s, black currant at UT: 4.5 s), n = 6.

## 4. Discussion

### 4.1. Impact of Liquid-Guided Ultrasound

In the experiments, the sonication occurred from the side through the bottle wall, whereby the waves propagated mostly below the foam due to the rising liquid height. It was reasoned that attaching the ultrasonic transducer to the bottle wall would overcome the severe acoustic impedance differences and high energy requirements of arrangements that included an air gap between the transducer and foam. In this respect, the required power of 15 W was far less than the 200 W of the airborne ultrasonic systems used to date [14–16]. This type of sonication differs from the common method, not only from the actuator design and power, but also from the defoaming behavior. The applied frequencies between 42 and 168 kHz might specifically induce resonant effects in the foam leading to higher drainage without atomization. These are well above the frequencies used by Winterburn [25], but within the range of McHardy [32].

The ratio of resonant bubble size to inserted frequency is simply assumed by the Rayleigh–Plesset frequency at first, assuming that the oscillation is linear and stable:

$$f_R = \frac{1}{2\pi r_R} \times \sqrt{\frac{3\gamma \left( p_0 + \frac{2\sigma}{r_R} \right) - \frac{2\sigma}{r_R}}{\rho_{Fluid}}} \tag{7}$$

The resonance radii are 76, 68, 57, 32, and 19 µm for $f_R$ = [42, 56, 85 and 168 kHz] and surface tension σ of the corresponding juice, respectively. This range of resonant radii are to be considered in the first percentile when comparing bubble size distributions of juice foams after similar filling tests and bottle geometries [33]. Since foams become coarser towards the top, it can be assumed that such bubbles are more likely to be found on the bottom, inside Plateau-channels, or during early foam formation. Comparing the juices' surface tension at a particular frequency used in Equation (7), the shift of resonant radii is ±1 µm, which means that the ultrasound/bubble interactions depend on the present bubble sizes rather than on the physical properties of juice. Therefore, the sonication is more effective when the ultrasound can reach the spot of resonant bubbles in the bottle. This effect is particularly evident above a filling height of 30 mm, where the actuator can first sonicate the bottom of the foam and the foam curves deviate from the reference measurements. This ultrasound method is most effective during the early phase of foam formation, when the bubbles were still spherical and freely moving. It can be assumed that, in this phase, the bubbles resonate by twice the usual resonant frequency due to their interacting acoustic influence [29]. The bubbles create surface waves during resonance enhance the flow around them. This causes a better drainage and coalescence already during foam development, resulting in coarser and more unstable foams by themselves.

In already existed foams, a subsequent sonication accelerates the decay (see Figure 8). Together, this results in a faster decay of the foam volume than in other experiments with subsequent sonication of already existing foams [32]. Because of the high reflections at and inside the foam, the sonication into the bottom of the foam may only penetrate into the first layers, where the liquid is removed more quickly. A downward sinking liquid gradient is created, which, in contrast to sonication from above, tends to remove the liquid from the plateau areas at the top [22].

The second influence is the standing wave formed within the liquid. It also showed a foam-avoiding effect in that smaller entrained bubbles passed through the pressure antinodes according to the applied frequency. In the presence of ultrasound in a gas-containing liquid, the bubbles can undergo either stable or transient cavitation. The bubbles undergoing stable oscillations grow to the resonance size or twice the resonance size by rectified diffusion or by coalescence due to the Bjerknes forces [34]. The bubble enlargements observed in Figure 7 can be explained by the fact that, at the applied sound pressures, the stably oscillating bubbles tend to grow by rectified diffusion and reach the resonance size [35,36]. This has a positive impact on foam prevention and should be achieved as early as possible in the process, for example by attaching the actuator to the bottle at a low level.

However, the applied acoustic pressures were not sufficient to subsequently destabilize the bubbles at resonance to the degree that they decayed into additional smaller bubbles and created transient cavitation (except for the run with 42 kHz and black currant nectar) [37–39]. Varying the test sequence for applied frequencies also showed no temporal effect on foam formation, which proved reversible effects of the sonication. Above the aforementioned sound pressures, defoaming effects slowly diminish or even reverse. It can be assumed that with increasing gas saturation of the liquid, this upper pressure limit decreases. Fruit juices have a very low saturation during production before bottling, which is why they have a higher upper pressure limit.

### 4.2. Industrial Feasibility

The method developed is for purely mechanical foam destruction using ultrasound to minimize product loss and filling times. The advantage over previous methods of foam prevention is that foam is already reduced in the bottle without the actuator being in direct contact with the product. In practice, rotary fillers and linear fillers inject the respective juice into bottles by moving the bottles to the filling spout via a height-adjustable positioning table. The ultrasonic actuator can be mounted at the positioning table and be in contact with the bottle via a waveguiding coupling material. Compared to an airborne

ultrasonic defoamer with 200 W, 13 of such actuators can already control the foam in small fillers [15]. The actuator used generates low amplitudes, which meant that no abrasion could be detected on the bottle. The energy efficiency can still be optimized by switching to PET bottles with better transmission properties into the liquid or by improving the ultrasonic system. In the experimental setup presented, a parasitic series resistor was connected between the amplifier and the transducer. This compensated for the deviating electrical impedance mismatch of both components via thermal dissipation. An improved electrical circuit can lower that dissipation and improve the electrical efficiency of the ultrasonic system.

## 5. Conclusions

The results show the influence of sonication with ultrasound via the bottle wall on foam development during hot beverage filling at 70 °C. This type of sonication method differs from previous airborne ultrasonic systems due to higher frequencies and modest pressure amplitudes. The resonance effects in the bottom of the foam enhance the drainage and the decay, respectively. By avoiding transient cavitations, the juice is not degraded thermally or mechanically. The comparatively high frequencies between 42–168 kHz have an enhanced effect on wet, dense foams containing a large proportion of small resonant bubbles. A low power of 15 W was sufficient to change the rise of the entrained bubbles and minimize foam development right at the beginning. Compared to industrial airborne ultrasonic defoamers, the proposed method reduces foaming already during filling with 7.7% of the electrical power. Power dependence between 10–20 W and the harmonic relationship of the three most effective frequencies suggests a much stronger role of the natural frequencies of the liquid films or bubbles. The defoaming effects occurred mainly once the liquid level was above the ultrasonic actuator. The advantage of sonication over the liquid is that it more easily dries out, coarsens, and destabilizes wet foams. During filling, foams can be kept low and dry enough in a relatively short time window of a few seconds that little or no liquid can escape from the bottle and contaminate the plant.

The bottling experiments showed total foam reductions of more than 50% and shorter half-lives of the remaining foam of up to 55% in apple juice and currant nectar. The lowest effects of foam destruction occurred in orange juice, mainly due to the preventive effect pulp particles have on drainage. The experiments have been carried out on a commercial bottle geometry, proving its easy adaptability to industrial processes. However, further experiments on the influence of ultrasound on juices with higher viscosity, e.g., at lower temperature, or on other geometries should follow.

# References

1. Voigt, T. *Neue Methoden für den Einsatz der Informationstechnologie bei Getränkeabfüllanlagen*; TU Munich: Munich, Germany, 2004.
2. Durand, M.; Langevin, D. Physicochemical approach to the theory of foam drainage. *Eur. Phys. J. E* **2002**, *7*, 35–44. [CrossRef]
3. Saint-Jalmes, A. Physical chemistry in foam drainage and coarsening. *Soft Matter* **2006**, *2*, 836–849. [CrossRef]
4. Koehler, S.A.; Stone, H.A.; Brenner, M.P.; Eggers, J. Dynamics of foam drainage. *Phys. Rev. E* **1998**, *58*, 2097. [CrossRef]
5. Cantat, I.; Cohen-Addad, S.; Elias, F.; Graner, F.; Höhler, R.; Pitois, O.; Rouyer, F.; Saint-Jalmes, A. *Foams: Structure and Dynamics*; OUP Oxford: Oxford, UK, 2013.
6. Boos, J.; Drenckhan, W.; Stubenrauch, C. On how surfactant depletion during foam generation influences foam properties. *Langmuir* **2012**, *28*, 9303–9310. [CrossRef]
7. Arteaga, G.; Nakai, S. Predicting protein functionality with artificial neural networks: Foaming and emulsifying properties. *J. Food Sci.* **1993**, *58*, 1152–1156. [CrossRef]
8. López-Barajas, M.; López-Tamames, E.; Buxaderas, S.; Suberbiola, G.; De la Torre-Boronat, M.C. Influence of wine polysaccharides of different molecular mass on wine foaming. *Am. J. Enol. Vitic.* **2001**, *52*, 146–150.
9. Dickinson, E. Food emulsions and foams: Stabilization by particles. *Curr. Opin. Colloid Interface Sci.* **2010**, *15*, 40–49. [CrossRef]
10. Murray, B.S.; Durga, K.; Yusoff, A.; Stoyanov, S.D. Stabilization of foams and emulsions by mixtures of surface active food-grade particles and proteins. *Food Hydrocoll.* **2011**, *25*, 627–638. [CrossRef]
11. Pahl, M.H.; Franke, D. Schaum und Schaumzerstörung–ein Überblick. *Chem. Ing. Tech.* **1995**, *67*, 300–312. [CrossRef]
12. Pugh, R. Foaming, foam films, antifoaming and defoaming. *Adv. Colloid Interface Sci.* **1996**, *64*, 67–142. [CrossRef]
13. Büttiker, A.; Keller, H.; Bates, D. Sonotrode and Device for Reducing and Eliminating Foaming of Liquid Products. U.S. Patent No 8,758,492, 24 June 2014.
14. Gallego-Juárez, J.A.; Rodriguez, G.; Acosta, V.; Riera, E. Power ultrasonic transducers with extensive radiators for industrial processing. *Ultrason. Sonochemistry* **2010**, *17*, 953–964. [CrossRef] [PubMed]
15. Riera, E.; Gallego-Juárez, J.A.; Mason, T.J. Airborne ultrasound for the precipitation of smokes and powders and the destruction of foams. *Ultrason. Sonochemistry* **2006**, *13*, 107–116. [CrossRef]
16. Sandor, N.; Stein, H.N. Foam Destruction by Ultrasonic Vibrations. *J. Colloid Interface Sci.* **1993**, *161*, 265–267. [CrossRef]
17. Bhakta, A.; Ruckenstein, E. Decay of standing foams: Drainage, coalescence and collapse. *Adv. Colloid Interface Sci.* **1997**, *70*, 1–124. [CrossRef]
18. Ross, S.; McBain, J. Inhibition of foaming in solvents containing known foamers. *Ind. Eng. Chem.* **1944**, *36*, 570–573. [CrossRef]
19. Sun, S.C. Destruction of flotation froth with intense high frequency sound. *Min. Eng.* **1951**, *3*, 865–867.
20. Dorsey, A.E. Control of foam during fermentation by the application of ultrasonic energy. *J. Biochem. Microbiol. Technol. Eng.* **1959**, *1*, 289–295. [CrossRef]
21. Boucher, R.; Weiner, A. Foam control by acoustic and aerodynamic means. *Br. Chem. Eng.* **1963**, *8*, 808–812.
22. Dedhia, A.C.; Ambulgekar, P.V.; Pandit, A.B. Static foam destruction: Role of ultrasound. *Ultrason. Sonochemistry* **2004**, *11*, 67–75. [CrossRef]
23. Morey, M.; Deshpande, N.; Barigou, M. Foam destabilization by mechanical and ultrasonic vibrations. *J. Colloid Interface Sci.* **1999**, *219*, 90–98. [CrossRef] [PubMed]
24. Chendke, P.; Fogler, H.S. Second-order sonochemical phenomena—Extensions of previous work and applications in industrial processing. *Chem. Eng. J.* **1974**, *8*, 165–178. [CrossRef]
25. Winterburn, J.B.; Martin, P.J. Mechanisms of ultrasound foam interactions. *Asia-Pac. J. Chem. Eng.* **2009**, *4*, 184–190. [CrossRef]
26. Leroy, V.; Fan, Y.; Strybulevych, A.L.; Bellido, G.G.; Page, J.H.; Scanlon, M.G. Investigating the bubble size distribution in dough using ultrasound. *Bubbles Food* **2008**, *2*, 51–60.
27. Gómez, M.; Oliete, B.; García-Álvarez, J.; Ronda, F.; Salazar, J. Characterization of cake batters by ultrasound measurements. *J. Food Eng.* **2008**, *89*, 408–413. [CrossRef]
28. McClements, D.J.; Gunasekaran, S. Ultrasonic characterization of foods and drinks: Principles, methods, and applications. *Crit. Rev. Food Sci. Nutr.* **1997**, *37*, 1–46. [CrossRef]
29. Leroy, V.; Strybulevych, A.; Scanlon, M.G.; Page, J.H. Transmission of ultrasound through a single layer of bubbles. *Eur. Phys. J. E* **2009**, *29*, 123–130. [CrossRef]
30. Leighton, T. *The Acoustic Bubble*; Academic Press: Cambridge, MA, USA, 2012.
31. Vo, H.S.; Kentish, S.; Ashokkumar, M. The enhancement of foam generated by low power ultrasound and its application to foam fractionation. *Colloids Surf. A Physicochem. Eng. Asp.* **2011**, *380*, 35–40. [CrossRef]
32. McHardy, C.; Thünnesen, J.; Horneber, T.; Kostova, J.; Hussein, M.A.; Delgado, A.; Rauh, C. Active control of foams by physically based destruction mechanisms. *PAMM* **2018**, *18*, e201800351. [CrossRef]
33. Panckow, R.P.; McHardy, C.; Rudolph, A.; Muthig, M.; Kostova, J.; Wegener, M.; Rauh, C. Characterization of fast-growing foams in bottling processes by endoscopic imaging and convolutional neural networks. *J. Food Eng.* **2021**, *289*, 110151. [CrossRef]
34. Vaidya, H.A.; Ertunç, Ö.; Lichtenegger, T.; Hachmann, J.; Delgado, A.; Skupin, A. High-speed visualization of acoustically excited cavitation bubbles in a cluster near a rigid boundary. *J. Vis.* **2017**, *20*, 359–368. [CrossRef]
35. Eller, A.; Flynn, H. Generation of subharmonics of order one-half by bubbles in a sound field. *J. Acoust. Soc. Am.* **1969**, *46*, 722–727. [CrossRef]
36. Crum, L. Acoustic cavitation series: Part five rectified diffusion. *Ultrasonics* **1984**, *22*, 215–223. [CrossRef]

37.  Neppiras, E.A. Acoustic cavitation. *Phys. Rep.* **1980**, *61*, 159–251. [CrossRef]
38.  Neppiras, E. Subharmonic and other low-frequency emission from bubbles in sound-irradiated liquids. *J. Acoust. Soc. Am.* **1969**, *46*, 587–601. [CrossRef]
39.  Sato, M.; Shibuya, N.; Okada, N.; Tou, T.; Fujii, T. Oscillation mode conversion and energy confinement of acoustically agitated bubbles. *Phys. Rev. E* **2002**, *65*, 046302. [CrossRef] [PubMed]

# Multi-Point Shape Optimization of a Horizontal Axis Tidal Stream Turbine

**Hassan el Sheshtawy *** [iD]**, Ould el Moctar** [iD] **and Satish Natarajan** [iD]

Institute of Ship Technology Ocean Engineering and Transport Systems, University of Duisburg-Essen, Bismarckstrasse 69, 47057 Duisburg, Germany; ould.el-moctar@uni-due.de (O.e.M.); satish.natarajan.1111@gmail.com (S.N.)
***** Correspondence: hassan.el-sheshtawy@stud.uni-due.de

**Abstract:** A method was developed to perform shape optimization of a tidal stream turbine hydrofoil using a multi-objective genetic algorithm. A bezier curve parameterized the reference hydrofoil profile NACA 63815. Shape optimization of this hydrofoil maximized its lift-to-drag ratio and minimized its pressure coefficient, thereby increasing the turbines power output power and improving its cavitation characteristics. The Elitist Non-dominated Sorting Genetic Algorithm (NSGA-II) was employed to perform the shape optimization. A comparative study of two- and three-dimensional optimizations was carried out. The effect of varying the angle of attack on the quality of optimized results was also studied. Predictions based on two-dimensional panel method results were also studied. Predictions based on a two-dimensional panel method and on a computational fluid dynamics code were compared to experimental measurements.

**Keywords:** hydrofoil optimization; NSGA-II; CFD; XFOIL; NACA 63815; bezier curve

## 1. Introduction

The demand for renewable energy has been increasing over the previous decade. Renewable energy resources are available in various forms; one of them is tidal energy. Tidal energy can be harnessed in a number of ways, one being the use of tidal stream turbines. Therefore, the shape of the subject hydrofoil itself was optimized to improve the performance. The optimization process can be thought of as making the best use of resources under given constraints. Many researchers performed a two-dimensional shape optimization of airfoils for aircraft [1–3] and others for wind turbines [4,5]. For wind turbines, airfoils can be operated at a high angle of attack (AOA), even under stall conditions, which is not the case for aircraft. The lift-to-drag ratio (L/D) or glide ratio (GR) are the most important criteria for shape optimization of wind turbine airfoil [5]. The same methods and algorithms that optimize airfoils can be used to optimize hydrofoils. However, the parameters to be optimized differ because every turbomachine has its own optimization targets that should be taken into account. Besides the L/D ratio, cavitation characteristics should be considered to perform shape optimization of hydrofoil ([6]). Liu and Veitch [7] developed a code to predict the strength of wind turbine blades and to evaluate the maximum bald thickness. They validated their code against model and full-scale experimental measurements. Besides increasing the strength of the turbine blade, Goundar and Ahmed [8] designed hydrofoils for different sections of an HATST to maximize its L/D ratio and to reduce the possibility of the occurrence of cavitation.

Paolo et al. [9] optimized marine propeller hydrofoils by taking into account the effect of cavitation. They parameterized the hydrofoil using B-Splines [10] and carried out shape optimization using NSGA-2 [11] to widen the cavitation bucket of parent hydrofoil and to maximize L/D ratio. To this end, they employed two different flow solver approaches: viscid and inviscid. The author investigated the influence of optimization algorithms such as NSGA-2 and Particle Swarm Optimization (PSO) [12,13] on the performance of

marine propeller design. Three modified optimization algorithms were proposed, two of which were based on NSGA-2 algorithm that was modified by a meta-model and one of which is non-dominated PSO algorithm that is a combination of NSGA-2 algorithm and standard PSO algorithm. Tests were carried out on a real-life propeller for naval application. The propeller geometry was described by Rolls-Royce distribution curves which can be controlled by input parameters such as blade area ratio, skew, or rake at the blade tip. Further, B-splines were used to change the geometry of the blade locally. Among the two algorithms presented that utilize a meta-model, the performance of SANA NSGA-2 was comparably better and the fastest and strictest convergence towards the Pareto front was observed in Non-dominated sorting PSO algorithm. The most important criterion for the marine propellers is to reduce the possibility of cavitation occurrence and to maximize the L/D ratio. Some researchers performed shape optimization of hydrofoils for marine applications. Ching-Yeh et al. [14] and Ouyang et al. [15] performed shape optimization for increasing the GR. Apart from GR, the authors of [16] also evaluated cavitation characteristics of the optimized hydrofoil using Computational Fluid Dynamics (CFD). Litvinov [17] optimized the shape of a hydrofoil, which moves with low velocity in a viscous incompressible fluid, to minimize the power needed to move it by imposing constraints on lift force, area, and geometry of the hydrofoil. Goundar et al. [18] studied the hydrodynamic characteristics of HF-Sx hydrofoil experimentally and numerically, and compared their performance to other hydrofoils of marine current turbines. Furthermore, they designed a hydrofoil to work well at TSRs from 3.0 to 4.0 without decreasing the possibility of cavitation. Tahani and Babayan [19] carried out shape optimization for horizontal axis turbines to increase their power coefficient by employing an Ant Colony Optimization (ACO) algorithm, and they calculated the power coefficient using Blade Element Momentum Theory (BEMT). Luo et al. [6] optimized the shape of a hydrofoil for marine current turbines to increase its L/D ratio and to improve its cavitation performance over a wide range of AOA. They parameterized the hydrofoil using Bezier curve and carried out optimization by employing NSGA-2. Furthermore, they developed a FORTRAN code and coupled it with CFD solver. The optimization method proposed in the paper effectively improves the L/D performance of the hydrofoil from AOA 0° to 12°. The cavitation performance was mainly improved from AOA 6° to 12° with little improvement from 0° to 6°.

In this paper, we carried out shape optimization of a tidal stream turbine hydrofoil to improve its output power. We developed a Matlab code to optimize the shape of its hydrofoil and coupled our code with a flow solver to determine the turbine's performance. We chose NSGA-2 as the optimization algorithm and carried out a comparative two- and three-dimensional optimization study. One of the salient features of the NSGA-2 algorithm is that if a situation arises wherein the optimization algorithm can select only one of the two good solutions for the next generation, it makes its selection on the basis of its distance from its neighbors on the Pareto-optimal graph. The solution with the greatest distance from its neighbors is selected. This ensures that the optimized hydrofoils are diverse. Many researchers used CFD as a flow-solver, which is time-intensive. Here, we coupled our optimization code with the 2D panel, which is computationally less time-intensive and accurate. Furthermore, we investigated the effect of angle of attack on the quality of optimization results.

## 2. Optimization Methodology

Optimization algorithms can be categorized into two groups: gradient-based methods and heuristic algorithms. Gradient-based methods usually require derivatives of the objective functions to guide the search process. Although efficient, these methods do not always converge to global optima [5]. The convergence of the gradient-based methods depends on the initial guess. Heuristic algorithms, on the other hand, are slower, but they converge to global optima regardless of the initial guess. Objective functions alone suffice for the optimization. A Genetic Algorithm (GA) is the most popular of the heuristic

algorithms. The GAs fall into two categories: single- and multi-objective GAs. For practical engineering applications, it is unrealistic to perform optimizations based on only one objective and to ignore the others.

It is prudent to consider all relevant parameters [20]. The GAs are optimization techniques that solve nonlinear or non-differentiable optimization problems. They start with an initial set of randomly generated solutions, which are then compared with the objective function. Subsequent generations are generated from previous generations. We developed a new code for two- and three-dimensional shape optimizations based on the Elitist Non-dominated Sorting Genatic Algorithm (NSGA-2), using the code of [21] as the starting point. We incorporated another set of codes that parameterized the reference hydrofoil to constrain the shape of the generated hydrofoils. This limited the upper and lower bounds of the generated hydrofoils and evaluated the crowding distance for the three-dimensional optimization. Then, we modified the conditions for a non-dominating sorting of the optimized hydrofoils. Finally, we coupled this code with the flow solver and added a postprocessing library to rearrange the results to meet our requirements. Figure 1 schematically depicts our optimization method.
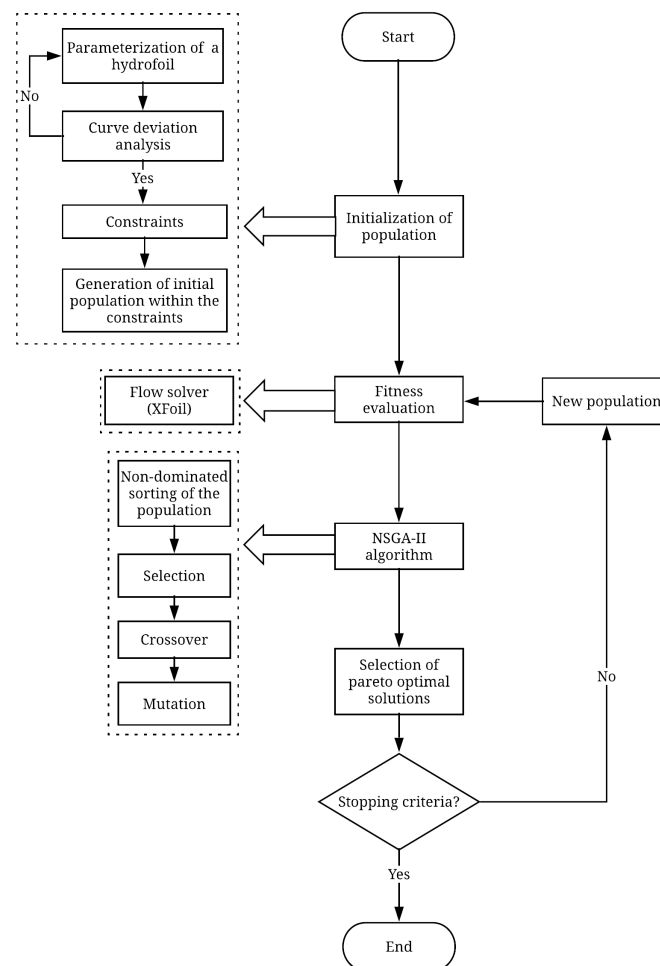


**Figure 1.** Flowchart of optimization methodology.

## 2.1. Initialization of Population

The initialization of population was done in three steps: parameterization of the reference hydrofoil, definition of constraints, and generation of initial population. Here, parameterization refers to the process of representing the hydrofoil shape mathematically and determining its control points, which in turn can vary and control the shape of the

hydrofoils. In this study, the reference hydrofoil was parameterized using Bezier curve. A Bezier curve of order *n* is defined by

$$P(t) = \sum_{i=0}^{n} B_{i,n} Q_i = \sum_{i=0}^{n} C_n^i (1-t)^{n-i} t^i Q_i \qquad (1)$$

where $P(t)$ is a point on the Bezier curve, $t \in [0,1]$ is the weight coefficient of the point on the Bezier curve, $B_{i,n}$ is the Bernstein basis function, $Q_i$ is the *i*th control vertex of the Bezier curve, and $C_n^i$ is the combinatorial symbol [6]. The hydrofoil NACA 63815 [22] was selected as the reference hydrofoil as it was assessed to be cavitation free through experiments [23]. In Figure 2, the original and parametrized hydrofoil are shown. It can be seen that the Bezier curve represents the hydrofoil accurately, and by moving the control points, the shape of the hydrofoil can be varied and controlled.

In order to determine mathematically how effectively the reference hydrofoil was parameterized, curve deviation analysis [6] was performed using the formula below:

$$\delta = \frac{\Delta_i}{l} \qquad (2)$$

where $\delta$ is the normalized deviation of the Y-coordinates of the Bezier curve from that of the reference hydrofoil at any point $i$, $\Delta_i$ is the deviation value of the parameterized hydrofoil with respect to the reference hydrofoil at $i$, and $l$ is the chord length of the original hydrofoil.
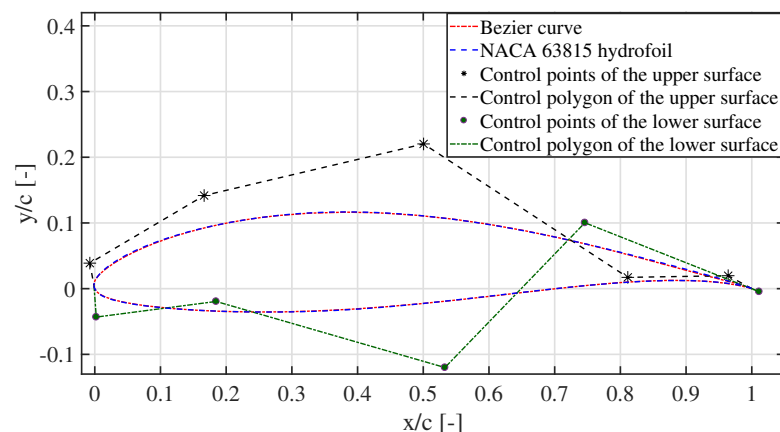


**Figure 2.** Original and Bezier curve parameterized NACA 63815 hydrofoil, where X/C and Y/C are the normalized distance in x and y directions to the hydrofoil chord length ratio.

It can be seen in Figure 3 that $\delta$ of the upper and the lower surfaces over X/C are less than $1 \times 10^{-3}$ and $2.2 \times 10^{-3}$, respectively. The deviation is negligible, and thus the parameterization technique is accurate enough to represent the reference hydrofoil and to carry out shape optimization.

To maintain a realistic shape of the hydrofoil, constraints on its thickness are employed as shown in Figure 4. A major concern with using flow solvers is making the hydrofoil shape smooth enough, which helps to get fast converge solution. Another concern is that if the upper and the lower bounds are too wide or too narrow, the hydrofoil may not be continuous at the leading edge. Therefore, the constraints must be chosen with care. The maximum thicknesses of both the pressure and suction sides are allowed to increase and decrease by 40% and 30%, respectively, with respect to its chord. Here, the upper and lower bounds were chosen through trial and error on flow solver. The constraints have been particularly limited near the trailing edge to ensure that the curves of the pressure and suction sides do not intersect during optimization. The leading and trailing edge coordinates of the hydrofoil are fixed. Further, the x-axis coordinates of all the control

points are fixed. Only the y-coordinates of the ten control points shown in Figure 2 are allowed to vary during optimization.
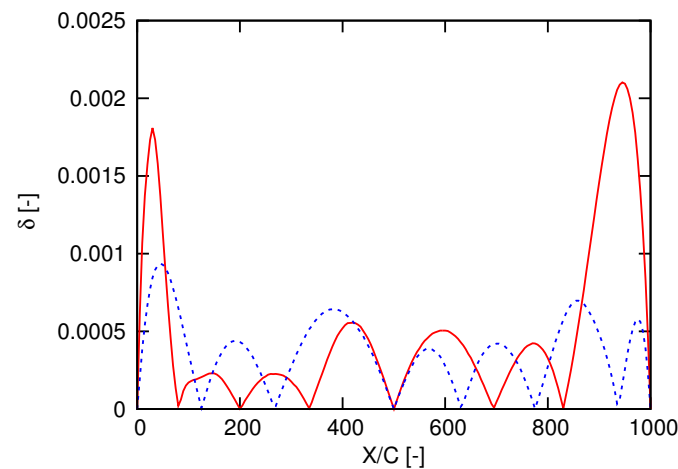


**Figure 3.** Curve deviation analysis of the parameterized hydrofoil, where $\delta$ denotes the normalized deviation of the Y-coordinates of the parameterized hydrofoil from that of the reference hydrofoil and X/C denotes the distance in x-direction normalized to the hydrofoil chord length. The dotted blue line represents the deviation in the suction side and the solid red line represents the deviation in the pressure side.
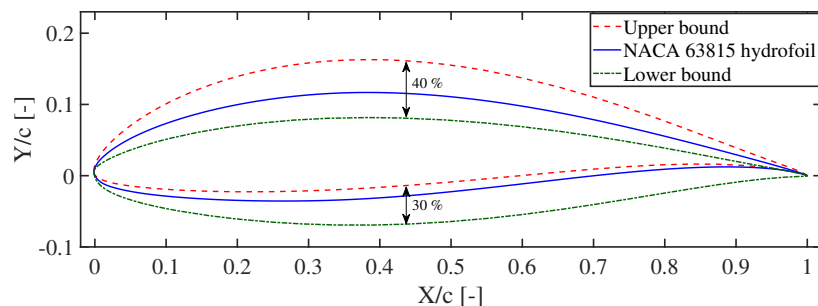


**Figure 4.** Optimized hydrofoil with upper and lower bounds for the optimization.

*2.2. The Flow Solver*

Some researchers use Computational Fluid Dynamics and many others use 2-D panel code, for instance, XFOIL [24], as the flow solver. In this study, we have compared the results of both the methods with the experimental data. The experimental results of performance characteristics of NACA 63815 were obtained from [25] for a Reynolds number of $0.8 \times 10^6$ and for a flow velocity of 3.492 m/s. The computational domain for evaluating the performance characteristics of the hydrofoil using CFD simulation was set as described in [6]. The CFD simulation was performed in a 2D environment. The mesh size was set to 198,000 cells with a $y^+$ value of 0.88. The $k - \omega$ SST turbulence model [26] was chosen to close RANS equations. The SIMPLE algorithm [27] was used for pressure–velocity coupling. Fifteen different Angles Of Attack (AOA) ranging from $-10°$ to 12.5° were considered for the CFD simulations. The Reynolds number in XFOIL was set to $8 \times 10^{+5}$ and the maximum number of iterations for the analysis of each of the hydrofoils was set to 500. The performance characteristics were evaluated in viscous mode. As shown in Figure 5, the CFD results are closer to the experimental measurements than that of XFOIL. However, the computational time using CFD was approximately 4 h, while using XFOIL it were merely 90 s. We tested almost 2480 optimized hydrofoils (30 generations and each generation has 80 hydrofoils), and to this end we needed a method that gives acceptable results within a short span of time. Furthermore, the performance predictions

of XFOIL were also found to match the experimental results closely and it is also simple to use. Therefore, XFOIL was chosen as the flow solver [8,16,18,23,28].
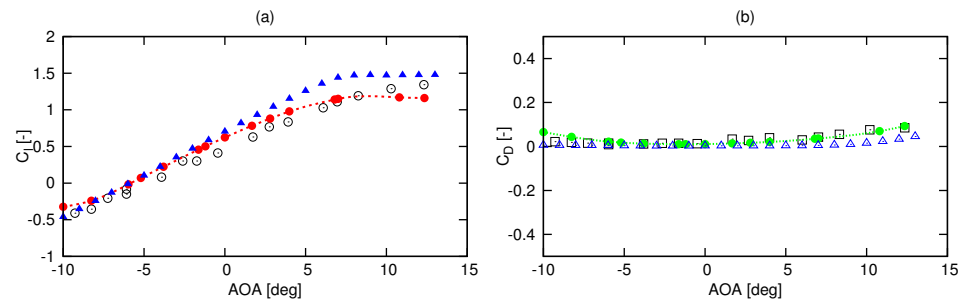


**Figure 5.** Comparison of RANS (red and green dots) and XFOIL (blue and green triangles) results with the experimental values, where (**a**) represents lift coefficients (Cl) vs. angles of attack (AOA) and (**b**) represents drag coefficients (Cd) vs. AOA.

### 2.3. Optimization Algorithm

In this algorithm, the solutions are sequenced based on their degree of non-domination and sent to the next generation. A simple GA consists of three genetic operators: selection, crossover, and mutation. Selection is a process in which pairs of candidate solutions are selected to reproduce. The pairs are selected based on their fitness scores. This operator is an artificial version of Darwinian's natural selection. The tournament selection method has been used to perform the selection operation. This method was chosen as it has better convergence and computational time compared to a lot of other methods. In this method, two candidate solutions are picked randomly and their ranks are compared. The candidate solution that has better rank is selected and the other one is rejected. This process continues until all the candidate solutions are analyzed. Here, two hydrofoils are picked randomly and their lift, drag, and minimum pressure coefficients are analyzed. In GA, the next generation's solutions are created by exchanging information among strings of the previous generation. This process is known as crossover and it is equivalent to reproduction in evolutionary biology. The crossover operation was performed by Simulated Binary Crossover (SBX). Mutation is the process in which the strings of the offspring are occasionally altered. Its main purpose is to maintain diversity in the population. Here, the polynomial mutation was performed [29].

### 2.4. Selection of Pareto Optimal Solutions

The optimization algorithm generated a lot of optimized hydrofoils with varied GR and cavitation characteristics. Each of the generated hydrofoils is better off with regard to one objective but is worse off with regard to another objective. Therefore, they are called Pareto-optimal solutions. The hydrofoil with the highest GR over a wide range of AOA was selected to model the optimized turbine.

### 2.5. The Stopping Criterion

In GAs, either of the two following approaches is employed to set the stopping criterion. First, the algorithm is set to stop if the solutions of the next generation are insignificant compared to that of its previous generation. Second, the maximum number of generations is set to a predefined value. Here, the second approach was adopted and the stopping criterion was set to 30 generations. It is a straightforward approach and was employed by taking into account the computational time. The optimized hydrofoils were generated until the stopping criterion was met. For more details, refer to the work in [30].

### 3. Hydrofoil Shape Optimization Strategies

The 2-D shape optimization of hydrofoil was done by considering three parameters: $C_L$, $C_D$, and minimum pressure coefficient ($C_{Pmin}$). Here, optimization was carried out

to maximize $C_L$ and minimize $C_D$ in order to increase the output power of a tidal stream turbine. The optimization algorithm also sought to increase the $C_{Pmin}$ to improve its cavitation performance [6]. Considering different AOA in the optimization algorithm generates optimized hydrofoils with different performance characteristics. Mukesh et al. [4] carried out the shape optimization for an AOA of 5°. However, the authors of [6] argued that the shape optimization done at a single AOA often results in poor performance at off-design conditions. Therefore, the authors of [6] performed the optimization for AOA of 0°, 6°, and 12°. The AOA in the optimization algorithm must therefore be chosen with care, and therefore its effect was studied in the following sections. The effect of number of objectives and its implication on the solutions was also studied.

### 3.1. Three Dimensional (3-D) Optimization

The main idea behind 3-D optimization was to handle each of the objectives individually, i.e., $C_L$, $C_D$ and $C_{Pmin}$. Two cases were considered to study the effect of AOA. In the first case a wide range of five AOA were considered and in the second case three AOA were considered. In the first case, the objective is to maximize the average $C_L$ and $C_{Pmin}$ and to minimize the average $C_D$ at AOA of 0°, 3°, 6°, 9°, and 12° for the objective functions shown in Equations (3)–(5). $C_L$, $C_D$, and $C_{p_{min}}$ were averaged to give equal weightage to all the AOA considered.

$$\frac{\Sigma C_{L,AOA}}{N} \tag{3}$$

$$\frac{\Sigma C_{D,AOA}}{N} \tag{4}$$

$$\frac{\Sigma C_{Pmin,AOA}}{N} \tag{5}$$

where $N$ is the number of different AOA considered. Table 1 shows the comparison of performance characteristics of the hydrofoils that were generated in the first case.

**Table 1.** Comparison of optimization results at different generations.

| Generation | Mean $C_L$ | Mean $C_D$ | Mean $C_{Pmin}$ | Number of Optimized Hydrofoils |
|---|---|---|---|---|
| 1 | 1.1337 | 0.014 | −2.3882 | 11 |
| 5 | 1.1664 | 0.012 | −2.115 | 18 |
| 10 | 1.1866 | 0.0123 | −2.0462 | 24 |
| 15 | 1.1789 | 0.0122 | −2.0458 | 22 |
| 20 | 1.1721 | 0.012 | −2.0645 | 18 |
| 25 | 1.1656 | 0.012 | −2.0568 | 20 |
| 30 | 1.1713 | 0.0119 | −2.0899 | 20 |

Every generation consists of 80 hydrofoils. The mean values were taken in order to assess the efficacy of the optimization algorithm in generating hydrofoils with higher $C_L$ and $C_{Pmin}$ and lower $C_D$. The hydrofoils with $C_L$ and $C_{Pmin}$ greater than the reference hydrofoil and $C_D$ lesser than the reference hydrofoil were designated as optimized hydrofoils in each generation. The mean $C_L$ and $C_{Pmin}$ of the hydrofoils in the last generation is higher than that of the first generation and the mean $C_D$ is lower, which indicates that the optimization algorithm could generate better hydrofoils as the execution of the optimization code progresses.

The 3D Pareto-optimal graphs of the first case for different generations are shown in Figure 6. The hydrofoil characteristics are color-coded according to their $C_L$ values. Figure 6a shows performance characteristics of optimized hydrofoils obtained in the 5th generation. Seven hydrofoils are shown within a dotted black ellipse, which have poor cavitation characteristics relative to the others in the same generation. Their $C_{Pmin}$ ranges from −2.438 to −2.359, $C_D$ ranges from 0.117 to 0.012, and $C_L$ ranges from 1.19 to 1.22. However, as some hydrofoils overlap another, only four of them are visible. In the 10th

generation, as shown in Figure 6b, the seven hydrofoils were excluded and were replaced with hydrofoils having better cavitation characteristics. Three hydrofoils that are shown within dotted black rectangles in Figure 6b–e having $C_L$ ranging from 1.35 to 1.4 and $C_D$ ranging from 0.014 to 0.0142 were retained until the 25th generation. Although they have the worst $C_D$ values, they were retained until the 25th generation owing to their high $C_L$ values. They were eventually replaced with hydrofoils having lower $C_D$ and higher $C_{Pmin}$ in the 30th generation as shown in Figure 6f. In course of the execution of the optimization code, the $C_D$ of the generated hydrofoils gradually became lesser compared to its preceding generations.
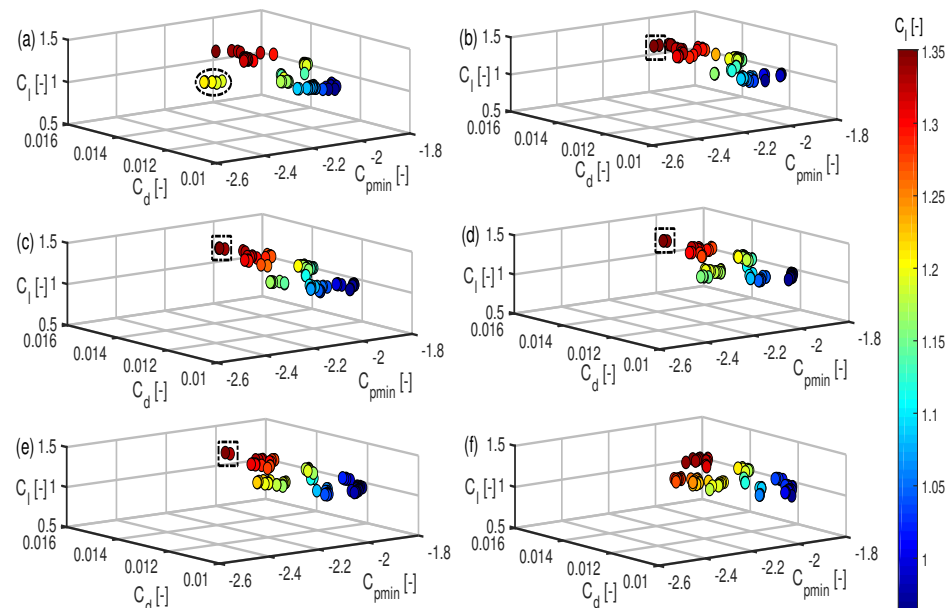


**Figure 6.** 3-D Pareto optimal graph of the different generations, where (**a**) represents the 5th, (**b**) represents the 10th, (**c**) represents the 15th, (**d**) represents the 20th, (**e**) represents the 25th, and (**f**) represents the 30th generations.

Four hydrofoils with the highest GR are shown in Figure 7. Figure 7a shows an optimized hydrofoil ( HF-3D-1-1), which has maximum thickness ($t_{max}$) of 17.09% at 33.7% of the chord length from the leading edge (0.337 $c$). Here, HF refers to hydrofoil, 3D refers to three-dimensional optimization, the first 1 refers to the first case, and the second 1 refers to the first optimized hydrofoil. It can be seen in Figure 8a that the HF-3D-1-1 has a higher GR from AOA = 0° to almost 2° and from AOA = 6.1° to 12°. Figure 9a shows that $C_{Pmin}$ of the optimized hydrofoil has been improved at higher AOA ranging from 5° to 12°. Figure 7b shows the second optimized hydrofoil (HF-3D-1-2), which has $t_{max}$ of 18.26% at 0.347 $c$. It can be seen from Figure 8b that the optimized hydrofoil has a higher GR from AOA = 0° to 1° and from AOA = 6.2° to 12°. Figure 9c shows that the cavitation performance of the optimized hydrofoil is closer to that of the reference hydrofoil at lower AOA and from AOA of 5° the cavitation performance has been greatly improved. Figure 7c shows the third optimized hydrofoil (HF-3D-1-3), which has $t_{max}$ of 17.58% at 0.339$c$. It can be seen from Figure 8c that the optimized hydrofoil has a higher GR from AOA = 0° to 1° and from AOA = 6.2° to 12°. Figure 9b shows that the cavitation performance of the optimized hydrofoil is closer to the reference hydrofoil at lower AOA and from AOA = 5° the cavitation performance has been greatly improved. Figure 7d shows the fourth optimized hydrofoil (HF-3D-1-4) has $t_{max}$ of 17.5% at 0.337 $c$. Although Figures 8d and 9d show that the GR and $C_{p_{min}}$ were not improved at lower AOA, it performs well at higher AOA.
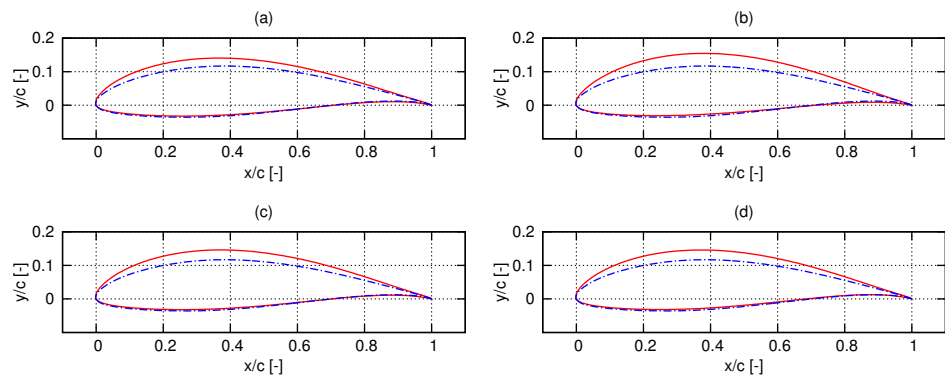
**Figure 7.** Reference and optimized hydrofoils of the first case of 3D optimization, where (**a**) is optimized hydrofoil HF-3D-1-1, (**b**) is optimized hydrofoil HF-3D-1-2, (**c**) is optimized hydrofoil HF-3D-1-3 and (**d**) is optimized hydrofoil HF-3D-1-4. The red solid lines denote the optimized hydrofoil and the blue dash-dot lines denote the original hydrofoil.
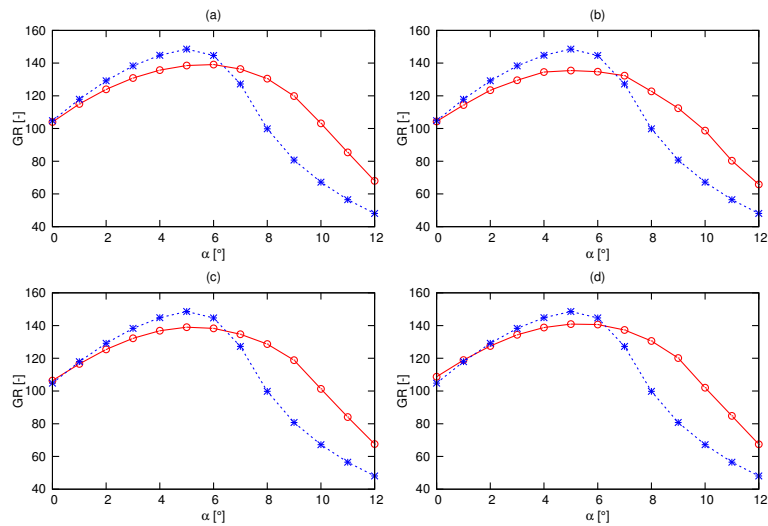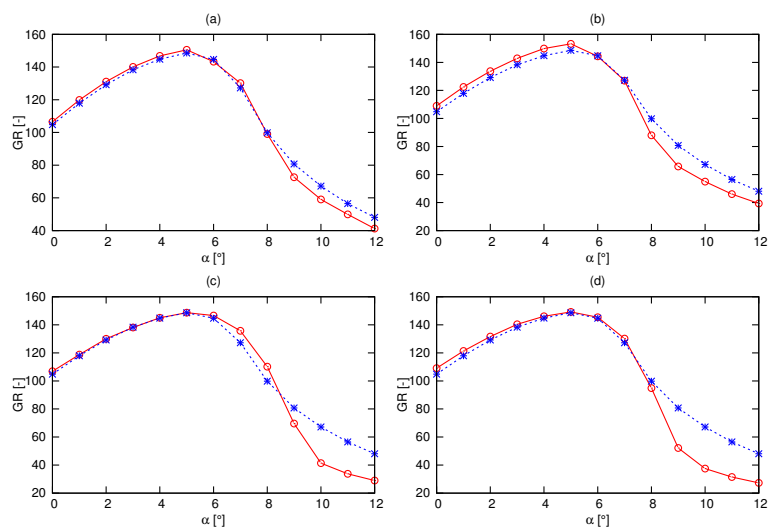


**Figure 8.** GR of the reference and optimized hydrofoils of the first case of 3D optimization, where (**a**) is GR of HF-3D-1-1, (**b**) is GR of HF-3D-1-2, (**c**) is GR of HF-3D-1-3 and (**d**) is GR of HF-3D-1-4. The red and blue lines denote optimized and reference hydrofoils, respectively.
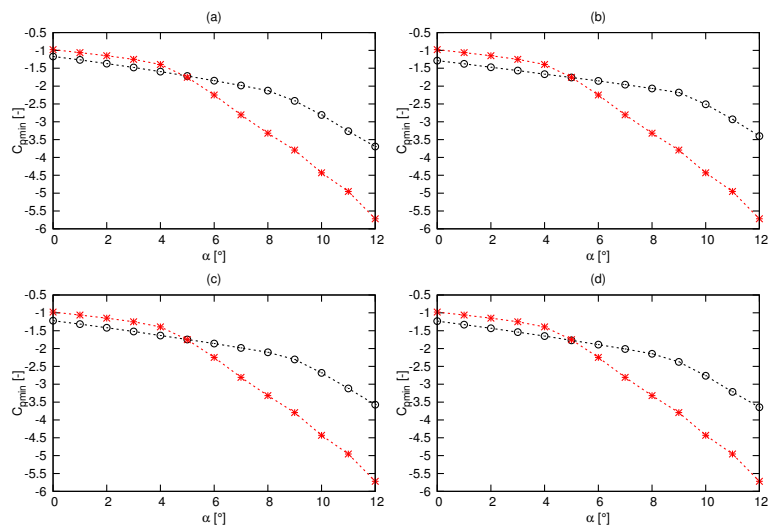


**Figure 9.** $C_{pmin}$ of the reference and optimized hydrofoils of the first case of 3-D optimization, where (**a**) is $C_{pmin}$ of HF-3D-1-1, (**b**) is $C_{pmin}$ of HF-3D-1-2, (**c**) is $C_{pmin}$ of HF-3D-1-3, (**d**) is $C_{pmin}$ of HF-3D-1-4. The red and black lines denote optimized and reference hydrofoil, respectively.

In the first case, hydrofoils were produced that performed consistently well at wide range of AOA. The GR and the cavitation performance were improved at higher AOA. However, the maximum GR of the optimized hydrofoils were lower than the reference hydrofoil, which is undesirable.

As the maximum GR of the reference and the optimized hydrofoils in the previous case lie between AOA of 4° and 6°, the optimization was done at these AOA with a focus on increasing the maximum GR. Equations (3), (4), and (6) are calculated for AOA of 4°, 5°, and 6°.

It can be seen in Figure 10 that the optimized hydrofoils of this case have higher GR at smaller AOA than the hydrofoils obtained in the first case (Figure 8), wherein the optimized hydrofoils had better performance characteristics only at higher AOA. As shown in Figure 11, the cavitation performance is better than in the first case (Figure 9). It, however, turned out that the Pareto optimal solutions are not as diverse as the first case.



**Figure 10.** GR of the reference and optimized hydrofoils of the second case of 3-D optimization, where (**a**) is GR of HF-3D-2-1, (**b**) is GR of HF-3D-2-2, (**c**) is GR of HF-3D-2-3 and (**d**) is GR of HF-3D-2-4. The red and blue lines denote optimized and reference,respectively.
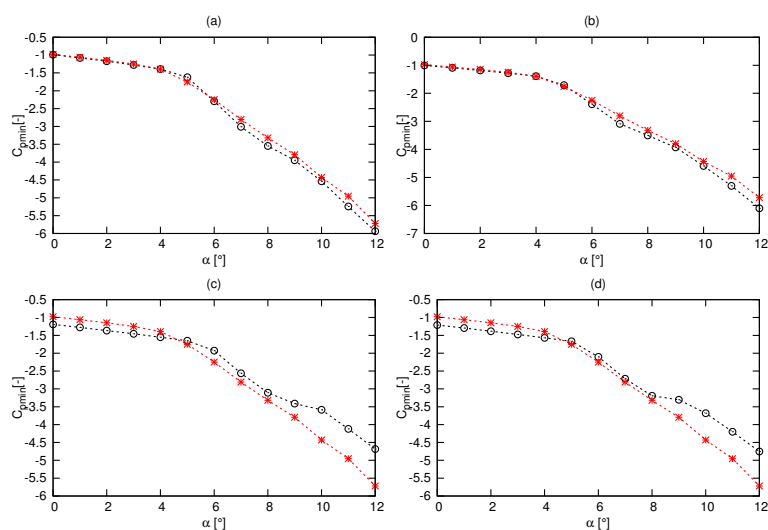


**Figure 11.** $C_{pmin}$ of the reference and optimized hydrofoils of the second case of 3-D optimization, where (**a**) is $C_{pmin}$ of HF-3D-2-1, (**b**) is $C_{pmin}$ of HF-3D-2-2; (**c**) is $C_{pmin}$ of HF-3D-2-3, (**d**) is $C_{pmin}$ of HF-3D-2-4. The red and black lines denote optimized and reference hydrofoil,respectively.

### 3.2. Two-Dimensional (2-D) Optimization

It is possible that the process of evolution of the optimized hydrofoils was slowed down on account of it being a high-dimensional multi-objective optimization problem. The number of objectives was, therefore, reduced to two by considering $C_{p_{min}}$ and GR as the optimization objectives as shown Equations (5) and (6) respectively.

$$\frac{\Sigma GR_{,AOA}}{N} \tag{6}$$

where $\Sigma GR_{,AOA}$ is the summation of GR at different AOA. In this case, the summation of GR and $C_{Pmin}$ are calculated at AOA of $0°$, $3°$, $6°$, $9°$, and $12°$. Table 2 shows the comparison of the performance characteristics of the hydrofoils in different generations. The mean GR and $C_{Pmin}$ values show only the general behaviour of the optimization algorithm, which sporadically increase and decrease for two reasons. First, optimization was done by considering two conflicting objectives, i.e., improvement in one objective is achieved at the expense of the other objective. Second, the crowding distance operator was used to diversify the solutions at the end of each generation. In Table 2, it can be seen that in the 5th generation the number of optimized hydrofoils is higher than all the other generations. For better clarity refer to Figure 12a, which shows that the hydrofoil characteristics of this generation are bunched together meaning that some of the optimized hydrofoils are almost similar. The duplication of hydrofoils is avoided by using the crowding distance operator whose primary task is to diversify the Pareto optimal solutions as shown in Figure 12. The distribution of the performance characteristics of hydrofoils in the 30th generation (Figure 12f) is the most diverse compared to the other generations (Figure 12a–e). With diverse solutions, the designer will have a lot of hydrofoils to choose from based on their specific requirements.
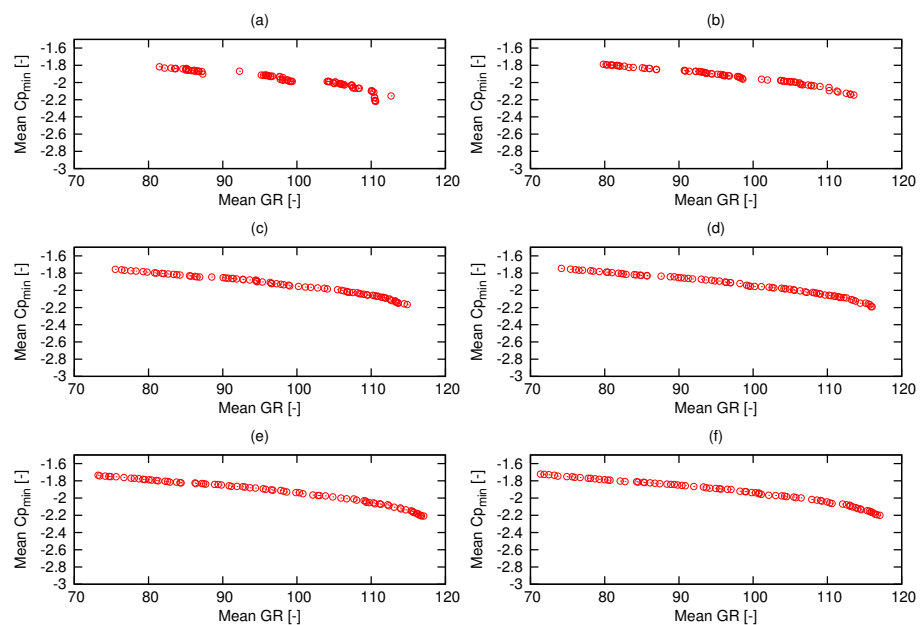


**Figure 12.** Graphs for the different generations in the first case of 2D optimization, where (**a**) represents the 5th, (**b**) represents the 10th, (**c**) represents the 15th, (**d**) represents the 20th, (**e**) represents the 25th, and (**f**) represents the 30th generations.

**Table 2.** Comparison of optimization results at different generations in the first case of 2D optimization.

| Generation | Mean Glide Ratio | Mean $C_{p_{min}}$ | Number of Optimized Hydrofoils |
|:---:|:---:|:---:|:---:|
| 1 | 94.4015 | −2.1861 | 21 |
| 5 | 98.1213 | −1.9701 | 34 |
| 10 | 96.5394 | −1.9297 | 27 |
| 15 | 97.9764 | −1.9464 | 33 |
| 20 | 97.6847 | −1.9418 | 33 |
| 25 | 96.7179 | −1.937 | 30 |
| 30 | 96.5764 | −1.9312 | 30 |

Four hydrofoils with the highest GR in the 30th generation are shown in Figure 13. The GR of hydrofoil (HF-2D-1-1) whose $t_{max}$ is 15.09% at 0.303 $c$ is shown Figure 13a. The GR of this hydrofoil is higher than that of the reference hydrofoil from AOA = 0° to 2° and from AOA = 6.2° to 12° as shown in Figure 14a. The cavitation performance is better from AOA greater than 5° as shown in Figure 15a. The second optimized hydrofoil (HF-2D-1-2) in Figure 13b has $t_{max}$ of 17.09% at 0.337$c$. The GR of this hydrofoil is higher than the reference hydrofoil from AOA = 0° to 2° and from AOA = 6.2° to 12° as shown in Figure 15b. The cavitation performance is better from AOA greater than 5° as shown in Figure 14b. The third optimized hydrofoil (HF-2D-1-3) shown in Figure 13c has $t_{max}$ of 17.5% at 0.339$c$ as shown in Figure 15c. The GR of the hydrofoil is higher than the reference hydrofoil from AOA = 0° to 2° and from AOA = 6.2° to 12°, and the cavitation performance is better from AOA greater than 5° as shown in Figure 15c. The fourth optimized hydrofoil (HF-2D-1-4) shown in Figure 13d has $t_{max}$ of 17.5% at 0.337$c$ as shown in Figure 15d. The maximum GRs of the hydrofoils obtained in this case are lower than the reference hydrofoil, but they have better cavitation characteristics at higher AOA as shown in Figure 14d. By reducing the number of objectives to two, a marked improvement in the performances of the hydrofoils could be noticed.
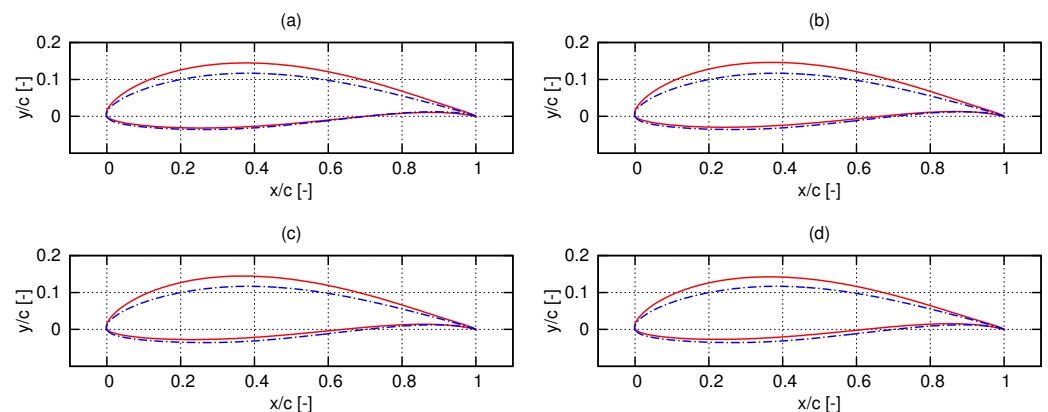


**Figure 13.** Reference and optimized hydrofoils of the first case of 2-D optimization, where (**a**) is optimized hydrofoil HF-2D-1-1, (**b**) is optimized hydrofoil HF-2D-1-2, (**c**) is optimized hydrofoil HF-2D-1-3 and (**d**) is optimized hydrofoil HF-2D-1-4. The red solid lines denote the optimized hydrofoil and the blue dash-dot lines denote the original hydrofoil.

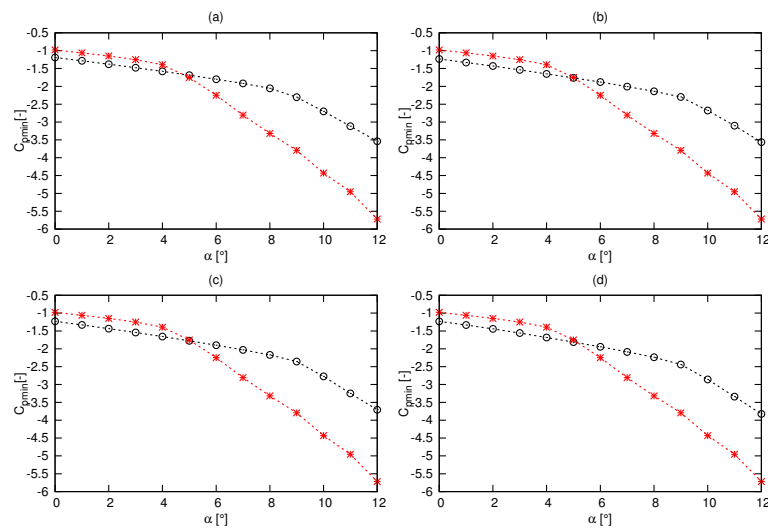**Figure 14.** $C_{p_{min}}$ of the reference and optimized hydrofoils of the first case of 2D optimization, where (**a**) is $C_{p_{min}}$ of HF-2D-1-1, (**b**) is $C_{p_{min}}$ of HF-2D-1-2, (**c**) is $C_{p_{min}}$ of HF-2D-1-3, (**d**) is $C_{p_{min}}$ of HF-2D-1-4. The red and black lines denote optimized and reference hydrofoil, respectively.
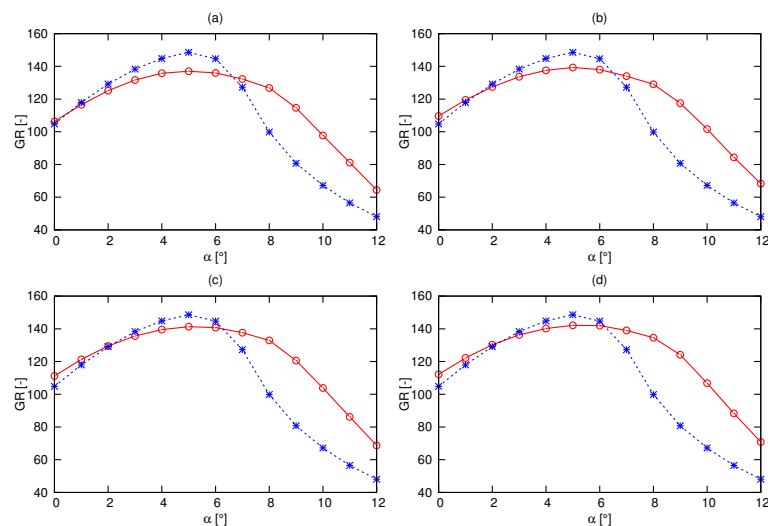


**Figure 15.** GR of the reference and optimized hydrofoils of the first case of 2D optimization, where (**a**) is GR of HF-2D-1-1, (**b**) is GR of HF-2D-1-2, (**c**) is GR of HF-2D-1-3 and (**d**) is GR of HF-2D-1-4. The red and blue lines denote optimized and reference hydrofoils, respectively.

Further, the maximum GRs lie between AOA 4 and 6. Therefore, 2D optimization was done for these AOA. Equations (5) and (6) are calculated for AOA of $4°$, $5°$, and $6°$. Four hydrofoils with the highest GR are shown in Figure 16. The GR of HF-2D-2-1 that is shown in Figure 16a is higher than that of the reference hydrofoil from AOA $0°$ to $8°$ and the $C_{Pmin}$ is greater from AOA greater than $5°$ as shown in Figure 17a and Figure 18a, respectively. The GR of the second hydrofoil (HF-2D-2-2) is higher than that of the reference hydrofoil from AOA $0°$ to $9.8°$ and the cavitation performance is better from AOA greater than $5°$ as shown in Figure 17b and Figure 18b, respectively. The third optimized hydrofoil (HF-2D-2-3) whose $t_{max}$ is 17.13% at $0.347c$ is shown in Figure 13c. The GR of this hydrofoil is significantly higher than the reference hydrofoil from AOA $0°$ to almost $10°$ and the cavitation performance was better than the reference hydrofoil from AOA greater than $5°$. The performance characteristics of the fourth hydrofoil (HF-2D-2-4) are not as good as the previous two hydrofoils but still are better than the reference hydrofoil, especially with regard to the GR from AOA $0°$ to $8°$. The GRs of the hydrofoils obtained in this case are also higher than that of all the optimized hydrofoils in the previous three cases. On the

other hand, the cavitation performances are not as good as that of the optimized hydrofoils obtained in the first cases of 2D and 3D optimization.
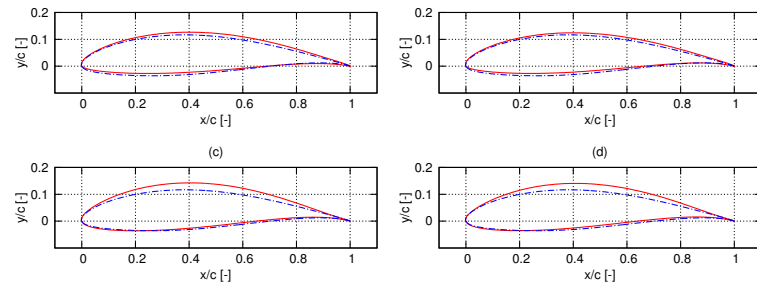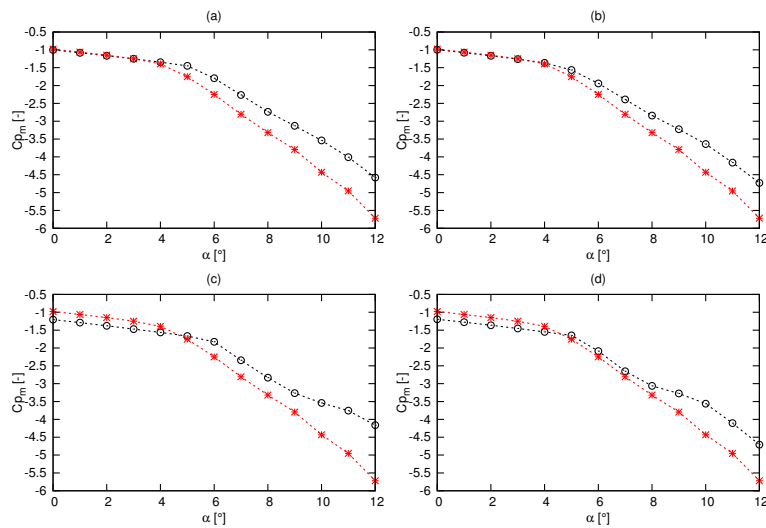


**Figure 16.** Reference and optimized hydrofoils of the second case of 2-D optimization, where (**a**) is optimized hydrofoil HF-2D-2-1, (**b**) is optimized hydrofoil HF-2D-2-2, (**c**) is optimized hydrofoil HF-2D-2-3 and (**d**) is optimized hydrofoil HF-2D-2-4. The red solid lines denote the optimized hydrofoil and the blue dash-dot lines denote the original hydrofoil.



**Figure 17.** $C_{pmin}$ of the reference and optimized hydrofoils of the second case of 2D optimization, where (**a**) is $C_{pmin}$ of HF-2D-2-1, (**b**) is $C_{pmin}$ of HF-2D-2-2, (**c**) is $C_{pmin}$ of HF-2D-2-3, (**d**) is $C_{pmin}$ of HF-2D-2-4. The red and black lines denote optimized and reference hydrofoil, respectively.
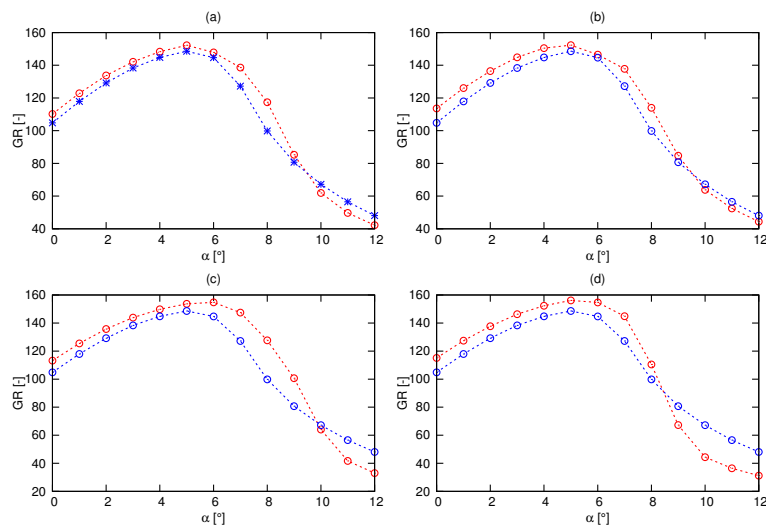


**Figure 18.** GR of the reference and optimized hydrofoils of the second case of 2D optimization, where (**a**) is GR of HF-2D-2-1, (**b**) is GR of HF-2D-2-2, (**c**) is GR of HF-2D-2-3 and (**d**) is GR of HF-2D-2-4. The red and blue lines denote optimized and reference hydrofoils, respectively.

Some of the optimized hydrofoils obtained in this case have GRs greater than the reference hydrofoil from AOA ranging from 0° to 10°. Further, the GRs of the hydrofoils are higher than those obtained in the other optimization cases. On the other hand, the cavitation characteristics were almost the same from 0° to 5° and after that they were improved until AOA of 12°. Although hydrofoils were generated with higher $C_L$, lower $C_D$, and higher $C_{Pmin}$, the evolution of optimized hydrofoils was slow. Further, the influence of the crowding distance operator was apparently insignificant in the 3D optimization.

## 4. Conclusions

The onus is on the designer of the tidal stream turbines to select hydrofoils based on their specific requirements.

An in-house MATLAB code was developed based on [11] to perform shape optimization of hydrofoils by employing NSGA-2 algorithm and coupled with a flow solver. The hydrofoil was parameterized using 6th order Bezier curves. In that case, NACA 63815 was taken as a reference hydrofoil. The maximum normalized deviation of the Bezier parameterized reference hydrofoil from the original hydrofoil was accurate enough to carry out shape optimization.

For the chosen solver, the 2D panel method XFOIL solver and CFD results of Lift and drag coefficients of the reference hydrofoil were compared with the experimental measurements. The results for CFD is more accurate but also more time consuming, thus the XFOIL solver was chosen on account of its simplicity and its ability to give acceptable results within a short span of time compare to CFD.

A comparative study on the optimization was done by considering two methods. For 3D optimization, three objective, i.e., $C_L$, $C_D$ and $C_{p_{min}}$, were the optimization objectives, and for 2D optimization, the number of objectives was reduced to two by unifying $C_L$ and $C_D$ as GR. The GR of the hydrofoils was further improved by performing optimization at AOA of 4°, 5° and 6° than on performing at AOA of 0°, 3°, 6°, 9°, and 12°.

This was because the maximum glide ratio of reference hydrofoil fell within the range of 4° and 6°. However, the cavitation characteristics were better at higher AOA in the first case.

The multi-objective optimization performed in the present work effectively improved the GR of the hydrofoils from AOA of 0° to almost 10°. The cavitation performance was mainly improved from AOA 5° to 12°. The solutions of optimization are not unique. Different hydrofoils with different performance characteristics were obtained, but a lot of them performed better than NACA 63815.

## References

1. Nemec, M.; Zingg, D.W.; Pulliam, T.H. Multipoint and multi-objective aerodynamic shape optimization. *AIAA J.* **2004**, *6*, 1057–1065. [CrossRef]
2. Peigin, S.; Epstein, B. Robust optimization of 2d airfoils driven by full Navier–Stokes computations. *Comput. Fluids* **2004**, *9*, 1175–1200. [CrossRef]
3. Shahrokhi, A.; Jahangirian, A. Airfoil shape parameterization for optimum Navier–Stokes design with genetic algorithm. *Aerosp. Sci. Technol.* **2007**, *6*, 443–450. [CrossRef]
4. Mukesh, R.; Lingadurai, K.; Selvakumar, U. Airfoil shape optimization using non-traditional optimization technique and its validation. *J. King Saud Univ.-Eng. Sci.* **2014**, *2*, 191–197. [CrossRef]
5. Ribeiro, A.; Awruch, A.M.; Gomes, H.M. An airfoil optimization technique for wind turbines. *Appl. Math. Model.* **2012**, *10*, 4898–4907. [CrossRef]

6.  Luo, X.-Q.; Zhu, G.-J.; Feng, J.-J. Multi-point design optimization of hydrofoil for marine current turbine. *J. Hydrodyn. Ser. B* **2014**, *5*, 807–817. [CrossRef]

7.  Liu, P.; Veitch, B. Design and optimization for strength and integrity of tidal turbine rotor blades. *Energy* **2012**, *1*, 393–404. [CrossRef]

8.  Goundar, J.N.; Ahmed, M.R. Design of a horizontal axis tidal current turbine. *Appl. Energy* **2013**, *111*, 161–174. [CrossRef]

9.  Papadrakakis, M.; Papadopoulos, V.; Stefanou, G.; Plevris, V.A. Framework for the design by optimization of hydrofoils under cavitating conditions. In Proceedings of the 7th European Congress on Computational Methods in Applied Sciences and Engineering, ECCOMAS 2016, Crete, Greece, 5–10 June 2016.

10. De Boor, C.; Mathematicien, E.-U.; De Boor, C. *A Practical Guide to Splines*; Springer: New York, NY, USA, 1978; Volume 27.

11. Deb, K.; Pratap, A.; Agarwal, S.; Meyarivan, T.A.M.T. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Trans. Evol. Comput.* **2002**, *2*, 182–197. [CrossRef]

12. Clerc, M. *Particle Swarm Optimization*; John Wiley & Sons:Hoboken, NJ, USA 2010.

13. Kennedy, J.; Eberhart, R. Particle swarm optimization. *Int. Conf. Neural Netw.* **1995**, *5*, 1942–1948.

14. Hsin, C.-Y.; Wu, J.-L.; Chang, S.-F. Design and optimization method for a two-dimensional hydrofoil. *J. Hydrodyn.* **2006**, *5*, 323–329. [CrossRef]

15. Ouyang, H.; Weber, L.J.; Odgaard, A.J. Design optimization of a two-dimensional hydrofoil by applying a genetic algorithm. *Eng. Optim.* **2014**, *5*, 529–540. [CrossRef]

16. Zhang, D.-S.; Chen, J.; Shi, W.-D.; Shi, L.; Geng, L.-L. Optimization of hydrofoil for tidal current turbine based on particle swarm optimization and computational fluid dynamic method. *Therm. Sci.* **2016**, *3*, 907–912. [CrossRef]

17. Litvinov, W. On the optimal shape of a hydrofoil. *J. Optim. Theory Appl.* **1995**, *2*, 325–345. [CrossRef]

18. Goundar, J.N.; Ahmed, M.R.; Lee, Y.-H. Numerical and experimental studies on hydrofoils for marine current turbines. *Renew. Energy* **2012**, *42*, 173–179. [CrossRef]

19. Tahani, M.; Babayan, N. Optimum section selection procedure for horizontal axis tidal stream turbines. *Neural Comput. Appl.* **2017**, *31*, 1211–1223. [CrossRef]

20. Konak, A.; Coit, D.W.; Smith, A.E. Multi-objective optimization using genetic algorithms: A tutorial. *Reliab. Eng. Syst. Saf.* **2006**, *91*, 992–1007. [CrossRef]

21. Baskar, S.; Tamilselvi, S.; Varshini, P. MATLAB Code for Constrained NSGA II. Available online: https://de.mathworks. com/matlabcentral/fileexchange/49806-matlab-code-for-constrained-nsga-ii-dr-s-baskar--s-tamilselvi-and-//mdx.plm. automation.siemens.com/star-ccm-plus/ (accessed on 12 October 2017). .

22. Abbott, I.H.; von Doenhoff, A.E. *Theory of Wing Sections, Including a Summary of Airfoil Data*; Courier Corporation: Mineola, NY, USA,1959.

23. Batten, W.; Bahaj, A.; Moll, A.; Chaplin, J. Hydrodynamics of marine current turbines. *Renew. Energy* **2006**, *31*, 249–256. [CrossRef]

24. Drela, M. *Xfoil: An Analysis and Design System for Low Reynolds Number Airfoils*; Springer:Berlin/Heidelberg, Germany, 1989; Volume 111, pp. 1–12.

25. Moll, A.F.; Bahaj, A.S.; Chaplin, J.R.; Batten, W.M.J. Measurements and predictions of forces, pressures and cavitation on 2-D sections suitable for marine current turbines. *Proc. Inst. Mech. Eng. Part M* **2004**, *2*, 127–138.

26. Menter, F. Zonal two equation kw turbulence models for aerodynamic flows. In Proceedings of the 23rd Fluid Dynamics, Plasma Dynamics, and Lasers Conference, Orlando, FL, USA, 6–9 July 1993.

27. Patankar, S.V. *Computational Methods for Fluid Dynamics*; Taylor& Francis: Germantown, NY, USA, 2016.

28. Batten, W.M.J.; Bahaj, A.S.; Moll, A.F.; Chaplin, J.R. Experimentally validated numerical method for the hydrodynamic design of horizontal axis tidal turbines. *Ocean. Eng.* **2007**, *7*, 1013–1020. [CrossRef]

29. Goldberg, D. *Genetic Algorithms in Search, Optimization, and Machine Learning*; Addison-Wesley Publishing Company: Boston, MA, USA, 1989.

30. Deb, K. *Multi-Objective Optimization Using Evolutionary Algorithms*; Wiley: Hoboken, NJ, USA, 2009.