

Review

AI Control of Power Converters Under Semiconductor Constraints: A Critical Review of Deployment Readiness

Sangyoon Woo , Gyeongsu Sim , Hoejin Jung , Soyeon Park , Wonchil Choi  and Won-Gyu Bae * 

Department of Electrical Engineering, Soongsil University, Seoul 06978, Republic of Korea; wsy04018@soongsil.ac.kr (S.W.); hyundai6013@soongsil.ac.kr (G.S.); wjdghlwlsl234@soongsil.ac.kr (H.J.); syee20221239@soongsil.ac.kr (S.P.); dnjsclf1456@soongsil.ac.kr (W.C.)

* Correspondence: wgbae@ssu.ac.kr

Abstract

Wide-bandgap (WBG) power converters impose stringent requirements, including high-frequency switching, strongly nonlinear dynamics, and limited computational time, which constrain conventional control and artificial intelligence (AI)-based approaches and hinder their practical deployment. Existing review studies have primarily focused on algorithmic structures or performance, while systematic analyses from a deployment feasibility perspective under hardware constraints remain limited. This paper examines AI applications in WBG power converters from a system-level deployment perspective and analyzes existing studies based on implementation feasibility. After outlining the physical characteristics and control requirements of WBG devices, it reviews AI-based modeling, AI-assisted model predictive control (MPC), and reinforcement learning (RL)-based direct control. These approaches are evaluated in terms of computational complexity, real-time feasibility, out-of-distribution (OOD) generalization, and integration with conventional control frameworks. Key deployment challenges, including safety-constrained RL, sim-to-real transfer, and field-programmable gate array (FPGA)/embedded implementation, are treated as core analytical dimensions. To support this assessment, this review introduces an AI Deployment Readiness framework organized around four analytical dimensions: (1) modeling accuracy, (2) safety assurance, (3) sim-to-real transfer capability, and (4) hardware implementability. Using this framework, prior studies are reassessed, and its applicability is further discussed for applications such as fault diagnosis and remaining useful life (RUL) prediction. The analysis identifies key bottlenecks and clarifies deployment-relevant considerations for high-frequency WBG systems.

Keywords: WBG power converters; power electronics control; AI-based control; real-time control; model predictive control; reinforcement learning; hardware implementation; sim-to-real transfer



Academic Editor: Pedro J. Villegas

Received: 11 June 2026

Revised: 22 July 2026

Accepted: 22 July 2026

Published: 28 July 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Power electronics technology plays an essential role in modern energy systems, and approximately 80% of global electricity is projected to be processed through various power electronic systems by 2030 [1]. In particular, the efficiency, volume, and cost of application systems such as photovoltaic inverters, electric vehicles and ultra-fast chargers, industrial automation, and smart grids are fundamentally determined by the performance of power semiconductor devices. This directly leads to the demand for next-generation device technologies and corresponding control methodologies [2,3].

Recently, power electronic systems have entered a new transition phase owing to the commercialization of wide-bandgap (WBG) semiconductor devices represented by silicon carbide (SiC) and gallium nitride (GaN) [2]. Unlike conventional silicon (Si)-based metal–oxide–semiconductor field-effect transistors (MOSFETs), which typically operate at switching frequencies below 100 kHz, WBG devices enable ultra-high-speed switching ranging from several tens of kHz to the MHz regime. Consequently, both miniaturization and high efficiency of power converters can be simultaneously achieved [3]. However, these improvements in physical device characteristics impose new challenges on control systems. In particular, under MHz-range switching environments, the sampling period of the controller is reduced to below the microsecond scale, which severely constrains the computational time required by conventional control methods such as proportional–integral (PI) controllers and model predictive control (MPC) [4].

Although the conventional linear control approach based on PI control offers advantages such as simple structure and ease of implementation, it exhibits high dependence on parameter tuning and cannot guarantee optimal performance in power converter systems with nonlinear dynamic characteristics [5]. To address these limitations, finite control set model predictive control (FCS-MPC) has been proposed, providing significant advantages in explicitly incorporating control objectives and handling system constraints. However, FCS-MPC suffers from an inherent structural limitation in which computational complexity increases exponentially with the number of switching states [4,6]. In practice, for multiphase and multilevel converters, the computational burden required by the controller has been reported to exceed real-time processing capability, thereby forcing the switching frequency to be limited to approximately 2.5–10 kHz [6]. This limitation fundamentally conflicts with the high-speed control requirements of WBG-based systems.

Against this background, artificial intelligence (AI)-based control methods have emerged as promising alternatives by transferring complex optimization procedures to offline training while requiring only relatively simple inference operations during real-time control [7]. In particular, neural network-based control and reinforcement learning (RL) have demonstrated the potential to overcome the limitations of conventional controllers through their high representational capability for nonlinear systems, even without accurate mathematical models [8,9]. Existing reviews have provided valuable taxonomies of AI algorithms, control strategies, and application domains in power converter systems [10]. However, deployment-related issues such as constrained hardware resources, stringent real-time requirements, sim-to-real transfer, and safety assurance are often discussed as separate topics rather than as interdependent requirements for practical implementation. Consequently, it remains difficult to assess from existing reviews whether a reported AI controller is merely effective under a given validation setting or is actually ready for deployment in WBG power converter systems.

As shown in Table 1, the present review differs from the selected prior review and perspective papers by treating WBG-specific deployment readiness across modeling, safety, sim-to-real transfer, and hardware implementation as the primary analytical axis. This perspective is necessary because an inherent gap exists between the theoretical ability of AI models to approximate complex nonlinear systems through sufficient training and the deployment readiness required in practical power electronic systems [8,11]. Power converter systems are characterized by constrained data availability, variable operating conditions, physical safety constraints, and microsecond-scale real-time computational requirements. Consequently, approximation capability alone cannot guarantee practical field deployment. For example, under operating conditions outside the training data distribution, prediction errors may increase abruptly; during RL exploration, unsafe actions may directly induce critical hardware failures such as overcurrent or overvoltage condi-

tions [12,13]. Furthermore, AI controllers validated in simulation environments frequently fail to achieve identical performance in real hardware because parasitic components, sensor noise, and delays induce steady-state error [11,14]. Moreover, the high-frequency switching characteristics of WBG-based systems severely constrain the allowable inference time for AI models, thereby requiring consideration not only of algorithmic structure optimization but also of hardware implementation feasibility on resource-constrained platforms such as field-programmable gate arrays (FPGAs) and microcontrollers [15].

Table 1. Comparison of analytical scopes in representative review and perspective papers.

Reference	Primary Scope	Main Topics Covered	Main Organizing Principle	Position Relative to This Review
WBG technology outlook [1]	WBG power electronic devices and circuits	WBG materials and devices, ARPA-E power electronics programs, modules, circuits, system integration, and application spaces such as motor drives, automotive systems, data centers, aerospace systems, and distributed energy resources	Technology roadmap and power-electronics value chain	Establishes the system-level motivation for WBG power electronics, but does not focus on AI-based controller deployment
WBG semiconductor perspective [2]	Power electronics based on SiC and GaN semiconductors	Material properties, device-level advantages, technology maturity, reliability issues, market penetration, cost considerations, and application-level comparisons with Si-based implementations	Material–device–market perspective	Clarifies WBG device feasibility and application benefits, but AI control and deployment readiness are outside its main scope
GaN/SiC device review [3]	Industrial state-of-the-art of GaN and SiC power devices	Commercially available SiC and GaN transistors, device structures, voltage ranges, application fields, reliability issues, and future technology perspectives	Device technology, application, and reliability perspective	Provides device-level constraints relevant to WBG systems, but does not analyze AI controller maturity or implementation pathways
AI controller overview [10]	AI techniques for improving controllers in power converter-based systems	Supervised, unsupervised, and reinforcement learning; online/offline training; AI-aided linear and nonlinear controllers; controller surrogate models; and application areas in converter-based systems	AI methodology and controller-architecture perspective	Reviews AI-based controller performance enhancement, but deployment-related issues are not treated as the primary analytical axis
Very-high-frequency converter review [16]	Very-high-frequency power conversion	Frequency scaling, magnetic component miniaturization, semiconductor parasitics, device loss mechanisms, topology requirements, control methods, and VHF experimental examples	Frequency-scaling and converter-design perspective	Provides background on high-frequency hardware constraints, but AI-based control and deployment readiness are not its focus

Accordingly, this review reassesses existing AI-based power converter studies from a deployment-readiness perspective. It synthesizes previously recognized deployment issues into a structured analytical framework for examining the maturity of AI-based control methods in WBG power converter applications. It is organized around four analytical dimensions: (1) modeling accuracy, (2) safety assurance, (3) sim-to-real transferability, and (4) hardware implementability.

These dimensions are sequentially dependent in deployment analysis, such that limitations in an earlier dimension constrain subsequent ones. Insufficient modeling accuracy weakens the reliability of safety analysis, and simulation-domain safety guarantees do not automatically remain valid after transfer to real hardware. Even a controller that transfers successfully cannot be practically deployed unless its inference meets the timing and resource constraints of the target hardware. The framework therefore helps identify deployment bottlenecks across existing studies.

The literature survey was conducted as a targeted critical review and deployment-oriented engineering synthesis. Relevant studies were identified using Web of Science,

Scopus, and Google Scholar with domain-specific search terms related to AI-based power converter control, model predictive control, reinforcement learning, safe reinforcement learning, WBG power converters, SiC/GaN converters, sim-to-real transfer, and embedded implementation. Studies were considered within the scope of this review when they addressed at least one deployment-relevant aspect of AI-based power electronic control, including AI-based modeling, AI-assisted MPC, RL-based converter control, safety assurance, sim-to-real validation, real-time implementation, or hardware-constrained deployment.

Based on this deployment-readiness perspective, this review examines AI-based power converter control technologies from the viewpoint of practical system deployment feasibility and analyzes the major technical challenges and potential solution directions associated with each dimension. Through this approach, the review aims to clarify the requirements that should be considered for AI-based control to become a practical alternative for next-generation WBG-based power electronic systems. In practical terms, the framework helps engineers identify the dominant deployment bottleneck and prioritize validation and implementation efforts. It also provides safety and security practitioners with a structured basis for identifying unverified assumptions and implementation-dependent risks across modeling, safety assurance, sim-to-real transfer, and hardware implementation before physical deployment.

2. AI Control Strategies for WBG Power Converters

2.1. Physical Characteristics of WBG Devices and Control System Requirements

WBG semiconductor devices represented by GaN and SiC possess material characteristics fundamentally different from those of Si-based devices. Owing to their wide bandgap energy, high critical electric field, and superior electron mobility, WBG devices exhibit lower on-resistance and smaller parasitic capacitance [2]. These characteristics significantly reduce switching losses while enabling ultra-high-speed switching operation ranging from several hundred kHz to the MHz regime [17]. Such ultra-high-speed switching directly reduces the size of passive components, and reductions in passive component volume compared with Si-based systems at the same power density have been reported [16,18].

However, the ultra-high-speed switching characteristics of WBG devices impose fundamentally new design constraints on the control systems of power converters. As illustrated in Figure 1, SiC and GaN devices operate in frequency regions far beyond the operational limits of conventional Si-based devices, thereby fundamentally tightening the computational time available to the controller. In digital control environments, the controller must complete measurement, computation, and gate signal generation within each switching cycle. When the switching frequency reaches 1 MHz, the total allowable time for this entire process is theoretically reduced to within 1 μ s. Since measurement, computation, and gate signal generation must all be completed within a single control cycle at the switching frequency f_{sw} , the effective time budget available for control computation, denoted as T_{budget} , can be formulated as follows:

$$T_{\text{budget}} \leq \frac{1}{f_{sw}} - T_{\text{sense}} - T_{\text{gate}} \quad (1)$$

Here, T_{sense} represents the analog-to-digital converter (ADC) sampling and signal propagation delay, whereas T_{gate} denotes the propagation delay of the gate driver. However, unlike SiC devices or Si insulated-gate bipolar transistors (IGBTs), GaN high-electron-mobility transistors (HEMTs) exhibit a short-circuit withstand time (SCWT) T_{sc} of only several hundred nanoseconds. Consequently, under overcurrent conditions, device failure may occur before the controller can respond. In standard industrial designs, overcurrent protection is implemented via dedicated analog hardware circuits (e.g., desaturation detec-

tion) that operate independently of the digital control algorithm; consequently, T_{sc} does not directly constrain AI inference time in such architectures. However, in emerging *integrated control–protection architectures*—where the AI controller is also responsible for state monitoring and protective gate-signal termination—the controller must complete both inference and any protective response before device destruction occurs. Therefore, T_{sc} constitutes a conservative upper bound on T_{budget} , and the practical safe computational budget under this architecture assumption can be expressed as:

$$T_{budget} \leq \min\left(\frac{1}{f_{sw}}, T_{sc}\right) - T_{sense} - T_{gate} \quad (2)$$

Substituting $f_{sw} = 1$ MHz, $T_{sc} = 300$ ns, and $T_{sense} + T_{gate} \approx 50$ ns yields $T_{budget} \leq 250$ ns under the integrated architecture assumption. For conventional designs with hardware-separated protection, the bound relaxes to $1/f_{sw} - T_{sense} - T_{gate} \approx 950$ ns at 1 MHz. The resulting 250 ns budget is therefore treated as an architecture-dependent reference case rather than as a general requirement for all WBG converters. It represents a restrictive GaN scenario in which control computation and protective decision-making are integrated in the digital controller. When overcurrent protection is implemented by dedicated analog or hardware circuits, the AI controller is mainly constrained by the switching-cycle budget rather than directly by T_{sc} . Accordingly, the Stage 4 analysis distinguishes the feasibility reported for each study’s own platform and operating frequency from feasibility under this stricter integrated GaN reference case.

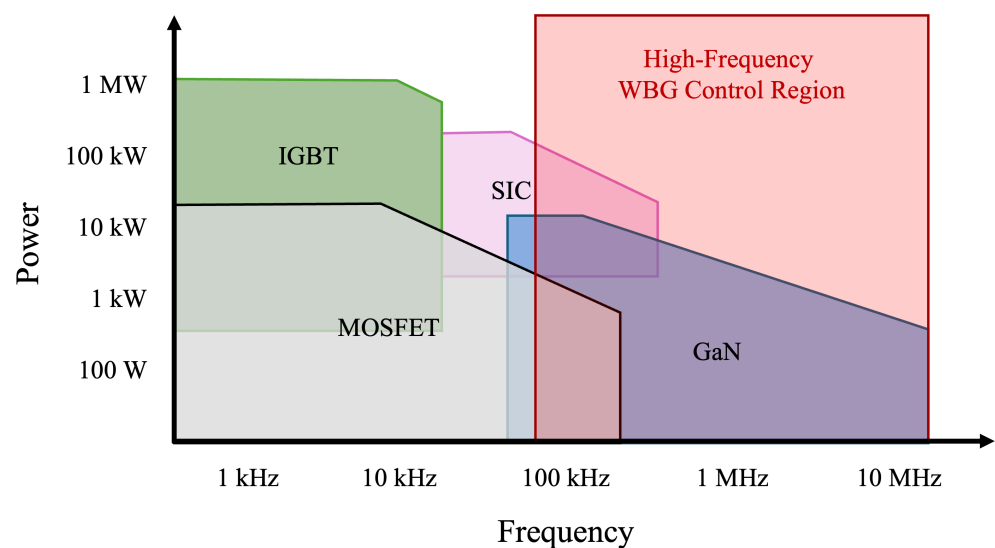


Figure 1. Operating regions of high-voltage power semiconductor devices in the power–frequency domain. The shaded red region indicates the high-frequency WBG control zone (SiC and GaN), where switching frequencies impose microsecond-level control budgets and necessitate AI-based control strategies.

Under the integrated GaN reference case, Equation (2) constitutes the physical foundation of Stage 4 (hardware implementation feasibility), because computational latency becomes a safety-relevant design constraint directly linked to both control computation and protective response.

The curves in Figure 2 were generated by evaluating Equation (2) pointwise over the plotted switching-frequency range using $T_{sc} = 300$ ns for GaN, $T_{sc} = 5$ μ s for SiC, and $T_{sense} + T_{gate} = 50$ ns. The FP32 and INT16 horizontal latency lines were taken directly from the FPGA implementation results reported in [7] and were not estimated in this review.

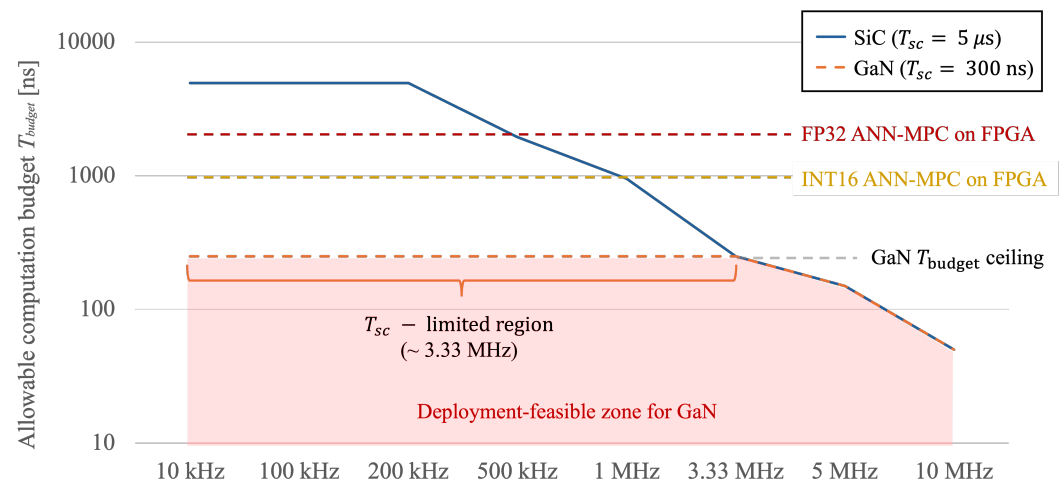


Figure 2. Allowable computation budget T_{budget} as a function of switching frequency for GaN ($T_{\text{sc}} = 300$ ns) and SiC ($T_{\text{sc}} = 5$ μ s) devices, derived from Equation (2) with $T_{\text{sense}} + T_{\text{gate}} = 50$ ns. The GaN T_{budget} is bounded at 250 ns for all switching frequencies below 3.33 MHz. Inference latencies of representative AI controller implementations are superimposed as horizontal reference lines [7]. The shaded region denotes the deployment-feasible zone under the integrated GaN control-protection assumption and should not be interpreted as a universal feasibility boundary for converter architectures employing independent hardware protection.

Figure 2 illustrates T_{budget} as a function of switching frequency for GaN ($T_{\text{sc}} = 300$ ns) and SiC ($T_{\text{sc}} = 5$ μ s) devices, with the inference latencies of representative AI controller implementations superimposed [7]. The GaN T_{budget} remains bounded at 250 ns for all switching frequencies below 3.33 MHz, whereas 32-bit floating-point (FP32) and 16-bit integer (INT16) implementations of artificial neural network (ANN)-based MPC on FPGA exhibit inference latencies of 2100 ns and 1000 ns, respectively—exceeding this ceiling by $8.4\times$ and $4.0\times$. This visualization quantitatively illustrates that the reported FP32 and INT16 ANN-MPC implementations are infeasible under the integrated GaN control-protection assumption. This result should not be generalized to all WBG converter systems; rather, it indicates that AI deployment feasibility depends jointly on the protection architecture, switching frequency, semiconductor device, and target hardware platform.

Under such conditions, conventional PI controllers remain widely employed in industrial applications because of their simple structure and low computational burden. However, PI control is designed based on the assumption of linear time-invariant systems, making it difficult to guarantee sufficient robustness against the switching nonlinearity, parameter variations, and abrupt load changes inherently present in power converters. In particular, when nonlinear loads exhibiting negative impedance characteristics, such as constant power loads (CPLs), are connected, PI control alone is fundamentally limited in securing a stable operating region [19]. Furthermore, gains optimized for a single operating point can result in severe performance degradation or instability when operating conditions vary. In WBG-based high-speed switching systems, such parameter variations appear over a much broader frequency range, thereby further increasing the complexity of PI gain tuning compared with conventional Si-based systems.

To overcome these limitations of PI control, FCS-MPC has attracted considerable attention in the field of power electronics because it can simultaneously achieve multi-objective control and system constraint handling through an explicit cost function. FCS-MPC operates by predicting the future behavior corresponding to all possible switching states at every sampling period and selecting the switching pattern that minimizes the cost function. However, this approach inherently suffers from structural limitations. In

multilevel converters, the number of possible switching states increases exponentially with the number of levels [4], frequently causing the required computational burden to exceed real-time processing capability. Indeed, in multiphase and multilevel power converters, the computational constraints of FCS-MPC have been reported to force the switching frequency to be limited to approximately 2.5–10 kHz [6], which fundamentally conflicts with the operational frequency range of WBG devices extending from several hundred kHz to the MHz regime.

For a y -phase, x -level flying capacitor (FC) converter, where each phase leg contains $(x - 1)$ independently controllable binary switches, the number of switching states K evaluated by FCS-MPC at each sampling period and the total computational complexity N can be expressed as:

$$K = 2^{(x-1)y}, \quad N_{\text{FCS-MPC}} = K \cdot (n_p + n_c) \quad (3)$$

Here, n_p and n_c denote the computational complexity of the prediction function and cost function, respectively [6]. Specifically, FCS-MPC requires 816 operations per cycle for a single-phase five-level flying capacitor converter, and this number increases to 4992 operations when extended to a three-phase three-level flying capacitor (FC) converter [6]. Various approaches for reducing computational complexity, including candidate vector partial search, region separation, and sphere decoding, have been proposed [4]. However, these methods fundamentally provide suboptimal solutions or introduce additional design parameters, and therefore do not completely bridge the gap with the high-speed control requirements imposed by WBG environments.

In this context, AI-based approaches have emerged as structurally suitable alternatives for WBG environments by transferring computationally intensive optimization procedures to offline training while performing only simple neural network inference during real-time control. In a real-time operating environment at 20 kHz using a standard Si-based DC-DC buck converter, model-free deep RL (DRL) controllers have been experimentally demonstrated to reduce per-step execution time by approximately five-fold compared with MPC by directly executing offline-trained policies at each sampling instant [11]. The key insight is that AI-based controllers can replace the stepwise rolling optimization process of MPC with parameterized function approximation, thereby maintaining computational complexity at a level independent of the number of switching states. This structural characteristic is particularly relevant to the sub-microsecond control periods that can arise in MHz-range WBG converter operation and forms the common design philosophy underlying the AI control strategies discussed in the following sections.

2.2. AI-Enhanced MPC

Research on AI-assisted MPC in WBG environments has primarily evolved along three directions: (1) imitation-based architectures in which ANNs replace the optimization procedure of MPC itself, (2) learning-based schemes in which AI is used to automatically tune or design the weighting matrices of the MPC cost function, and (3) hybrid frameworks in which RL compensates for model uncertainties in an online manner. The following sections review each direction in sequence.

2.2.1. ANN-Based MPC Imitation: Inference Acceleration Through Offline Training

The most direct approach to AI-enhanced MPC is to collect the input–output pairs of FCS-MPC offline and train an ANN to reproduce the same control outputs using only simple forward-pass operations during the real-time inference stage. In this structure, the computational complexity required for real-time inference of ANN-MPC, denoted as N_{ANN} , depends solely on the network architecture:

$$N_{\text{ANN}} = \sum_{l=1}^L n_{l-1} \cdot n_l \quad (4)$$

Here, L denotes the number of layers, and n_l represents the number of neurons in the l -th layer. Compared with Equation (3), N_{ANN} does not scale directly with K , although it may increase as the required input, hidden-layer, and output dimensions grow with converter complexity. The key premise of this structure is that the prediction-optimization process of FCS-MPC is entirely deterministic. Given identical measurement vectors and cost functions, the output switching pattern is always identical, making it possible to approximate this input–output relationship using a neural network.

Figure 3 directly quantifies this structural difference across five representative converter configurations, including conventional two-level and flying-capacitor multilevel cases. Across these configurations, the estimated FCS-MPC operation count increases from approximately 100 operations per cycle for the 1P2L converter to approximately 210,000 operations per cycle for the 3P5L FC converter, primarily because the number of candidate switching states K increases exponentially with the converter phase number and level. In contrast, ANN-MPC complexity is determined by the selected network dimensions rather than directly by K . Equation (4) represents the leading-order interlayer weight-operation count, whereas the values plotted in Figure 3 use the expanded arithmetic-operation count described below. Over the configurations considered, ANN-MPC complexity therefore increases substantially more slowly than the switching-state-dependent complexity of FCS-MPC. This structural independence from K constitutes the primary motivation for ANN-based MPC approximation in WBG environments.

For transparency, Figure 3 distinguishes values reported directly in [6] from analytical estimates introduced in this review. For FCS-MPC, the values of 816 operations per cycle for the 1P5L FC configuration and 4992 operations per cycle for the 3P3L FC configuration were taken directly from [6]. The remaining FCS-MPC values were estimated using Equation (3) and a fixed reference cost of 51 operations per candidate, obtained from 816/16 for the reported 1P5L FC case, to isolate the effect of switching-state growth. This procedure yields approximately 100, 400, and 210,000 operations per cycle for the 1P2L, 3P2L, and 3P5L FC configurations, respectively.

For ANN-MPC, the 1P5L FC and 3P3L FC counts of 386 and 1362 operations per cycle were taken directly from [6]. The remaining values were estimated using network dimensions (d, m, n) of $(3, 8, 1)$, $(5, 15, 3)$, and $(14, 78, 12)$, where d , m , and n denote the input, hidden-layer, and output dimensions, respectively. The hidden-layer dimensions for the 1P2L and 3P2L cases preserve the ratio $m/(d+n) = 17/9$ of the reported 1P5L FC network, whereas the 3P5L FC case preserves the ratio $m/(d+n) = 42/14$ of the reported 3P3L FC network. Applying the expanded arithmetic-operation count $2dm + 2mn + 4m + 3n$ reported in [6] gives 99, 309, and 4404 operations per cycle for the three estimated configurations. These estimates are intended to illustrate computational scaling rather than optimized network designs or measured execution latency.

For series resonant inverters, a framework has been proposed in which complex nonlinear MPC policies are trained offline using deep neural networks and executed in real time on an FPGA [7]. Under a 16-bit fixed-point arithmetic implementation, the inference latency was only 1 μs , while the lookup table (LUT) utilization and digital signal processor (DSP) utilization were reduced by approximately 9.3 times and 5 times, respectively, compared with floating-point implementations. Notably, even under conditions with 15% model parameter error, the zero-voltage switching (ZVS) violation rate remained lower than that of accurate nonlinear MPC, at 0.0784% versus 0.3922%. This result suggests that neural network approximation can exhibit a smoothing effect against model errors.

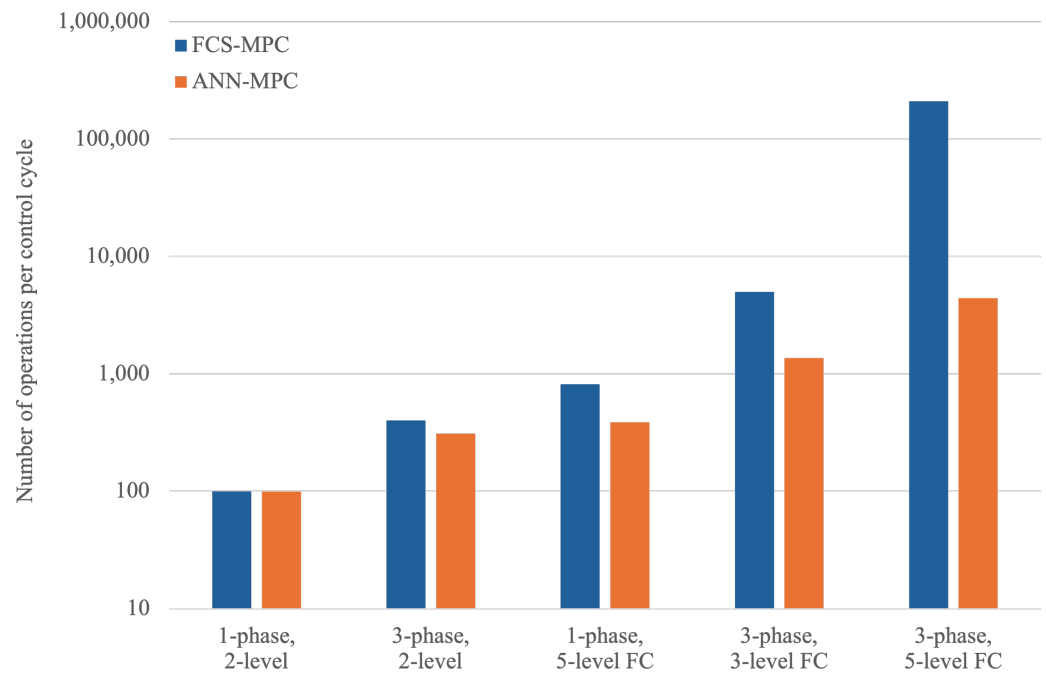


Figure 3. Arithmetic-operation comparison between FCS-MPC and ANN-MPC across five representative converter configurations, including conventional two-level and flying-capacitor multilevel cases. Values for the 1P5L FC and 3P3L FC configurations were taken directly from [6], whereas the remaining values are analytical estimates obtained using the procedure described in the text. The comparison is intended to illustrate computational scaling and does not represent measured execution latency or a hardware-normalized benchmark.

An additional approach has been proposed to compress the piecewise-affine (PWA) control function generated by explicit MPC into a neural network representation [15]. The original PWA function consisted of 70 regions, making embedded implementation infeasible. However, through the PWA-NN structure, implementation on an 80 MHz microcontroller was achieved, and the entire control algorithm was verified to execute within approximately 27.5 μ s. This result demonstrates that constraint-aware predictive control can be realized on low-cost general-purpose microcontrollers without requiring expensive FPGAs, thereby improving deployment feasibility for WBG-based low-power applications.

2.2.2. AI-Based Automatic Weight Design: Automation of MPC Cost Function Tuning

The cost function of FCS-MPC consists of weighted sums of multiple control objectives, and the determination of each weighting factor has traditionally relied on iterative empirical tuning. This process strongly depends on the designer's expertise and suffers from the structural limitation that fixed weighting factors cannot guarantee optimal performance under varying operating conditions.

An ANN-based weight tuning method has been proposed in which operating conditions, including the reference current magnitude and target average switching frequency, are provided as inputs, and the optimal weighting factors are generated in real time [20]. By applying integer quantization, real-time implementation on a low-cost field-programmable system-on-chip (FPGa) platform was successfully demonstrated, and transient response under varying operating points was significantly improved compared with conventional fixed-weight approaches. However, precision degradation in sensitive operating regions caused by quantization was also reported, which can be understood as an inevitable tradeoff associated with low-precision hardware implementations.

From an RL perspective, a method has been investigated in which a deep deterministic policy gradient (DDPG) agent optimizes the finite-set MPC weighting factors of a voltage-source converter in an uninterruptible power supply system online within a continuous action space [21]. By continuously updating the weighting factors using total harmonic distortion (THD) as the primary reward signal, superior waveform quality was achieved compared with manually tuned fixed weighting factors. The practical significance of RL-based approaches lies in structurally reducing dependence on expert knowledge by transforming weight optimization into an iterative exploration process.

2.2.3. RL-MPC Hybrid Structures: Online Compensation for Uncertainty

Hybrid structures combining RL and FCS-MPC in real time, without requiring offline training or separate parameter design procedures, have also been proposed to directly compensate for model uncertainty and disturbances online.

A method integrating actor–critic RL into a structure combining a neural network-based predictor and FCS-MPC has been applied to a three-phase neutral-point-clamped (NPC) converter [22]. In this approach, the control policy was learned online without prior knowledge of model parameters or predefined weighting factors. Uniform ultimate boundedness stability of the closed-loop signals was theoretically proven based on Lyapunov function analysis, and robust current tracking performance under parameter uncertainty was experimentally validated. This result is significant in demonstrating that RL can serve not merely as a performance enhancement tool, but also as a framework compatible with theoretical stability guarantees in control systems.

As a data-driven approach, iterative learning predictive control, which updates control inputs solely using input–output data without requiring mathematical models or parameter information of the system, has been applied to a three-level NPC converter [23]. Under conditions where the load inductance varied from 10 mH to 7 mH, the performance of conventional FCS-MPC degraded noticeably, whereas the proposed approach maintained robustness. However, the additional computational burden associated with iterative learning introduces a tradeoff with control performance, and when the number of iterations (N_g) exceeds 10, the computational cost may exceed the allowable range for real-time implementation.

A structure combining neural predictors with Lyapunov control theory to reconstruct the cost function itself from a stability perspective has also been proposed [24]. In this approach, both multi-objective optimization and robustness guarantees were simultaneously achieved without manual weight tuning, and experimental validation was conducted on an active front-end modular multilevel converter system.

2.2.4. Event-Based RL-MPC: Active Reduction of Switching Frequency

Although the high-speed switching capability of WBG devices provides superior power quality, excessively high switching frequencies directly lead to increased switching losses and electromagnetic interference (EMI). To address this issue, an approach combining the adaptive capability of RL with an event-triggered mechanism in FCS-MPC has been proposed for actively controlling switching frequency [25]. Through the integration of actor–critic RL and a fuzzy multi-criteria decision-making technique, THD improvement of up to 37.19% was achieved in a three-phase NPC converter under parameter mismatch conditions compared with conventional event-based FCS-MPC. By updating the control input only when an event occurs, this structure distributes computational burden non-uniformly over time and can therefore be interpreted as a practical strategy for improving implementation feasibility in WBG environments where sub-microsecond computational margins are extremely constrained.

2.3. Direct DRL Control

Unlike the AI-MPC approaches discussed in Section 2.2, which utilize AI as an auxiliary tool while preserving the conventional MPC framework, direct DRL control eliminates the prediction-optimization process of MPC itself and instead enables a neural network policy to directly learn control outputs through iterative interactions with the environment. The fundamental advantage of this approach lies in its ability to adapt to nonlinear dynamic characteristics without requiring a mathematical model of the system. Furthermore, since the computations required during policy inference consist primarily of simple matrix multiplications, the approach is structurally compatible with the real-time constraints of WBG environments. Direct DRL control formulates the power converter control problem as a Markov decision process, denoted by $\langle \mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R}, \gamma \rangle$. At each time step t , the agent observes the state $s_t \in \mathcal{S}$ and selects an action $a_t \in \mathcal{A}$ according to the policy $\pi(a_t|s_t)$, where the objective function is defined as the expected discounted cumulative reward:

$$\pi^* = \arg \max_{\pi} \mathbb{E}_{\pi} \left[\sum_{k=0}^{\infty} \gamma^k r_{t+k} \mid s_0 \right] \quad (5)$$

Here, $\gamma \in [0, 1)$ denotes the discount factor, and $r_t = \mathcal{R}(s_t, a_t)$ represents the instantaneous reward. In power electronics control problems, the action space $\mathcal{A}$ corresponds either to a continuous duty ratio or a discrete set of switching states. Policy inference is performed via a neural network forward pass of the form $\pi_{\xi} : s_t \mapsto a_t$, and its leading-order computational complexity is characterized by Equation (4).

In the following section, direct DRL-based control is reviewed by categorizing approaches according to the nature of the control objectives.

2.3.1. Direct Voltage–Current Control: Policy Learning in CPL Environments

The most actively investigated area of direct DRL-based control is voltage stabilization under CPL conditions. CPLs exhibit a negative impedance characteristic, where the current increases as the output voltage decreases, making it difficult for linear control gains to guarantee stability over a wide operating range. In such environments, DRL provides a structural advantage by enabling the exploration of optimal control policies across the entire state space without requiring an explicit system model.

A proximal policy optimization (PPO)-based approach has been proposed for stabilizing the output voltage of a DC-DC buck–boost converter [26]. In this formulation, the PPO agent takes a state vector composed of output voltage, tracking error, and error derivative as input, and directly outputs the optimal switching action (MOSFET on/off signal) at each sampling instant. In addition to the pure DRL configuration, hybrid architectures combining PI controllers and filtering stages were also evaluated.

From an online learning perspective, a DDPG-based RL controller has been proposed for real-time training of AI controllers in DC-DC switching converters [27]. Based on an actor–critic architecture, the controller achieved a reduction in settling time of at least 65% and a reduction in overshoot of 42.8% compared to a neural-network-based proportional–integral–derivative controller under a 2 MHz switching frequency environment, as validated through simulations. The ability of a single agent architecture to exhibit adaptability across multiple types of DC-DC converters is also a notable characteristic of this approach.

2.3.2. DRL-Classical Control Hybridization: Robustness Enhancement

Offline-trained DRL policies inherently suffer from performance degradation or instability in real hardware due to mismatched disturbances arising from discrepancies between simulation environments and physical systems. To address this issue, a composite control

structure has been proposed in which an Extended State Observer (ESO) is integrated as an inner loop of the DRL controller [28]. The ESO estimates external disturbances and unmodeled dynamics as a lumped disturbance, which is then compensated in real time via feedforward correction of the DRL control output, thereby partially mitigating the gap between offline training environments and real-world systems.

Experimental validation demonstrated that the proposed method maintains stable voltage tracking even under inductance variations ranging from 0.5 to 2 times the nominal design value ($L = 0.5$ mH to 2 mH), whereas conventional PI controllers exhibit pronounced steady-state oscillations under identical conditions. Notably, the ESO-based augmentation is not restricted to a specific DRL algorithm, making it broadly applicable across different RL frameworks. However, the requirement of minimal system model information for ESO gain tuning indicates that the approach partially reintroduces model dependence, thereby limiting the extent of a fully model-free paradigm.

2.3.3. Direct Switching Control in Multilevel Converters

Direct DRL control can also be extended to systems with complex switching state spaces, such as multilevel converters. A model-free switching control method based on Deep Q-Networks (DQN) has been proposed for a three-level NPC converter, achieving simultaneous voltage balancing and current tracking [5]. In this approach, the DQN agent directly defines the 27 admissible switching states of the converter as a discrete action space and selects optimal switching patterns without requiring a separate pulse-width modulation (PWM) modulator or auxiliary control loop.

Simulation and experimental results show that the output current total harmonic distortion is maintained at 3.62%, satisfying regulatory limits. Under uncertainty conditions such as DC-link capacitance variations and sensor noise, conventional MPC becomes unstable, whereas the proposed DQN-based method preserves performance. However, this approach requires extensive training time and carefully engineered reward functions, which the authors explicitly identify as remaining barriers to practical deployment.

2.3.4. Extension to Stability-Oriented Multi-Objective Optimization

A more advanced direction of direct DRL control extends beyond voltage and current tracking to include stability margins derived from small-signal impedance characteristics as part of the optimization objective. In weak grid-connected conditions, the impedance profile of power converters varies significantly with operating points, and fixed-gain controllers optimized for a specific condition may induce instability under others.

To address this issue, a multi-objective framework has been proposed in which stability metrics derived from small-signal impedance models are explicitly embedded into the reward function of the DRL agent [29]. A DDPG agent dynamically adjusts PI gains, enabling stable operation under both strong and weak grid conditions. Hardware-in-the-loop (HIL) experiments implemented on a DSP platform (TMS320F28379) further confirmed that step changes in power setpoints are tracked within 20 ms. The key contribution of this approach lies in partially overcoming the black-box nature of DRL by integrating interpretable control-theoretic constructs, namely impedance-based stability analysis, directly into the reward design.

Taken together, the approaches reviewed in Sections 2.2 and 2.3 reveal no uniformly superior control architecture. ANN-MPC imitation provides the clearest reduction in online optimization burden and has demonstrated feasibility on FPGA- and microcontroller-based platforms, but its performance remains dependent on the teacher MPC policy, training-data coverage, and the fidelity of the offline model. AI-based weight tuning and RL-MPC hybrid approaches preserve the explicit predictive structure, constraint-handling capability, and, in

some cases, stability-analysis compatibility of MPC; however, they retain partial model dependence and introduce additional online computation or adaptation. Event-based RL-MPC can reduce the average control-update and switching burden by distributing computation non-uniformly over time, although its worst-case event-handling latency must still satisfy the converter-level deadline. Direct DRL provides greater freedom from explicit models and online optimization, but places a larger burden on reward design, safety assurance, sim-to-real transfer, steady-state accuracy, and hardware validation. Therefore, the practical choice among these approaches represents a trade-off among computational latency, model dependence, explicit constraint handling, adaptability, and validation burden rather than a universal ranking of algorithmic performance. For near-term WBG deployment, architectures retaining explicit constraints may be more straightforward to validate, whereas direct DRL offers greater adaptive potential but currently involves higher deployment uncertainty. These trade-offs motivate the cross-cutting structural limitations examined in Section 2.4.

2.4. Structural Limitations in Practical Deployment

As discussed in the previous sections, AI-enhanced MPC and direct DRL control strategies have demonstrated control performance comparable to or exceeding that of conventional control methods in simulation and laboratory environments. However, it should be noted that many of these results were obtained under idealized conditions, including accurate simulation models, sufficient exploration time, and high-performance computational environments. In practical deployment scenarios of AI control for WBG-based power converters, such conditions are often unavailable. Consequently, a distinct gap exists between algorithm-level performance and practical deployment feasibility. This gap is not an isolated or incidental issue, but can instead be systematically categorized into four fundamental limitations originating from the structural characteristics of AI-based control.

Most RL algorithms employ feedforward ANNs as policy approximators. This corresponds to a stateless structure in which control actions are generated solely from the current state at each sampling instant, lacking the accumulated error compensation mechanism corresponding to the integral term of PI control. As a result, RL-based controllers often exhibit fast tracking characteristics during transient periods while structurally producing non-vanishing steady-state error under steady-state conditions [14]. This phenomenon can be expressed mathematically as follows. In the stateless structure where the RL policy π_{ξ} receives only the current state s_t as input, the steady-state error between the reference value v^* and the output $v(t)$ is given by:

$$e_{ss} = \lim_{t \rightarrow \infty} (v^* - v(t)) = v^* - \pi_{\xi}(s_{\infty}) \neq 0 \quad (6)$$

The integral term in PI control structurally guarantees $e_{ss} \rightarrow 0$. However, according to Equation (6), eliminating e_{ss} requires that s_{∞} inherently contain complete information regarding v^* . The integral action state augmentation (IASA) method proposed in [13] compensates for this issue by extending s_t to $[s_t, \int_0^t e, d\tau]$. However, this extension increases n_0 in Equation (4) by the number of augmented integral states—typically one or two additional dimensions—thereby increasing the first-layer parameter count from $n_0 \cdot n_1$ to $(n_0 + n_{\text{aug}}) \cdot n_1$ and imposing additional pressure on the computational budget of Stage 4.

This study explicitly analyzed that steady-state error in power electronic systems originates fundamentally from the approximation nature of RL control policies and finite training duration. Furthermore, it demonstrated that the Integral Action State Augmentation (IASA) technique can reduce steady-state error by up to 52%. However, because IASA augments the *actor input* (i.e., the state vector fed into the policy network π_{ξ}) rather than

the actor output, the first-layer weight matrix expands accordingly, increasing the total inference complexity N_{ANN} in proportion to $(n_0 + n_{\text{aug}})$. This growth in input dimensionality imposes an additional burden on the feasibility of real-time inference in high-speed WBG environments.

RL policies are generally trained in simulation environments, where factors such as parasitic inductance, dead time, sensor noise, and temperature-dependent resistance variations of actual hardware are either neglected or simplified. Consequently, when policies validated in simulation are deployed to real systems, unexpected performance degradation or instability may occur, resulting in the so-called simulation-to-real gap (sim-to-real gap). This issue is particularly severe in power electronic systems because the control periods of typical drive control systems are only 10–200 μs , making direct exploratory training on real-time hardware practically infeasible [8]. In this study, a transfer pipeline from simulation pretraining to a real experimental testbench was constructed for permanent magnet synchronous motor control, and experimental results confirmed that control performance significantly deteriorated immediately after transfer because of model mismatch, thereby necessitating additional online learning on the physical system. As illustrated in Figure 4, the mean absolute error (MAE) during simulation pre-training converges smoothly to near-zero within approximately 20 training steps, whereas adaptation on real hardware exhibits substantially higher initial error and greater volatility throughout the learning process, requiring twice as many steps to reach comparable accuracy. In particular, the authors' assessment that current data-driven RL controllers merely achieve performance comparable to MPC methodologies developed over more than 20 years reflects the still immature state of RL control for practical power electronic systems.

The learning mechanism of RL fundamentally relies on trial-and-error-based exploration. Although such exploration is freely permitted in software simulations, random exploratory actions in real power converters can immediately induce overcurrent or over-voltage conditions, resulting in irreversible damage to semiconductor devices or power circuits. This issue cannot be resolved solely through reward functions based on soft penalties. In penalty-based approaches, the agent must still experience constraint-violating states during training in order to learn their associated risks, which represents a fundamental incompatibility in power electronic environments where such violations must not merely be discouraged but rendered impossible [30]. While standard RL algorithms frequently violate voltage stability limits during training, power system applications have demonstrated that structures incorporating safety filters based on control barrier functions (CBFs) can mathematically guarantee constraint satisfaction even during exploration [31]. However, such safety-guaranteed structures require prior knowledge of the control-affine dynamics of the system, which still limits their applicability in fully model-free environments. Furthermore, strengthening stability-guarantee conditions narrows the feasible exploration space, thereby introducing a tradeoff in which policy optimality may degrade [13].

Whereas the preceding three limitations primarily arise at the algorithmic level, the fourth limitation originates from constraints imposed by physical time scales. In GaN-based converters operating at switching frequencies in the MHz range, the inference time allowed for the controller is below the microsecond level. Within this interval, forward-pass computations of multilayer neural networks consisting of tens to hundreds of neurons must be completed. When implementing a standard floating-point DDPG actor on a general-purpose microcontroller, inference cannot be completed within the control period unless the neural network size is sufficiently reduced, and such compression inevitably degrades the representational capacity of the policy [32]. This study experimentally verified that the network size of a DDPG-based neural network must be substantially reduced for implementation on a general-purpose DSP (TI DSP F28379D), emphasizing this result as an

important precedent for moving beyond conventional FPGA-dependent implementations. Nevertheless, because the determination of the tradeoff point between neural network size and control performance still depends heavily on designer experience, systematic automation of this process remains an unresolved challenge.

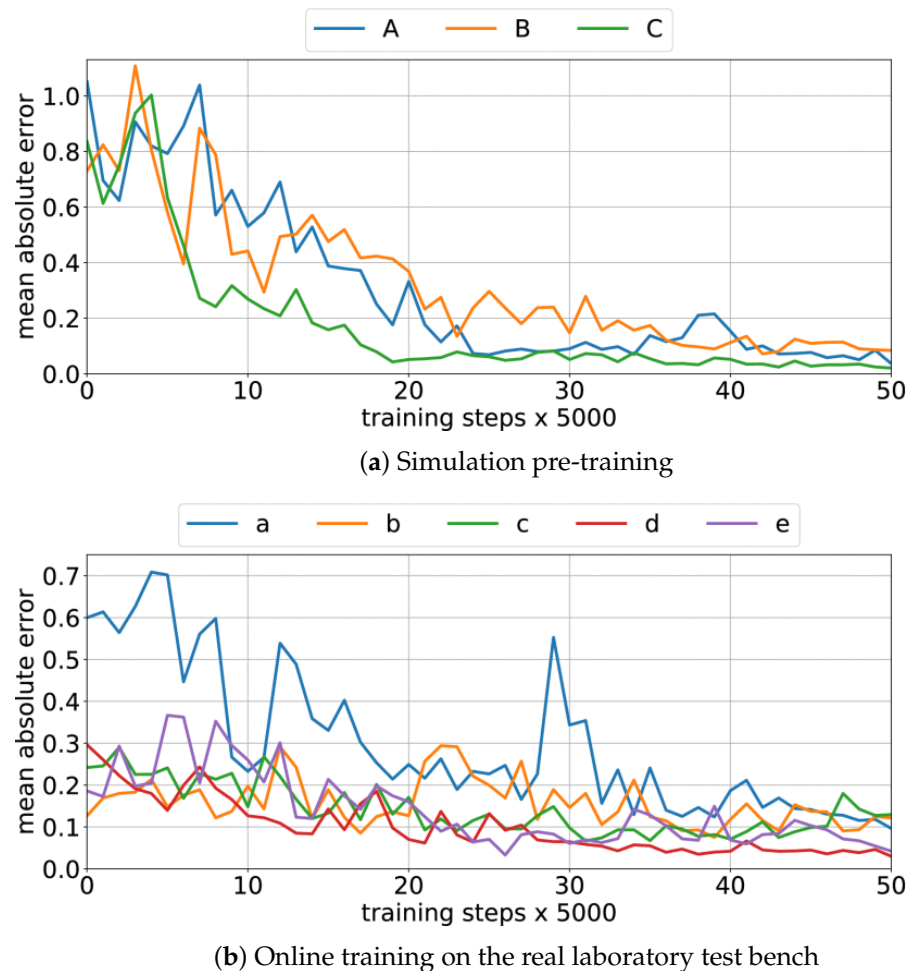


Figure 4. Pre-training and real-hardware adaptation performance of DDPG-based RL controllers for permanent-magnet synchronous motor (PMSM) current control. (a) MAE convergence during simulation pre-training (software-in-the-loop (SIL) model); (b) MAE convergence during online training on the real laboratory test bench. The substantially higher initial error, greater volatility, and slower convergence observed in (b) relative to (a) quantitatively illustrate the sim-to-real gap: policies that converge well in simulation require considerable additional online adaptation before achieving comparable performance on physical hardware. Reproduced from [8], CC BY 4.0.

The four limitations discussed above are not independent, but instead form a sequential dependency structure. Without accurate system modeling, the foundation for designing safety constraints cannot be established. Without guaranteed safety, exploratory training on real hardware becomes infeasible, and the simulation-to-real gap cannot be safely characterized or bridged. Finally, even if all these conditions are satisfied, deployment in high-frequency WBG systems remains infeasible unless inference satisfies the timing and resource constraints of the target converter, which may fall below the microsecond scale in MHz-range operation. These dependencies motivate an integrated evaluation of deployment feasibility rather than separate optimization of individual performance metrics. Section 3 therefore formalizes the four conditions as staged analytical requirements.

3. The AI Deployment Readiness Framework

The four structural limitations identified in Section 2.4 are treated here as interdependent deployment concerns. The resulting framework examines how modeling accuracy, safety assurance, sim-to-real transfer, and hardware implementation feasibility jointly affect practical deployment. These concerns define four analytical dimensions: (1) modeling accuracy, (2) safety assurance, (3) sim-to-real transfer, and (4) hardware implementation feasibility.

Within this structure, limitations in an earlier dimension can constrain subsequent dimensions, while cyclic co-design dependencies may also arise, as discussed in Section 3.5. Insufficient modeling accuracy weakens the foundation for safety analysis, and simulation-domain safety assurance does not automatically carry over to real hardware. Unresolved sim-to-real mismatch can also prevent performance preservation after transfer. Even when these conditions are satisfied, deployment remains infeasible if inference exceeds the timing or resource limits of the target hardware. Figure 5 summarizes these dimensions and their interdependencies.

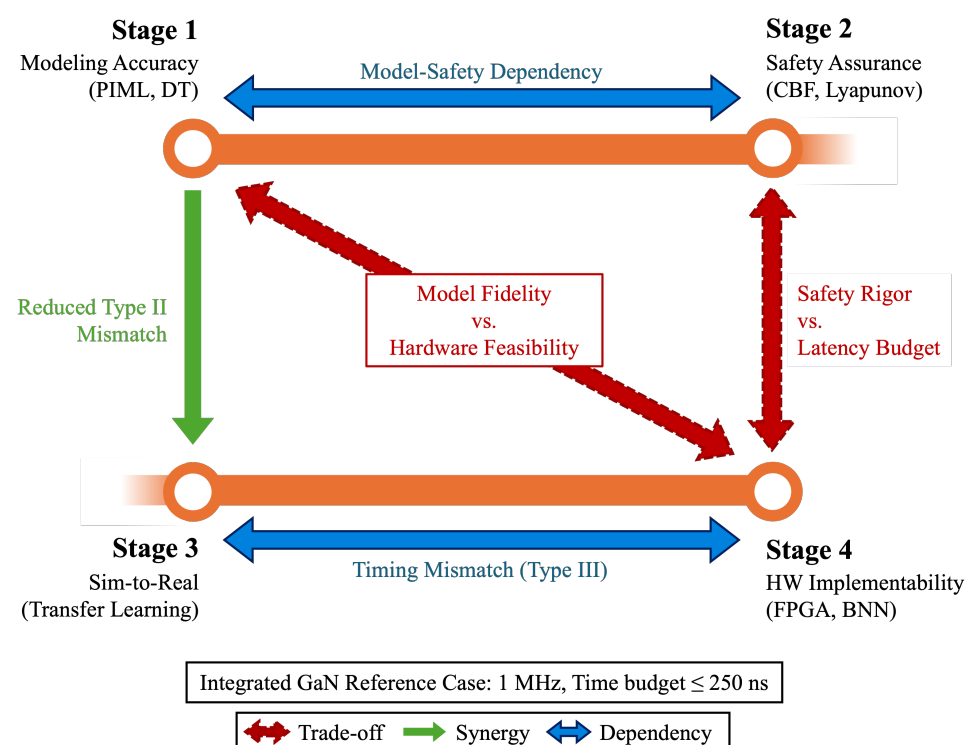


Figure 5. The proposed AI Deployment Readiness Framework consisting of analytical dimensions: (1) Modeling Accuracy, (2) Safety Assurance, (3) Sim-to-Real Transferability, and (4) Hardware Implementability. Arrows indicate three inter-stage relationships: sequential dependencies (blue), synergistic connections (green), and structural trade-offs (red). For the most restrictive integrated GaN control-protection case considered in this review, Stage 4 is further constrained by an approximately 250 ns inference budget at 1 MHz switching frequency. In architectures with independent hardware protection, the digital-control budget may be less restrictive.

3.1. Stage 1: Modeling Accuracy and Digital Twins

The central question of Stage 1 is not simply whether an AI model can accurately reproduce converter behavior. A more fundamental question is whether a modeling strategy exists that can simultaneously satisfy three objectives: accuracy, data efficiency, and OOD generalization. These objectives are competing, and any design decision favoring one objective inevitably compromises the other two. Because the design choices made in Stage 1 directly affect both Stage 3 (sim-to-real transfer) and Stage 4 (hardware implementation feasibility), understanding the tradeoffs among these three objectives constitutes the starting point for discussing the entire framework.

Purely data-driven approaches partially satisfy the first objective, namely accuracy, because they can achieve high approximation precision when sufficient data are available. However, this approach sacrifices the remaining two objectives simultaneously. From the perspective of data efficiency, analyses have shown that approximately one million data points are required to precisely map current stress under hybrid modulation strategies in dual active bridge (DAB) converters [33], which is difficult to achieve realistically in WBG systems with limited experimental environments. A more fundamental issue is the failure of OOD generalization. Purely data-driven models provide unreliable predictions once the operating conditions deviate beyond the distribution boundaries of the training data. This limitation becomes critical in the context of Stage 3, where distribution shift inevitably exists between simulation environments and physical systems. In other words, even if purely data-driven models satisfy the requirements of Stage 1, their inherent vulnerability inevitably reappears in Stage 3.

To address these limitations, physics-informed machine learning (PIML) approaches have emerged by incorporating circuit physics into AI models. However, an important distinction must be made here: where physical laws are embedded fundamentally determines the extent to which the three objectives can be achieved. The dominant approach in current research, referred to as the physics-in-loss (PL-PINN) method, incorporates physical residuals derived from circuit differential equations as additional terms in the loss function. In this case, the loss function is formulated as follows:

$$\mathcal{L}_{\text{PL}} = \mathcal{L}_{\text{data}} + \alpha \cdot \mathcal{L}_{\text{phys}}, \quad \mathcal{L}_{\text{phys}} = \left\| \frac{d\hat{x}}{dt} - f(\hat{x}, u; \theta) \right\|^2 \quad (7)$$

Here, α denotes the weighting factor associated with the physical constraint, and $f(\hat{x}, u; \theta)$ represents the right-hand side of the circuit differential equation. The selection of the weighting factor α determines the tradeoff between convergence and physical consistency. If this value is excessively small, $\mathcal{L}_{\text{phys}}$ becomes ineffective; conversely, if it is excessively large, the training process fails to converge.

Although this approach encourages compliance with physical laws during training, the loss function merely provides a guiding direction throughout the training process, meaning that physical consistency may still be violated during inference. A more critical limitation is that complete retraining is required whenever new operating conditions are introduced. This prevents online adaptation in WBG systems, where operating conditions vary rapidly, and inherently limits the post-transfer adaptation capability required in Stage 3. Furthermore, the selection of the physical loss weighting factor still relies on trial-and-error procedures. Consequently, the method inherently suffers from a tuning problem in which excessively small weighting factors render physical constraints ineffective, whereas excessively large weighting factors prevent convergence [34].

In contrast, when circuit differential equations are directly embedded into the architecture through forward Euler discretization, the following recursive relationship becomes encoded into the neural network structure itself [33]:

$$\hat{x}(k+1) = \hat{x}(k) + T_s \cdot f(\hat{x}(k), u(k); \theta) \quad (8)$$

For example, for the inductor current i_L and output voltage v_o of a buck converter:

$$i_L(k+1) = \left(1 - \frac{R_L T_s}{L}\right) i_L(k) + \frac{T_s}{L} (S \cdot V_{in}(k) - v_o(k)) \quad (9)$$

In Equation (9), physical consistency is not a byproduct of the loss function but rather a design constraint embedded directly into the architecture itself, and is therefore structurally guaranteed under arbitrary operating conditions. The essential distinction between the

two approaches is that, unlike Equation (7), where $\mathcal{L}_{\text{phys}}$ merely serves as a guiding direction during training, the physical relationships in Equation (8) remain identically satisfied even during inference.

However, this structural guarantee holds strictly only when the architectural weights faithfully encode the exact physical parameters. Specifically, in Equation (9), the coefficients $(1 - R_L T_s / L)$ and (T_s / L) must be represented without approximation for Kirchhoff's laws to remain satisfied. Once Stage 4 quantization replaces these weights with their fixed-point approximations $\tilde{w}$, the encoded physical relationships are corrupted:

$$\tilde{w}_1 = Q_b \left(1 - \frac{R_L T_s}{L} \right) \neq 1 - \frac{R_L T_s}{L} \quad (10)$$

and the physical consistency guarantee is therefore no longer structurally enforced during inference. This introduces a critical asymmetry: the physics-in-architecture approach achieves the strongest Stage 1 guarantee precisely because physical parameters are directly encoded as weights, yet this same property makes it uniquely sensitive to Stage 4 quantization error. In other words, the approach offering the most rigorous Stage 1 compliance is simultaneously the most fragile under Stage 4 compression—a specific and severe instantiation of the Stage 1 $\leftrightarrow$ Stage 4 cyclic dependency that has not been explicitly analyzed in prior work.

However, the computational structure embedded within the architecture in Equation (8) may conflict with the T_{budget} requirement imposed by Equation (2), and this cyclic dependency forms the core tradeoff between Stage 1 and Stage 4.

Furthermore, predictions under new operating conditions can be performed solely through parameter adjustment without requiring retraining, and this on-demand prediction capability fundamentally reduces model transfer costs in Stage 3. A study demonstrating high-accuracy estimation of the inductance, equivalent series resistance and on-resistance of a buck converter using only 360 training samples [35], as well as a data-light pipeline employing differentiated data strategies for loss components with varying acquisition costs [34], can both be interpreted as practical implementations of this physics-in-architecture paradigm. However, this advantage is conditional on Stage 4 co-design: as established in Equation (10), Stage 4 quantization directly corrupts the embedded physical parameters rather than merely reducing general representational capacity, eliminating the structural consistency guarantee that constitutes the primary advantage of the approach. Physics-in-architecture models should therefore be adopted only when the target hardware precision is determined concurrently with training via QAT.

However, there exists a broader insight that extends beyond the simple conclusion that “physics-in-architecture is superior.” As the depth of physics integration increases — that is, as the model achieves a more balanced satisfaction of the three objectives — additional computational structures for solving circuit differential equations are embedded within the architecture, thereby increasing inference latency. This directly conflicts with the sub-microsecond inference-time constraint required in Stage 4 (hardware implementation feasibility). In other words, better models in Stage 1 inherently require more difficult implementations in Stage 4. This cyclic dependency implies that the two stages cannot be optimized independently. Nevertheless, studies explicitly quantifying this tradeoff or attempting integrated optimization remain difficult to find in the field of power electronics. This represents one of the most unresolved gaps within the overall framework.

Beyond the static modeling paradigm of Stage 1, DTs provide a dynamic modeling framework that continuously synchronizes with real-time sensor data from physical systems, thereby reflecting device degradation and changing operating conditions online. The key significance of this approach lies in blurring the boundary between Stage 1 and Stage 3.

Whereas static models follow a sequential process in which learning occurs in Stage 1 and transfer occurs in Stage 3, DTs continuously update the model using feedback from the physical system, causing the sim-to-real gap to decrease progressively during operation. In a 75 kW, 60 kHz SiC-based multiphase interleaved boost converter, a study combining a state-space-model-based DT with metaheuristic optimization demonstrated non-invasive real-time estimation of inductor and capacitor health indicators [36]. Experimental results showed that the decoupled structure improved estimation accuracy by 9.4% while reducing embedded computational requirements by 22%. The explicit observation in this study that high-speed switching of WBG devices accelerates passive component degradation demonstrates that the accuracy requirement in Stage 1 constitutes a dynamic problem that evolves over time.

However, the online updating capability of DTs itself consumes computational resources, thereby introducing another tradeoff between DT real-time operation and the inference-time constraints of Stage 4. Ultimately, the fundamental question raised in Stage 1 extends beyond mere accuracy to encompass computational lightness and robustness to varying operating conditions. To date, no systematic methodology capable of simultaneously satisfying these combined requirements has been comprehensively established in the field of WBG power electronics.

3.2. Stage 2: Safety Assurance

The central question of Stage 2 is not merely how RL can be made safe, a question already addressed by dozens of prior studies. A more incisive question concerns what type of safety guarantee holds under which assumptions and what must be sacrificed in return. Viewed from this perspective, the progression of existing safe RL methodologies—from soft-penalty approaches, to Lyapunov-based methods, and ultimately to CBF-based frameworks—is revealed not as a simple improvement in performance, but as a systematic tradeoff between the strength of guarantees and the amount of required prior knowledge. This tradeoff conflicts with the deployment feasibility of AI-based power converter control in a uniquely critical manner.

Incorporating safety constraints into the reward function as penalty terms is widely adopted because of its implementation simplicity and compatibility with virtually any RL algorithm. However, the failure of this approach is not merely a parameter-tuning issue, but rather a structural limitation. It has been theoretically established that even when soft penalties appear to achieve safety during training, they do not guarantee safety during the testing (deployment) phase [30,37]. Empirically, standard RL policies have been shown to exhibit non-negligible unsafe rates during deployment evaluation, whereas barrier-certificated structures reduce constraint violations by more than an order of magnitude [31]. The significance of this result lies in the fact that, in AI-based power converter control, the training phase corresponds to simulation, whereas the testing phase corresponds to real hardware deployment. In other words, policies trained using soft penalties fail to satisfy the requirements of Stage 2 by definition.

The problem of selecting the penalty coefficient λ is also structural in nature. If λ is excessively small, safety violations are not sufficiently suppressed; conversely, if it is excessively large, convergence is hindered. No principled method currently exists for automatically determining this operating region, and even dynamic tuning cannot guarantee safety under all conditions. This implies that soft-penalty approaches are incapable of satisfying the core requirement of Stage 2, namely, that safety violations must never occur in real deployment environments.

Methods that integrate Lyapunov theory into safe RL are not monolithic; rather, the nature of the guarantees fundamentally depends on how the integration is performed. Understanding this distinction is essential for properly evaluating Stage 2.

The first interpretation employs Lyapunov functions as constraint generators within constrained Markov decision processes [37]. Through a linear-programming-based algorithm that decomposes global safety guarantees into a set of local linear constraints, this approach guarantees that a safe policy is always returned, and once the policy enters the safe set, all subsequent policy updates remain safe. In terms of guaranteeing safety throughout the entire training process, this approach surpasses soft-penalty methods. In power-system voltage control applications, approaches utilizing explicitly constructed Lyapunov functions as stability certificates achieved voltage stability consistently while reducing transient control cost by more than 30% and shortening response time by approximately one-third compared with linear policies [38]. This stands in clear contrast to conventional RL methods, which frequently violate voltage stability constraints. However, constructing suitable Lyapunov functions is itself non-trivial, and the scale of the associated linear programs grows exponentially in large state spaces.

The second interpretation employs violations of Lyapunov conditions as regularization terms during policy updates [39]. Because penalties are imposed only when actions violate the Lyapunov decrease condition ($\dot{V}(x) < 0$), this approach is more physically grounded than conventional soft penalties. However, it still operates as a guidance-based mechanism and therefore cannot completely eliminate violations. Its practical advantage lies in the fact that the Lyapunov function can be learned using neural networks rather than manually designed analytically. In transient stability control for power systems, this approach was shown to improve stabilization capability while simultaneously reducing power losses compared with RL methods without Lyapunov regularization.

The third interpretation adopts the most structural approach by designing the neural-network controller architecture itself to satisfy Lipschitz constraints [40]. Because the controller is structurally designed to satisfy stabilization conditions, exponential stability can be guaranteed without additional runtime verification, while decentralized learning using only local data at each bus becomes feasible. The central advantage of this approach is that the additional computational burden associated with safety guarantees is concentrated entirely during training and does not arise during inference. This characteristic minimizes conflicts with Stage 4 (hardware implementation feasibility).

When these three interpretations are examined collectively, a clear pattern emerges. As the strength of guarantees increases (from the first to the third interpretation), the design of Lyapunov functions or architectural constraints must become progressively more sophisticated, and all such approaches require prior knowledge of system dynamics. This forms a cyclic dependency between Stage 1 and Stage 2. If the model established in Stage 1 is inaccurate, the Lyapunov function designed on the basis of that model cannot guarantee safety conditions in the actual system. In the direction of acquiring such prerequisite information, a method integrating physics-informed neural networks (PINNs) directly into power-system swing equations to simultaneously identify system inertia and damping coefficients in a data-driven manner has been proposed [41]. This study explicitly demonstrates that the model information established in Stage 1 directly constitutes the design foundation for Lyapunov- and CBF-based methods in Stage 2.

CBF-based methods provide stronger real-time safety guarantees than Lyapunov-based methods by projecting or filtering control inputs so that the system state never leaves the safe set at any time instant t . In a power-system application combining neural-network barrier certificates with RL policies, standard RL policies repeatedly violated

constraints during training, whereas the CBF-filtering structure successfully prevented such violations [31].

The core mechanism of the CBF-based safety filter is to project the RL policy output u_{RL} into the safe set in real time at every control step through the following quadratic programming (QP) problem:

$$u^* = \arg \min_u \|u - u_{RL}\|^2 \quad \text{s.t.} \quad \dot{h}(x, u) + \gamma h(x) \geq 0 \quad (11)$$

Here, $h(x) \geq 0$ denotes a continuously differentiable function defining the safe set $\mathcal{C} = \{x \mid h(x) \geq 0\}$, and $\gamma > 0$ is the coefficient of an extended class- $\mathcal{K}$ function. If the constraint condition in Equation (11) is satisfied, then $h(x(t)) \geq 0$ is guaranteed for any arbitrary $t \geq 0$, ensuring that the state trajectory never leaves the safe set [31]. The solution time required for this QP, denoted by T_{QP} , depends on the state dimension n_s and the number of constraints, thereby forming the following tension relationship with the timing budget T_{budget} in WBG systems:

$$T_{QP}(n_s) \geq T_{\text{budget}}(f_{sw}) \quad (12)$$

Equation (12) represents the mathematical expression of the Stage 2 $\leftrightarrow$ Stage 4 tension discussed in Section 3.5.

However, two conditions must be satisfied for CBFs to be directly applicable to power converter control. First, the system must possess a control-affine structure in order to compute the time derivative of $h(x, u)$, and accurate knowledge of this structure is required. Nonlinear switching converters satisfy this requirement only approximately. Second, solving a QP problem is required at every control step to verify the CBF condition and perform the projection operation, and this additional computation directly consumes the sub-microsecond timing budget of WBG systems. In other words, although CBFs satisfy Stage 2 most strongly, they also pose the greatest threat to Stage 4. Within the literature examined in this review, this conflict has not yet been systematically resolved.

A comprehensive examination of the existing literature on Stage 2 reveals a common characteristic: the most mathematically rigorous safety guarantees have been validated primarily at the power-system level, including voltage control, frequency control, and transient stability applications. In these applications, the characteristic timescales of safety-related dynamics range from several milliseconds to several seconds. In contrast, the switching cycles of WBG power converters occur on sub-microsecond timescales, while device damage due to overcurrent may arise within several hundred nanoseconds [17]. This discrepancy in timescales is not merely a matter of implementation speed, but instead implies that the very definition of the safe set changes fundamentally. In power systems, the safe set is defined relatively smoothly, such as through permissible voltage magnitude ranges. In converters, however, safety boundaries are determined by instantaneous current limits that vary at every switching cycle. Among the existing Lyapunov- and CBF-based safe RL methods, no experimentally validated case operating at the converter switching timescale was identified within the scope of the literature considered in this work. Furthermore, re-validation of safety properties on real hardware following sim-to-real transfer—here designated Stage 2'—constitutes an implicit requirement of the sequential framework, but has not been systematically addressed in the studies examined in this review.

Although a standardized Stage 2' procedure has not yet been established, two complementary studies indicate how AI controllers can be evaluated on physical hardware with reduced risk. In [12], a computationally light one-step MPC safety policy was employed to reject DQN-selected switching states predicted to violate the converter-current limit; the controller was trained in a matched simulation environment and subsequently validated on

a practical two-level voltage-source converter (VSC). In contrast, direct online RL training was demonstrated on a PMSM-drive test bench using a recursive least-squares (RLS)-based safeguard that excluded switching actions predicted to violate current or voltage constraints [42]. The hardware validation proceeded in three steps: separate stress-testing of the safeguard using random switching actions at low- and high-speed operating points, safeguarded online training, and final controller-performance testing with exploration disabled. These studies suggest a practical minimum procedure consisting of independent safeguard validation, conservative margins between nominal and absolute limits, safeguarded learning or deployment, and separate post-training performance evaluation. Nevertheless, neither study establishes formal Lyapunov/CBF-level guarantees or demonstrates hardware feasibility under the 250 ns integrated GaN control-protection reference case considered in this review.

3.3. Stage 3: Sim-to-Real Transferability

The dynamics of the real system, denoted by $\tilde{f}$, can be represented as the simulation model f superimposed with three heterogeneous mismatch terms:

$$\tilde{f}(x, u) = f(x, u) + \underbrace{\delta_I(x, u)}_{\text{Type I: parameter}} + \underbrace{\delta_{II}(x, u)}_{\text{Type II: architecture}} + \underbrace{\delta_{III}(k)}_{\text{Type III: timing}} \quad (13)$$

Here, δ_I represents temperature- and manufacturing-induced variations in inductance, capacitance, and on-resistance; δ_{II} denotes structurally unmodeled effects such as parasitic inductance, dead time, and switching ringing; and δ_{III} corresponds to simulation inaccuracies arising from ADC sampling, computation, and PWM update delays.

Equation (13) explicitly establishes the fundamental connection between Stage 2 and Stage 3. For analytical convenience, the aggregate mismatch is defined as $\delta(x, u, k) \triangleq \delta_I(x, u) + \delta_{II}(x, u) + \delta_{III}(k)$. For a Lyapunov function $V(x)$ designed based on the simulation model f , the Lyapunov condition in the real system can be decomposed exactly as follows due to the linearity of $\dot{V} = \nabla V(x)^\top \dot{x}$:

$$\dot{V}(\tilde{f}) = \nabla V(x)^\top \tilde{f}(x, u) = \underbrace{\nabla V(x)^\top f(x, u)}_{\dot{V}(f)} + \nabla V(x)^\top \delta(x, u, k) \quad (14)$$

As long as $\delta \neq 0$, the condition $\dot{V}(\tilde{f}) < 0$ is not guaranteed even if $\dot{V}(f) < 0$ holds. Since δ_{III} originates from discrete-time timing effects, the continuous-time formulation in Equation (14) is valid only under a continuous-time approximation of $\delta_{III}(k)$. The corresponding discrete-time decomposition is:

$$\Delta V_{\tilde{f}}(x_k) = \underbrace{[V(x_{k+1}^f) - V(x_k)]}_{\Delta V(f)} + \underbrace{[V(x_{k+1}^{\tilde{f}}) - V(x_{k+1}^f)]}_{\leq L_V \|\delta_k\|} \quad (15)$$

where L_V is the Lipschitz constant of V and $\delta_k \triangleq \delta(x_k, u_k, k)$. Equation (15) shows that stability is preserved in the real system if and only if $\Delta V(f) < -L_V \|\delta_k\|$, i.e., the Lyapunov decrease under the simulation model must exceed the Lipschitz-weighted mismatch magnitude. For stochastic Type III timing jitter, bounding $\|\delta_k\|$ requires statistical characterization of ADC and PWM delay distributions—information rarely available in WBG simulation environments and constituting an open problem beyond the scope of the present analysis.

Equation (14) mathematically represents the mechanism through which Stage 2 safety evidence may not remain valid during Stage 3, and constitutes one of the primary reasons why the proposed framework requires the two stages to be verified independently.

The sim-to-real gap is not a single phenomenon, but rather a composite manifestation of three distinct types of mismatch. First, parameter mismatch (Type I) arises when circuit parameters such as inductance, capacitance, and on-resistance deviate from their simulation design values due to temperature variation, aging, and manufacturing tolerances. Second, structural mismatch (Type II) originates from physical structures that are omitted or excessively simplified in circuit simulations, including parasitic inductance, dead time, switching ringing, and PCB layout effects. Third, computational timing mismatch (Type III) results from the inaccurate modeling of ADC sampling delay, computation delay, and PWM update delay in simulation environments. This distinction is important because most existing sim-to-real approaches focus almost exclusively on Type I mismatch, often leading to the misconception that the overall gap has been fully resolved.

The most direct approach for mitigating parameter mismatch is to compensate for the altered input–output relationship that emerges when an offline-trained policy is deployed on the real system. Duty Ratio Mapping (DRM) enables direct policy transfer without retraining by offline identification of the relationship between the action space in simulation and the effective duty ratio in actual hardware [11]. However, because the mapping itself is obtained under specific operating points and parameter ranges, even slight variations in system parameters require remapping, thereby making DRM inherently a form of conditional transfer.

Unlike DRM, which assumes a single-degree-of-freedom (1-DoF) duty-ratio space, the action space expands to three dimensions when triple-phase-shift modulation is adopted in Dual-Active-Bridge (DAB) DC-DC converters. A study that directly transferred an offline-trained RL policy to real hardware for power conversion efficiency optimization demonstrated that the offline-learning/direct-transfer pipeline remains feasible even in more complex action spaces [43]. However, the study simultaneously confirmed that transfer errors accumulate as the action dimensionality increases, thereby amplifying the limitations of conditional transfer under identical conditions. These findings clarify the structural challenges faced by transfer methodologies beyond DRM, while also suggesting that sim-to-real transfer research in high-dimensional action spaces remains at an early stage.

An alternative approach incorporates an Extended State Observer (ESO) into the inner control loop, thereby relaxing this constraint by estimating and compensating for the lumped disturbance between the offline training environment and the real system in real time [28]. Although both approaches improve robustness against Type I mismatch, ESO gain design still requires at least partial model information and, in principle, cannot capture Type II and Type III mismatches.

Studies that explicitly analyze Type II mismatch (structural unmodeled dynamics) remain limited. However, in a study that applied the PPO algorithm to both continuous conduction mode and discontinuous conduction mode operation of a buck converter, non-ideal hardware characteristics—including IGBT and diode dead time, real-time computational delay, and chattering suppression—were explicitly incorporated into the reward function design and state-feature construction. Experimental results confirmed that neglecting such effects constitutes one of the primary causes of post-transfer performance degradation [44]. In particular, the study emphasized that neural network architecture optimization and input feature design influence practical deployment feasibility as significantly as algorithm selection itself, thereby suggesting that improving the fidelity of the simulation model should precede algorithm-level approaches for mitigating Type II mismatch.

An even more noteworthy observation is that direct transfer fails to hold even for passive predictive models that are substantially simpler than control policies, namely power-loss estimation models. The discrepancy between power-loss models constructed solely from simulation data and experimentally measured values is both systematic and

reproducible. To address this issue, a three-stage transfer-learning pipeline was proposed: a source-domain model is first trained using large-scale simulation data (233,280 samples), followed by fine-tuning in the target domain using only a small amount of experimental data (45 samples) [45]. This approach reduced the average power-loss estimation error by 50% compared with conventional methods. The implication of this result for Stage 3 analysis is clear. If even a physically deterministic loss model cannot achieve sufficient accuracy without domain adaptation, then direct transfer of a control policy obtained through exploration-based learning is likely to induce substantially more severe mismatch. The fact that this study experimentally demonstrated the necessity of transfer learning for loss models paradoxically establishes a lower bound on the difficulty of domain adaptation required for control-policy transfer.

Whereas existing sim-to-real transfer approaches predominantly focus on post-hoc error compensation, meta-learning has recently emerged as a promising alternative for structurally reducing the transfer cost itself. In a study applying the model-agnostic meta-learning algorithm to broadband impedance identification of photovoltaic converters for maximum power point tracking control, a pretrained initialization model accurately predicted broadband impedance under unseen weather conditions using only a few frequency-point measurements. Compared with conventional ANN-based approaches, the amplitude and phase prediction errors were reduced by more than one order of magnitude, while robust generalization performance was maintained even under extreme weather conditions [46]. The essential contribution of meta-learning lies in learning parameter initializations optimized for rapid adaptation to new domains, which directly translates into a structural capability for minimizing the domain-adaptation cost required in Stage 3 without full retraining. Extending this concept to direct power converter control applications, an adaptive meta-DDPG (AM-DDPG)-based controller was applied to voltage regulation of an interleaved DC-DC boost converter. The meta-controller dynamically adjusted critical hyperparameters, including the learning rate, noise scale, and stability coefficient, in real time, thereby achieving voltage stabilization within 10 ms and output-voltage deviation below 0.1% even under rapid irradiance variation conditions reaching 200 W/m²/s [47].

Collectively, these studies indicate that meta-learning represents a promising pathway for structurally reducing Stage 3 transfer costs. However, both studies remain limited to simulation-based validation, while the adaptation cost within the microsecond timescale of high-speed WBG switching environments has not yet been quantitatively characterized.

One of the most easily overlooked structural issues in the context of Stage 3 is that the safety guarantees established in Stage 2 may become invalidated during Stage 3. Since Lyapunov functions and CBFs are designed based on the simulation-model dynamics $f(\mathbf{x}, \mathbf{u})$, the corresponding safety conditions are no longer guaranteed under the real-system dynamics $\tilde{f}(\mathbf{x}, \mathbf{u}, \mathbf{k}) = f(\mathbf{x}, \mathbf{u}) + \delta(\mathbf{x}, \mathbf{u}, \mathbf{k})$, where δ denotes the model mismatch term. This non-propagation problem provides the rationale for independently validating Stage 2 and Stage 3, and constitutes one of the primary design motivations for the sequential organization of the proposed framework. At present, the power-electronics field lacks a standardized procedure for revalidating safety on real hardware following simulation-based safety verification, thereby creating the risk that simulation-domain safety evidence may be overinterpreted as sufficient for deployment.

Stage 1 and Stage 3 are logically interconnected within the proposed framework. As the physical fidelity of the simulation environment improves—equivalently, as Stage 1 quality increases—the structural mismatch (Type II) between simulation and reality correspondingly decreases. In particular, the structural compliance of physics-in-architecture approaches with circuit physical laws provides the potential to fundamentally reduce adaptation costs in Stage 3 by enabling on-demand prediction under new operating conditions

without retraining [33]. However, existing Stage 1 and Stage 3 studies have largely been conducted independently, and no prior work has experimentally validated this relationship. To date, no study has quantitatively analyzed the extent to which improvements in Stage 1 model fidelity translate into reductions in Stage 3 transfer error, making this a critical gap that must be addressed from the perspective of the proposed framework.

The high dv/dt and di/dt induced by the high-speed switching characteristics of WBG devices render Type II mismatch qualitatively different from that observed in Si-based systems. Voltage slew rates on the order of several hundred V/ns interact with parasitic inductance within the circuit, generating switching ringing and EMI, both of which are difficult to reproduce accurately using standard circuit simulators. From the perspective of control-loop design, when these ringing components are embedded in current and voltage sensor signals, the AI controller inevitably receives state vectors containing high-frequency noise patterns that were absent during training, thereby inducing OOD conditions. Unlike Type I mismatch, this form of structural discrepancy cannot be resolved through additional parameter measurement or compensation. Consequently, the sim-to-real problem in AI-based WBG power converter control is more challenging than in conventional Si-based applications.

3.4. Stage 4: Hardware Implementability and Inference Latency

Even if the conditions of Stages 1–3 are satisfied sequentially, deployment fundamentally fails if the AI controller cannot complete inference within the WBG switching period. Stage 4 addresses the conditions required to satisfy this physical deadline, and it is the only stage in the framework for which algorithmic circumvention is impossible. The clock speed of a semiconductor device is not negotiable. However, existing AI implementation studies in power electronics reveal a common structural issue in how this constraint is addressed. Most studies adopt a reactive approach in which the control algorithm is first designed and only afterward compressed to satisfy hardware limitations. This section analyzes the limitations of that paradigm and argues for the necessity of a proactive design philosophy in which hardware constraints are embedded into the algorithm design process from the outset.

The most direct approach to accelerating neural network inference is model compression, which can broadly be categorized into quantization and pruning. Quantization converts floating-point weights and activations (typically 32-bit) into lower-precision representations. As the bit width decreases, multiply–accumulate operations are increasingly replaced by bitwise operations, thereby simultaneously reducing computational complexity and memory usage. In an application of a Binary Neural Network (BNN), the extreme form of quantization, to MMC open-circuit fault diagnosis, the parameter storage requirement was reduced by 96.87% (238.93 kB $\rightarrow$ 7.48 kB) relative to a standard 1-D CNN, while floating-point operations were entirely replaced with bitwise computations [48]. However, this compression incurred a classification accuracy degradation of 4.33%, which may appear numerically small but can correspond to violations of safety thresholds in control applications. More fundamentally, this degradation partially invalidates the modeling accuracy requirements established in Stage 1. Under truncation-based quantization, which is commonly used in hardware implementations, the maximum representation error of a real-valued weight w under a b -bit uniform quantizer $Q_b(\cdot)$ is bounded by the quantization step size Δ_b :

$$|\epsilon_q| = |w - Q_b(w)| \leq \frac{w_{\max} - w_{\min}}{2^b - 1} \triangleq \Delta_b \quad (16)$$

When the quantization error of each of the n_w weights is bounded by Δ_b , the aggregate error vector satisfies $|\epsilon_q|_2 \leq \sqrt{n_w} \cdot \Delta_b$. Since the neural network output $\hat{u}$ is a nonlinear

function of the weight vector $\mathbf{w}$, the influence of the weight quantization error ϵ_q on the output can be approximated via a first-order Taylor expansion as follows:

$$\|\hat{u}_{\text{quant}} - \hat{u}_{\text{float}}\| \leq \sqrt{n_w} \Delta_b \cdot \|\nabla_w f_{\text{NN}}\|_F \quad (17)$$

where n_w denotes the total number of weights in the neural network. This bound is derived from a first-order Taylor expansion and is therefore locally valid under the assumption that the quantization step Δ_b is sufficiently small relative to the curvature of f_{NN} . For networks with ReLU activations, the Jacobian $\nabla_w f_{\text{NN}}$ is piecewise constant and changes discontinuously across activation boundaries; consequently, a single Jacobian evaluation may not accurately represent the global sensitivity of the network output to weight perturbations. In practice, this bound provides a reliable approximation for moderate precision levels (e.g., $b \geq 4$), but becomes increasingly loose as b decreases. In the extreme case of binary neural networks (BNNs, $b = 1$), the quantization step $\Delta_b = w_{\text{max}} - w_{\text{min}}$ spans the full weight range, violating the small-perturbation premise of the Taylor expansion entirely; for such cases, alternative frameworks based on empirical sensitivity analysis or Vapnik-Chervonenkis (VC) dimension bounds are more appropriate [49].

The condition under which Stage 4 does not invalidate Stage 1 is that the right-hand side of Equation (17) remains bounded by the allowable control error ϵ_{ctrl} :

$$\sqrt{n_w} \Delta_b \cdot \|\nabla_w f_{\text{NN}}\|_F \leq \epsilon_{\text{ctrl}}$$

By substituting $\Delta_b = (w_{\text{max}} - w_{\text{min}})/(2^b - 1)$ and rearranging with respect to b , since

$$2^b - 1 \geq \sqrt{n_w} (w_{\text{max}} - w_{\text{min}}) \|\nabla_w f_{\text{NN}}\|_F / \epsilon_{\text{ctrl}},$$

the lower bound on the bit width is:

$$b \geq \log_2 \left(\frac{\sqrt{n_w} (w_{\text{max}} - w_{\text{min}}) \cdot \|\nabla_w f_{\text{NN}}\|_F}{\epsilon_{\text{ctrl}}} + 1 \right) \quad (18)$$

Equations (17) and (18) are intended as analytical design guidelines derived from first-order perturbation analysis, rather than as universally validated empirical relationships. Equation (18) provides a sufficient condition for selecting the bit width such that the estimated quantization-induced output deviation remains within the allowable control error. Its validity is therefore limited to the local small-perturbation regime assumed in Equation (17). For moderate precision levels, such as $b \geq 4$, the relation can be used as an initial design criterion for determining the bit width backward from ϵ_{ctrl} . However, the resulting bound may be conservative for well-trained networks whose weight distributions are concentrated near zero, or overly optimistic for networks exhibiting large activation variance near quantization boundaries. For aggressive compression regimes, including $b \leq 2$ and binary neural networks, the Taylor-approximation premise underlying Equation (17) is no longer reliable, and Equation (18) should not be applied. In such cases, hardware-in-the-loop or physical-hardware testing is required to verify whether the quantized controller satisfies the target accuracy, stability, and real-time implementation requirements.

In PTQ, Δ_b is fixed after training, and thus the condition in Equation (18) cannot be guaranteed. In contrast, QAT simulates Δ_b during the training process and optimizes the satisfaction of this condition [49].

This trade-off is also highly sensitive to the timing of quantization application. Post-Training Quantization (PTQ), which applies quantization after training completion, exhibits substantial precision degradation, particularly at low bit widths where prediction performance deteriorates rapidly. In contrast, Quantization-Aware Training (QAT), which simulates quantization during training so that the network adapts to low-precision represen-

tations, maintains significantly higher accuracy at the same bit width [49]. The implication of this difference for Stage 4 design is clear. To apply QAT, the hardware precision constraints must already be known during the algorithm design stage, implying that the target deployment hardware should be selected concurrently with model training.

A solution to a structurally identical problem to that of Stage 4 has already been demonstrated in the field of particle physics, which appears entirely unrelated to power electronics. The Level-1 trigger system of the CERN LHC must process proton collisions at 40 MHz while selecting events at 100 kHz, where the allowable time for each decision is constrained to a fixed latency of 1 μ s. To satisfy this requirement, the hls4ml library automatically converts neural network models into FPGA firmware. By employing QAT and pruning, it achieved a 97% reduction in critical FPGA resources without accuracy degradation and reduced the inference latency of CNNs to approximately 5 μ s [49]. The key aspect of this approach is that model compression is performed not as a post-processing step, but as an integral part of the training process itself. Hardware constraints, including bit width, LUT budget, and latency targets, are incorporated into the loss function design and network architecture selection from the outset.

Another example illustrating how the co-design paradigm internalizes hardware constraints from the training stage can be found outside the power electronics domain. MicroNAS, a neural architecture search framework for time-series classification on microcontroller environments, directly incorporates MCU execution latency and peak memory consumption into the optimization objective function. By combining LUT-based latency estimation with quantization-aware retraining, it automatically generates neural network architectures that satisfy user-defined hardware constraints [50]. Experimental results demonstrated that the architectures discovered by MicroNAS achieved accuracy comparable to desktop-class models while operating on MCUs, and that the LUT-based latency estimation improved precision by more than one order of magnitude compared with FLOPs-based estimation, with an average error of approximately ± 1.59 ms. When transferred to AI control of WBG power converters, this methodology could serve as the foundation for a neural architecture co-design pipeline in which FPGA or MCU clock constraints are explicitly incorporated from the pre-training stage.

Notable progress has also been achieved toward providing formal error guarantees for quantization bit-width selection. An approach was proposed to automate mixed-precision fixed-point quantization for deep feedforward neural network controllers using mixed-integer linear programming, enabling precision allocation that minimizes bit width while mathematically guaranteeing a user-specified upper bound on rounding error [51]. In FPGA High-Level Synthesis (HLS) compiler-generated implementations, the required machine cycle count was reduced by approximately 67% on average compared with conventional uniform-precision methods, while optimization for large-scale neural networks containing thousands of parameters was completed within several tens of seconds. Whereas the bit-width lower bound derived from Equation (18) provides a theoretical sufficient condition, this approach offers a practical pathway for achieving that condition in an automated manner, thereby concretely realizing, at the implementation level, the central argument of this section that the bit width should be determined backward from the Stage 1 modeling accuracy requirement ϵ_{ctrl} . Nevertheless, since this methodology was not developed specifically for power electronics applications, its validity in environments involving the nonlinear switching dynamics of WBG converters remains to be verified.

It is noteworthy that this paradigm has not yet been adopted for AI control of WBG power converters. Existing studies on AI hardware implementation in power electronics have primarily focused on measuring inference latency on FPGAs or compressing network size for MCUs. However, no study has yet demonstrated the application of a

hardware-aware co-design pipeline, in which target hardware constraints are internalized from the training stage, to power converter control. This gap does not arise from technical infeasibility, but rather from the absence of methodology transfer between the two research domains.

Unlike approaches that complete the entire neural network inference process on a single processor, distributing computations across multiple embedded controllers within the system introduces a new strategy for Stage 4 [48]. In the case of Modular Multilevel Converters (MMCs), a structure was proposed in which the layer-wise computations of a Binary Neural Network (BNN) are assigned to different submodule controllers, thereby reusing existing computational resources. This strategy possesses theoretical appeal because it aligns the physical architecture inherent to distributed power systems with the computational architecture. However, communication latency between controllers accumulates into the overall inference latency, creating the risk of exceeding the timing budget of high-speed WBG switching environments. Here again, the timescale mismatch problem reemerges, implying that precise quantification of communication overhead must precede the practical adoption of distributed inference as a Stage 4 strategy.

In IoT-based power converter control, online DDPG-based adaptation of active disturbance rejection control (ADRC) coefficients has been demonstrated in a real-time embedded environment, while the fundamental incompatibility between IoT communication latency (ranging from hundreds of milliseconds to seconds) and microsecond-level control cycles remains an unresolved challenge [52]. The rationale of this approach is straightforward. Since training does not impose real-time requirements, computationally intensive processes are executed in resource-rich environments, whereas only inference is deployed at the edge to satisfy latency constraints. However, for this separation to remain valid, additional learning after policy convergence must be unnecessary during the inference stage. In practical systems with continuously varying operating conditions, particularly in environments requiring post-deployment adaptation as discussed in Stage 3 (Sim-to-Real), purely separating training and inference cannot guarantee timely policy updates. IoT environments, where communication latency ranges from several hundred milliseconds to multiple seconds, are fundamentally incompatible with control cycles on the order of microseconds.

Table 2 consolidates the inference latencies of representative AI controller implementations reviewed in this section and evaluates their deployment feasibility against the GaN and SiC timing budgets derived from Equation (2). Three observations emerge from this comparison. First, no reviewed implementation satisfies the 250 ns GaN timing budget: the closest result is the DRL-based controller deployed on a standard DSP at approximately 447 ns [11], which still exceeds the ceiling by a factor of 1.8. Second, fixed-point quantization yields a meaningful but insufficient improvement—the INT16 ANN-MPC reduces latency from 2100 ns to 1000 ns relative to its FP32 counterpart [7], a $2.1\times$ reduction—yet this gain does not cross the GaN boundary and leaves the implementation marginally feasible only for SiC environments at or below 1 MHz. Third, feasibility in SiC systems is strongly frequency-dependent: at 1 MHz the timing budget relaxes to approximately 950 ns, making only the DRL implementation conditionally feasible, whereas at 100 kHz the budget expands to nearly 10 μ s, enabling most implementations except the MCU-based PWA-NN [15]. Collectively, Table 2 makes concrete what Equation (2) and Figure 2 establish theoretically: the deployment gap under the integrated GaN reference case is not marginal but structural, and cannot be resolved through incremental post-training compression of existing architectures. This observation motivates the hardware-aware co-design paradigm in which timing constraints are internalized from the training stage rather than imposed as a post-hoc compression target. These findings indicate that Stage

4 cannot be evaluated independently of modeling accuracy and safety requirements. The resulting cross-stage tradeoffs, including the accuracy–latency relationship, are examined systematically in Section 3.5.

Table 2. Inference latency benchmarks of representative AI and predictive controller implementations evaluated against the GaN (250 ns) and SiC (950 ns) timing budgets from Equation (2). ✓: feasible; △: marginally feasible; ×: infeasible. No implementation satisfies the GaN constraint; the closest result, DRL on DSP [11] at 447 ns, still exceeds the ceiling by $1.8\times$. INT16 quantization yields a $2.1\times$ reduction over FP32 [7] but remains GaN-infeasible, and neither the MCU-based [15] nor the FSM-HDL FPGA [53] implementation closes the gap, confirming that the barrier is structural rather than incremental.

Study	Platform	Precision	Latency	GaN Feasible (250 ns)	SiC Feasible (1 MHz)	Ref.
ANN-MPC, Resonant Conv.	FPGA	FP32	2100 ns	×	×	[7]
ANN-MPC, Resonant Conv.	FPGA	INT16	1000 ns	×	△	[7]
DRL, Buck Conv.	DSP	FP32	447 ns	×	✓	[11]
MPC, Buck Conv.	DSP	FP32	2239 ns	×	×	[11]
PWA-NN, DC-DC Conv.	MCU (80 MHz)	Fixed-point	27,500 ns	×	×	[15]
FSM HDL, Predictive	FPGA (50 MHz)	—	3200 ns	×	×	[53]

3.5. Cross-Stage Analysis: Interdependencies and Deployment Readiness Assessment

Sections 3.1–3.4 analyzed the technical barriers and existing solutions associated with each of the four stages. Because these dimensions are sequentially dependent, a limitation in one area can constrain subsequent validation. Separate analysis therefore cannot fully capture the resulting cross-stage effects. This section examines their interactions and derives structural insights that are not apparent from individual studies.

Viewed as a whole, the framework also reveals cyclic interdependencies that cut across this sequential order. Three relationships are particularly prominent.

The first is the modeling–hardware cycle, Stage 1 ↔ Stage 4. As the degree of physics-informed integration increases, the computational burden associated with solving circuit differential equations during inference also increases, making it more difficult to satisfy hardware timing requirements [33,49]. Conversely, compressing the model to meet inference deadlines reduces representational capacity, thereby degrading modeling accuracy [48]. This degradation is particularly acute for physics-in-architecture models because quantization directly corrupts embedded physical parameters rather than merely reducing general representational capacity. The guarantee of structural physical consistency is therefore the first property lost under compression, and no mitigation strategy specific to WBG power converter applications is currently available. This cycle implies that modeling fidelity and hardware implementability cannot be optimized independently, and it remains unverified whether a Pareto-optimal point even exists within the physical timing budget of WBG systems.

The second is the non-propagation from safety assurance to real-hardware deployment, Stage 2 → Stage 3. Safety guarantees validated in simulation environments are no longer valid in real systems where the model mismatch term $\delta(x, u, k)$ exists. This unidirec-

tional non-propagation introduces the risk that simulation-domain safety evidence may be misused as sufficient evidence for deployment, constituting one of the principal reasons why the framework requires independent validation of both stages.

The third is the safety–latency tension, Stage 2 $\leftrightarrow$ Stage 4. CBF-based safety guarantees require solving a quadratic programming problem at every control step, and this additional computational overhead directly conflicts with the sub- μ s timing budget of WBG systems. The strongest existing safety mechanisms therefore structurally impose the greatest threat to hardware implementability. This tension has not been explicitly resolved in the studies examined in this review, indicating that methodologies capable of simultaneously satisfying strong safety guarantees and stringent real-time hardware constraints remain underdeveloped for WBG power converters.

Based on the preceding analysis, Table 3 classifies representative primary AI-control studies across the four analytical dimensions. This mapping operationalizes the framework by showing which requirements were explicitly addressed, partially considered, or left unexamined. It is intended to reveal the distribution of reported evidence rather than rank individual methods.

To reduce subjectivity in the classification, stage assignment was based only on evidence explicitly reported in the cited studies. Stage 1 was assigned when a study developed or used an AI-based model, surrogate model, neural predictor, digital twin, physics-informed model, or parameter-estimation mechanism to improve modeling accuracy or generalization. Stage 2 was assigned when safety, stability, constraint satisfaction, safe exploration, Lyapunov-based analysis, control barrier functions, or protection-oriented validation was explicitly considered. Stage 3 was assigned when simulation-to-hardware transfer, domain adaptation, post-transfer real-hardware retraining, or explicit validation under simulation-to-real model discrepancy was investigated. Direct training on physical hardware without simulation-to-hardware transfer was not classified as Stage 3. Stage 4 was assigned when real-time execution, FPGA/DSP/MCU implementation, HIL validation, measured inference latency, or hardware-resource feasibility was reported. When a study addressed a stage only indirectly, theoretically, or without the full deployment constraints considered in this review, the corresponding entry was marked as partial rather than complete. Table 3 reports the resulting classification, from which three collective observations emerge.

First, no study within the literature considered in this review completely and jointly validates all four stages. Furthermore, as indicated by the $\dagger$ notation in Table 3, the Stage 4 assessments reflect each study's own target switching frequency; under the 250 ns GaN timing budget of Equation (2), none of the reviewed studies with reported timing data demonstrates Stage 4 compliance, consistent with the deployment-feasibility analysis in Figure 2. Most classified studies concentrate on one or two analytical dimensions, while the remaining dimensions are unaddressed or treated only partially. Even when three dimensions are considered, they are not jointly validated as an integrated deployment pipeline. This pattern demonstrates the framework's diagnostic value for revealing structural gaps in existing research.

Second, the study reported in [42] provides a notable empirical example of the joint realization of Stage 2 and Stage 4. Its RLS-based action safeguard was validated during real-world online RL training, while the FPGA inference time of 8.94 μ s satisfied the 50 μ s sampling period of the target motor-drive platform. However, the safeguard was limited to current and voltage constraints and remained dependent on one-step model-prediction accuracy. Moreover, its inference time exceeds the 250 ns GaN budget considered in this review by approximately 36 times. Therefore, this study demonstrates empirical safety-hardware co-realization at a motor-drive timescale, but not formal Lyapunov/CBF-level safety or Stage 4 feasibility for MHz-range WBG converters.

Third, Stage 3 is the most isolated stage. Studies explicitly addressing Stage 3 rarely consider the other stages, while studies focused on the remaining stages typically leave Stage

3 unexamined. This reflects the fact that Sim-to-Real transfer research and algorithm design research have historically evolved within largely independent communities, suggesting that their integration constitutes a key direction for future work.

Table 3. Framework-Based Classification of Representative AI Control Studies According to the Four Deployment-Readiness Dimensions. ✓: explicitly addressed and experimentally validated; △: partially addressed, theoretically discussed, or validated without full deployment constraints; —: not addressed. [†] Stage 4 is evaluated relative to each study’s own target platform and switching frequency, not the 250 ns GaN integrated control–protection budget in Equation (2) and Figure 2. The Validation/Platform/Timing column reports the highest validation domain, target implementation platform, and measured inference or execution time explicitly reported in each study. N/R denotes that an isolated inference or execution latency was not reported; sampling periods and monitoring-cycle times are identified separately and should not be interpreted as neural-network inference latency. Laboratory-scale hardware and HIL validation are reported separately and should not be interpreted as evidence of industrial field deployment unless such deployment was explicitly documented. The classifications reflect the authors’ evaluation criteria and indicate the coverage of reported deployment evidence rather than an objective ranking of the reviewed methods.

Study	Validation/Platform/Timing	Stage 1 Modeling	Stage 2 Safety	Stage 3 Sim-to-Real	Stage 4 [†] Hardware	Ref.
NN-MPC, Cascaded H-bridge (CHB) Converter, FPGA-HIL	FPGA-HIL; Cyclone V FPGA; 3.03–10.94 microseconds	✓	△	—	✓	[4]
Model-Free DQN, NPC Converter, Experimental Testbed	Physical NPC testbed; dSPACE 1202; 100-microsecond sampling, latency N/R	—	△	△	△	[5]
ANN-MPC, Flying-Capacitor Multilevel Converter, FPGA-HIL	Physical multilevel-converter hardware; FPGA; 1.12 microseconds	✓	△	—	✓	[6]
Deep Learning MPC, Resonant Converter	FPGA-HIL; 1.0 microsecond with INT16 and 2.1 microseconds with FP32	✓	△	—	✓	[7]
RL Sim-to-Real, Electric Motor Drive	Simulation-to-hardware transfer; physical motor testbed; dSPACE; isolated latency N/R	△	△	✓	—	[8]
Safety-Enhanced Online RL, Two-Level VSC, Practical Testbed	Physical VSC testbed; IMPERIX platform; 20-microsecond sampling, latency N/R	—	✓	△	△	[12]
PWA-NN Embedded Predictive Control	Physical buck-converter prototype; STM32L476 MCU; execution time below 30 microseconds	✓	△	—	✓	[15]
ANN-Tuned FCS-MPC Weighting Factor, Low-Cost Platform	Physical prototype; low-cost FPGa; ANN latency of 8 clock cycles	△	△	—	✓	[20]
DRL + ESO, Buck Converter	Physical buck-converter testbed; dSPACE 1202; latency N/R	—	—	✓	—	[28]
Stability-Oriented DRL, Power Converter	Control hardware-in-the-loop; TMS320F28379 DSP; latency N/R	—	△	—	△	[29]
Barrier-Certificated RL, Power System	Simulation case studies; no embedded target hardware; latency N/R	—	✓	—	—	[31]
Practical DDPG Control, Buck Converter, Microcontroller	Physical buck-converter hardware; TMS320F28379D; 5-kHz control, latency N/R	—	—	△	✓	[32]
Physics-in-architecture recurrent neural network (PA-RNN) Temporal Modeling, Power Converter, 1-kW Hardware	1-kW physical DAB hardware; dSPACE 1202; latency N/R	✓	—	△	—	[33]
PIML Parameter Estimation, Buck Converter	Physical parameter-estimation experiment; CPU and cloud-edge concept; real-time latency N/R	✓	—	—	—	[35]
Self-Evolving DT, SiC Interleaved Boost Converter	75-kW SiC hardware and HIL; DSP and FPGA; 28.3-second monitoring cycle	✓	—	△	△	[36]
Stability-Constrained RL, Voltage Control	Simulation; Southern California Edison 56-bus pandapower model; latency N/R	—	✓	—	—	[38]
Safeguarded Online RL, PMSM Drive, FPGA/Edge Testbed	Physical online-learning PMSM testbed; FPGA and edge computing; 8.94 microseconds	△	✓	—	✓	[42]

The practical role and scope of the proposed framework can be clarified through comparison with an existing domain-specific review structure and a broader maturity framework. A previous overview of AI-assisted power-converter control [10] organizes the literature according to learning paradigm, online/offline implementation, linear and nonlinear controller structure, design and control roles, and application area. This structure provides a broad view of how AI techniques are used to enhance controller performance, but does not systematically assess whether multiple deployment-related requirements

have been jointly evidenced. At a broader system level, the Machine Learning Technology Readiness Levels (MLTRL) framework [54] evaluates the maturation of ML technologies across research, prototyping, productization, system integration, deployment, and post-deployment monitoring. In contrast, the four stages used in this review are neither controller categories nor ordinal maturity levels; rather, they serve as WBG-specific analytical dimensions for mapping explicitly reported evidence concerning modeling accuracy, safety assurance, sim-to-real transferability, and hardware implementability. This narrower structure provides a domain-specific diagnostic perspective by revealing which technical deployment requirements remain insufficiently evidenced under device-, protection-, and timing-dependent constraints. However, it does not provide the broad methodological coverage of the review in [10] or the complete lifecycle, organizational, and post-deployment assessment of MLTRL. These perspectives should therefore be regarded as complementary.

Several additional limitations follow from this narrower scope. Because the stage classifications are based on the evidence explicitly reported in each study, deployment-related aspects that were implemented but not quantitatively documented may be underestimated. In addition, the reviewed studies differ substantially in converter topology, switching frequency, power level, control objective, and hardware platform. The resulting classifications should therefore be interpreted as indicators of the deployment requirements addressed by each study, rather than as a direct ranking of controller performance. Moreover, the assessment is limited to technical deployment considerations and does not directly account for broader factors such as regulatory approval, technology maturity, manufacturing cost, maintainability, or long-term field reliability. Similarly, the timing-related assessments remain dependent on the target semiconductor device, protection architecture, switching frequency, and implementation platform, and should not be interpreted as universal thresholds applicable to all power converter systems.

4. Industrial Applications and Condition Monitoring

Although fault diagnosis and remaining useful life prediction do not directly generate switching commands at each control step, they remain closely connected to converter control through supervisory decision-making. Fault-diagnosis outputs can trigger protection actions, fault-tolerant control, controller reconfiguration, or converter shutdown, whereas RUL estimates can support condition-dependent derating, thermal management, operating-point adjustment, and maintenance scheduling. These functions generally operate at a slower supervisory timescale than the inner voltage- or current-control loop, but they modify the constraints, references, and admissible operating modes under which the real-time controller operates. Accordingly, Sections 4.1–4.3 examine fault diagnosis and RUL prediction as supervisory extensions of converter control and assess whether the proposed deployment-readiness dimensions remain applicable when AI is used to support safe and reliable converter operation over the system lifetime.

4.1. Fault Diagnosis: Reconfigured Bottleneck Structure

Section 3 established the deployment-readiness framework primarily in the context of real-time control applications; however, the same logical structure can be extended to AI-based fault diagnosis. Nevertheless, fault diagnosis exhibits a fundamentally different temporal structure compared to real-time control, and as a result, the nature of bottlenecks at each of the four stages changes systematically. Analyzing these differences provides a substantive validation of the framework's generality beyond a simple enumeration of applications.

The most immediate distinction arises at Stage 4 (hardware feasibility). In real-time control, AI inference must complete within a physical deadline on the order of sub-

microseconds, constrained by the switching period. In contrast, fault diagnosis inference operates over signal patterns spanning hundreds to thousands of switching cycles, without a strict per-cycle completion constraint. Indeed, in modular multilevel converter (MMC) open-circuit fault diagnosis, it was feasible to distribute binary neural network (BNN) layers across multiple submodule controllers precisely because this temporal slack provides a structural degree of freedom [48]. In other words, Stage 4 is not a critical bottleneck in fault diagnosis applications, implying that computational resources can be reallocated to strengthen requirements in other stages. The extent to which this slack is transferred depends on the outcome structure of fault diagnosis itself.

If, in real-time control, Stage 2 addresses the RL-specific question of whether physical constraints are violated during exploration, then its analogue in fault diagnosis shifts to whether the classification error rate can be mathematically bounded. This distinction is crucial, as high classification accuracy alone is insufficient for safety-critical deployment assessment unless accompanied by statistical guarantees on error rates. A direct realization of this requirement is achieved by combining deep neural network (DNN)-based posterior density estimation with Sequential Probability Ratio Test (SPRT) formulations [55]. SPRT imposes explicit upper bounds on Type I and Type II error probabilities, thereby expressing classifier reliability in terms of statistical guarantees rather than empirical test accuracy. The study further demonstrated that this approach is implementable on low-power embedded platforms such as Arduino 33 BLE Sense, while maintaining detection latency within 25 s. This concretely shows that satisfying Stage 2-like requirements without compromising Stage 4 feasibility is achievable in the fault diagnosis domain. However, this statistical guarantee is contingent on Stage 1 (modeling accuracy), preserving the same sequential dependency structure observed in control applications.

A survey of algorithmic approaches in fault diagnosis reveals particularly active research on multilevel inverter systems. For a T-type three-level inverter, a multi-modal deep residual filtering network that jointly utilizes voltage and current signals was proposed, achieving 99.18% accuracy in diagnosing multiple open-circuit switch faults, with an on-line inference time of approximately 13.5 ms for single-fault detection [56]. From a Stage 4 perspective, this result is significant because the relaxed temporal constraints of fault diagnosis allow for a reallocation of design resources toward improving Stage 1 quality, such as multi-modal feature extraction instead of single-signal processing. In more complex scenarios involving inverter-driven induction motor systems, where multiple faults such as IGBT open-circuit failures and current sensor faults may coexist, a Random Vector Functional Link (RVFL) network-based approach capable of classifying 30 fault modes with 98.83% accuracy and achieving inference within 22 ms in hardware-in-the-loop (HIL) experiments further confirms the practical viability of data-driven methods in multilevel fault classification [57].

The Stage 3 analogue, namely cross-system transferability of fault diagnosis models across inverter systems with differing parameters, poses an independent technical challenge. A transferable fault diagnosis framework combining geodesic flow kernel-based manifold feature learning with extreme learning machines has been proposed, where a source-trained diagnostic model is adapted to a target system via maximum mean discrepancy minimization [58]. On a 22-label fault classification task across inverters with identical topology but different parameters, the transferred model achieved 96.76% accuracy, maintaining robustness under varying speed conditions in real-time experiments. Although this approach partially addresses the Stage 3 gap in fault diagnosis, its reliance on richly labeled source data and at least minimal unlabeled target data introduces practical constraints in WBG system deployment.

PINN-based approaches further contribute at the intersection of fault diagnosis and state monitoring. By embedding the state-space equations of a buck converter directly into a PINN, a non-invasive identification of circuit parameters such as L , C , and R using only front-end measurements was demonstrated [59]. Compared with conventional PINN methods, the neural network size was reduced from 49 kB to 2 kB (over 95% reduction), while maintaining estimation errors within 7% for inductance and 5% for capacitance. Moreover, by analyzing the training loss landscape to pre-classify parameter-wise estimation confidence, the method suppresses unnecessary maintenance alarms at the source. This result explicitly demonstrates that Stage 1 accuracy requirements and the Stage 2-like requirement of statistically bounded classification error can be unified within a single PINN-based pipeline.

In architectures where the discrepancy (residual) between the DT and the physical system response is used as the sole basis for fault detection, any degradation in DT accuracy directly leads to increased overlap between the residual distributions of normal and faulty operating conditions, thereby reducing detection reliability [60]. In particular, in the case of voltage and current sensor fault diagnosis, the DT provides surrogate signals that replace physically measured current estimates. Consequently, a reduction in Stage 1 DT accuracy does not merely degrade diagnostic performance but can directly result in the failure of fault-tolerant control itself. This explicitly demonstrates that Stage 1 functions as a foundational requirement even in fault diagnosis applications.

Meanwhile, as component aging progresses, the intrinsic signal characteristics of normal operation evolve, leading to cumulative distribution shifts between the nominal distribution acquired during the training phase and the actual distribution observed during deployment. In the fault diagnosis context, this phenomenon corresponds to the Stage 3 problem, where the reference model trained under initial operating conditions gradually loses validity as degradation advances. A reduced-order DT implemented on FPGA hardware, which executes in parallel with the physical asset and continuously computes real-time residuals, addresses this gap by functioning not as an offline reference model but as a self-aware dynamic baseline that incrementally absorbs distributional drift [61]. This architecture provides a concrete realization of how Stage 1 and Stage 3 can be dynamically coupled in fault diagnosis systems.

In summary, deployment-readiness evaluation in fault diagnosis applications follows a reconfigured bottleneck structure characterized by Stage 4 relaxation, Stage 2 (statistical guarantees) reinforcement, and increased dependence on Stage 1 (DT accuracy). This reconfiguration indicates that the framework functions as a flexible analytical tool whose emphasis shifts depending on the application domain. For fault-diagnosis-specific deployment-readiness assessment, the most urgent challenge is the formal definition and measurement methodology of the Stage 2 analogue in terms of statistical guarantees.

4.2. Remaining Useful Life Prediction: Extended Stage 1 and Domain-Shift Challenges

RUL prediction, similarly to fault diagnosis, is not constrained by the microsecond-level timing budget characteristic of real-time control. However, it imposes fundamentally different modeling requirements on AI systems. While fault diagnosis constitutes a classification problem between the current state and predefined fault patterns, RUL prediction requires temporal extrapolation from the present state to the future failure point. This extrapolative nature is the key reason why RUL prediction imposes fundamentally extended modeling accuracy requirements compared to static estimation tasks.

The structural failure of purely data-driven neural network approaches manifests in a particularly concrete form in RUL prediction. Although such methods may achieve satisfactory training accuracy, they frequently produce physically inconsistent estimates

during inference, such as non-monotonic or negative RUL values, as demonstrated across both recurrent neural network (RNN) and long short-term memory (LSTM) variants. This arises because the network does not structurally enforce monotonic physical constraints in regions outside the training distribution, particularly as the system approaches end-of-life conditions [62]. Figure 6 directly illustrates this failure mode: the unconstrained RNN ($\alpha = 0$) produces physically impossible spike predictions near end-of-life across all four test devices, while progressively strengthening the monotonic decreasing constraint (increasing α) systematically eliminates these violations, bringing predictions into structural compliance with the physical requirement that RUL must decrease monotonically.

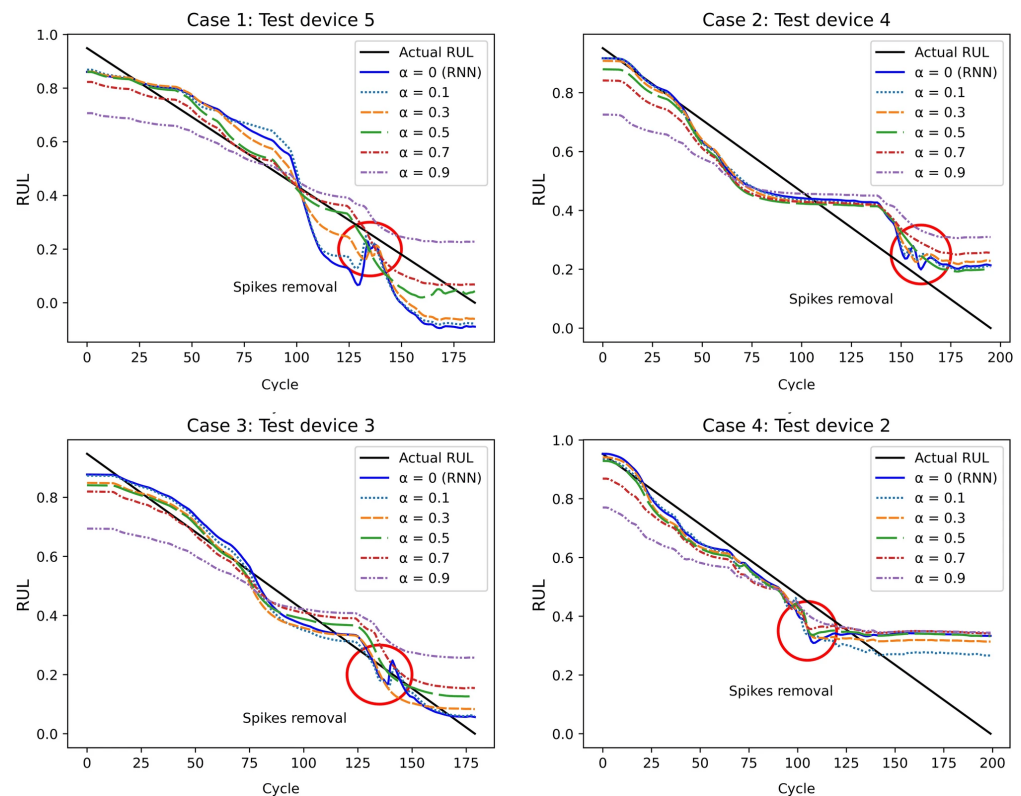


Figure 6. Effect of the monotonic decreasing physical constraint on RUL estimation for IGBT devices under out-of-sample conditions (NASA IGBT accelerated aging dataset). Each panel shows the actual RUL trajectory (black solid line) alongside predictions from physics-informed RNN (PI-RNN) with varying constraint weight α (0 = no constraint; 0.9 = strongest constraint). Red circles highlight physically inconsistent spike predictions produced by the unconstrained network ($\alpha = 0$), in which RUL values temporarily become non-monotonic or negative near end-of-life. As α increases, the monotonic decreasing condition is progressively enforced and spikes are eliminated, demonstrating that physics-informed regularization is necessary to ensure structurally valid RUL predictions in regions outside the training distribution. Reproduced from [62], CC BY 4.0.

PINN-based approaches address this issue by embedding two physical constraints into the loss function as regularization terms: (i) RUL must monotonically decrease, and (ii) RUL must converge to zero at the terminal state. When applied to the NASA IGBT accelerated aging dataset, the physics-informed LSTM (PI-LSTM) achieved an average 13.9% improvement in mean squared error under out-of-sample conditions compared with a baseline LSTM, while simultaneously ensuring structural monotonicity and non-negativity of the predictions.

This result introduces an important contrast connected to the physics-in-loss versus physics-in-architecture discussion in Section 3.3. In real-time control modeling, physics-in-loss formulations are inferior because physical consistency may be violated during inference

and retraining is required under new operating conditions [33]. In RUL prediction, however, this limitation is less critical. The physical constraints governing RUL can be expressed as simple inequalities rather than complex circuit differential equations, and the absence of microsecond-level real-time constraints prevents Stage 4 conflicts arising from increased architectural complexity. Furthermore, in environments where degradation data are costly to acquire, the ability to enforce physical constraints after a single offline training phase provides a significant practical advantage. This contrast clearly indicates that optimal Stage 1 design choices are strongly dependent on the temporal structure of the application domain.

While the studies reviewed in Section 4.2 primarily focus on point estimation, industrial safety standards require that confidence intervals of RUL predictions be directly linked to deterministic failure rate constraints. A probabilistic framework based on Gaussian process (GP) regression for thermally aged IGBTs constructs RUL estimates from entropy-based health indicators derived from collector–emitter voltage signals, while naturally providing 90% confidence intervals [63]. The key contribution of this approach is not merely improved predictive accuracy, but the provision of a quantitative basis for determining when predictions are unreliable in fault-tolerant management systems. However, GP regression suffers from a computational complexity of $O(N^3)$ with respect to the number of training samples N , limiting its scalability to large-scale aging datasets.

To address this scalability limitation, a self-attention and Monte Carlo dropout (SA-MCD) framework was proposed for IGBT RUL prediction. This method employs a multi-head self-attention encoder to select the most sensitive subset of 21 candidate features derived from transient thermal impedance, while MC dropout sampling during inference is used to construct calibrated confidence intervals [64]. The framework enables high-precision RUL estimation from the initial stage using only limited online data and incorporates an adaptive mechanism that updates the model with minimal data, thereby exhibiting potential to address Stage 3-like issues in RUL prediction, namely distribution shifts between training and deployment conditions. A comparison between GP regression and SA-MCD reveals a fundamental trade-off: GP regression provides a strong theoretical foundation for uncertainty quantification but lacks scalability, whereas SA-MCD offers scalability and adaptability at the cost of confidence interval quality, which depends on the number of MC samples.

In RUL prediction, domain shift manifests differently from control applications. While the sim-to-real gap in control arises from physical discrepancies between simulation and real hardware, the analogous gap in RUL prediction originates from systematic inconsistencies between accelerated aging tests (Power Cycling Tests, PCT) and actual operating conditions. It has been repeatedly observed that dominant failure modes vary across different PCT conditions and module types, and that models trained under one PCT condition may produce inaccurate predictions when applied to others [65]. This introduces a more fundamental issue: failure modes themselves may differ between the source domain (PCT conditions) and the target domain (real operating conditions).

This mismatch can be partially mitigated through transfer learning frameworks that leverage insufficient target-domain degradation data. A two-stage RUL prediction framework based on a Dual-Stage Convolutional Neural Network with Dual Transfer Attention Mechanism (DSCN-DTAM) adaptively determines the first prediction time via a dual-alarm mechanism and dynamically selects transferable degradation features to minimize domain discrepancy between source and target distributions [66]. Experimental results across multiple transfer tasks under varying operating conditions show that DSCN-DTAM achieves the lowest prediction error among six comparative methods, including domain adaptation and adversarial learning approaches. However, the presence of negative transfer in certain scenarios confirms that cross-condition RUL transfer remains an incompletely resolved problem. The absence of a unified pipeline that jointly integrates transfer learning-based

RUL prediction with prior fault mode classification represents a significant gap in current research, consistent with observations in Section 4.2.

To bridge this gap, a pre-classification stage identifying the expected failure mode under current operating conditions is required before applying RUL models. An approach based on convolutional neural networks that classifies four fault categories (e.g., solder fatigue and bond wire degradation) from precursor parameters such as $V_{ce,on}$ and $V_{ge,th}$ should therefore be reinterpreted not merely as a fault diagnosis tool, but as a prerequisite validation stage for reducing the Stage 3 gap in RUL prediction. In this formulation, reliable application of a fault-specific RUL model becomes conditional on prior identification of the failure mode. The absence of a unified pipeline integrating these two stages remains one of the explicit research gaps in RUL prediction literature.

4.3. Application-Specific Deployment Requirements

The analyses in the previous two sections demonstrated that bottleneck structures within the deployment-readiness framework are reconfigured across application categories. This section extends the analysis to a broader level, showing that even within the same control domain, the dominant bottleneck varies depending on the specific application context. This domain-dependent differentiation is not merely a taxonomy of applications, but functions as a design principle indicating where research resources should be prioritized.

4.3.1. EV Chargers and GaN-Based High-Frequency Systems

The commercialization of GaN HEMTs shifts switching frequencies into the multi-megahertz regime, compressing the Stage 4 timing budget to sub-100 ns scales. At the same time, the SCWT of GaN HEMTs (approximately 300 ns to 1 μ s) shifts the definition of Stage 2 safety constraints to a timescale that conventional overcurrent protection circuits cannot directly address. In this regime, the role of AI is redefined not as direct per-cycle control signal generation, but as an intelligent gate driver that adaptively shapes switching transients according to operating conditions. Experimental results demonstrate that this formulation reduces switching losses by 51% at a 200 A load compared with fixed-resistance gate driving, while simultaneously relaxing hardware timing constraints [17]. The key mechanism enabling the mitigation of Stage 4 bottlenecks is this structural redefinition of AI as a transient-shaping operator rather than an instantaneous decision maker within a single switching cycle.

In parallel, a reduced-order DT implemented on FPGA that operates alongside the physical asset and continuously computes the thermal safe operating area (SOA) provides a hierarchical architecture in which safety constraints are delegated to a slower DT layer. This effectively decouples safety enforcement from real-time switching decisions and alleviates the conflict between Stage 2 and Stage 4 constraints [61].

4.3.2. Motor Drive Systems (SiC-Based)

In SiC-based motor drive systems, the dominant bottleneck in AI control shifts to sim-to-real transfer. These systems constitute strongly coupled nonlinear dynamics involving both electrical and mechanical subsystems, where key mechanical parameters such as inertia and friction coefficients, as well as EMI effects, are difficult to accurately reproduce in simulation. A simulation-pretraining-to-real-testbed transfer pipeline revealed significant performance degradation immediately after deployment and the necessity of additional online adaptation, with RL-based controllers achieving performance only comparable to MPC methods that have been developed over two decades [8]. This indicates that research priorities in this domain must move beyond merely minimizing Sim-to-Real error toward meta-learning or few-shot adaptation strategies capable of updating policies with minimal real-world data.

4.3.3. Grid-Connected Inverters and Microgrids

In grid-connected applications, the joint satisfaction of Stage 1 (modeling accuracy) and Stage 2 (safety guarantees) presents a distinct challenge. Embedding stability metrics derived from small-signal impedance models directly into the reward function of a DRL agent demonstrates that Stage 1 accuracy requirements extend beyond internal converter parameter estimation to real-time modeling of converter-grid interactions [29]. Furthermore, a hybrid approach combining an MPC-based safety policy with a DQN agent constrained the converter current during simulation-based training and, after transfer to a practical two-level VSC, achieved a total harmonic distortion of 2.19%, compared with 1.97% for conventional FCS-MPC [12]. This result provides experimental evidence that a model-based safety filter can support simulation-to-hardware deployment of a learned controller; however, online self-learning directly on the physical converter was not demonstrated.

As summarized in Table 4, each domain exhibits distinct characteristics in terms of the dominant bottleneck, the role of AI, and the primary constraints. The emphasis on joint safety-hardware co-design in GaN-based EV systems versus the focus on minimizing transfer cost in motor drives reflects distinct research priorities dictated by application-specific constraints.

Table 4. Application-domain-differentiated deployment readiness: binding bottleneck stages and representative studies for AI control in WBG power converters. * Stage 2 analog for fault diagnosis denotes a statistical guarantee on classification error rates (Type I/II), not RL constraint satisfaction.

Application Domain	Critical Stage	AI Role	Key Constraint	Ref.
EV Charger (GaN, >1 MHz)	Stage 2 + Stage 4	Intelligent gate driver; DT-based SOA monitoring	SCWT: 300 ns–1 μ s; sub- μ s inference budget	[17,61]
Motor Drive (SiC-based)	Stage 3	Direct DRL policy	Electromechanical coupling; mechanical parameter mismatch	[8]
Grid-Tied Inverter/Microgrid	Stage 1 + Stage 2	Impedance-aware DRL; safe online RL	Dynamic grid impedance variation; overcurrent during exploration	[12,29]
Fault Diagnosis (General)	Stage 1 + Stage 2 *	DT-based residual monitor	Distribution shift over degradation lifecycle	[55,60]
RUL Prediction (IGBT)	Stage 1	PINN-based prognostic model	Failure-mode-dependent lifetime model selection	[62,65]

5. Open Challenges and Future Directions

The analyses in Sections 3 and 4 identify both the evidence established at each stage and the gaps that remain. These recurring issues are consolidated here into four core challenges and one forward-looking research outlook. Rather than independent barriers, they represent interconnected bottlenecks arising from cross-stage dependencies.

5.1. Challenge 1: Real-Time AI Inference Limits in MHz GaN Environments

For GaN HEMT-based converters pursuing integrated digital control and protection, the allowable inference budget can be bounded at approximately 250 ns under the worst-case assumption established in Equation (2). This value is used here as the most stringent forward-looking hardware reference case, rather than as a universal industrial requirement.

This constraint is not merely a restriction on neural network size, but reflects a more fundamental issue in which the dynamics that must be modeled by AI are inherently coupled with this timescale. The switching voltage and current behavior of GaN HEMTs exhibits rise times on the order of 10 ns and includes parasitic effects that are not accurately captured by standard circuit simulators. In a study constructing GaN HEMT behavioral models using long short-term memory and gated recurrent unit recurrent neural networks and validating

them across three commercial devices, current prediction error was significantly lower than manufacturer-provided models; however, total training time reached 75–120 min (approximately 3–4 s per epoch), making online adaptation structurally infeasible [67]. This result indicates that, even when Stage 1 requirements are satisfied at the device-model level, no established pathway exists for deploying such models in real-time inference under Stage 4 timing budgets.

To address this challenge, deeply pipelined and parallel FPGA implementations represent a near-term platform for reducing AI inference latency, particularly when fixed-point arithmetic and application-specific data paths are employed. An FSM-based HDL design methodology has demonstrated resource-efficient FPGA implementation of finite-state predictive current control, completing the full algorithm within 160 clock cycles at 50 MHz through the inherent parallelism of FPGA architectures [53]. However, this approach remains primarily applicable to explicit predictive-control formulations and does not directly resolve the implementation of neural-network controllers requiring online weight updates.

Beyond incremental FPGA optimization, future progress is likely to require hardware-aware co-design in which network architecture, numerical precision, latency, and resource constraints are optimized jointly from the training stage. Promising directions include FPGA-SoC integration, application-specific neural accelerators, sparsity-aware architectures, and power-electronics-specific neural architecture search. However, these platforms have not yet been sufficiently validated under MHz-range WBG converter conditions. Their development will therefore require standardized GaN-oriented benchmark models, worst-case latency measurements, and closed-loop hardware validation under representative protection and switching constraints.

5.2. Challenge 2: Absence of Public WBG-Specific Datasets

As identified in the Stage 1 analysis in Section 3.1, most AI-based power converter modeling studies construct proprietary datasets tailored to specific topologies and operating conditions. Early work proposing Bayesian-regularized artificial neural network and random forest-based modeling for buck and boost converters explicitly noted that, despite the potential extension of these methods to DTs, circuit automation, and software-defined power grids, each converter still requires an independent data acquisition pipeline [68]. This fragmentation problem becomes more severe in the WBG domain. Switching ringing and EMI characteristics in GaN-based converters vary significantly depending on device characteristics, PCB layout, and operating conditions, making direct transfer of datasets across experimental setups infeasible.

In the context of DTs, this limitation can be expressed through more concrete technical requirements. For a power converter DT to be practically useful, it must simultaneously satisfy real-time executability, accurate capture of large-signal dynamics across all operating modes, and portability enabling direct deployment on embedded platforms using standard programming languages [69]. Although the study addressed these requirements by replacing numerical integration of state equations with dynamic ANNs, it still remains constrained by the dependence of training data coverage on individual experimental setups. This gap can only be systematically resolved through the establishment of a WBG-specific public dataset, analogous to how NASA IGBT accelerated aging datasets function as standard benchmarks in the RUL prediction domain.

5.3. Challenge 3: Limited Formal Verification for Industrial Assurance

The analysis in Section 3.2 identified Stage 2 (safety guarantees) as one of the least mature components in AI-based WBG power converter control. Existing Lyapunov-based and CBF-based safe RL methods have been validated primarily at grid-level timescales (ms–s), and no systematic analysis has yet addressed the implications of transferring these methods to the switching timescales of WBG converters.

A relevant theoretical development addressing this gap has been proposed in the machine learning literature. The Krasovskii constrained RL (KCRL) framework embeds stability constraints derived from Krasovskii functions into the learning process and theoretically guarantees finite-time stabilization for unknown nonlinear dynamical systems [70]. The key mechanism is the incorporation of Jacobian-based linear matrix inequality stability constraints directly into the policy optimization problem, ensuring that stability is structurally enforced throughout learning. This is consistent with the Stage 4 compatibility condition discussed in Section 3.2, namely that safety enforcement should be concentrated during training and not impose additional computational burden during inference. However, it remains unverified whether the sample complexity bounds of KCRL are achievable under high-frequency switching conditions in WBG converters, and whether the framework can be directly extended to hybrid systems with inherent switching nonlinearities.

From a broader perspective, industrial qualification and assurance processes generally require reproducible evidence that failure risks remain within application-specific limits, whereas many current AI safety methods provide probabilistic guarantees whose adequacy depends on the target system and assurance framework. Within the power-electronics literature reviewed here, systematic methodologies that jointly address AI-based controller design and industrial assurance requirements remain limited.

5.4. Challenge 4: Integration of Design-Time and Runtime AI Optimization

The role of AI considered in this work has primarily been limited to real-time control and condition monitoring of power converters. However, system-level performance and deployment readiness are fundamentally determined at the design stage through parameters such as device selection, inductance, capacitance, and switching frequency. A study combining DNN-based surrogate models with Soft Actor–Critic-based DRL demonstrated optimal component sizing across 11 converter topologies, achieving a 36-fold reduction in optimization time (from 18 h to 30 min) compared with conventional simulation-based approaches [71].

Data-driven design automation has also been demonstrated using experimentally trained statistical models to predict converter efficiency and temperature across multiple power-converter-array configurations [72]. Together, these studies indicate that AI-assisted design-space exploration is already being investigated through both simulation-based optimization and experimentally grounded design methodologies.

From this analytical perspective, the implications of these results extend beyond design efficiency. Design-time choices such as switching frequency and passive-component sizing directly determine the Stage 4 inference budget and the required modeling fidelity at Stage 1. Although design-stage AI precedes the four stages considered here, the reviewed literature rarely integrates it with control-stage AI spanning Stages 1–4. Such integration could extend the present assessment into a complete methodology from design to field operation.

5.5. Outlook: Potential Role of Foundation Models

The four challenges outlined above share a common structural issue: data fragmentation, domain specificity, and limited generalization capability. Foundation models, which encode diverse domain patterns through large-scale pre-training, offer a potential long-term solution. Pre-training on heterogeneous datasets covering multiple topologies and operating conditions, followed by fine-tuning with limited domain-specific data, could substantially reduce the adaptation cost associated with Sim-to-Real gaps analyzed in Section 3.3. However, realizing this direction requires prior establishment of the WBG-specific public datasets discussed in Challenge 2, as well as systematic investigation into whether inference complexity can be reduced to remain compatible with the microsecond-level timing constraints identified in

Challenge 1. This direction should therefore be regarded as a speculative long-term research opportunity rather than an evidence-supported deployment pathway.

6. Conclusions

The central finding is that AI-based control for WBG power converters has reached considerable algorithmic maturity, but this progress remains structurally decoupled from hardware deployment feasibility. As summarized in Table 3, no study in the reviewed literature jointly validates all four dimensions. This gap becomes visible only when the literature is assessed across the four WBG-specific dimensions adopted here rather than through any single methodological lens.

A second non-trivial insight emerging from the analysis is the existence of cyclic interdependencies among the four stages. The sub-microsecond computational budget of Stage 4 constrains the complexity of admissible safety mechanisms in Stage 2, while Stage 2 safety guarantees, in turn, impose strict prerequisites on Stage 1 modeling accuracy. This dependency structure implies that sequentially addressing each stage is not an effective strategy, as improvements in one stage necessarily reshape the requirements of others. In particular, Stage 2 (safety guarantees) is identified as the least mature stage in the WBG converter domain. Existing Lyapunov-based and CBF-based methods have been validated primarily in power system timescales (ms–s), leaving their compatibility with WBG switching timescales an open problem.

Application-level analysis further yields an additional practical insight. Deployment challenges are not uniformly distributed across domains. In GaN-based EV chargers, the dominant bottleneck lies in the conflict between Stage 2 and Stage 4; in SiC-based motor drives, it arises from Stage 3 failure-mode mismatch; and in grid-connected applications, Stage 1 dynamic system modeling becomes the limiting factor. This observation implies that treating “WBG AI control” as a monolithic research problem is not appropriate, and that application-specific domain selection must precede any rational allocation of research resources.

For practical deployment studies, researchers should report the target converter topology and switching frequency together with the validation domain (simulation, HIL, or physical hardware). They should also specify the enforced safety constraints and observed violation rates, the sim-to-real or hardware-adaptation procedure, and the measured worst-case inference latency and resource utilization on the target platform. Consistent reporting of these elements would allow readiness to be evaluated without conflating algorithmic performance with hardware feasibility.

Ultimately, this review calls for a methodological shift. Current AI research in power electronics largely optimizes performance within isolated stages, thereby overlooking cross-stage constraints. Making AI a practical industrial standard for WBG converters will require co-optimization across the full pipeline, from design-parameter selection to field deployment. The proposed analytical structure provides a common language for this shift.

Author Contributions: Conceptualization, S.W. and G.S.; data curation, S.P.; writing—original draft preparation, S.W. and H.J.; writing—review and editing, S.W., G.S. and H.J.; visualization, S.P.; supervision, W.-G.B. and W.C.; project administration, S.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Acknowledgments: During the preparation of this manuscript, the authors used Claude-Sonnet 4.6 for the purposes of translation. The authors have reviewed and edited the output and take full responsibility for the content of this publication.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ADC	Analog-to-digital converter
ADRC	Active disturbance rejection control
AI	Artificial intelligence
ANN	Artificial neural network
ANN-MPC	Artificial neural network-based model predictive control
ARPA-E	Advanced Research Projects Agency–Energy
BNN	Binary neural network
CBF	Control barrier function
CHB	Cascaded H-bridge
CNN	Convolutional neural network
CPL	Constant power load
CPU	Central processing unit
DAB	Dual-active-bridge
DC	Direct current
DDPG	Deep deterministic policy gradient
DNN	Deep neural network
DQN	Deep Q-network
DRL	Deep reinforcement learning
DRM	Duty ratio mapping
DSCN-DTAM	Dual-stage convolutional neural network with dual transfer attention mechanism
DSP	Digital signal processor
DT	Digital twin
EMI	Electromagnetic interference
ESO	Extended state observer
EV	Electric vehicle
FC	Flying capacitor
FCS-MPC	Finite control set model predictive control
FLOPs	Floating-point operations
FP32	32-bit floating-point
FPGA	Field-programmable gate array
FSM	Finite-state machine
GaN	Gallium nitride
GP	Gaussian process
HDL	Hardware description language
HEMT	High-electron-mobility transistor
HIL	Hardware-in-the-loop
HLS	High-level synthesis
IASA	Integral action state augmentation
IGBT	Insulated-gate bipolar transistor
INT16	16-bit integer
KCRL	Krasovskii constrained reinforcement learning
LSTM	Long short-term memory
LUT	Lookup table
MAE	Mean absolute error
MC	Monte Carlo

MCU	Microcontroller unit
MLTRL	Machine Learning Technology Readiness Levels
MMC	Modular multilevel converter
MOSFET	Metal–oxide–semiconductor field-effect transistor
MPC	Model predictive control
NPC	Neutral-point-clamped
OOD	Out-of-distribution
PA-RNN	Physics-in-architecture recurrent neural network
PCB	Printed circuit board
PCT	Power cycling test
PI	Proportional–integral
PIML	Physics-informed machine learning
PINN	Physics-informed neural network
PMSM	Permanent-magnet synchronous motor
PPO	Proximal policy optimization
PTQ	Post-training quantization
PWA	Piecewise affine
PWM	Pulse-width modulation
QAT	Quantization-aware training
QP	Quadratic programming
RL	Reinforcement learning
RLS	Recursive least squares
RNN	Recurrent neural network
RUL	Remaining useful life
RVFL	Random vector functional link
SA-MCD	Self-attention Monte Carlo dropout
SCWT	Short-circuit withstand time
Si	Silicon
SiC	Silicon carbide
SIL	Software-in-the-loop
SOA	Safe operating area
SPRT	Sequential probability ratio test
THD	Total harmonic distortion
VHF	Very high frequency
VSC	Voltage-source converter
WBG	Wide-bandgap
ZVS	Zero-voltage switching

References

1. Kizilyalli, I.C.; Xu, Y.A.; Carlson, E.; Manser, J.; Cunningham, D.W. Current and future directions in power electronic devices and circuits based on wide band-gap semiconductors. In Proceedings of the 2017 IEEE 5th Workshop on Wide Bandgap Power Devices and Applications (WiPDA), Albuquerque, NM, USA, 30 October–1 November 2017; p. 417.
2. Iannaccone, G.; Sbrana, C.; Morelli, I.; Strangio, S. Power electronics based on wide-bandgap semiconductors: Opportunities and challenges. *IEEE Access* **2021**, *9*, 139446–139456. [[CrossRef](#)]
3. Buffolo, M.; Favero, D.; Marcuzzi, A.; De Santi, C.; Meneghesso, G.; Zanoni, E.; Meneghini, M. Review and outlook on GaN and SiC power devices: Industrial state-of-the-art, applications, and perspectives. *IEEE Trans. Electron Devices* **2024**, *71*, 1344–1355. [[CrossRef](#)]
4. Simonetti, F.; D’Innocenzo, A.; Cecati, C. Neural network model-predictive control for CHB converters with FPGA implementation. *IEEE Trans. Ind. Inform.* **2023**, *19*, 9691–9702. [[CrossRef](#)]
5. Qashqai, P.; Babaie, M.; Zgheib, R.; Al-Haddad, K. A model-free switching and control method for three-level neutral point clamped converter using deep reinforcement learning. *IEEE Access* **2023**, *11*, 105394–105409. [[CrossRef](#)]
6. Wang, D.; Shen, Z.J.; Yin, X.; Tang, S.; Liu, X.; Zhang, C.; Wang, J.; Rodriguez, J.; Norambuena, M. Model predictive control using artificial neural network for power converters. *IEEE Trans. Ind. Electron.* **2021**, *69*, 3689–3699. [[CrossRef](#)]

7. Lucia, S.; Navarro, D.; Karg, B.; Sarnago, H.; Lucia, O. Deep learning-based model predictive control for resonant power converters. *IEEE Trans. Ind. Inform.* **2020**, *17*, 409–420. [[CrossRef](#)]
8. Book, G.; Traue, A.; Balakrishna, P.; Brosch, A.; Schenke, M.; Hanke, S.; Kirchgässner, W.; Wallscheid, O. Transferring online reinforcement learning for electric motor control from simulation to real-world experiments. *IEEE Open J. Power Electron.* **2021**, *2*, 187–201. [[CrossRef](#)]
9. Alfred, D.; Czarkowski, D.; Teng, J. Reinforcement learning-based control of a power electronic converter. *Mathematics* **2024**, *12*, 671. [[CrossRef](#)]
10. Gao, Y.; Wang, S.; Dragicevic, T.; Wheeler, P.; Zanchetta, P. Artificial intelligence techniques for enhancing the performance of controllers in power converter-based systems—an overview. *IEEE Open J. Ind. Appl.* **2023**, *4*, 366–375. [[CrossRef](#)]
11. Cui, C.; Yang, T.; Dai, Y.; Zhang, C.; Xu, Q. Implementation of transferring reinforcement learning for DC–DC buck converter control via duty ratio mapping. *IEEE Trans. Ind. Electron.* **2022**, *70*, 6141–6150. [[CrossRef](#)]
12. Wan, Y.; Xu, Q.; Dragičević, T. Safety-enhanced self-learning for optimal power converter control. *IEEE Trans. Ind. Electron.* **2024**, *71*, 15229–15234. [[CrossRef](#)]
13. Berkenkamp, F.; Turchetta, M.; Schoellig, A.; Krause, A. Safe model-based reinforcement learning with stability guarantees. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 908–918.
14. Weber, D.; Schenke, M.; Wallscheid, O. Steady-state error compensation for reinforcement learning-based control of power electronic systems. *IEEE Access* **2023**, *11*, 76524–76536. [[CrossRef](#)]
15. Maddalena, E.T.; Specq, M.W.; Wisniewski, V.L.; Jones, C.N. Embedded PWM predictive control of DC-DC power converters via piecewise-affine neural networks. *IEEE Open J. Ind. Electron. Soc.* **2021**, *2*, 199–206. [[CrossRef](#)]
16. Perreault, D.J.; Hu, J.; Rivas, J.M.; Han, Y.; Leitermann, O.; Pilawa-Podgurski, R.C.; Sagneri, A.; Sullivan, C.R. Opportunities and challenges in very high frequency power conversion. In Proceedings of the 2009 Twenty-Fourth Annual IEEE Applied Power Electronics Conference and Exposition, Washington, DC, USA, 15–19 February 2009; pp. 1–14.
17. Henn, J.; Lüdecke, C.; Laumen, M.; Beushausen, S.; Kalker, S.; van der Broeck, C.H.; Engelmann, G.; de Doncker, R.W. Intelligent gate drivers for future power converters. *IEEE Trans. Power Electron.* **2021**, *37*, 3484–3503. [[CrossRef](#)]
18. Eull, M.; Preindl, M.; Emadi, A. Analysis and design of a high efficiency, high power density three-phase silicon carbide inverter. In Proceedings of the 2016 IEEE Transportation Electrification Conference and Expo (ITEC), Dearborn, MI, USA, 27–29 June 2016; pp. 1–6.
19. Hajihosseini, M.; Andalibi, M.; Gheisarnejad, M.; Farsizadeh, H.; Khooban, M.H. DC/DC power converter control-based deep machine learning techniques: Real-time implementation. *IEEE Trans. Power Electron.* **2020**, *35*, 9971–9977. [[CrossRef](#)]
20. Vazquez, S.; Marino, D.; Zafra, E.; Peña, M.D.V.; Rodríguez-Andina, J.J.; Franquelo, L.G.; Manic, M. An artificial intelligence approach for real-time tuning of weighting factors in FCS-MPC for power converters. *IEEE Trans. Ind. Electron.* **2021**, *69*, 11987–11998.
21. Wan, Y.; Dragicevic, T.; Mijatovic, N.; Li, C.; Rodriguez, J. Reinforcement learning based weighting factor design of model predictive control for power electronic converters. In Proceedings of the 2021 IEEE International Conference on Predictive Control of Electrical Drives and Power Electronics (PRECEDE), Jinan, China, 20–22 November 2021; pp. 738–743.
22. Liu, X.; Qiu, L.; Fang, Y.; Wang, K.; Li, Y.; Rodríguez, J. Predictive control of voltage source inverter: An online reinforcement learning solution. *IEEE Trans. Ind. Electron.* **2023**, *71*, 6591–6600. [[CrossRef](#)]
23. Wu, W.; Qiu, L.; Liu, X.; Guo, F.; Rodriguez, J.; Ma, J.; Fang, Y. Data-driven iterative learning predictive control for power converters. *IEEE Trans. Power Electron.* **2022**, *37*, 14028–14033. [[CrossRef](#)]
24. Liu, X.; Qiu, L.; Rodriguez, J.; Wu, W.; Ma, J.; Peng, Z.; Wang, D.; Fang, Y. Data-driven neural predictors-based robust MPC for power converters. *IEEE Trans. Power Electron.* **2022**, *37*, 11650–11661. [[CrossRef](#)]
25. Liu, X.; Qiu, L.; Fang, Y.; Rodriguez, J. Reinforcement learning-based event-triggered FCS-MPC for power converters. *IEEE Trans. Ind. Electron.* **2023**, *70*, 11841–11852. [[CrossRef](#)]
26. Prag, K.; Woolway, M.; Celik, T. Data-driven model predictive control of DC-to-DC buck-boost converter. *IEEE Access* **2021**, *9*, 101902–101915. [[CrossRef](#)]
27. Shi, X.; Chen, N.; Wei, T.; Wu, J.; Xiao, P. A reinforcement learning-based online-training AI controller for DC-DC switching converters. In Proceedings of the 2021 6th international conference on integrated circuits and microsystems (ICIM), Nanjing, China, 22–24 October 2021; pp. 435–438.
28. Yang, T.; Cui, C.; Zhang, C.; Yang, J. Robustness enhancement of DRL controller for DC–DC buck converters fusing ESO. *J. Control Decis.* **2025**, *12*, 101–110.
29. Jiang, S.; Zeng, Y.; Zhu, Y.; Pou, J.; Konstantinou, G. Stability-oriented multiobjective control design for power converters assisted by deep reinforcement learning. *IEEE Trans. Power Electron.* **2023**, *38*, 12394–12400. [[CrossRef](#)]
30. Emam, Y.; Notomista, G.; Glotfelter, P.; Kira, Z.; Egerstedt, M. Safe reinforcement learning using robust control barrier functions. *IEEE Robot. Autom. Lett.* **2022**, *10*, 2886–2893. [[CrossRef](#)]
31. Zhao, T.; Wang, J.; Yue, M. A barrier-certificated reinforcement learning approach for enhancing power system transient stability. *IEEE Trans. Power Syst.* **2023**, *38*, 5356–5366. [[CrossRef](#)]

32. Kwon, S.; Yoon, C.; Lee, Y.I. Practical implementation method of reinforcement learning for power converter. *IFAC-PapersOnLine* **2022**, *55*, 437–441. [[CrossRef](#)]
33. Li, X.; Lin, F.; Wang, H.; Zhang, X.; Ma, H.; Wen, C.; Blaabjerg, F. Temporal modeling for power converters with physics-in-architecture recurrent neural network. *IEEE Trans. Ind. Electron.* **2024**, *71*, 14111–14123. [[CrossRef](#)]
34. Li, X.; Lin, F.; Zhang, X.; Ma, H.; Blaabjerg, F. Data-light physics-informed modeling for the modulation optimization of a dual-active-bridge converter. *IEEE Trans. Power Electron.* **2024**, *39*, 8770–8785. [[CrossRef](#)]
35. Zhao, S.; Peng, Y.; Zhang, Y.; Wang, H. Parameter estimation of power electronic converters with physics-informed machine learning. *IEEE Trans. Power Electron.* **2022**, *37*, 11567–11578. [[CrossRef](#)]
36. Choksi, K.; Mirza, A.B.; Zhou, A.; Singh, D.; Hijikata, M.; Luo, F. Self-evolving digital twin-based online health monitoring of multiphase boost converters. *IEEE Trans. Power Electron.* **2023**, *38*, 16100–16117. [[CrossRef](#)]
37. Chow, Y.; Nachum, O.; Duenez-Guzman, E.; Ghavamzadeh, M. A lyapunov-based approach to safe reinforcement learning. In Proceedings of the Advances in Neural Information Processing Systems 31, Montreal, QC, Canada, 3–8 December 2018.
38. Shi, Y.; Qu, G.; Low, S.; Anandkumar, A.; Wierman, A. Stability constrained reinforcement learning for real-time voltage control. In Proceedings of the 2022 American Control Conference (ACC), Atlanta, GA, USA, 8–10 June 2022; pp. 2715–2721.
39. Cui, W.; Zhang, B. Lyapunov-regularized reinforcement learning for power system transient stability. *IEEE Control Syst. Lett.* **2021**, *6*, 974–979. [[CrossRef](#)]
40. Cui, W.; Li, J.; Zhang, B. Decentralized safe reinforcement learning for voltage control. *arXiv* **2021**, arXiv:2110.01126.
41. Stiasny, J.; Misyris, G.S.; Chatzivasileiadis, S. Physics-informed neural networks for non-linear system identification for power system dynamics. In Proceedings of the 2021 IEEE Madrid PowerTech, Madrid, Spain, 28 June–2 July 2021; pp. 1–6.
42. Schenke, M.; Haucke-Korber, B.; Wallscheid, O. Finite-set direct torque control via edge-computing-assisted safe reinforcement learning for a permanent-magnet synchronous motor. *IEEE Trans. Power Electron.* **2023**, *38*, 13741–13756. [[CrossRef](#)]
43. Tang, Y.; Hu, W.; Xiao, J.; Chen, Z.; Huang, Q.; Chen, Z.; Blaabjerg, F. Reinforcement learning based efficiency optimization scheme for the DAB DC–DC converter with triple-phase-shift modulation. *IEEE Trans. Ind. Electron.* **2020**, *68*, 7350–7361.
44. Mazaheri, N.; Santamargarita, D.; Bueno, E.; Pizarro, D.; Cobrecas, S. A deep reinforcement learning approach to dc-dc power electronic converter control with practical considerations. *Energies* **2024**, *17*, 3578. [[CrossRef](#)]
45. Xiao, Z.; Jiang, Y.; Sun, T.; Wu, Y.; Tang, Y. Refining power converter loss evaluation: A transfer learning approach. *IEEE Trans. Power Electron.* **2024**, *39*, 4313–4324. [[CrossRef](#)]
46. Peng, Y.; Wang, Y.; Liu, Y.; Gao, K.; Yin, T.; Yu, H. Few-shot learning based multi-weather-condition impedance identification for MPPT-controlled PV converters. *IET Renew. Power Gener.* **2022**, *16*, 1345–1353. [[CrossRef](#)]
47. De Silva, W.M.P.; Krishnan, T.N.H.; Ho, P.W.; Sarimuthu, C.R. A Novel Meta-Learning-Based Reinforcement Controller for Voltage Regulation of an Interleaved Boost Converter in Solar Photovoltaic Systems. *IET Renew. Power Gener.* **2025**, *19*, e70156. [[CrossRef](#)]
48. Tong, L.; Chen, Y.; Xu, T.; Kang, Y. Fault diagnosis for modular multilevel converter (MMC) based on deep learning: An edge implementation using binary neural network. *IEEE J. Emerg. Sel. Top. Power Electron.* **2022**, *11*, 5553–5568. [[CrossRef](#)]
49. Aarrestad, T.; Loncar, V.; Ghielmetti, N.; Pierini, M.; Summers, S.; Ngadiuba, J.; Petersson, C.; Linander, H.; Iiyama, Y.; Di Guglielmo, G.; et al. Fast convolutional neural networks on FPGAs with hls4ml. *Mach. Learn. Sci. Technol.* **2021**, *2*, 045015. [[CrossRef](#)]
50. King, T.; Zhou, Y.; Röddiger, T.; Beigl, M. MicroNAS for memory and latency constrained hardware aware neural architecture search in time series classification on microcontrollers. *Sci. Rep.* **2025**, *15*, 7575. [[CrossRef](#)] [[PubMed](#)]
51. Lohar, D.; Jeangoudoux, C.; Volkova, A.; Darulova, E. Sound mixed fixed-point quantization of neural networks. *ACM Trans. Embed. Comput. Syst.* **2023**, *22*, 136. [[CrossRef](#)]
52. Gheisarnejad, M.; Khooban, M.H. IoT-based DC/DC deep learning power converter control: Real-time implementation. *IEEE Trans. Power Electron.* **2020**, *35*, 13621–13630. [[CrossRef](#)]
53. Sankar, D.; Syamala, L.; Chembathu Ayyappan, B.; Kallarackal, M. FPGA-Based Cost-Effective and Resource Optimized Solution of Predictive Direct Current Control for Power Converters. *Energies* **2021**, *14*, 7669. [[CrossRef](#)]
54. Lavin, A.; Gilligan-Lee, C.M.; Visnjic, A.; Ganju, S.; Newman, D.; Ganguly, S.; Lange, D.; Baydin, A.G.; Sharma, A.; Gibson, A.; et al. Technology readiness levels for machine learning systems. *Nat. Commun.* **2022**, *13*, 6039. [[CrossRef](#)] [[PubMed](#)]
55. Markovic, N.; Stoetzel, T.; Staudt, V.; Kolossa, D. Hybrid condition monitoring for power converters: Learning-based methods with statistical guarantees. *IEEE Access* **2023**, *11*, 31855–31865. [[CrossRef](#)]
56. Xing, Z.; He, Y.; Zhang, W. An online multiple open-switch fault diagnosis method for T-type three-level inverters based on multimodal deep residual filter network. *IEEE Trans. Ind. Electron.* **2022**, *70*, 10669–10679.
57. Gou, B.; Xu, Y.; Xia, Y.; Deng, Q.; Ge, X. An online data-driven method for simultaneous diagnosis of IGBT and current sensor fault of three-phase PWM inverter in induction motor drives. *IEEE Trans. Power Electron.* **2020**, *35*, 13281–13294. [[CrossRef](#)]
58. Xia, Y.; Xu, Y. A transferrable data-driven method for IGBT open-circuit fault diagnosis in three-phase inverters. *IEEE Trans. Power Electron.* **2021**, *36*, 13478–13488. [[CrossRef](#)]

59. Ou, S.; Sahoo, S.; Sangwongwanich, A.; Blaabjerg, F.; Hassanifar, M.; Votava, M.; Langwasser, M.; Liserre, M. Physics-informed neural network for parameter identification: A buck converter case study. In Proceedings of the 2025 IEEE Energy Conversion Congress & Exposition Asia (ECCE-Asia), Bengaluru, India, 11–14 May 2025; pp. 1–6.
60. Sharida, A.; Kamal, N.F.; Alnuweiri, H.; Bayhan, S.; Abu-Rub, H. Digital-twin-based diagnosis and tolerant control of T-type three-level rectifiers. *IEEE Open J. Ind. Electron. Soc.* **2023**, *4*, 230–241. [[CrossRef](#)]
61. Nwoke, J.; Milanese, M.; Viola, J.; Chen, Y.; Visioli, A. A Reduced-Order digital twin FPGA-Based implementation with Self-Awareness capabilities for power electronics applications. *IEEE J. Radio Freq. Identif.* **2024**, *8*, 493–505. [[CrossRef](#)]
62. Lu, Z.; Guo, C.; Liu, M.; Shi, R. Remaining useful lifetime estimation for discrete power electronic devices using physics-informed neural network. *Sci. Rep.* **2023**, *13*, 10167. [[CrossRef](#)] [[PubMed](#)]
63. Ismail, A.; Saidi, L.; Sayadi, M.; Benbouzid, M. Remaining useful lifetime prediction of thermally aged power insulated gate bipolar transistor based on Gaussian process regression. *Trans. Inst. Meas. Control* **2020**, *42*, 2507–2518. [[CrossRef](#)]
64. Xiao, D.; Qin, C.; Ge, J.; Xia, P.; Huang, Y.; Liu, C. Self-attention-based adaptive remaining useful life prediction for IGBT with Monte Carlo dropout. *Knowl.-Based Syst.* **2022**, *239*, 107902. [[CrossRef](#)]
65. Yang, X.; Zhang, Y.; Wu, X.; Liu, G. Failure mode classification of IGBT modules under power cycling tests based on data-driven machine learning framework. *IEEE Trans. Power Electron.* **2023**, *38*, 16130–16141. [[CrossRef](#)]
66. Cheng, H.; Kong, X.; Wang, Q.; Ma, H.; Yang, S. The two-stage RUL prediction across operation conditions using deep transfer learning and insufficient degradation data. *Reliab. Eng. Syst. Saf.* **2022**, *225*, 108581. [[CrossRef](#)]
67. Hari, N.; Ahsan, M.; Ramasamy, S.; Sanjeevikumar, P.; Albarbar, A.; Blaabjerg, F. Gallium nitride power electronic devices modeling using machine learning. *IEEE Access* **2020**, *8*, 119654–119667. [[CrossRef](#)]
68. Krishnamoorthy, H.S.; Aayer, T.N. Machine learning based modeling of power electronic converters. In Proceedings of the 2019 IEEE Energy Conversion Congress and Exposition (ECCE), Baltimore, MD, USA, 29 September–3 October 2019; pp. 666–672.
69. Wunderlich, A.; Santi, E. Digital twin models of power electronic converters using dynamic neural networks. In Proceedings of the 2021 IEEE Applied Power Electronics Conference and Exposition (APEC), Virtual, 14–17 July 2021; pp. 2369–2376.
70. Lale, S.; Shi, Y.; Qu, G.; Azizzadenesheli, K.; Wierman, A.; Anandkumar, A. Kcr1: Krasovskii-constrained reinforcement learning with guaranteed stability in nonlinear dynamical systems. *arXiv* **2022**, arXiv:2206.01704.
71. Bui, V.H.; Chang, F.; Su, W.; Wang, M.; Murphey, Y.L.; Da Silva, F.L.; Huang, C.; Xue, L.; Glatt, R. Deep neural network-based surrogate model for optimal component sizing of power converters using deep reinforcement learning. *IEEE Access* **2022**, *10*, 78702–78712. [[CrossRef](#)]
72. 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](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.