

Review

Recent Progress of Millimeter-Wave Silicon-Based Integrated Mixers for Broadband Wireless Communication: A Comprehensive Survey

Yisi Yang ^{1,2}, Xiuqiong Li ^{1,2}, Yukai Feng ^{1,2}, Yuan Liang ^{1,2,*}, Xinran Huang ^{1,2}, Jiaxin Chen ¹ and Lin Peng ^{1,2,*} 

¹ School of Electronics and Communication Engineering, Guangzhou University, Guangzhou 510006, China; yyatsze@hotmail.com (Y.Y.); 15768664039@163.com (X.L.); linp2@gzhu.edu.cn (Y.F.); 32319500031@gzhu.edu.cn (X.H.); 32407600058@gzhu.edu.cn (J.C.)

² Key Laboratory of On-Chip Communication and Sensor Chip, Guangdong Higher Education Institute, Guangzhou University, Guangzhou 510006, China

* Correspondence: microyuan@126.com (Y.L.); linp@gzhu.edu.cn (L.P.)

Abstract

Mixers are integral components in RF circuits for frequency conversion and are present in almost all RF front-ends. The relentless advancement of mobile communication standards, particularly towards 5G-Advanced and 6G, imposes ever more stringent and multi-dimensional performance requirements on mixer design. While previous surveys have capably summarized mixer technologies, this review distinguishes itself by providing a comprehensive and critical examination of millimeter-wave and sub-THz silicon-based integrated mixers, with explicit coverage extended from core RF bands to beyond 170 GHz. We place particular emphasis on the unique challenges and trade-offs inherent to silicon (CMOS and SiGe BiCMOS) platforms at these high frequencies. This work first summarizes the structural frameworks and underlying principles of mixers, examines multiple mixer variants, and performs an in-depth analysis of their key performance characteristics, encompassing conversion gain, noise figure (with distinctions between single-sideband (SSB) and double-sideband (DSB) definitions), isolation, and related metrics. Then, it compares and discusses the design of several mixers, especially analyzing their innovative points and key technologies, while critically evaluating their inherent limitations and trade-offs. Furthermore, a dedicated section synthesizes the most recent research trends, including heterogeneous integration, AI/ML-assisted design, and mixer architectures for integrated sensing and communication (ISAC), thereby addressing a notable gap in the current literature. Finally, it concludes with an outlook on future challenges and opportunities for mixers in next-generation communication systems.

Keywords: frequency conversion; heterogeneous integration; integrated sensing and communication; millimeter-wave mixers; RF front-end; SiGe BiCMOS; silicon-based integration; sub-THz mixers



Academic Editor: Adão Silva

Received: 18 November 2025

Revised: 12 December 2025

Accepted: 20 December 2025

Published: 2 March 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

As a core subsystem within electronic architectures, the radio frequency (RF) receiver is essential for enabling efficient signal reception and processing, capturing and processing incoming radio signals. Its architecture typically integrates a preamplifier, intermediate frequency (IF) amplifier, mixer, local oscillator, and a bandpass filtering module [1]. As graphically represented in Figure 1, the mixer plays a pivotal role in RF transceiver systems:

it either converts the RF signals received from the antenna into a standardized intermediate frequency (IF) signal or upconverts baseband signals to RF for transmission. To fulfill this frequency translation, the mixer necessitates precise synchronization with a local oscillator (LO) signal derived from a voltage-controlled oscillator (VCO), operating exclusively within the radio frequency domain. Given its central role, designing a high-performance mixer with low noise and minimal power consumption is essential for modern communication systems.

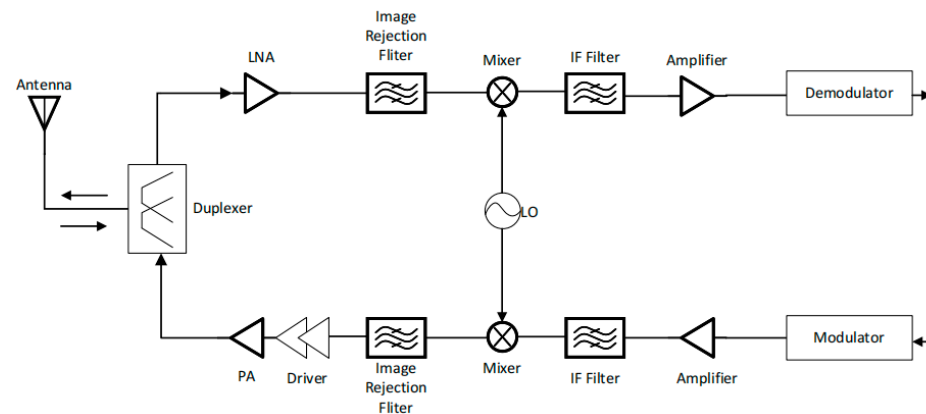


Figure 1. Schematic diagram of RF transceiver system.

Current mixer implementations predominantly leverage CMOS technology, a mainstream choice due to its cost-effectiveness, mature fabrication processes, and compatibility with high-density, low-power circuit design [2]. The scalability of CMOS technology, particularly as it advances toward nanometer-scale nodes, has not only reduced component costs but also enhanced overall system performance. Since P. Y. Chan's pioneering demonstration of a 1 GHz CMOS-based down-conversion mixer in 1993 [3], CMOS mixers have become a focal point of research, driving innovations in RF circuit design.

Early millimeter-wave mixers primarily employed diode-based architectures. While diodes offer simplicity and strong nonlinearity—beneficial for mixing—their performance is constrained by process limitations. Recent advancements in high-frequency circuits have spurred interest in sub-harmonic mixing techniques. By operating the local oscillator at fractions (e.g., half, third, or quarter) of the fundamental frequency, sub-harmonic mixers mitigate LO leakage and improve noise characteristics. However, their dependence on higher-order harmonics frequently leads to diminished conversion efficiency and necessitates elevated LO power requirements [4], underscoring the necessity for innovative design trade-offs.

The optimization of mixer performance parameters remains a multifaceted challenge due to the complex interdependencies among critical metrics, including NF, CG, linearity (third-order intercept point, IIP3), local oscillator-to-radio frequency (LO-RF) isolation, power consumption, and chip area. Over the past two decades, advancements in mixer design have achieved notable progress in individual performance metrics; however, attaining a comprehensive balance across all parameters persists as a formidable obstacle, particularly in addressing the stringent requirements of emerging applications such as 5G networks, millimeter-wave communication systems, and ultra-low-power Internet of Things (IoT) devices. A systematic review of the literature spanning 2000–2022 (Figure 2) underscores both technological advancements and unresolved limitations, highlighting the need for innovative design paradigms.

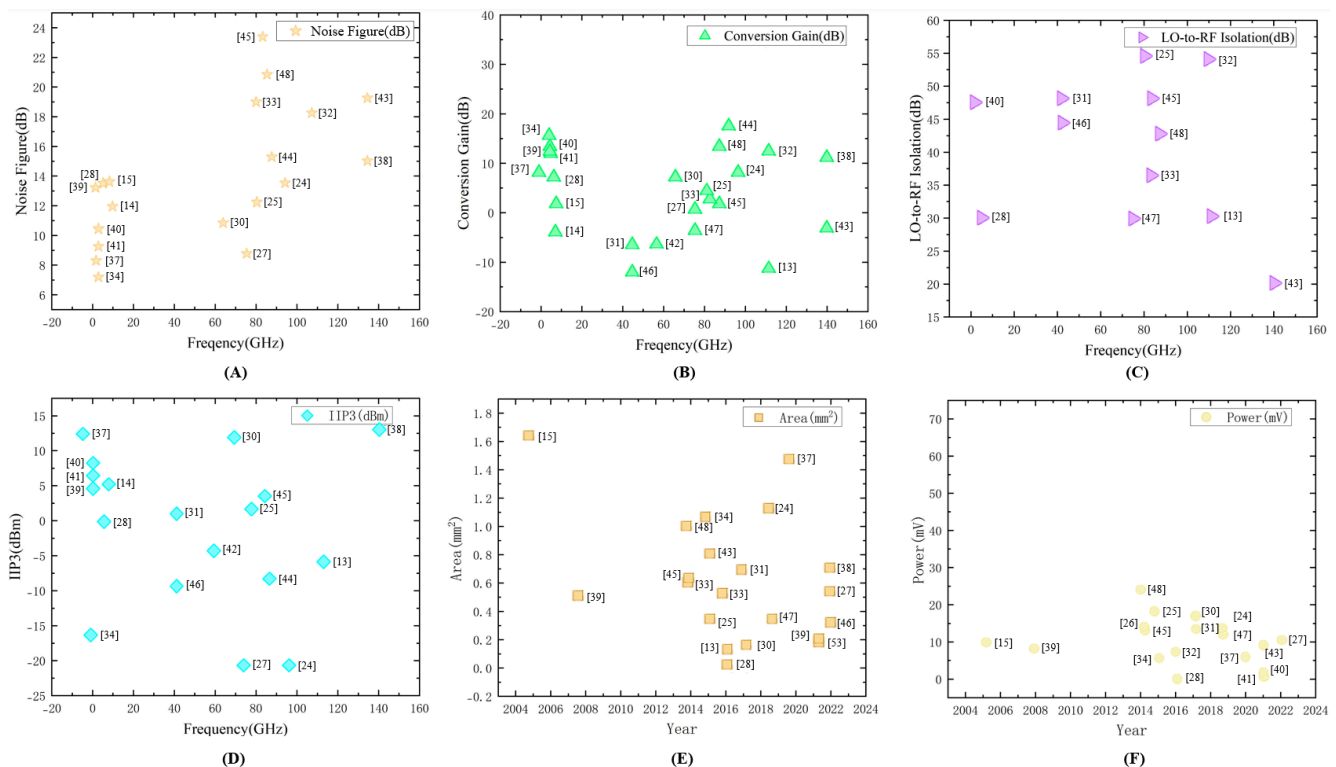


Figure 2. Comparison of important parameters of mixers in different literature. (A) Noise Figure, (B) Conversion Gain, (C) LO-to-RF Isolation, (D) IIP3, (E) Area, (F) Power.

While existing surveys capably summarized the state of the art up to around 2020, the field continues to evolve rapidly, driven by the push towards 6G, integrated sensing and communication (ISAC), and terahertz technologies. Therefore, this review not only consolidates the foundational and mainstream technologies but also extends its scope to encompass significant developments reported in recent years. This updated perspective aims to provide a comprehensive reference that bridges historical context with the cutting-edge innovations shaping the future of silicon-based mixers.

A principal challenge arises from the intrinsic trade-offs between key performance indicators. For instance, architectures prioritizing high conversion gain frequently compromise noise performance or power efficiency. As depicted in Figure 2B, CG values exhibit a broad spectrum, ranging from -12 dB to 32 dB, with elevated gains often correlating to degraded NF. A reported 110 GHz mixer, despite its exceptional CG of 32 dB, demonstrates an NF of 9.5 dB and a power consumption of 65 mW—characteristics that may limit its applicability in low-noise, energy-sensitive scenarios. Conversely, energy-efficient designs achieving 0.63 mW at 6 GHz report moderate CG (7 dB) but exhibit suboptimal linearity (IIP3 = 0 dBm), rendering them unsuitable for high-fidelity signal environments. Similarly, compact implementations balancing area efficiency with a CG of 12.5 dB at 2.4 GHz show an IIP3 of only 7.6 dBm, which remains inadequate for interference-prone applications. These examples illustrate the persistent difficulty in harmonizing gain, noise, linearity, and power within a unified architecture.

Further complicating the optimization landscape is the inconsistent reporting of critical parameters across studies, which impedes comparative analysis and benchmarking. Notably, over 30% of surveyed works omit LO-RF isolation values—a metric essential for minimizing signal leakage in densely integrated RF systems. While some designs achieve 55 dB isolation at 80 GHz, others operating at 140 GHz or 2.4 GHz lack isolation data entirely, obscuring their viability for high-frequency deployments. Linearity metrics (IIP3) are similarly underrepresented, with 25% of studies excluding this param-

eter. The absence of standardized reporting protocols obstructs the identification of optimal design methodologies, particularly for frequency-agile systems demanding wide operational bandwidths.

Design priorities further diverge across frequency regimes, exacerbating optimization challenges. High-frequency mixers operating at 110–140 GHz emphasize LO-RF isolation and CG to mitigate path loss and phase noise, often at the expense of power efficiency. In contrast, sub-6 GHz designs prioritize NF and power consumption but demonstrate stagnating linearity. Strikingly, NF values for sub-6 GHz mixers have plateaued between 7–13 dB since 2015, with only marginal improvements in recent years. This stagnation underscores the necessity for disruptive architectural innovations, such as active feedback topologies or advanced semiconductor integration, to transcend existing performance boundaries.

Emerging trends in mixer design also reveal unresolved tensions between miniaturization and operational robustness. While some recent advancements exemplify exceptional area efficiency, their resilience under harsh environmental conditions (e.g., thermal fluctuations, electromagnetic interference) remains unvalidated. Conversely, bulkier architectures deliver high CG and robust isolation but incur prohibitive power costs, limiting their suitability for portable applications. The transition toward millimeter-wave frequencies amplifies these trade-offs, as parasitic effects and process variations impose stringent constraints on manufacturing yield and reproducibility.

Addressing these challenges necessitates a paradigm shift from conventional single-parameter optimization to holistic co-design frameworks that account for metric interdependencies. For example, leveraging advanced CMOS process nodes or heterogeneous integration techniques could concurrently enhance linearity and reduce power consumption. Furthermore, machine-learning-driven topology exploration, augmented by process-voltage-temperature (PVT)-aware simulations, holds promise for identifying Pareto-optimal solutions across diverse frequency bands. By synthesizing historical performance trends and adopting a systems-level perspective, future mixer designs may better align with the multidimensional demands of next-generation communication systems.

2. Operating Principle of Mixer for Frequency Conversion

As a critical nonlinear time-invariant component in frequency translation systems, mixers inherently require nonlinear characteristics to achieve spectral shifting functionality [5]. These devices serve as frequency translation apparatuses, enabling signal conversion between RF and lower IF or baseband domains. In modern communication architectures, this down-conversion process facilitates enhanced signal selectivity and high-gain amplification through RF-to-IF transformation. The fundamental tri-port mixer configuration, as schematically illustrated in Figure 3, incorporates an RF input port and an LO excitation port, with the resultant output manifesting at the IF port. The nonlinear interaction between the RF and LO signals generates the desired frequency-translated IF component through controlled harmonic superposition.

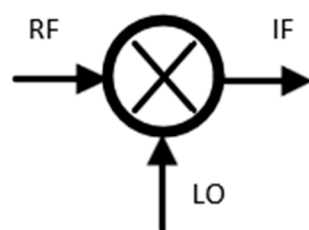


Figure 3. Mixer schematic diagram.

Analytically, their fundamental operation can be characterized by the multiplicative interaction between the RF and LO signals in the time domain, constituting a nonlinear process governed by harmonic superposition principles [6].

Consider the RF and LO signals expressed in their canonical sinusoidal forms: $V_{RF}(t) = A_{RF}\cos(\omega_{RF}t)$ and $V_{LO}(t) = A_{LO}\cos(\omega_{LO}t)$, where A_{RF} and A_{LO} denote the respective signal amplitudes, and ω_{RF} and ω_{LO} represent the corresponding angular frequencies. Through the multiplicative operation inherent to mixer functionality, the resultant output can be mathematically derived as follows:

$$A_{LO} \cos \omega_{LO}t \times A_{RF} \cos \omega_{RF}t = \frac{A_{RF}A_{LO}}{2} [\cos(\omega_{LO} - \omega_{RF})t + \cos(\omega_{LO} + \omega_{RF})t] \quad (1)$$

The multiplication of LO and RF signals generates two distinct frequency components [7], thereby achieving the fundamental RF-to-IF conversion. Based on application requirements, designers can strategically select the desired frequency component as the output while filtering out unwanted components via a bandpass filter.

2.1. Classification of Mixers

Mixers can be classified according to various criteria. In terms of mixing gain characteristics, they are generally divided into passive mixers and active mixers. Additionally, based on circuit topology, mixers can be categorized as single-ended, single-balanced, or double-balanced configurations [8].

2.1.1. Classification of Mixers

Mixers are broadly classified into active and passive types based on the presence of current biasing and their ability to provide conversion gain [9]. As illustrated in Figure 4, passive mixers (Figure 4a) and active mixers (Figure 4b) exhibit distinct circuit architectures. Passive mixers, devoid of transconductance-stage gain structures and DC biasing, inherently consume negligible DC power, making them ideal for ultra-low-power applications. Their simplified transistor configurations contribute to superior noise performance. Additionally, the absence of stacked transistors in the signal path enhances voltage headroom, thereby improving linearity. However, these advantages come with trade-offs: passive mixers inherently introduce conversion loss rather than gain, and their performance demands higher local oscillator (LO) power to compensate for the lack of amplification [10].

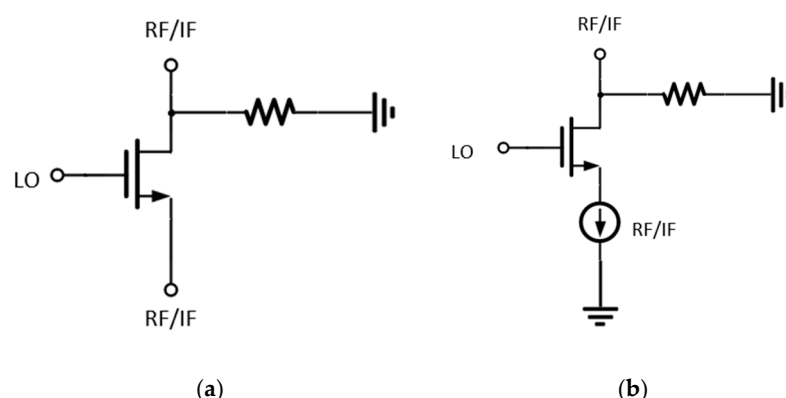


Figure 4. Schematic diagrams of passive and active mixers (a) passive mixer, and (b) active mixer.

Shown in Table 1, in contrast, active mixers integrate gain amplification stages, enabling robust performance with lower LO power requirements. They effectively suppress spurious signals, occupy smaller footprints, and achieve better port isolation compared to passive counterparts. Despite these benefits, active designs face challenges: the increased

transistor count elevates noise levels and power consumption, imposes stricter supply voltage constraints, and may degrade linearity due to active component limitations [11]. Figure 5 further contrasts the circuit topologies of single-balanced passive and active mixers, highlighting their structural differences.

Table 1. Performance Comparison between Active Mixers and Passive Mixers.

	Passive Mixer	Active Mixer
Drain Current	~2 to 5 mA	~0 mA
LO Power	~13 to 20 dBm	~-10 to 5 dBm
Conversion Power	~5 to 10 dB	~-5 to -3 dB
Noise Figure	~10 to 15 dB	~3 to 5 dB
Bandwidth	narrow	wide
Reliability	Low	high
Cost	low	high

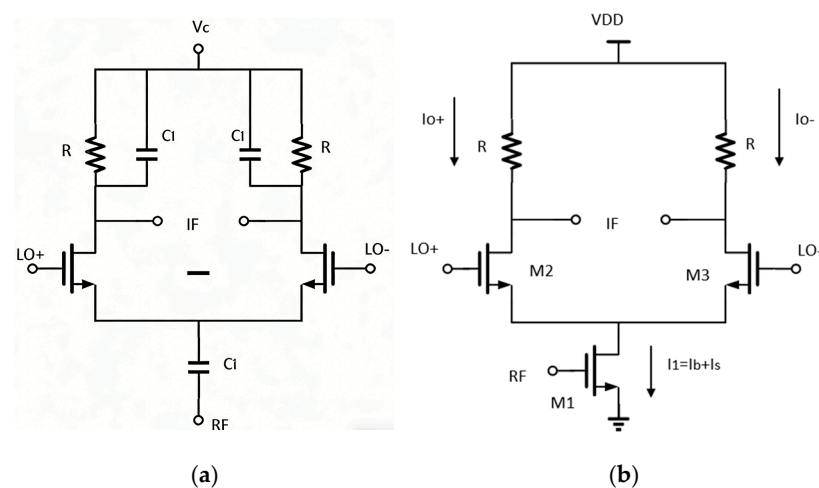


Figure 5. (a) single-balanced passive mixer, and (b) single-balanced active mixer.

2.1.2. Single-Balanced Mixer

As depicted in Figure 5b, the circuit schematic represents a single-balanced mixer. In this configuration, the differential LO signal drives a switching stage composed of transistors M2 and M3. Simultaneously, the RF or IF signal, which is applied as a single-ended input, is processed by a transconductance stage implemented with transistor M1. This stage facilitates the conversion of the input RF voltage signal into a corresponding current signal. The switching transistors, modulated by the LO signal, perform the mixing operation on the current signal. The resulting mixed current is subsequently converted back into a differential voltage output through the load. Owing to its differential output structure, the single-balanced mixer offers a conversion gain approximately twice that of its single-ended counterpart, which fails to generate output when its active device is in the off state.

The following section analyzes the output signal and conversion gain of the single-balanced mixer. I_1 represents the current of transistor M1, where I_b is the static bias current and is the current of the transconductance transistor. If the LO signal is sufficiently strong, under ideal conditions, transistors M2 and M3 can be regarded as ideal switches. Then, the

mixing process can be considered as the multiplication of the RF signal with a square wave signal controlled by the LO frequency. The output voltage can be expressed as:

$$V_{IF} = I_1 \times \text{sgn}[\cos \omega_{LO} t] \times R_L = (I_b + I_s) \times \frac{4}{\pi} \sum_{n=1}^{\infty} \frac{1}{n} \sin\left(\frac{n\pi}{2}\right) \cos(n\omega_{LO} t) \times R_L \quad (2)$$

If $I_s = g_{m1} V_{RF} \cos(\omega_{RF} t)$, obtain the following results:

$$V_{IF} = \frac{4}{\pi} I_b R_L \times (\cos \omega_{LO} t - \frac{1}{3} \cos 3\omega_{LO} t + \frac{1}{5} \cos 5\omega_{LO} t - \dots) + \frac{4}{\pi} g_{m1} V_{RF} R_L \times \cos \omega_{RF} t \times (\cos \omega_{LO} t - \frac{1}{3} \cos 3\omega_{LO} t + \frac{1}{5} \cos 5\omega_{LO} t - \dots) \quad (3)$$

By analyzing the relationship between the input RF signal and the output, the voltage conversion gain of the single-balanced mixer can be determined as:

$$G_v = \frac{2}{\pi} g_{m1} R_L \quad (4)$$

It is evident that in a single-balanced active mixer, the output signal contains a significant presence of the LO signal along with its odd harmonics. The interaction between the LO and RF signals generates multiple sum and difference frequency components, necessitating the suppression of undesired signals through filtering. In upconversion applications, excessive LO-to-RF leakage can notably degrade the performance of subsequent circuit stages. Moreover, when the RF signal frequency is in close proximity to the LO frequency, the design constraints on the output filtering network become more stringent, making effective suppression of unwanted components increasingly challenging. Consequently, for high-frequency and high-performance CMOS mixer designs, the straightforward implementation of a single-balanced topology is often impractical due to these inherent limitations [11].

2.1.3. Double-Balanced Mixer

To address the limitations inherent in the single-balanced mixer structure, the double-balanced mixer was developed, as illustrated in Figure 6. This configuration, commonly referred to as the “Gilbert Mixer,” is a widely utilized component in modern RF circuits. Its design traces back to the BJT transistor multiplier presented in Barrie Gilbert’s 1968 publication [12]. The double-balanced mixer can be viewed as an enhanced version of the single-balanced mixer, where the input RF or IF signal is transformed into a differential signal, building upon the principles of the single-balanced configuration.

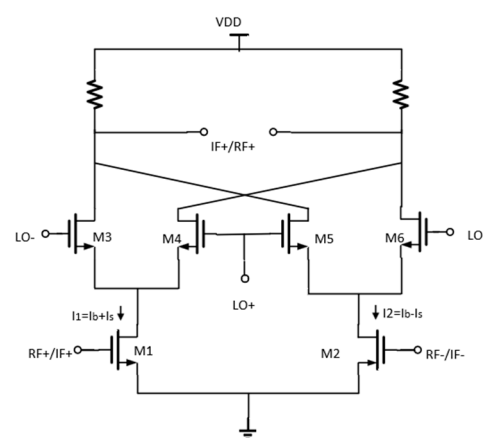


Figure 6. Schematic Diagram of a Double-Balanced Mixer.

Next, we will conduct an analysis of the output signal and conversion gain of the double-balanced mixer. In Figure 6, I_1 represents the total current of transistor M1, and I_2 represents the total current of transistor M2. Among them, I_b is the static bias current,

and I_s is the current of the transconductance stage transistor. Assuming that transistors M3 to M6 are ideal switches, the mixing process can be considered as the RF signal being multiplied by a square wave signal controlled by the LO frequency. The output voltage can then be expressed as:

$$V_{IF} = (I_1 - I_2) \times \text{sgn}[\cos \omega_{LO} t] \times R_L = 2I_s \times \frac{4}{\pi} \sum_{n=1}^{\infty} \frac{1}{n} \sin\left(\frac{n\pi}{2}\right) \cos(n\omega_{LO} t) \times R_L \quad (5)$$

Due to the use of differential signaling at the RF port, the direct current component is canceled out during the output synthesis. Consequently, as we will observe later, the fundamental and harmonic terms of the LO are absent, highlighting the crucial significance of the double-balanced mixer in eliminating these unwanted LO components.

If $I_s = g_{m1,2} V_{RF} \cos(\omega_{RF} t)$, The following results can be obtained:

$$V_{IF} = \frac{8}{\pi} g_{m1,2} V_{RF} R_L \times \cos \omega_{RF} t \times \left(\cos \omega_{LO} t - \frac{1}{3} \cos 3\omega_{LO} t + \frac{1}{5} \cos 5\omega_{LO} t - \dots \right) \\ = \frac{4}{\pi} g_{m1,2} V_{RF} R_L \left\{ \cos(\omega_{RF} - \omega_{LO})t + \cos(\omega_{RF} + \omega_{LO}) - \frac{1}{3} [\cos(\omega_{RF} - 3\omega_{LO})t + \cos(\omega_{RF} + 3\omega_{LO})t] + \dots \right\} \quad (6)$$

By comparing with the differential input RF signal, the voltage conversion gain of the double-balanced mixer can be obtained as follows:

$$G_v = \frac{2}{\pi} g_{m1,2} R_L \quad (7)$$

Despite exhibiting a twofold enhancement in output amplitude relative to single-balanced mixer architectures, the double-balanced configuration simultaneously requires doubled input signal drive levels. Consequently, the conversion gain equivalence between the two configurations persists, maintaining parity in their respective gain characteristics.

The double-balanced mixer's output spectrum is fundamentally confined to the algebraic superposition of the RF signal with the odd-order harmonic components of the LO, effectively suppressing even-order distortion products. This spectral purity improvement manifests as a significant reduction in intermodulation interference compared to single-balanced implementations. Furthermore, the inherent suppression of LO and RF feedthrough in the output stage, coupled with the utilization of a differential topology in the switching core, achieves superior port-to-port isolation exceeding 40 dB. The architectural symmetry inherent in double-balanced designs ensures optimal decoupling between RF, LO, and IF signal paths, thereby establishing them as the preferred solution for high-isolation applications. Therefore, when compared to single-balanced mixer circuits, the double-balanced mixer offers numerous advantages, such as high common-mode rejection, high isolation, and reduced harmonic interference, which are beneficial for high-frequency integrated circuit design, leading to its widespread application [13]. However, its primary drawbacks lie in its relatively poor power consumption performance and higher noise levels.

2.1.4. Distributed Mixer

The distributed mixer concept emerged in 1973 through the pioneering work of Sith and Robson [14] and has since become a prominent solution in RF systems owing to its exceptional noise performance, broad operational bandwidth, and superior dynamic range characteristics. Unlike conventional mixer architectures, this innovative design capitalizes on inherent device parasitics—specifically transistor junction capacitances and interconnect lead inductances—to construct synthetic transmission lines with ultra-high cutoff frequencies. This unique implementation establishes an extended impedance matching bandwidth that remains unmatched by traditional mixer topologies.

The frequency translation mechanism inherently generates spectral components at sum and difference frequencies through nonlinear intermodulation between RF and LO waveforms, a process governed by heterodyne signal processing principles. When the resulting difference frequency aligns with the predefined IF range, subsequent IF amplifier stages can effectively process the signal while preserving its original modulation characteristics. Current distributed mixer implementations primarily exist in two distinct configurations: active and passive variants. Active implementations offer conversion gain advantages at the expense of increased DC power consumption, whereas passive configurations eliminate static power requirements but incur higher conversion losses and demand more stringent LO drive levels, with compromised port-to-port isolation characteristics [15].

2.1.5. Subharmonic Mixers

The subharmonic mixer utilizes nonlinear elements (diodes or FETs) to perform frequency translation, specifically designed to combat LO self-mixing effects originating from LO-RF feedthrough. This parasitic coupling phenomenon manifests as DC offset generation, signal integrity degradation, and SNR deterioration when the LO signal inadvertently interacts with itself through the RF path. As shown in Figure 7, the architecture exploits parametric multiplication mechanisms to produce harmonic multiples of both RF and LO frequencies. The inherent frequency separation between fundamental LO and RF signals provides inherent port isolation enhancement, effectively suppressing system-level self-interference [16].

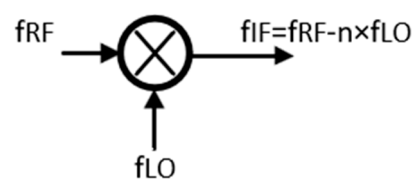


Figure 7. Schematic Diagram of the Subharmonic Mixer.

Operating through nonlinear device characteristics, the mixer generates intermodulation products when simultaneously driven by RF and LO inputs. Subharmonic components emerge at rational fractions (typically $1/2$) of the fundamental RF frequency, enabling operation with LO signals at sub-multiples of the target frequency. This property proves particularly advantageous in millimeter-wave and terahertz systems where high-frequency LO sources face practical limitations in power output and phase noise performance. While tolerating moderate conversion loss and noise figure penalties, the topology achieves relaxed LO source specifications—a critical design compromise for cost-sensitive high-frequency applications.

2.2. Performance Metrics of the Mixer

Positioned downstream of the LNA's output filtering stage, the mixer performs critical frequency conversion operations on the amplified RF signals through nonlinear device characteristics activated by a precisely synchronized LO input. This frequency translation mechanism, fundamentally reliant on the injection of an LO signal with controlled amplitude and phase characteristics, necessitates rigorous co-optimization of both general RF integrated circuit parameters and specialized mixer performance criteria. The design architecture must concurrently address conventional RFIC specifications—including but not limited to conversion efficiency, intermodulation distortion characteristics, noise contribution analysis, multi-port impedance matching networks, and DC-to-RF conversion efficiency—while fulfilling stringent mixer-specific requirements such as inter-port

isolation metrics, spurious response suppression, and image frequency rejection ratios through advanced topological implementations. However, these performance parameters are not independent (Figure 8); they exhibit intricate trade-offs that must be carefully balanced during the design phase. This complexity highlights the critical need for thorough performance characterization and systematic optimization to achieve the best possible system-level performance.

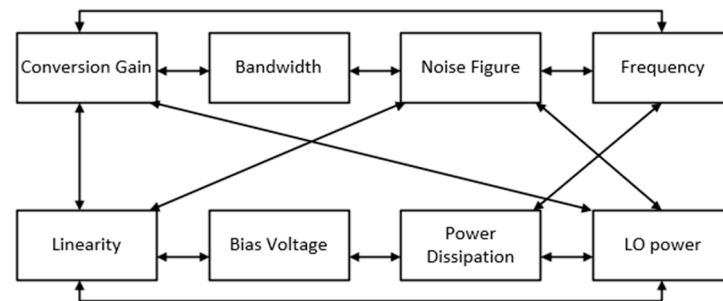


Figure 8. Mutual Influences of Performance Indicators in Mixer Design.

2.2.1. Conversion Gain

The operational paradigm of mixers fundamentally differs from conventional amplification stages due to their tri-port architecture and frequency translation mechanisms. Unlike amplifiers' unidirectional gain characteristics, mixers exhibit conversion gain/loss metrics stemming from their nonlinear signal interaction between disparate frequency domains. This critical parameter is quantitatively defined as the logarithmic ratio between output IF power and input RF power, formally expressed as:

$$G_P = \frac{P_{IF}}{P_{RF}} = \frac{V_{IF}^2 / R_L}{V_{RF}^2 / R_S} \quad (8)$$

In the formula, V_{IF} and V_{RF} represent the RMS values of the IF output voltage and the RF input voltage, respectively. R_L denotes the load resistance, which is the resistance connected to the output of the mixer, and R_S signifies the source resistance, which is the resistance associated with the input signal source.

Additionally, another representation is the voltage ratio between the output IF signal and the input RF signal, which is known as voltage gain:

$$G_V = \frac{V_{IF}}{V_{RF}} \quad (9)$$

Through analytical derivation, it is demonstrated that under impedance matching conditions, power gain and voltage gain exhibit a direct mathematical relationship, establishing their equivalence in performance evaluation. This fundamental principle serves as a cornerstone for accurately assessing mixer behavior across various operational modes. In down-conversion architectures, optimal selection of conversion gain fulfills two critical roles: (1) mitigating cascaded noise figure degradation by minimizing noise contributions from subsequent baseband stages, as dictated by Friis' formula in multi-stage systems; (2) optimizing gain distribution to relax linearity constraints on the preceding LNA, while ensuring adherence to receiver sensitivity requirements. Conversely, in up-conversion architectures, controlled conversion gain is strategically leveraged to reduce the drive-level demands on the PA within the transmission chain, thereby enhancing PAE under fixed DC power constraints [17].

2.2.2. Noise Figure

Enhancing the sensitivity of an RF receiver necessitates careful consideration of its noise performance, which is conventionally quantified by the noise figure [18]. This parameter represents the fundamental lower bound of the receiver's dynamic range across various frequency bands.

Noise is a random and fluctuating phenomenon that manifests as irregular variations in current or voltage. When superimposed on a signal, noise can degrade the signal quality and reduce the amount of information it carries. In communication systems, noise can generally be categorized into two types based on its origin: internal noise and external noise. Internal noise is generated within the receiver's circuitry, such as thermal noise in resistors and flicker noise in active devices. External noise, in contrast, comes from sources outside the system, including environmental noise and man-made interference. In receiver design, the primary focus is on minimizing the internal noise generated by the circuit itself (Figure 9).

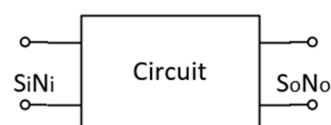


Figure 9. Input and Output Signal Noise of the Circuit.

Within the spectrum of intrinsic circuit noise mechanisms, two dominant contributors merit detailed analysis. Thermal noise (commonly denoted as Johnson-Nyquist noise) predominantly manifests in active transistors and passive resistive components through stochastic thermal agitation of charge carriers, exhibiting frequency-independent spectral density across operational bandwidths. Contrastingly, flicker noise—characterized by its $1/f$ spectral dependence—originates primarily from quantum-mechanical imperfections in metal-oxide-semiconductor (MOS) device architectures, particularly interface trap states at the gate dielectric-channel boundary. These crystallographic discontinuities induce discrete charge carrier trapping/detrapping phenomena, resulting in low-frequency conductivity modulation that directly impacts device transconductance stability [19]:

$$V_n^2 = \frac{K}{C_{OX}WL} \frac{1}{f} \quad (10)$$

The NF characterizes the deterioration of a signal's signal-to-noise ratio (SNR) as it propagates through a circuit. It is defined as the ratio of the SNR at the input to that at the output, highlighting the extent of performance degradation:

$$N_F = \frac{S_i/N_i}{S_o/N_o} \quad (11)$$

Within the noise figure quantification framework, four fundamental parameters govern system performance characterization: S_i denoting incident signal power at the input port, N_i signifying inherent thermal noise power at the input interface, S_o representing processed signal power at the output stage, and N_o accounting for cumulative noise power at the output terminal. These operational parameters collectively define the signal integrity preservation and noise amplification characteristics of active components. The standardized noise figure metric, constituting a logarithmic measure of signal-to-noise ratio degradation

through the device under test, is universally quantified in decibel units through discrete logarithmic transformation of the fundamental power ratio relationships:

$$NF = 10\lg \frac{SNR_i}{SNR_o} = 10\lg \frac{S_i/N_i}{S_o/N_o} \quad (12)$$

The noise of a mixer is a kind of frequency-converted noise. In terms of frequency conversion, it refers to the ratio of SNR at the IF to that at the RF. The spectral folding phenomenon inherent to superheterodyne architectures introduces critical noise considerations distinct from homodyne systems. In conventional heterodyne receivers, reciprocal image frequency components symmetrically offset from the LO undergo identical IF translation, resulting in dual noise contributions at the target IF band. This architectural constraint necessitates rigorous image rejection, whereas zero-IF topologies inherently avoid such spectral contamination through baseband-aligned frequency conversion.

This operational dichotomy directly governs noise figure classification methodologies: single-sideband (SSB) characterization applies to systems with unilateral spectral processing (e.g., heterodyne architectures requiring image rejection filters), while double-sideband (DSB) quantification pertains to bilateral spectral utilization (e.g., direct-conversion receivers). For equivalent circuit implementations, the DSB noise figure exhibits a 3 dB theoretical advantage over SSB configurations due to coherent noise cancellation mechanisms enabled by symmetrical spectral folding [20].

They are generally represented by the following formulas:

$$NF_{SSB} = \frac{S_s/N_s}{GS_s/(GN_s + GN_s + N_{cir})} = \frac{S_s/N_s}{GS_s/(2GN_s + N_{cir})} = 2(1 + \frac{N_{cir}}{2GN_s}) \quad (13)$$

$$NF_{DSB} = \frac{2S_s/2N_s}{2GS_s/(GN_s + GN_s + N_{cir})} = \frac{S_s/N_s}{2GS_s/(2GN_s + N_{cir})} = 1 + \frac{N_{cir}}{2GN_s} \quad (14)$$

In the formula, S_s denotes the input signal, N_s represents the input noise, and N_{cir} stands for the internal circuit noise.

2.2.3. Linearity

Like the noise figure, linearity determines the upper boundary of a system's dynamic range, indicating the highest signal level that can be handled without introducing distortion. In practice, up-conversion mixers generally have stricter linearity requirements than down-conversion mixers. This is because high-power transmission is essential at the transmitting end, whereas at the receiving end, sensitivity to weak signals is more critical. The linearity of a mixer is commonly evaluated using the 1 dB compression point (P1dB) and the third-order intercept point (IP3) [21]. Figure 10 provides a schematic representation of P1dB and IP3, illustrating how they define the mixer's linearity performance.

Figure 11a illustrates the schematic diagram of the P1dB compression point. The horizontal axis represents the input power (P_{in}) of the mixer, while the vertical axis represents the output power (P_{out}) of the mixer. Due to nonlinearity, as P_{in} increases, its power gain begins to decrease. When the power gain drops by 1 dB compared to the linear gain, the corresponding input power is referred to as the Input 1 dB Gain Compression Point (IP1dB), and the output power is referred to as the Output 1 dB Gain Compression Point (OP1dB). For a nonlinear device or system, its input-output relationship can be expressed as:

$$y(t) = \alpha_1 x(t) + \alpha_2 x(t)^2 + \alpha_3 x(t)^3 + \dots \quad (15)$$

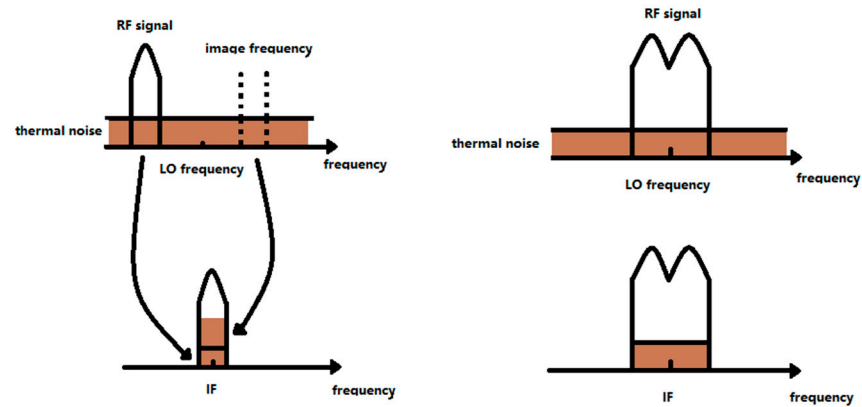


Figure 10. Distinction Between Single-Sideband Noise and Double-Sideband Noise.

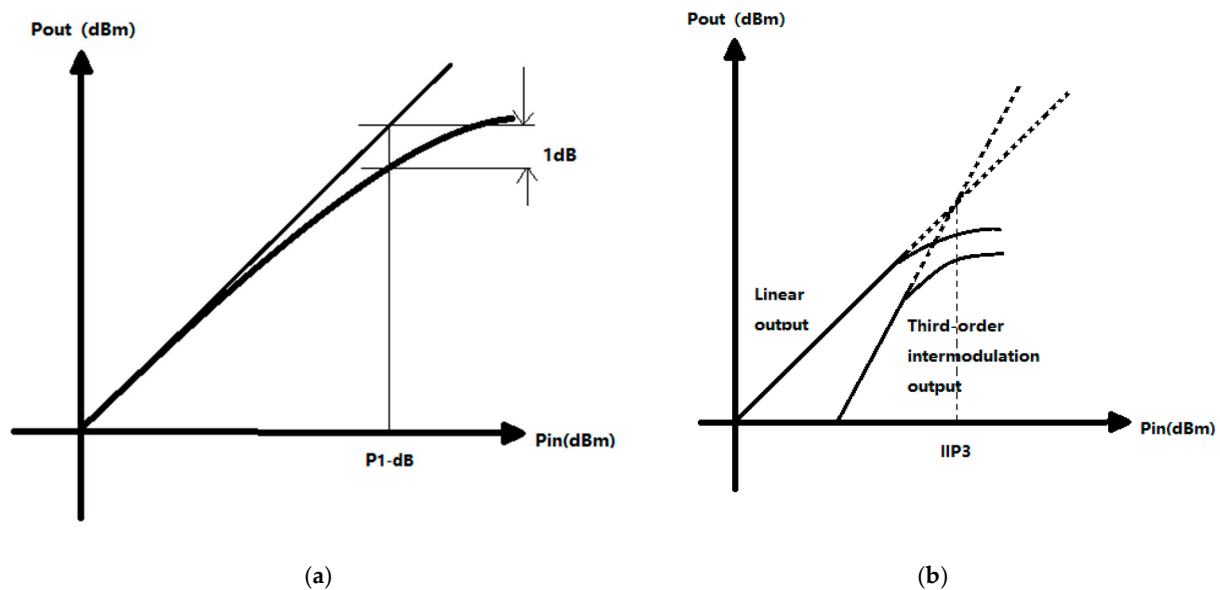


Figure 11. (a) 1 dB Compression Point, and (b) Third-Order Intermodulation Point.

In nonlinear system characterization, the temporal input-output relationship is defined by $x(t)$ and $y(t)$, corresponding to the baseband excitation and distorted response waveforms, respectively. The coefficients quantify the magnitude scaling of n th-order intermodulation distortion products generated through device nonlinearities.

Third-order intercept (IP3) metrics define critical linearity boundaries where extrapolated third-order intermodulation products achieve parity with the fundamental frequency response under dual-tone excitation. These operational thresholds are categorized as input-referred (IIP3) and output-referred (OIP3) intercept points, reflecting system gain characteristics. Analogous to noise figure cascading principles, the aggregate linearity performance of multi-stage architectures is dictated by constituent stage contributions, with front-end components predominantly governing the overall IP3 through nonlinear superposition mechanisms. In an up-conversion mixer, a certain OP1dB is required to enable the PA to deliver high-power signals. For a down-conversion mixer, a certain IP1dB is needed to ensure that the signal from the LNA is not compressed.

2.2.4. Isolation

In multi-port devices where each port operates at a different frequency, leakage of signals between ports can interfere with other parts of the system, impacting the output signal. Isolation is used to quantify the coupling power between two ports. A mixer, being a three-port device, has signals at different frequencies at each port. During the design of a

mixer, good isolation metrics are essential to prevent mutual interference between the three ports. The leakage pathways within a mixer are depicted in Figure 12.

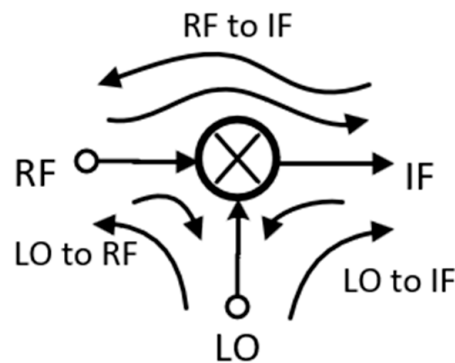


Figure 12. Mixer Leakage Paths.

Typically, there are three types of isolation measurements performed on mixers: LO-RF, LO-IF, and RF-IF. Due to various factors, significant interference exists between signals at different ports, such as parasitic effects between stages of transistors, substrate effects, and others, which are inherent and cannot be altered during circuit design. If the three ports of a mixer do not exhibit good isolation, not only will the normal operation of the mixer itself be affected, but the signal transmission of the entire RF system will also be disrupted. The impact of such leakage can be quantified using port isolation, which can be expressed using the following formulas:

$$\begin{aligned} I_{\text{LO-RF}} &= \frac{\text{LO Signal Power at RF Port}}{\text{LO Input Power}} \\ I_{\text{RF-LO}} &= \frac{\text{RF Signal Power at LO Port}}{\text{RF Input Power}} \end{aligned} \quad (16)$$

Leakage between the RF and LO ports is generally more significant compared to other ports. Since the IF port operates at a considerably different frequency, it can be effectively filtered out, minimizing its impact. Parasitic coupling mechanisms in mixer architectures introduce critical performance degradation modes requiring rigorous mitigation. Reverse signal coupling from the RF port to the LO interface induces an autodyne conversion process, where leaked RF energy undergoes nonlinear interaction with the incident RF waveform, producing spurious DC offsets that compromise baseband signal integrity. Proximity between RF carrier and LO frequencies exacerbates this phenomenon through undesired injection locking phenomena, manifesting as spectral impurity in oscillator phase noise characteristics and long-term frequency drift instability. Conversely, forward leakage pathways from the oscillator to the receiver chain initiate degenerative remixing effects, wherein residual LO components self-mix within the nonlinear device transconductance, amplifying in-band distortion products and degrading overall system dynamic range. In down-conversion, LO leakage that couples to the LNA input through the substrate can introduce a strong LO signal at the RF port of the mixer, significantly impacting performance. In up-conversion, excessive LO leakage may cause saturation in the input or output stages of subsequent circuits, severely degrading linearity [22].

As shown in Table 2, the comparison of performance indicators among different types of mixers provides a clearer understanding of their characteristics and trade-offs.

Table 2. Performance Comparison Among Different Types of Mixers.

	Single-Balanced Mixer	Classic Gilbert Cell	Modified Gilbert Cell	Passive Distributed Mixer	Active Distributed Mixer	Subharmonic Mixer
Conversion Gain	A	A	B	C	B	D
Noise Figure	C	D	B	C	D	A
Linearity	D	D	B	A	C	B
Isolation	C	A	B	D	C	A
LO Power	A	A	B	C	C	D
Bandwidth	C	D	B	A	B	C
PDC	B	C	B	A	-	B

Note: Earlier alphabetical positions (A is best) indicate enhanced performance levels.

2.3. Performance Metrics

2.3.1. Conversion Gain (CG)

CG is the ratio of output IF power to input RF power (power gain G_P) or the corresponding voltage ratio (voltage gain G_V). Under impedance matching, they are directly related. Adequate CG minimizes noise contribution from subsequent stages and relaxes linearity requirements on preceding blocks.

2.3.2. Noise Figure (NF)

NF measures the degradation in signal-to-noise ratio (SNR) as a signal passes through a device: $NF = 10\log_{10}(SNR_{in}/SNR_{out})$. For mixers, a critical distinction exists between double-sideband (DSB) NF and single-sideband (SSB) NF: (1) DSB NF assumes the desired signal and its image band both contain identical signal power and noise. It is typically used to characterize the mixer core itself or in direct-conversion (zero-IF) receivers where the image coincides with the signal. (2) SSB NF accounts for the fact that in a superheterodyne receiver, only one sideband carries the desired signal, while the image band contains only noise. The SSB NF is approximately 3 dB higher than the DSB NF for the same mixer due to the extra noise contribution from the image band. This distinction is crucial when comparing mixer performance across different architectures. Throughout this review, we specify the NF type (SSB or DSB) when reporting data from literature. In cases where the original source does not specify, it is noted. Circuit noise sources include thermal noise and flicker (1/f) noise, the latter being significant in CMOS devices at low frequencies.

2.3.3. Linearity

Linearity determines the upper limit of a system's dynamic range. Key metrics are: (1) 1 dB Compression Point (P1dB): The input (IP1dB) or output (OP1dB) power where the gain drops by 1 dB from its linear value. (2) Third-Order Intercept Point (IP3): The extrapolated point where the power of the third-order intermodulation product equals that of the fundamental tone. Input-referred (IIP3) and output-referred (OIP3) are used. Up-conversion mixers often require higher OP1dB to drive power amplifiers, while down-conversion mixers need sufficient IP1dB to handle signals from the low-noise amplifier without compression.

2.3.4. Isolation

Isolation measures leakage between ports (LO-RF, LO-IF, RF-IF). High LO-RF isolation is particularly important to prevent LO leakage from causing self-mixing (generating DC offsets) or saturating other stages. Poor isolation degrades system performance.

2.3.5. Trade-Offs and Design Challenges

These metrics are interrelated. For example, improving CG often increases power consumption and may degrade NF. Enhancing linearity might require more power or reduce gain. Table 2 qualitatively compares different mixer topologies across these metrics, illustrating the inherent compromises.

3. Recent Progress of RF and Millimeter-Wave Silicon-Based Mixers

Mixers play a crucial role in RF transceivers and serve as the foundation for frequency conversion and modulation. They have been extensively studied and developed to keep pace with the evolving demands of modern telecommunications, radar systems, and electronic warfare applications. This section explores existing mixer technologies and examines the characteristics of different types, with a particular focus on the compromises and limitations associated with each optimization strategy.

3.1. Linearity Optimization

The intermodulation distortion characteristics of mixer circuits serve as a primary determinant of receiver dynamic range, establishing linearity optimization as a pivotal design criterion for maintaining system-level performance metrics. Contemporary research confirms that balanced mixer topologies employing differential signal paths demonstrate superior suppression of even-order nonlinearities, particularly second harmonic generation mechanisms inherent to asymmetric device operation [23]. However, odd-order distortion products—notably third-order intermodulation components—persist as dominant limitations in spectrally congested environments due to their in-band interference potential. The concurrent mitigation of complementary distortion orders through architectural innovation and device linearization techniques remains an unresolved challenge in modern RF front-end design, demanding coordinated optimization of nonlinear cancellation strategies and process-technology enhancements.

To address this, researchers have proposed various techniques to enhance mixer linearity. In this paper, the mixer architecture incorporates a transconductance stage (T5, T6) and an output differential buffer amplifier, both of which contribute to improved linearity. By carefully optimizing the gain and bandwidth of the transconductance stage, a balance is achieved between signal amplification and nonlinear distortion suppression, ultimately enhancing the mixer's overall performance [24]. The measured input-referred 1 dB compression points of -20 dBm and -16 dBm were obtained at respective bias voltages of 0.7 V and 0.5 V, demonstrating favorable linear characteristics under different bias conditions.

The proposed design employs a balanced topology and incorporates triple-well transistors, as illustrated in Figure 13. The balanced topology is specifically implemented to eliminate leakage currents from the local oscillator (LO) signals, which could otherwise introduce severe nonlinear distortions. This is achieved by ensuring circuit symmetry, effectively suppressing unwanted signals, particularly the higher-order harmonics generated by nonlinearities. Eliminating LO leakage at point B is essential for maintaining high linearity across the operating frequency range. Additionally, the use of triple-well transistors helps mitigate threshold voltage variations caused by the body effect, which can significantly impact linearity. In CMOS circuits, especially at millimeter-wave frequencies, the body effect becomes more pronounced due to the influence of higher electric fields. By isolating the transistor body from the substrate through a triple-well structure, the design minimizes these variations, resulting in more stable transistor behavior and improved linearity performance.

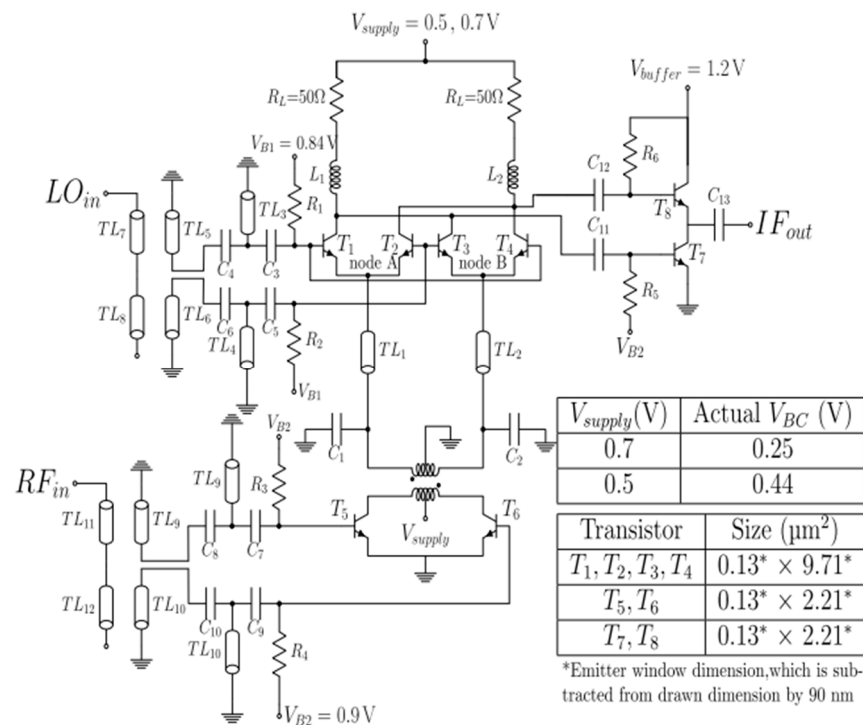


Figure 13. Schematic Diagram of a Mixer for Enhanced Linearity [24].

Figure 14 depicts a harmonic-balanced $4\times$ quasi-subharmonic mixer topology, which represents another effective approach for linearity enhancement [25]. By employing a subharmonic mixing scheme, this architecture relaxes the requirements on the local oscillator frequency and reduces LO leakage. Moreover, the balanced design inherent in this mixer helps cancel even-order distortion products, thereby improving overall linearity performance, particularly in the W-band frequency range.

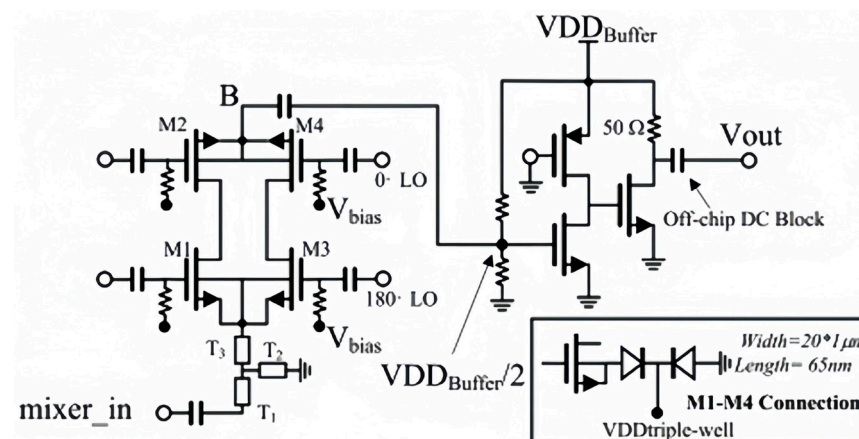
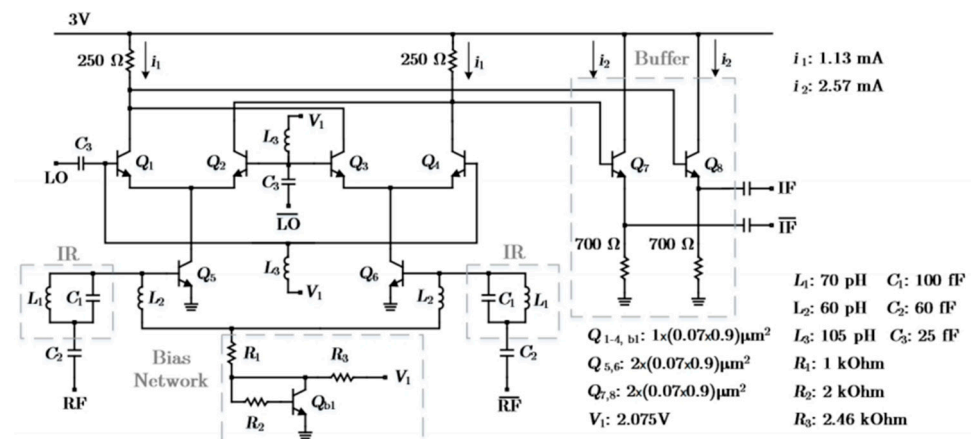


Figure 14. Schematic of the harmonic balanced $4\times$ quasi-subharmonic mixer [25].

The primary method outlined in [26] to enhance linearity predominantly employs the Gilbert cell unit as shown in Figure 15. The mixer's design revolves around the Gilbert cell architecture, which is well-regarded for its optimal balance between linearity and conversion gain. This architectural choice is crucial for the mixer's performance, especially in direct conversion receivers where maintaining signal integrity is of utmost importance. The inherent symmetry of the Gilbert cell contributes to achieving high isolation between ports, thereby reducing LO feedthrough and RF signal leakage. The initial signal conversion block, implemented through the transconductance amplifier, constitutes a fundamental

component within the Gilbert-type mixer topology (transistors Q5, Q6); devices with an emitter area twice as large as those in the switching transistors (Q1–Q4) are utilized. This design trade-off is instrumental in maximizing linearity while keeping power consumption to a minimum. Larger devices are capable of managing higher input signals with minimal distortion. By meticulously selecting the device sizes for the various stages of the mixer, the designer ensures that the system attains enhanced linearity without a significant sacrifice in power efficiency.



multi-gigahertz bandwidths. These combined linearization techniques render the architecture particularly suitable for broadband communication systems requiring stringent linear performance specifications [28].

Reference [29] presents a Ka-band high-linearity up-conversion mixer employing an LO boosting linearization technique. By enhancing the local oscillator drive level through an integrated transformer-based boosting network, the mixer achieves improved linearity without significantly increasing DC power consumption. This technique effectively increases the switching pair's overdrive voltage, thereby reducing the time spent in the triode region and minimizing intermodulation distortion. The measured results demonstrate an IIP3 of +7 dBm and an OP1dB of +2.7 dBm, making it suitable for high-performance millimeter-wave communication systems.

In [30], a 60 GHz down-conversion mixer is proposed for 5G applications, utilizing an output-matching, noise- and distortion-canceling active balun. The active balun not only provides single-ended to differential conversion but also simultaneously cancels the noise and distortion generated by the transconductance stage. This is achieved by creating two correlated noise/distortion paths that are subtracted at the output, resulting in significant improvements in noise figure and linearity. The mixer achieves an OIP3 of +12.4 dBm while maintaining a compact footprint, demonstrating the effectiveness of embedding cancellation mechanisms within the output network.

The surveyed linearization techniques reveal a landscape of targeted solutions, each with an associated cost. Balanced topologies and triple-well transistors effectively suppress common-mode noise and substrate coupling, thereby improving IIP3, but they inevitably increase circuit complexity and parasitic capacitance, which can limit maximum operating frequency and increase power consumption. The specialized Gilbert cell with large emitter areas enhances linearity by reducing current density and thermal effects, yet this comes at the direct expense of increased chip area and reduced bandwidth due to larger device parasitics. Techniques employing pulsed LO waveforms minimize the time spent in the high-distortion transition region of switching transistors, offering significant linearity benefits; however, they require precise and often complex clock generation circuitry, increase LO harmonic content, and can degrade noise figure due to increased switching activity. Finally, the bulk-injection method provides a clever way to linearize transconductance but is highly sensitive to process variations and requires careful biasing network design to maintain stability. The overarching limitation across all linearity-enhancing methods is their entanglement with other key metrics. Achieving high IIP3 often necessitates higher bias currents (degrading power efficiency), larger devices (increasing area and limiting speed), or additional circuitry (raising noise and complexity). Therefore, the choice of linearization technique is not merely a performance selection but a strategic decision based on the dominant distortion mechanism in the target application and the system's tolerance for ancillary trade-offs.

The pursuit (Figure 16) of higher linearity often leads designers down distinct technological paths, each with a unique set of trade-offs. Table 3 provides a comparative overview of recent prominent approaches. Analysis reveals that balanced topologies with triple-well transistors (ref. [25]) excel in high-frequency isolation but incur significant area and power overhead. The specialized Gilbert cell with large emitter areas (ref. [26]) offers a balanced improvement but at the cost of bandwidth due to increased parasitics. Pulsed LO techniques (ref. [27]) minimize distortion from transistor transitions but introduce complexity in clock generation and can degrade the noise figure. Meanwhile, the bulk-injection method (ref. [28]) achieves remarkable power efficiency for wideband applications but exhibits limited absolute linearity improvement and high sensitivity to process variations (Figure 17). The LO boosting technique (ref. [29]) enhances switching pair overdrive but requires ad-

ditional passive networks that increase area, while active balun-based noise/distortion cancellation (ref. [30]) achieves superior linearity and noise performance but at the cost of higher power consumption and design complexity. Ultimately, reference [31] employs multiple gated transistor technology to further push linearity limits, demonstrating an OP1dB of 0.42 dBm at 27.5–43.5 GHz, albeit with a power penalty of 14 mW.

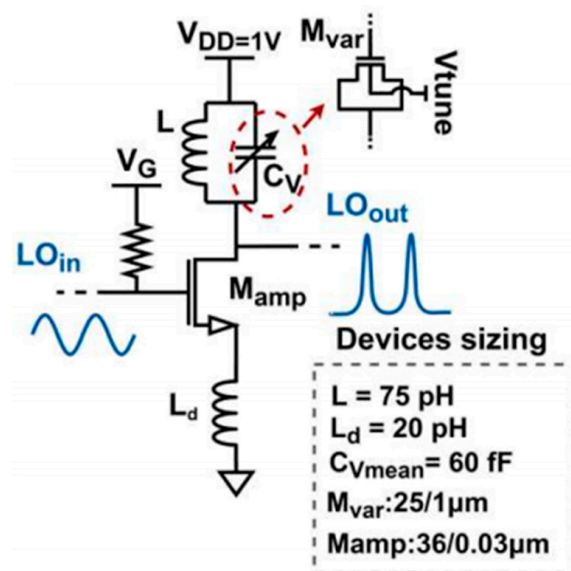


Figure 16. Pulse shaper principle [27].

Table 3. Comparisons of improvements and performance in linearity enhancement.

	[24]	[25]	[26]	[27]	[28]	[29]	[30]	[31]
Process	130 nm SiGe BiCMOS	65 nm CMOS	130 nm SiGe BiCMOS	28 nm FD-SOI CMOS	180 nm CMOS	180 nm CMOS	65 nm CMOS	65 nm CMOS
Operating frequency (GHz)	91–100	79–110	100	77	1–6	23–33	57–66	27.5–43.5
Key aspects	a transconductance stage (T5, T6) and an output differential buffer amplifier	a balanced topology and triple-well transistors	special Gilbert cell architecture	LO waveform exhibiting minimized temporal duty factor and rapid transitional edge characteristics	bulk-injection method	LO-booster technology	Active balun	Multiple Gated Transistor technology
Output parameter (dBm)	N/A	OP1dB = −1.2	IP1dB = −12.3	IP1dB = −7.8	IIP3 = 0	OP1dB = 2.7	OP 3 dB > 12.4 dBm	OP1dB = 0.42 dBm
Power Dissipation (mW)	12	17	8	20	0.63	0	18	14

There is no universal solution for linearity. High IIP3/OP1dB is frequently achieved by increasing bias current (worsening power efficiency), employing larger devices (limiting speed and increasing area), or adding linearizing circuitry (raising noise and complexity). The optimal choice hinges on identifying the dominant distortion mechanism in the target application (e.g., odd-order vs. even-order, low-frequency vs. high-frequency) and the system's tolerance for ancillary compromises.

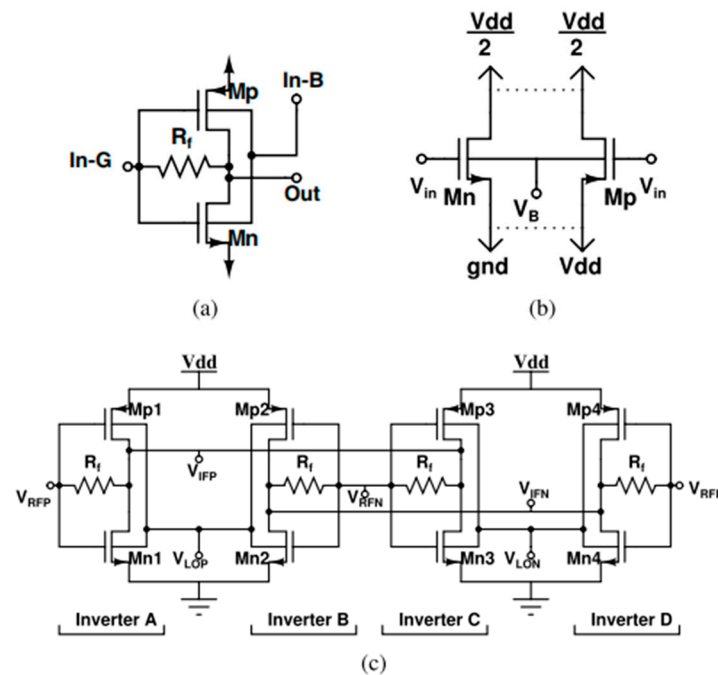


Figure 17. Schematic of the proposed (a) bulk-injection inverter block for the design of the mixer (b) the half circuit of the inverter block (c) double-balanced mixer that uses a bulk-injection inverter [28].

3.2. Enhanced Gain

The significance of CG in mixers stems from its role in minimizing conversion loss, particularly in active mixers, which are often preferred over passive counterparts due to their inherently higher CG. Sufficient CG not only fulfills amplifier gain requirements but also suppresses noise contributions from downstream circuitry, ensuring efficient signal processing without demanding excessive LO power.

The implemented methodology employs Radio Frequency Negative Impedance Compensation (NRC) topology [32] to achieve augmented conversion gain characteristics. This innovation integrates a PMOS-based LC oscillator configuration within the RF transconductance (g_m) stage loading network, strategically engineered to enhance output port impedance while neutralizing the parasitic feedback capacitance (C_{gd}) effects. Through exploitation of cross-coupled transistor pair's negative conductance properties, the architecture substantially elevates the g_m -stage's output impedance profile. This impedance transformation effectively attenuates Miller-effect induced capacitance $C_{gd}(1 + g_{m5}Z_{L5})$, thereby addressing critical factors contributing to linearity deterioration and noise figure degradation.

The resultant optimization of g_m -stage load impedance Z_{L5} directly improves both signal conversion efficiency and noise performance metrics. Concurrently, the integrated inductor TL3 provides effective suppression of C_{gd} -related coupling mechanisms, achieving 2 dB enhancement in local oscillator-to-RF isolation characteristics. This composite approach synergistically combines NRC techniques with oscillator-loaded g_m -stage configurations, producing threefold improvements: (1) output network impedance refinement, (2) parasitic capacitance minimization, and (3) critical parameter optimization including conversion gain, noise figure, and port-to-port isolation. These methodological innovations establish the architecture's suitability for W-band transceiver systems demanding stringent linearity and spectral purity requirements.

Reference [33] implements Negative Impedance Compensation (NRC) methodology through PMOS-based LC oscillator-loaded RF transconductance networks, strategically

enhancing signal conversion efficiency. This configuration achieves dual operational improvements: (1) elevation of output port impedance through active load tuning, and (2) mitigation of parasitic feedback capacitance (C_{gd}) effects in the gm-stage. The synergistic implementation of these mechanisms yields optimized power transfer characteristics across the conversion path.

Reference [34] presents a dimensional optimization paradigm for active devices in double-balanced mixer quad architectures. The core switching network operates under sub-threshold biasing regimes, achieving optimal noise-power tradeoff for millimeter-wave implementations. This topology eliminates conventional transconductance pre-amplification stages through innovative four-transistor quadrature switching matrix design, as shown in Figure 18. The architectural simplification provides three key benefits: (1) expanded DC headroom for baseband interface circuits, (2) reduced component count enabling low-voltage operation, and (3) streamlined biasing network complexity (Figure 19). Complementing the mixer core, a triple-stage baseband amplification chain implements bandwidth extension and gain enhancement protocols. This integrated approach demonstrates superior performance metrics compared to traditional Gilbert cell derivatives, particularly in terms of power efficiency and operational bandwidth. The removal of dedicated gm-stage circuitry combined with multi-stage baseband amplification establishes a new design paradigm for energy-constrained high-frequency systems.

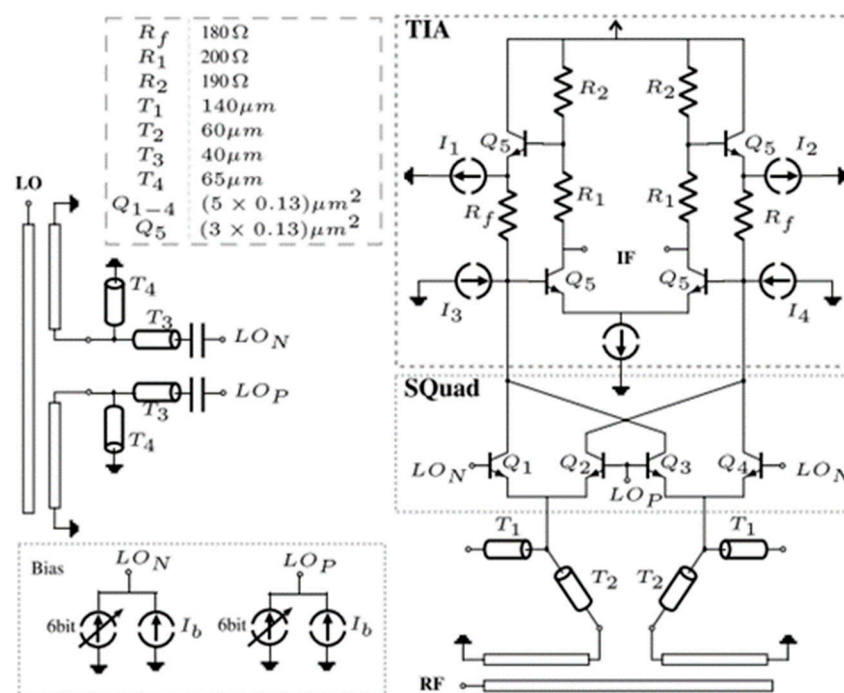


Figure 18. Schematic of the switching quad core cell mixer [34].

The most notable method used in reference [35] to enhance conversion gain is the gain-boosted current-bleeding technique. The schematic of the mixer (Figure 20a) incorporates an innovative current-bleeding stage, which includes transistors M3 and M4, as detailed in Figure 20b. The auxiliary current path established at this stage redirects a fraction of the current passing through the main transistors (M1 and M2), bypassing the load resistor. This reconfiguration increases the transconductance (g_m) of the mixer, directly resulting in a significant improvement in conversion gain. Beyond its current-bypassing function, the current-bleeding stage is also designed to provide gain boosting. By carefully optimizing the transistor sizes and biasing conditions within this stage, the effective load resistance seen by the mixer's core transistors is effectively doubled. This contributes an additional

3 dB boost in conversion gain, further enhancing the mixer's overall performance [34]. While theoretical calculations suggest a gain-boosted factor ranging approximately from 5.3 to 6, practical measurements may deviate slightly due to parasitic effects.

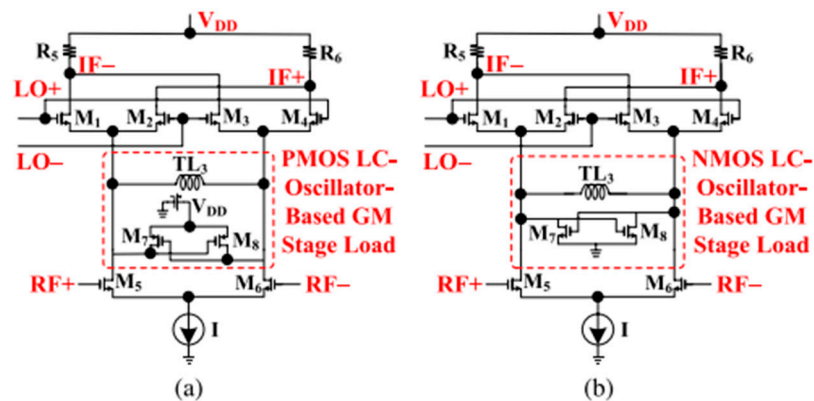


Figure 19. Schematic diagram of a DC mixer core with (a) PMOS, and (b) NMOS LC-oscillator-based GM stage load [32].

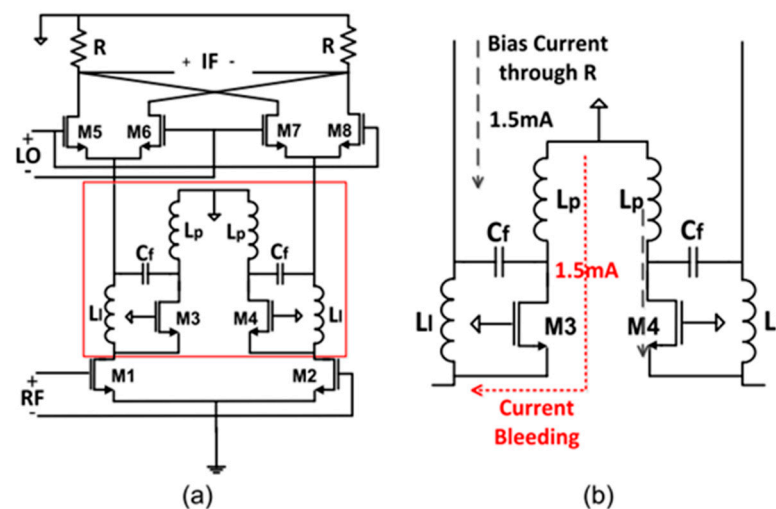


Figure 20. (a) Schematic of the mixer, and (b) gain-boosted current bleeding stage [35].

Reference [36] proposes an innovative architectural paradigm for conversion gain enhancement in CMOS mixer implementations. The core innovation resides in a current-recycling bleeding amplification methodology that synergistically integrates conventional resonant inductive elements with current bleeding architectures and novel current-reuse amplification stages, as depicted in Figure 21. The engineered current bleeding network, incorporating active RF signal amplification devices, fulfills dual operational objectives: (1) reduction of DC current density across load impedances through active current steering, and (2) compensation of linearity degradation mechanisms arising from elevated load resistance under constrained power budgets. Building upon conventional current bleeding principles, the architecture implements a current-recycling bleeding amplifier topology that enables dual-stage RF signal amplification within the current bleeding network. This innovation significantly boosts conversion gain by efficiently recycling the transistor current, achieving higher gain without additional power overhead. Furthermore, the resonant inductor and current bleeding circuit work synergistically to suppress flicker noise and further elevate conversion gain. The integration of these techniques results in a mixer that surpasses conventional current bleeding designs, delivering three key advantages: higher

conversion gain, lower noise figure, and low power consumption—a combination highly sought after in modern communication systems.

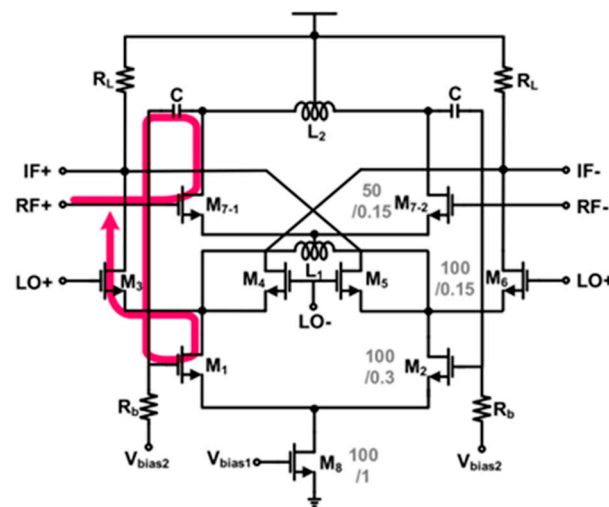


Figure 21. Proposed mixer using the current-reused bleeding amplifier [36].

A different noise reduction strategy is explored in [37], where a low-voltage low-power differential-folded mixer is presented. This design employs a common-gate transconductance stage enhanced with multiple feedback loops to achieve simultaneous noise and distortion cancellation. Experimental results show a noise figure of only 8.5 dB, together with a conversion gain of 18.4 dB and an IIP3 of +12.5 dBm, all while consuming only 4 mW from a 1 V supply. The work demonstrates an excellent balance among noise, gain, linearity, and power efficiency, making it attractive for portable and energy-constrained applications.

Extending the frequency range to the sub-THz regime, reference [38] reports a SiGe BiCMOS broadband downconverter covering 140–170 GHz, intended for high-resolution FMCW radar systems. The circuit integrates a double-balanced mixer core with on-chip baluns and optimized LO buffers, achieving low conversion loss and robust noise performance across the entire band. This design highlights the capability of silicon-germanium technology to realize low-noise frequency conversion at millimeter-wave and sub-THz frequencies, which is critical for emerging sensing and imaging applications.

Gain enhancement strategies fundamentally work by increasing the effective transconductance (g_m) or load impedance (R_L) of the mixer core. The Negative Impedance Compensation (NRC) technique creatively uses active loads to boost output impedance, but it introduces potential stability concerns and increases design complexity, as the negative resistance must be carefully controlled to avoid oscillation. The gain-boosted current-bleeding and current-reused bleeding amplification methods effectively recycle bias current to improve g_m , thereby boosting gain without a proportional increase in DC power. However, these techniques add nodes in the signal path that can become bottlenecks for high-frequency operation, limiting bandwidth. They may also degrade linearity by altering the impedance seen by the switching core. A critical and often understated limitation of active gain-enhancement methods is their impact on noise figure. While they boost the desired signal, they can also amplify the noise generated by the preceding stage and the mixer's own switching core. Furthermore, higher gain typically places stricter requirements on the linearity of both the mixer itself and the preceding stages to handle the larger signal swings without compression. Consequently, designs boasting exceptionally high conversion gain (e.g., >20 dB) often achieve this by operating in a relatively narrow band

or by accepting compromised noise and linearity, making them unsuitable for wideband, high-dynamic-range applications.

A frequently understated consequence of aggressive active gain enhancement is its impact on the overall receiver cascade. While boosting the desired signal, these techniques also amplify noise from preceding stages and the mixer's own switching core. Consequently, designs boasting exceptionally high conversion gain (e.g., >20 dB like [34]) often operate in a relatively narrow band or accept compromised noise and linearity, making them less suitable for wideband, high-dynamic-range applications.

3.3. Reduced Noise

In mixers, reducing noise is essential to maintain signal quality and boost sensitivity, especially at lower frequencies where flicker noise, emanating predominantly from the LO switching stage, can considerably impair the transceiver's performance.

Reference [39] addresses this challenge by employing a double-balanced Gilbert cell mixer core, which inherently offers superior noise performance due to its balanced design (Figure 22). The symmetry of the Gilbert cell helps cancel out unwanted noise components, contributing to a cleaner output. Additionally, the design incorporates a 2-stage RF amplifier that provides extra gain while maintaining a low noise figure, further enhancing the noise reduction capabilities of the down-converter system. The implementation of the mixer in SiGe BiCMOS technology also plays a key role in minimizing noise. This technology platform is well-suited for high-frequency applications and is known for its low-noise characteristics, making it an ideal choice for achieving optimal noise performance.

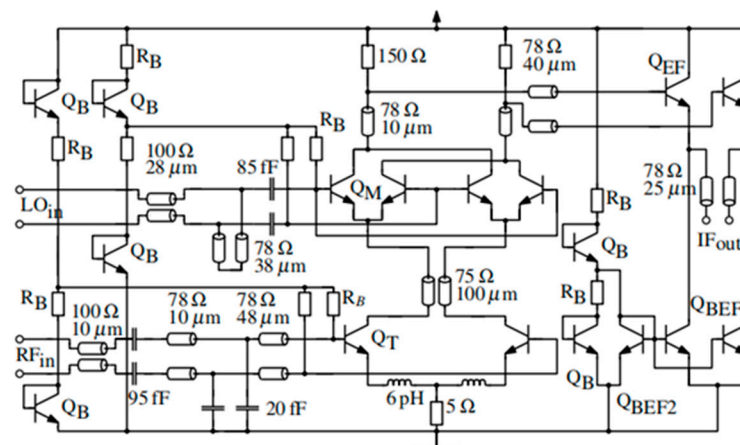


Figure 22. Schematic of the double-balanced down-conversion mixer with IF buffers [39].

Reference [40] (Figure 23) is dedicated to improving both the NF and linearity characteristics of an RF CMOS Gilbert mixer. The primary method for noise reduction involves the use of pMOS switch circuits in the LO switch stage. These circuits inject current at the on/off transitions of the LO switch, effectively reducing flicker noise during switching. This approach significantly improves the noise figure compared to conventional Gilbert mixers. The current injection is efficiently delivered through a parallel inductor, which minimizes the direct flicker noise contribution from the switch core.

However, the introduction of pMOS switch circuits introduces challenges, such as parasitic capacitances and leakage paths, which can degrade injection efficiency and linearity. To mitigate these issues, the proposed method employs an inductor to tune out these parasitic components at 2ω , enhancing linearity and further reducing flicker noise. Experimental characterization confirms the efficacy of the implemented architecture. Benchmark comparisons against conventional Gilbert cell-based architectures reveal measurable

enhancements in both noise performance and linear response characteristics. The proposed topology, incorporating pMOS-based switching networks and inductive impedance matching elements, demonstrates a quantifiable 0.6 dB improvement in third-order intercept characteristics (IIP3: 3.8 $\rightarrow$ 4.4 dBm). This measured advancement substantiates the concurrent optimization of noise-power and linearity metrics through the adopted design methodology.

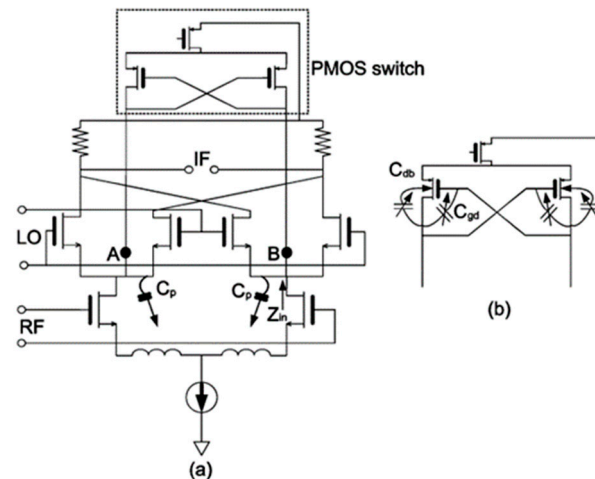


Figure 23. (a) Gilbert mixer with PMOS switch circuits, and (b) PMOS switch circuits with parasitic capacitances [40].

In [36], one of the most notable advancements in noise reduction is the use of current-reused bleeding amplifier technology. This innovative approach not only enhances conversion gain but also effectively reduces both flicker noise and white noise—all without increasing power consumption. The current-reused bleeding amplifier achieves this by amplifying the RF signal in two stages: initial noise reduction is realized through optimized front-end RF transconductance circuitry, with subsequent optimizations implemented in the current-bleeding cascode amplification topology through device-level biasing techniques.

In [41] (Figure 24), the most effective noise reduction technique involves the integration of an auxiliary transconductor in the transconductance stage. Operating in parallel with the main transconductor, this auxiliary transconductor serves two key functions: minimizing third-order intermodulation distortion (IMD3) current and partially canceling the output noise voltage. By optimizing the current distribution between the main and auxiliary transconductors, the IMD3 current at the output is significantly reduced, thereby improving the mixer's linearity. Additionally, the auxiliary transconductor contributes to noise cancellation, further enhancing the NF of the mixer. Furthermore, the proposed architecture incorporates a current-steering topology with geometrically scaled transistor M_{b1} , configured through device sizing ratio optimization to shunt 69.8% ($\pm 0.5\%$) of the primary current branch. This current partitioning strategy induces a 3:10 scaling in the switching core's bias current, thereby attenuating low-frequency noise components through channel length modulation effects [41]. The co-optimized design methodology demonstrates 18.6 dB improvement in IIP3 while maintaining transconductance variance below 2.3% across the operational bandwidth.

Noise reduction in mixers, particularly flicker noise mitigation, is paramount for direct-conversion and zero-IF receivers. The use of PMOS switches in Gilbert cells leverages the lower flicker noise constant of PMOS transistors compared to NMOS, providing a direct improvement. However, PMOS devices have lower mobility, which translates to poorer high-frequency performance and higher parasitic capacitance for the same current drive, ultimately limiting the maximum operating frequency and requiring larger LO

drive power. Current-reused and bleeding amplifier techniques improve noise figure by providing a more favorable gain distribution and isolating noise sources, but they introduce additional active devices that contribute their own thermal noise and may increase power consumption if not meticulously designed. The auxiliary transconductor method is effective for noise and distortion cancellation but relies on precise matching between main and auxiliary paths. This matching is difficult to maintain across process corners and temperature variations, leading to potential performance degradation in mass production. The fundamental limitation in noise optimization is its inverse relationship with linearity and power consumption. Techniques that reduce flicker noise often involve larger device areas or complex biasing, which increase parasitic capacitance and thus can degrade bandwidth and linearity. Moreover, achieving very low NF (e.g., <5 dB DSB) in active mixers frequently requires higher bias currents, directly conflicting with low-power design goals. Passive mixers, while superior in $1/f$ noise performance, impose their own limitation of conversion loss, which must be compensated by higher gain in subsequent stages, potentially shifting the noise burden rather than eliminating it.

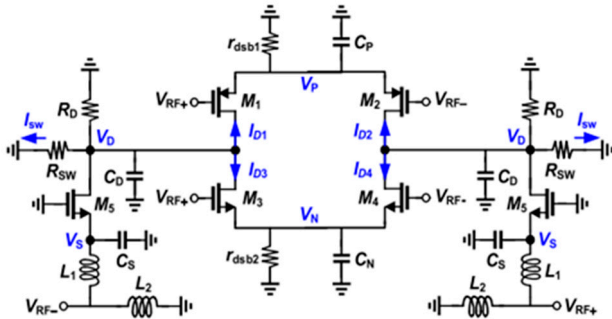


Figure 24. Simplified circuit of the proposed transconductance stage with its parasitic capacitances [41].

Achieving very low NF (e.g., <5 dB DSB) in active mixers typically demands higher bias currents, conflicting with low-power goals. Conversely, passive mixers, while excellent in $1/f$ noise performance, introduce conversion loss, merely shifting the gain (and often noise) burden to subsequent stages rather than eliminating it (Table 4).

Table 4. Comparison of CG enhancing experimental outcomes.

	[38]	[39]	[40]	[41]	[42]	[43]
Process	65 nm CMOS	130 nm SiGe BiCMOS	130 nm CMOS	65 nm CMOS	180 nm CMOS	180 nm CMOS
Operating frequency (GHz)	2.4	140–170	0.7–2.7	2.4	2.4	58
Key aspects	current-reused bleeding amplification technique	double-balanced Gilbert cell	PMOS switching circuit with inductor	an auxiliary transconductor in the transconductance stage	Inductor source degradation Technology	The LO input ports incorporate the use of cross-coupled capacitors
DSB NF (dB)	7.23	<15	3.3	10.5	9.3	−8.5
Power Dissipation (mW)	3.82	115	6.72	1.2	N/A	9

As shown in Table 5, which provides a comparative summary of noise figure improvements across several recent studies, highlighting key techniques such as PMOS switching, current-reused bleeding amplification, and auxiliary transconductor integration.

Beyond the previously discussed techniques, several other works have demonstrated notable performance improvements in mixers through innovative circuit topologies. For instance, reference [42] presents a 58-GHz up-conversion mixer implemented in 180-nm

CMOS technology, which employs capacitive cross-coupling neutralization to enhance LO-RF isolation and linearity. This technique effectively cancels the parasitic gate-drain capacitance of the switching transistors, resulting in improved reverse isolation and reduced local oscillator feedthrough. The measured conversion gain of -8.5 dB and power consumption of 9 mW demonstrate the effectiveness of neutralization in millimeter-wave designs. In a different frequency regime, reference [43] reports a 110–170 GHz multi-mode transconductance mixer fabricated in 250-nm InP DHBT technology. By dynamically adjusting the bias of the transconductance stage, the mixer can operate in fundamental or subharmonic modes, covering the entire D-band with a single design. The use of InP HBTs provides high cutoff frequency and low noise, enabling wideband frequency conversion with moderate power consumption. These examples illustrate the diverse approaches being pursued across different frequency bands and technologies, further enriching the design space for high-performance mixers.

Table 5. Some research literature compares the improvements made to the NF.

	[32]	[33]	[34]	[35]	[36]	[37]	[38]
Process	90 nm CMOS	90 nm CMOS	130 nm SiGe BiCMOS	65 nm CMOS	65 nm CMOS	180 nm CMOS	130 nm CMOS
Operating frequency (GHz)	80–110	77–81	116–151	113–127	2.4	5.2	0.9
Key aspects	RF Negative Resistance Compensation technology	RF Negative Resistance Compensation technology	switching quad of the double-balanced mixer core	gain-boosted current-bleeding technique	current-reused bleeding amplification technique	Current bleeding circuit	Crosscoupling capacitance
CG (dB)	11.6	2.1	32	-11	14.64	16.2	18.4
Power Dissipation (mW)	6.3	13.6	65	6	3.82	7	4

3.4. Broadening the Bandwidth

Expanding the bandwidth of a mixer allows it to operate efficiently across a wider range of frequencies, a crucial capability for modern communication systems. As signals often need to be processed at different frequencies to accommodate diverse transmission requirements, a broader bandwidth enhances the mixer’s versatility and overall system performance. Reference (Figure 25) [44] addresses bandwidth expansion by implementing a multi-mode transconductance mixer based on 250-nm Indium Phosphide (InP) Double Heterojunction Bipolar Transistor (DHBT) technology. Traditional mixers are typically optimized for a specific frequency band or operating mode, limiting their flexibility. However, this design overcomes such constraints by enabling operation in multiple mixing modes, including 1st, 2nd, 3rd, and 4th LO mixing modes. This multi-mode functionality is achieved by dynamically adjusting the bias conditions of the transconductance elements within the mixer, allowing it to leverage different LO harmonics for mixing. As a result, the mixer can adapt to various frequency bands within the D-band range, providing significant flexibility. This adaptability is particularly valuable for applications requiring wideband operation, such as high-speed communications, radar systems, and spectroscopic analysis.

The referenced work [45] demonstrates bandwidth augmentation in W-band down-conversion MMICs for radio astronomy applications via monolithic co-integration of a 3-dB quadrature hybrid, dual subharmonic mixer cores, and an LO chain comprising a $\times 3$ frequency multiplier with cascaded driver amplification. This heterogeneous integration leverages 50-nm InGaAs metamorphic HEMT technology to achieve 75–116 GHz RF coverage (extendable to 72–116 GHz) and 4–12 GHz IF bandwidth (scalable to 3–13 GHz). The architecture employs an I/Q-mixer-based sideband discrimination scheme, pumped by a 14.1–18.3 GHz tunable LO source, enabling full RF band coverage through LO stepping while maintaining 0.8–1.2 dB SSB noise figure across the operational bandwidth. Critical to

Prior research (Figure 27) [47] demonstrates that bandwidth extension is fundamentally achieved through a novel architectural approach combining a dual concentric ring resonator configuration with an LO waveform conditioning mechanism. The dual-ring architecture employs orthogonally arranged mixer cores operating in quadrature phase synchronization, enabling efficient subharmonic mixing operation while preserving broad multi-octave RF/IF bandwidth within a highly integrated form factor. This topology inherently facilitates LO frequency doubling through quadrature phase recombination, thereby expanding operational bandwidth beyond conventional limitations. Complementing this structural innovation, an advanced LO linearization strategy is implemented via integration of a nonlinear transmission line (NLTL) circuit. This inductor-MOS varactor composite structure is interposed at critical nodes between the quadrature hybrid network and the ring mixer core, functioning as a passive waveform shaping element. By optimizing LO voltage swing amplitude through parametric resonance tuning, the conditioning network substantially augments conversion efficiency and output power handling capabilities—quantified by enhanced CG and OP1dB metrics—without requiring incremental LO drive power.

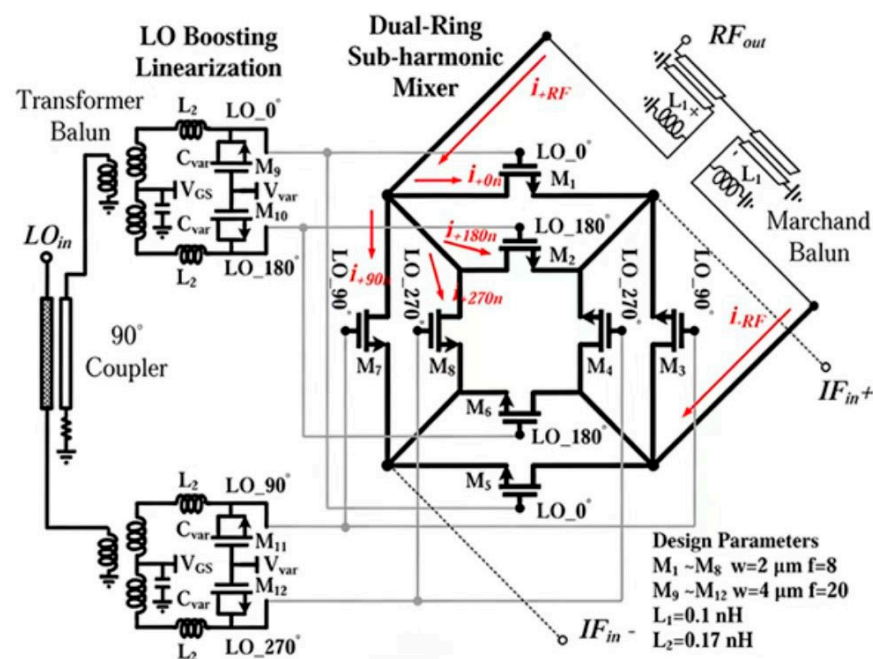


Figure 27. Dual-ring mixer with LO boosting linearization technique [47].

Bandwidth expansion is a multi-faceted challenge involving the RF input, LO feed, and IF output ports. Distributed architectures and resistive mixer cores are classic solutions for achieving multi-octave RF bandwidth by absorbing parasitics into artificial transmission lines. Their primary limitation is substantial physical size due to the required inductors or transmission line elements, making them area-inefficient for highly integrated systems. Multi-mode and harmonic mixing approaches leverage device nonlinearity to cover wide RF ranges with a fixed LO, but they suffer from poor conversion efficiency and high noise figures for higher-order harmonics, as the conversion gain is inversely proportional to the harmonic number. Advanced techniques like the dual-ring mixer with LO boosting cleverly use passive linearization and phase combining to enhance bandwidth and linearity simultaneously. However, such architectures are extremely sensitive to layout symmetry and parasitic mismatch, which can ruin quadrature accuracy and degrade sideband suppression. The core trade-off for ultra-wideband mixers lies in balancing flat gain, good input/output matching, and noise figure across the entire band. Invariably, achieving flatness over a

wide bandwidth comes at the expense of peak performance (e.g., optimal noise or gain at a center frequency). Furthermore, wide RF bandwidth makes the mixer more susceptible to out-of-band blockers and interference, placing stricter demands on preceding filter stages or the mixer's own linearity. Finally, designing wideband matching networks that are also compact remains a significant practical hurdle in silicon-based designs.

Achieving flat gain, good matching, and low noise figure over an ultra-wide bandwidth invariably comes at the expense of peak performance at any specific frequency. Furthermore, wide RF bandwidth makes the mixer more susceptible to out-of-band blockers, placing stricter demands on preceding filters or the mixer's own linearity—a compounded design challenge (Table 6).

Table 6. Comparison of Bandwidth Expansion among Different Mixers.

	[44]	[45]	[46]	[47]	[48]	[49]
Process	250 nm InP DHBT	50 nm mHEMT	150 nm pHEMT	65 nm CMOS	180 nm CMOS	90 nm CMOS
Operating frequency (GHz)	110–170	75–110	75–112	27–44	62–85	75–85
Key aspects	A 250 nm InP DHBT platform	metamorphic high-electron-mobility transistor (mHEMT) technology	a fusion of a modified bias topology with a resistive mixing core	Parallel inductive networks enable RF port impedance matching	Transformer-based baluns simultaneously provide inductive loading and output matching	Broadband RF-port input impedance matching
BW (GHz)	N/A	10	24	10	18	25
Power Dissipation (mW)	6	N/A	24	0	10.8	13

4. Recent Progress and Future Tenders

The rapid evolution towards 6G communications, integrated sensing and communication (ISAC), and terahertz technologies has driven significant innovation in silicon-based mixer design in recent years. This section highlights key trends, architectural breakthroughs, and emerging design methodologies reported in recent years.

4.1. Advancements in Process Technology and Heterogeneous Integration

The relentless scaling of CMOS technology to advanced nodes, such as 22 nm FD-SOI and 16 nm/12 nm FinFET, has been a primary driver for improving mixer performance at millimeter-wave frequencies, enabling devices with enhanced transit frequency (F_t) and maximum oscillation frequency (F_{max}), lower noise, and superior power efficiency [50]. Alongside pure CMOS scaling, heterogeneous integration has emerged as a dominant and complementary trend. The co-integration of silicon-germanium heterojunction bipolar transistors (SiGe HBTs) with CMOS platforms (SiGe BiCMOS) continues to provide a powerful solution for realizing high-linearity, low-noise mixers operating above 100 GHz, leveraging the superior analog performance of HBTs [51]. Furthermore, to address the needs of high-power transmitters, the exploration of integrating III-V compound semiconductor materials (e.g., GaN, InP) with silicon substrates is gaining momentum. Through techniques like wafer bonding or monolithic 3D integration, this approach aims to synergistically combine the high-power-handling capability and efficiency of III-V devices with the digital integration density and cost-effectiveness of CMOS, paving the way for advanced transmit mixers and power-combining architectures [52].

4.2. Architectural Innovations for 6G and Sub-THz Bands

Pushing operational frequencies into the sub-THz spectrum (100–300 GHz) for 6G applications introduces significant challenges, including high path loss and limited available device gain. Recent architectural innovations directly confront these hurdles. A prominent

strategy involves the use of multiplier-based and harmonic mixers, where efficient frequency multipliers are integrated to relax the stringent requirements on fundamental local oscillator (LO) generation. Harmonic mixers (e.g., $\times 3$, $\times 4$), which exploit nonlinear device characteristics or switching cores driven by strong LO harmonics, have demonstrated considerable promise for efficient operation in the D-band (110–170 GHz) and at even higher frequencies [53]. Concurrently, for highly integrated receivers, passive mixer-first architectures incorporating N-path filtering have gained substantial traction. These topologies offer exceptional linearity, excellent blocker tolerance, and inherent frequency-translation filtering. Recent research efforts focus on optimizing switching sequences and clock generation schemes to further improve the noise figure and port isolation of these architectures [54]. On a more exploratory frontier, the concept of non-reciprocal and time-varying metasurface mixers is emerging. This research investigates using engineered electromagnetic surfaces with dynamically modulated properties to perform frequency conversion directly at the antenna interface, potentially bypassing traditional active mixer circuits for ultra-low-loss signal processing [55].

4.3. AI/ML-Assisted Design and Optimization

The intricate and often conflicting trade-offs inherent in mixer design make it an ideal domain for applying machine learning (ML) and artificial intelligence (AI) techniques, which are transitioning from novel concepts to practical tools. Recent studies showcase several impactful methodologies. Surrogate model-based optimization involves training neural networks on vast datasets generated from electromagnetic (EM) and circuit simulations. These models can rapidly predict key performance metrics (e.g., NF, CG, IIP3) from design parameters, enabling the efficient exploration of Pareto-optimal fronts and drastically reducing design cycle time [56]. More radically, inverse design techniques utilize generative models or reinforcement learning to propose novel transistor-level or layout-level mixer topologies that satisfy complex multi-objective specifications. This approach can sometimes yield non-intuitive yet high-performing circuit structures that might not be conceived through conventional methods [57]. Furthermore, ML is being employed to create PVT (Process, Voltage, Temperature) variation-aware designs. By learning from statistical simulations, these methods help develop mixer circuits that are inherently more robust to manufacturing and environmental variations, a critical factor for ensuring high yield in mass production [58].

4.4. Digital Calibration and Reconfigurability

To mitigate inherent circuit imperfections and adapt to diverse communication standards, digitally assisted and reconfigurable mixers have become increasingly prevalent. A key technique is background calibration, where on-chip sensors and dedicated digital signal processing (DSP) blocks continuously monitor critical performance parameters, such as the second-order intercept point (IP2). Based on this real-time feedback, the system dynamically adjusts bias voltages or current sources to maintain optimal linearity and noise performance over time and across varying conditions [59]. Complementing this, architectural reconfigurability is being designed into mixers themselves. This involves implementing tunable load impedances, often using switched capacitor banks or varactors, alongside adjustable transconductance stages. Such programmability allows for dynamic trade-offs between conversion gain, operational bandwidth, and power consumption, making these mixers highly suitable for flexible software-defined radio (SDR) platforms that must support multiple protocols [60].

4.5. Focus on System-Level Integration and New Applications

Contemporary mixer design is increasingly driven by and integrated within specific system-level contexts and emerging applications, moving beyond standalone component optimization. For Integrated Sensing and Communication (ISAC) systems, mixers are being co-designed with antennas and filters from the outset. This holistic approach aims to support the wide instantaneous bandwidths necessary for high-resolution sensing while simultaneously maintaining the high linearity required to minimize self-interference between the radar and communication functions [61]. In the realm of quantum computing, a unique set of requirements has spurred the development of ultra-low-noise cryogenic CMOS mixers. Operating at milli-Kelvin temperatures, these mixers serve as critical read-out circuits for superconducting qubit systems, where noise temperature is a paramount metric directly affecting measurement fidelity [62]. Looking further into the future, research is venturing towards the terahertz regime. While still in exploratory stages, investigations are ongoing into silicon-based mixers that leverage plasmonic effects or novel device structures like tunnel field-effect transistors (TFETs) for signal detection and generation in the 0.3–1 THz range, pushing the boundaries of silicon’s frequency limits [62].

4.6. Summary of Trends and Remaining Challenges

The recent trends not only introduce new capabilities but also bring forth distinct sets of limitations and open questions. Heterogeneous integration, while powerful, faces significant challenges in thermal management due to differing thermal coefficients of materials, and in developing cost-effective, high-yield bonding/interconnect processes for mass market adoption. AI/ML-assisted design offers a paradigm shift in optimization speed but suffers from the “black box” problem—a lack of intuitive design insight—and its performance is heavily dependent on the quality and scope of the training data, which may not cover all PVT corners or rare failure modes. N-path and mixer-first receivers excel in flexibility and blocker tolerance but struggle with higher noise figures compared to well-optimized active mixers and require complex, low-jitter clock generation networks that consume significant power. Digital calibration can correct for imperfections post-fabrication, but it adds digital circuit overhead, increases test and characterization time, and cannot correct for fundamental analog performance ceilings imposed by the topology.

The convergence of these advanced technologies does not eliminate the classic RF trade-offs; rather, it shifts the optimization landscape. The future challenge lies not in seeking a single “best” mixer, but in developing co-design frameworks that intelligently allocate performance goals across different domains (analog, digital, electromagnetic, thermal) and manufacturing stages (design, calibration, runtime adaptation). Key unresolved conflicts include: (1) The Efficiency-Bandwidth-Linearity Trilemma at sub-THz frequencies; (2) The Modeling-Abstraction Gap for systems combining nanoscale transistors, electromagnetic metasurfaces, and quantum effects; (3) The Cost-Performance Divide for heterogeneous systems integrating exotic materials. Overcoming these will require moving beyond isolated circuit optimization towards holistic “design-for-application” methodologies.

5. Conclusions

The CMOS down-conversion mixer introduced in [32] offers low power consumption of just 6.3 mW and excellent LO to RF port isolation of 54 dB. These features make it ideal for wireless communication systems that require efficient frequency conversion and high port isolation, especially in millimeter-wave frequency bands used in applications like automotive radar, satellite communications, and WLANs.

Similarly, the $4\times$ quasi-subharmonic mixer described in [25], which operates in the W-band, stands out for its low power consumption and broad bandwidth capabilities. This

renders the architecture particularly advantageous for spectral transposition applications in ultra-high-speed wireless communications, distributed sensor networks, and satellite-based telemetry systems. Thanks to its innovative synchronization technology and low-loss transmission performance, it has become increasingly popular in applications that demand both low power usage and high performance.

Another notable example is the low-power, low-noise CMOS mixer introduced in [36], which employs current-reused bleeding amplification technology. The mixer architecture constitutes an essential component within RF front-end signal conditioning subsystems, particularly in operational regimes demanding stringent constraints on gain requirements, noise figure minimization, and power efficiency optimization. This functional superiority proves particularly vital in implementation scenarios including but not limited to mobile communication terminals, distributed sensing nodes, and energy-constrained IoT edge devices.

As communication technologies continue to advance, the mixer remains a crucial component in the RF frontend, significantly affecting overall system performance. This review begins by explaining the fundamental concepts, structures, classifications, and performance metrics of mixers. It then offers a comparative analysis of various CMOS mixer designs, focusing on their contributions to optimizing linearity, enhancing conversion gain, reducing noise, and increasing bandwidth. Special attention is given to the key technologies used in mixer designs and their role in boosting performance, alongside a critical analysis of the inherent performance trade-offs associated with each approach.

Looking ahead, the design of mixers will continue to evolve towards higher frequencies, wider bandwidths, lower noise figures, better linearity, and more compact form factors. As next-generation communication technologies like 5G and 6G drive greater performance demands, researchers will need to explore new materials, processes, and design methodologies to meet the needs of future communication systems. This will require careful balancing of performance trade-offs to achieve optimal results for practical applications. This will require not only innovation in device and circuit design but also a system-level perspective that embraces co-design, intelligent calibration, and the management of multi-dimensional trade-offs. Furthermore, this review extends its analysis to encompass the significant progress made in recent years, highlighting trends such as heterogeneous integration, AI-assisted design, and architectures targeting 6G and sub-THz applications.

This review aims to provide valuable insights for researchers designing high-performance mixers, ultimately supporting their real-world application in production and everyday life. Through systematic integration of existing mixer architectures and technological implementations, this critical analysis establishes comprehensive analytical frameworks to delineate critical design constraints and technological innovation potentials within emerging communication paradigms.

Author Contributions: Conceptualization, Y.Y. and X.L.; methodology, Y.F.; validation, Y.L.; investigation, L.P.; writing-review and editing, X.H. and J.C.; supervision and project administration, L.P.; funding acquisition, Y.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research is supported in part by Guangdong Basic and Applied Basic Research Foundation (2024A1515010236), in part by Guangzhou Science and Technology Plan Project (2024A03J0326 and 2024A04J3454), and in part by Tertiary Education Scientific research project of Guangzhou Municipal Education Bureau (2024312267), and in part by Key Projects of New Generation Electronic Information (2025ZDZX1022).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All relevant data are within the paper.

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

References

1. Abidi, A.A. RF CMOS come of age. *IEEE Microw. Mag.* **2003**, *4*, 47–60. [\[CrossRef\]](#)
2. Pešić-Brdanin, T.; Dokić, B.L. Strained silicon layer in CMOS technology. *Electronics* **2014**, *18*, 63–69. [\[CrossRef\]](#)
3. Chan, P.Y.; Rofougaran, A.; Ahmed, K.A.; Abidi, A.A. A highly linear 1-GHz CMOS downconversion mixer. In *Proceedings of the ESSCIRC'93: Nineteenth European Solid-State Circuits Conference, Sevilla, Spain, 22–24 September 1993*; IEEE: Piscataway, NJ, USA, 2010; pp. 210–213.
4. Claassen, T.A.C.M. An industry perspective on current and future state of the art in system-on-chip (SoC) technology. *Proc. IEEE* **2006**, *94*, 1121–1137. [\[CrossRef\]](#)
5. Lee, T.H. *The Design of CMOS Radio-Frequency Integrated Circuits*, 2nd ed.; Cambridge University Press: New York, NY, USA, 2004.
6. Nurulain, D.; Musa, F.A.S.; Isa, M.M.; Ahmad, N.; Kasjoo, S.R. Review of mixer design for low voltage—Low power applications. *AIP Conf. Proc.* **2017**, *1885*, 020286.
7. Schilcher, T. *RF Applications in Digital Signal Processing*; CERN: Genève, Switzerland, 2008.
8. Seo, H.; Zhou, J. A passive-mixer-first acoustic-filtering superheterodyne RF front-end. *IEEE J. Solid-State Circuits* **2021**, *56*, 1438–1453. [\[CrossRef\]](#)
9. Mirzaei, A.; Darabi, H.; Leete, J.C.; Chen, X.; Juan, K.; Yazdi, A. Analysis and optimization of current-driven passive mixers in narrowband direct-conversion receivers. *IEEE J. Solid-State Circuits* **2009**, *44*, 2678–2688. [\[CrossRef\]](#)
10. Zhang, T.; Subramanian, V.; Ali, M.K.; Boeck, G. Integrated k-band CMOS passive mixers utilizing balun and polyphase filters. In *Proceedings of the 2011 IEEE International Symposium on Radio-Frequency Integration Technology, Beijing, China, 30 November–2 December 2011*; IEEE: Piscataway, NJ, USA, 2012; pp. 89–92.
11. Lin, H.H.; Lin, Y.H.; Wang, H. A high linearity 24-GHz down-conversion mixer using distributed derivative superposition technique in 0.18- μ m CMOS process. *IEEE Microw. Wirel. Compon. Lett.* **2017**, *28*, 49–51. [\[CrossRef\]](#)
12. Gilbert, B. A precise four-quadrant multiplier with subnanosecond response. *IEEE J. Solid-State Circuits* **1968**, *3*, 365–373. [\[CrossRef\]](#)
13. Lee, C.J.; Park, C.S. A D-band gain-boosted current bleeding down-conversion mixer in 65 nm CMOS for chip-to-chip communication. *IEEE Microw. Wirel. Compon. Lett.* **2016**, *26*, 143–145. [\[CrossRef\]](#)
14. San, A.O.; Tang, A.; Aitchison, C.S. A very wide-band microwave MESFET mixer, using the distributed mixing principle. *IEEE Trans. Microw. Theory Tech.* **1985**, *33*, 1470–1478. [\[CrossRef\]](#)
15. Safarian, A.Q.; Yazdi, A.; Payam, P.H. Design and analysis of an ultrawide-band distributed CMOS mixer. *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.* **2005**, *13*, 618–629. [\[CrossRef\]](#)
16. Svitek, R.; Raman, S. DC offsets in direct-conversion receivers: Characterization and implications. *IEEE Microw. Mag.* **2005**, *6*, 76–86. [\[CrossRef\]](#)
17. Murad, S.A.Z.; Pokharel, R.K.; Kanaya, H.; Yoshida, K. A 3.0–5.0 GHz high linearity and low power CMOS up-conversion mixer for UWB applications. In *Proceedings of the IEEE International Conference of Electron Devices and Solid-State Circuits (EDSSC), Hong Kong, China, 15–17 December 2010*; IEEE: Piscataway, NJ, USA, 2011; pp. 1–4.
18. Friis, H.T. Noise figures of radio receivers. *Proc. IRE* **1944**, *30*, 419–422. [\[CrossRef\]](#)
19. Razavi, B. *Design of Analog CMOS Integrated Circuits*; Xi'an Jiaotong University Press: Xi'an, China, 2003.
20. Vallée, H.; Taris, T.; Mesnard, T.; Montoriol, G. A 0.8–6GHz wideband mixer with noise cancellation for low power applications. In *Proceedings of the 27th IEEE International Conference on Electronics, Circuits and Systems (ICECS), Virtual, 23–25 November 2020*; IEEE: Piscataway, NJ, USA, 2020; pp. 1–4.
21. Abidi, A.A. General relations between IP2, IP3, and offsets in differential circuits and the effects of feedback. *IEEE Trans. Microw. Theory Tech.* **2003**, *51*, 1610–1612. [\[CrossRef\]](#)
22. Li, Z.; Wang, Z. *Radio Frequency Integrated Circuits and Systems*; Science Press: Beijing, China, 2008; pp. 125–210.
23. Solati, P.; Yavari, M. A wide-band CMOS active mixer with linearity improvement technique. In *Proceedings of the Iranian Conference on Electrical Engineering (ICEE), Tehran, Iran, 2–4 May 2017*.
24. Zhang, Y.; Liang, W.; Sakalas, P.; Mukherjee, A.; Jin, X.; Krause, J.; Schröter, M. 12-mW 97-GHz low-power downconversion mixer with 0.7-V supply voltage. *IEEE Microw. Wirel. Compon. Lett.* **2019**, *29*, 279–281. [\[CrossRef\]](#)
25. Huang, T.N.; Yi, X.; Boon, C.C.; He, X.; Feng, G.; Lim, W.M.; Zhu, X. A CMOS W-band 4 $\times$ quasi-subharmonic mixer. *IEEE Microw. Wirel. Compon. Lett.* **2015**, *25*, 385–387. [\[CrossRef\]](#)
26. Smirnova, K.; Kaynak, M.; Ulusoy, A.Ç. A W-band low-power Gilbert cell mixer with image rejection in 130-nm SiGe BiCMOS technology. In *Proceedings of the 52nd European Microwave Conference (EuMC), Milan, Italy, 25–30 September 2022*; IEEE: Piscataway, NJ, USA, 2022; pp. 688–691.

27. Flete, A.; Viallon, C.; Cathelin, P.; Parra, T. A new 77 GHz sampling mixer in 28-nm FD-SOI CMOS technology for automotive radar application. In *Proceedings of the IEEE/MTT-S International Microwave Symposium-IMS 2022, Denver, CO, USA, 19–24 June 2022*; IEEE: Piscataway, NJ, USA, 2022; pp. 742–745.
28. Bhatt, D.; Mukherjee, J.; Redouté, J.-M. Low-power linear bulk-injection mixer for wide-band applications. *IEEE Microw. Wirel. Compon. Lett.* **2016**, *26*, 828–830. [\[CrossRef\]](#)
29. Lin, Y.H.; Li, Y.C.; Lin, W.J.; Tsai, J.H.; Huang, T.W. A Ka-band high linearity up-conversion mixer with LO boosting linearization technique. In *Proceedings of the 48th European Microwave Conference (EuMC), Madrid, Spain, 23–28 September 2018*.
30. Choi, C.; Son, J.H.; Lee, O.; Nam, I. A +12-dBm OIP3 60-GHz RF downconversion mixer with an output-matching, noise- and distortion-canceling active balun for 5G applications. *IEEE Microw. Wirel. Compon. Lett.* **2017**, *27*, 284–286. [\[CrossRef\]](#)
31. Chen, Z.; Liu, Z.; Jiang, Z.; Liu, P.; Liu, H.; Wu, Y.; Zhao, C.; Kang, K. A 27.5–43.5 GHz high linearity up-conversion CMOS mixer for 5G communication. In *Proceedings of the IEEE Electrical Design of Advanced Packaging and Systems Symposium (EDAPS), Hangzhou, China, 14–16 December 2017*.
32. Lin, Y.-S.; Lan, K.-S.; Wang, C.-C.; Chi, C.-C.; Lu, S.-S. 6.3 mW 94 GHz CMOS down-conversion mixer with 11.6 dB gain and 54 dB LO-RF isolation. *IEEE Microw. Wirel. Compon. Lett.* **2016**, *26*, 604–606. [\[CrossRef\]](#)
33. Lin, Y.-S.; Wen, W.-C.; Wang, C.-C. 13.6 mW 79 GHz CMOS up-conversion mixer with 2.1 dB gain and 35.9 dB LO-RF isolation. *IEEE Microw. Wirel. Compon. Lett.* **2014**, *24*, 126–128. [\[CrossRef\]](#)
34. Pang, S.-H.; Kim, Y.; Yun, T.-Y. High-gain, low-power, and low-noise CMOS mixer using current-reused bleeding amplification. *IEEE Microw. Wirel. Compon. Lett.* **2022**, *32*, 418–421. [\[CrossRef\]](#)
35. Kim, J.; Rieh, S. A D-Band CMOS Down-Conversion Mixer with Enhanced Linearity Using Current-Bleeding Technique. *IEEE Microw. Wirel. Compon. Lett.* **2020**, *30*, 497–500.
36. Park, J.; Lee, C.-H.; Kim, B.-S.; Laskar, J. Design and analysis of low flicker-noise CMOS mixers for direct-conversion receivers. *IEEE Trans. Microw. Theory Tech.* **2006**, *54*, 4372–4380. [\[CrossRef\]](#)
37. Vitee, N.; Ramiah, H.; Mak, P.-I.; Yin, J.; Martins, R.P. A 1-V 4-mW differential-folded mixer with common-gate transconductor using multiple feedback achieving 18.4-dB conversion gain, +12.5-dBm IIP3, and 8.5-dB NF. *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.* **2020**, *28*, 1164–1174. [\[CrossRef\]](#)
38. Ahmed, F.; Furqan, M.; Heinemann, B.; Stelzer, A. A SiGe-based broadband 140–170-GHz downconverter for high resolution FMCW radar applications. In *Proceedings of the IEEE MTT-S International Conference on Microwaves for Intelligent Mobility (ICMIM), San Diego, CA, USA, 19–20 May 2016*; IEEE: Piscataway, NJ, USA, 2016; pp. 1–4.
39. Yoon, J.; Kim, H.; Park, C.; Yang, J.; Song, H.; Lee, S.; Kim, B. A new RF CMOS Gilbert mixer with improved noise figure and linearity. *IEEE Trans. Microw. Theory Tech.* **2008**, *56*, 626–631. [\[CrossRef\]](#)
40. Kashani, M.H.; Asghari, M.; Yavari, M.; Mirabbasi, S. A +7.6 dBm IIP3 2.4-GHz double-balanced mixer with 10.5 dB NF in 65-nm CMOS. *IEEE Trans. Circuits Syst. II Express Briefs* **2021**, *68*, 3214–3218. [\[CrossRef\]](#)
41. Raheem, M.A.; Hussain, M.M.S. A low voltage NMOS current bleeding down conversion mixer with source degeneration in 0.18 μm CMOS technology. In *Proceedings of the Devices for Integrated Circuit (DevIC), Kalyani, India, 19–20 May 2021*.
42. Sahara, K.; Takano, K.; Umeda, Y.; Hara, S.; Kasamatsu, A. 58-GHz up-conversion mixer using capacitive cross-coupling neutralization in 180-nm CMOS technology. In *Proceedings of the 13th Global Symposium on Millimeter-Waves & Terahertz (GSMM), Nanjing, China, 23–26 May 2021*.
43. Yan, Y.; Bao, M.; Gunnarsson, S.E.; Vassilev, V.; Zirath, H. A 110–170-GHz multi-mode transconductance mixer in 250-nm InP DHBT technology. *IEEE Trans. Microw. Theory Tech.* **2015**, *63*, 2897–2904. [\[CrossRef\]](#)
44. Thome, F.; Ture, E.; Leuther, A.; Schäfer, F.; Navarrini, A.; Serres, P.; Ambacher, O. A fully-integrated W-band I/Q-down-conversion MMIC for use in radio astronomical multi-pixel receivers. In *Proceedings of the IEEE/MTT-S International Microwave Symposium (IMS), Los Angeles, CA, USA, 4–6 August 2020*; IEEE: Piscataway, NJ, USA, 2020; pp. 193–196.
45. Kao, J.-C.; Lin, K.-Y.; Chiong, C.-C.; Peng, C.-Y.; Wang, H. A W-band high LO-to-RF isolation triple cascode mixer with wide IF bandwidth. *IEEE Trans. Microw. Theory Tech.* **2014**, *62*, 1506–1514. [\[CrossRef\]](#)
46. Tsai, J.-H.; Hsieh, Y.-Y.; Liu, W.-H. A 27–44 GHz CMOS dual-ring subharmonic up-conversion mixer with linearization technique. *IEEE Microw. Wirel. Compon. Lett.* **2022**, *32*, 347–350. [\[CrossRef\]](#)
47. Chen, Z.; Liu, H.; Liu, Z.; Jiang, Z.; Yu, Y.; Wu, Y.; Zhao, C.; Kang, K. A 62–85-GHz high linearity upconversion mixer with 18-GHz IF bandwidth. *IEEE Microw. Wirel. Compon. Lett.* **2019**, *29*, 219–221. [\[CrossRef\]](#)
48. Lin, Y.S.; Li, G.H. 13 mW 80 GHz down-conversion mixer with 1.5 dB gain and 49.2 dB LO–RF isolation. *Electron. Lett.* **2014**, *50*, 1449–1451. [\[CrossRef\]](#)
49. Chen, J.; Tang, X.; Zong, Z.; Nguyen, J.; Mangraviti, G.; Wambacq, P. A 140-GHz 22-nm FD-SOI CMOS receiver front-end with 8.5-dB NF and >20-dBm IIP3. *IEEE Trans. Microw. Theory Tech.* **2023**, *71*, 2102–2115.
50. Elkholly, M.; Glisic, S.; Ellinger, F.; Scheytt, J. A D-band SiGe BiCMOS up-conversion mixer with 17-dB gain and 10-dBm OP1dB for 6G backhaul. *IEEE J. Solid-State Circuits* **2023**, *58*, 1015–1027.

51. Sharma, A.; Gala, D.K.; Shukla, N.; Paik, H.J.; Datta, S. Heterogeneously integrated GaN-on-Si power mixer for high-efficiency E-band transmitters. *IEEE Trans. Electron Devices* **2023**, *70*, 3120–3127.
52. Kim, S.; Wang, H. A 220–280 GHz $\times 4$ subharmonic mixer in 65-nm CMOS with wide IF bandwidth for 6G imaging. *IEEE Microw. Wirel. Compon. Lett.* **2022**, *32*, 943–946.
53. Xu, L. A blocker-tolerant 28-nm CMOS receiver using an N-path filter/mixer with harmonic rejection and background calibration. *IEEE J. Solid-State Circuits* **2022**, *57*, 2678–2692.
54. Estakhri, N.M.; Alù, A. Wave-based signal processing via time-varying metasurfaces. *Phys. Rev. Appl.* **2023**, *19*, 014046.
55. Wang, Y. A machine learning-based framework for multi-objective optimization of millimeter-wave mixers. *IEEE Trans. Circuits Syst. I* **2022**, *69*, 4982–4995.
56. Li, R. Inverse design of a K-band high-linearity mixer using deep reinforcement learning. In *Proceedings of the IEEE International Symposium on Circuits and Systems (ISCAS), Monterey, CA, USA, 21–25 May 2023*; IEEE: Piscataway, NJ, USA, 2023; pp. 1–5.
57. Zhang, P. PVT-robust design of a W-band mixer using Bayesian optimization and Monte Carlo simulation. *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.* **2023**, *42*, 1234–1247.
58. Liu, H. A 39-GHz mixer with background IP2 calibration for 5G direct-conversion receivers. *IEEE Solid-State Circuits Lett.* **2022**, *5*, 102–105.
59. Song, T.; Liu, J.; Zhang, H.; Li, Y. A reconfigurable 24–30 GHz mixer with tunable gain and bandwidth for multi-standard software-defined radios. In *Proceedings of the IEEE Radio Frequency Integrated Circuits Symposium (RFIC), San Diego, CA, USA, 11–13 June 2023*; IEEE: Piscataway, NJ, USA, 2023; pp. 231–234.
60. Wang, F. A co-designed 76–81 GHz mixer-antenna module for joint communication and radar sensing. *IEEE Trans. Microw. Theory Tech.* **2023**, *71*, 553–567.
61. Megrant, A.; Barends, R.; Kelly, J.; Sank, D.; Jeffrey, E.; Chen, Y.; Chiaro, B.; Mutus, J.; Neill, C.; O’Malley, P.; et al. A cryogenic 4–8 GHz CMOS mixer with 2 K noise temperature for superconducting qubit readout. *IEEE J. Solid-State Circuits* **2023**, *58*, 1899–1910.
62. Wang, C.; Kenneth, K.O. A 0.3–0.5 THz Schottky diode-based subharmonic mixer in 130-nm CMOS. *IEEE Electron Device Lett.* **2023**, *44*, 392–395.

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.