Sciencespress is a publisher of peer-reviewed scientific journals, established in 2018 with a mission to advance global research dissemination. Specializing in multidisciplinary fields such as life sciences, environmental research, and technology, the platform emphasizes rigorous peer review to maintain high academic standards.

Advancements in Equalization Techniques for Improving Data Throughput and Reliability in Next-Generation Communication Systems

Hao Lin[1] and Jun Wei[2]

[1]Liaoning University of Electronic Science, Department of Communication Engineering, 59 Beihuan Middle Road, Shenyang, Liaoning, China
[2]Hunan Institute of Technology and Telecommunication, Department of Electrical and Information Engineering, 102 Xiangjiang Avenue, Hengyang, Hunan, China

RESEARCH ARTICLE

OPEN ACCESS
Reproducible Model

Edited by
Associate Editor

Curated by
The Editor-in-Chief

Abstract
Digital communication links operating under aggressive spectral reuse, high mobility, and power-constrained architectures rely on equalization to mitigate intersymbol and intercarrier interference introduced by band-limited channels, radio-frequency front-end imperfections, and multiuser coupling. As carrier frequencies extend toward millimeter-wave and sub-terahertz bands, and as baseband sampling becomes coarsely quantized for power efficiency, equalization strategies must adapt to rapidly time-varying, frequency-selective, and hardware-impaired regimes. This paper develops a broad technical treatment of equalization advances for next-generation systems by analyzing linear, nonlinear, and learning-augmented detectors across single-carrier, multi-carrier, and delay–Doppler waveforms. A unified mathematical perspective is used to connect estimation-theoretic derivations, message-passing viewpoints, and optimization-based formulations, highlighting performance–complexity–energy trade-offs under realistic constraints such as low-resolution conversion, hybrid beamforming, and oscillator phase noise. Algorithmic robustness is examined under channel uncertainty, non-Gaussian disturbances, and structured interference arising in massive multiple-antenna and cell-free architectures. Emphasis is placed on stable numerical formulations, scalable preconditioning, and hardware-friendly updates that map efficiently to fixed-point pipelines, systolic arrays, and modern accelerators. The discussion integrates channel-shortening and sequence-detection ideas with iterative decoding, shows how state-evolution tools predict operating points under large-system limits, and outlines regimes where model-driven deep unrolling can improve convergence without sacrificing interpretability. The paper articulates modeling assumptions, identifies key operating regions for different equalizer classes, and provides implementation notes that clarify latency, memory footprint, and data movement bottlenecks. The overall goal is to present technically grounded guidance that helps map waveform, coding, and front-end design choices to equalization architectures capable of sustaining reliable and spectrally efficient links under stringent power and mobility conditions.

1 Introduction

Equalization in modern communication systems stands at the intersection of escalating architectural complexity and increasingly adverse propagation conditions [1]. The continual demand for higher spectral efficiency—driven by dense user deployments, aggressive frequency reuse, and ever-expanding bandwidth—forces receivers to operate at lower signal-to-noise ratios, where interference and channel distortion dominate performance. In such environments, mobility introduces Doppler shifts that render channels time-selective, while oscillator imperfections lead to carrier frequency and phase offsets that couple symbols across both time and frequency domains. These effects jointly erode the separability assumptions underpinning classical equalization theory, compelling the development of algorithms capable of managing spatiotemporal coupling and stochastic nonstationarity without sacrificing real-time operation.

Table 1. Key Challenges in Modern Equalization

Domain	Challenge / Description
Spectral Efficiency	Dense deployments, frequency reuse, and bandwidth expansion increase interference and ISI
Mobility Effects	Doppler shifts and time-varying channels destroy symbol orthogonality
Hardware Constraints	Low-resolution ADCs, nonlinear RF chains, and hybrid analog–digital structures distort received signals
Multiuser / MIMO Coupling	Spatially correlated interference requires joint multi-dimensional equalization

At the same time, front-end power and cost constraints impose significant limitations on data conversion and radio-frequency chain linearity. The adoption of low-resolution analog-to-digital converters, for instance, reduces power consumption but introduces coarse quantization effects that invalidate Gaussian noise models and degrade the performance of traditional linear equalizers. Hybrid analog–digital beamforming architectures, increasingly employed in millimeter-wave and massive MIMO systems, further complicate the problem by partitioning processing between analog phase shifters and digital baseband components. The analog domain imposes hardware-dependent constraints such as constant modulus and limited phase granularity, while the digital stage inherits correlated and distorted signal statistics [2]. Together, these front-end impairments necessitate equalization strategies that are aware of hardware nonidealities and capable of integrating calibration and compensation functions within the same inferential loop.

The rise of multiuser and multiantenna systems exacerbates these challenges by introducing structured interference patterns that cannot be treated as uncorrelated noise. In large-scale antenna arrays or distributed access networks, each user or access point may contribute interference that is spatially or temporally correlated with desired signals. Classical equalizers designed for small, uncoordinated channels break down under such coupling, as their underlying assumptions of independence and limited dimensionality no longer hold. Equalizers must now operate jointly across antennas, frequencies, and sometimes users, performing multi-dimensional filtering or detection to disentangle overlapping streams. This high-dimensional regime transforms equalization from a simple per-symbol correction into a large-scale inference problem, demanding scalable and statistically principled algorithms that balance optimality against computational viability.

Beyond these architectural considerations, practical equalization must coexist with other receiver functions—channel estimation, synchronization, demodulation, and decoding—under strict latency and complexity constraints. Because all of these tasks interact through shared signal models and timing references, the equalizer's role cannot be isolated: it must account for estimation uncertainty, timing errors, and decoding feedback in a unified framework. Joint design of equalization and channel estimation, for instance, allows iterative refinement where improved equalized outputs enhance channel tracking, and refined channel estimates, in turn, sharpen equalization accuracy. Similarly, synchronization loops that mitigate carrier and phase offsets rely on accurate equalized symbols for feedback, making the relationship between synchronization and equalization inherently circular. The challenge lies in orchestrating these modules such that convergence and stability are preserved without violating real-time processing budgets.

A coherent perspective emerges by recognizing equalization as a problem of statistical inference on a constrained observation model. The received signal can be viewed as a noisy, possibly nonlinear transformation of transmitted symbols, where the transformation encodes both the propagation channel and the front-end distortions. This viewpoint unifies linear filtering, decision feedback, and sequence detection within a common probabilistic or optimization-based formalism. Depending on the nature of the distortion and the signal-to-noise ratio, one may adopt different inference strategies: a linear minimum mean-squared error (LMMSE) equalizer when the model is approximately Gaussian and memoryless; a decision-feedback equalizer (DFE) when significant postcursor intersymbol interference can be cancelled using previously detected symbols; or full-sequence detection methods such as the Viterbi algorithm, belief propagation, or message passing when interference exhibits strong temporal or spatial memory.

In large-scale multiantenna or multicarrier systems, exact optimal equalization is computationally prohibitive, motivating approximations based on random matrix theory and large-system analysis [3]. These techniques provide deterministic equivalents for metrics such as mean-squared error or signal-to-interference-plus-noise ratio, enabling parameter tuning and algorithm design without exhaustive simulation. For example, in massive MIMO systems, asymptotic analyses predict performance trends under various power allocations, regularization parameters, and correlation structures, guiding the selection of linear precoding and equalization weights. The resulting insights allow practitioners to design scalable algorithms—such as conjugate-gradient or preconditioned iterative solvers—that achieve near-optimal performance with reduced arithmetic complexity and predictable latency. These solvers exploit the structure of the channel matrix, often leveraging Toeplitz, block-circulant, or sparsity properties to accelerate convergence.

When front-end impairments such as phase noise, IQ imbalance, or coarse quantization dominate, equalization cannot be treated independently from hardware calibration and compensation. Joint estimation and equalization frameworks become essential, wherein the equalizer parameters and impairment models are inferred simultaneously. For example, in systems with severe phase noise, the equalizer must account for random phase rotations across symbols, effectively embedding phase tracking within the equalization recursion. In low-resolution systems, equalizers may incorporate Bussgang decompositions or quantization-aware linearization techniques to maintain consistency between model assumptions and actual hardware behavior [4]. These co-designed methods prevent the onset of error floors that arise when front-end distortions are treated as stationary additive noise—a simplification that fails in modern, power-constrained transceivers.

The increasing integration of equalization with broader signal processing and communication system design underscores the shift from static, isolated modules to dynamic, interdependent inference networks. Modern receivers, especially those implemented in software-defined or reconfigurable architectures, can adapt equalizer structure and parameters in real time based on environmental conditions and hardware states. This adaptability extends beyond parameter tuning to encompass model reconfiguration—for instance, switching between linear and nonlinear equalizers as mobility or interference conditions change. As computational architectures evolve, equalization algorithms increasingly exploit parallelism, sparsity, and low-rank structure to achieve real-time performance, leveraging advances in matrix factorization, convex optimization, and machine learning–inspired estimation.

Designing equalizers for multicarrier waveforms requires attention to intercarrier interference generated by Doppler spreads and oscillator phase noise. Orthogonal frequency-division multiplexing simplifies frequency-selective equalization through diagonalization, yet doubly selective channels and phase noise destroy perfect orthogonality and motivate per-subcarrier interference cancellation and windowed time–frequency processing. For emerging delay–Doppler waveforms, sparsity in the scattering function can be exploited through structured transforms and sparse recovery, allowing equalization to leverage compact support in the delay–Doppler domain. [5]

Learning-augmented methods offer a complementary approach by embedding domain knowledge into neural architectures that respect known algebraic structure. Unrolled algorithms map iterative inference to trainable layers with guaranteed initialization and monotonic descent properties under mild conditions, reducing the risk of overfitting while improving convergence speed. When combined with error-correction decoding in iterative loops, these methods can adapt to residual impairments and refine soft information passed to the decoder.

The remainder of this paper develops a mathematical foundation for these perspectives, including precise problem formulations, solvability conditions, and algorithmic schematics. Emphasis is placed on transformations that reveal conditioning, such as whitening and unitary diagonalization, and on proximal operators that capture nonlinearities introduced by quantization and clipping. Complexity, memory footprint, and data movement are analyzed alongside numerical stability concerns. The presentation favors constructs that connect directly to hardware, including polyphase structures for channel shortening, pipelined factor-graph message passing, and matrix decompositions compatible with fixed-point arithmetic.

2 System and Channel Modeling Under Impairments

Table 2. Equalization Strategies under Different Conditions

Strategy	Operating Conditions / Remarks
LMMSE Equalizer	Near-Gaussian, memoryless environments; analytically tractable
Decision-Feedback Equalizer (DFE)	Strong postcursor ISI; leverages past symbol decisions
Sequence Detection / Message Passing	Channels with long memory or nonlinear distortions; higher complexity
Learning-Augmented / Unrolled Models	Data-driven refinement of inference; structure-preserving neural layers

Table 3. Representative System and Channel Impairments

Impairment	Model / Effect on Equalization
Phase Noise	Multiplicative distortion $e^{j\theta[n]}$; induces nonlinear observation
IQ Imbalance	Creates asymmetric complex baseband response
Quantization	Coarse ADC (b bits); requires Bussgang or nonlinear likelihood modeling
Hybrid Beamforming	Reduced baseband rank via analog precoder/combiner; tighter conditioning
Delay–Doppler Spread	2D convolution structure; sparse recovery in (ν, τ) domain

A baseband discrete-time observation for single-carrier transmission over a frequency-selective channel with length L can be expressed as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w},$$

where $\mathbf{y} \in \mathbb{C}^N$ is the received vector, $\mathbf{x} \in \mathbb{C}^N$ collects transmitted symbols, $\mathbf{H} \in \mathbb{C}^{N\times N}$ is a Toeplitz convolution matrix derived from taps $h[\ell]$ for $0 \le \ell < L$, and $\mathbf{w} \sim \mathcal{CN}(0, \sigma_w^2\mathbf{I})$. In high-mobility scenarios with normalized Doppler ν, a time-varying convolution emerges,

$$y[n] = \sum_{\ell=0}^{L-1} h[n, \ell]\, x[n-\ell] + w[n],$$

with $h[n, \ell]$ satisfying a wide-sense stationary uncorrelated scattering model. Multicarrier transmission with N_c subcarriers adopts an OFDM stacking where the frequency-domain equalizer ideally decouples tones; however, doubly selective channels create leakage described by

$$\tilde{y}[k] = \sum_{m=0}^{N_c-1} \Gamma_{k,m}\, \tilde{x}[m] + \tilde{w}[k],$$

where Γ has dominant diagonal and near-diagonal bands whose structure depends on Doppler spread and windowing.

Phase noise modeled as a discrete-time Wiener process $\theta[n]$ multiplies the baseband signal by $e^{j\theta[n]}$, yielding a nonlinear observation that, under small-angle approximations, admits a first-order linearization

$$y[n] \approx (1 + j\theta[n]) \sum_{\ell} h[\ell] x[n-\ell] + w[n],$$

with an additional colored noise term induced by $\theta[n]$. IQ imbalance and carrier-frequency offset introduce further asymmetries that can be folded into augmented models [6]. For low-resolution data conversion with b bits per in-phase and quadrature components, the quantizer $Q(\cdot)$ yields

$$\mathbf{z} = Q(\mathbf{y}) = Q(\mathbf{H}\mathbf{x} + \mathbf{w}),$$

and equalization proceeds with either a Bussgang decomposition or a generalized likelihood that retains saturation effects.

In multiantenna systems with M receive and K transmit dimensions, a block-fading model reads

$$\mathbf{Y} = \mathbf{H}\mathbf{X} + \mathbf{W},$$

with $\mathbf{H} \in \mathbb{C}^{M\times K}$ having entries that may exhibit spatial correlation and line-of-sight components. Hybrid analog–digital beamforming decomposes $\mathbf{H}$ through analog precoder $\mathbf{F}_{\mathrm{RF}}$ and combiner $\mathbf{W}_{\mathrm{RF}}$, so the effective baseband model becomes

$$\mathbf{Y}_{\mathrm{BB}} = \mathbf{W}_{\mathrm{RF}}^{\mathrm{H}}\mathbf{H}\mathbf{F}_{\mathrm{RF}}\mathbf{X}_{\mathrm{BB}} + \mathbf{W}_{\mathrm{RF}}^{\mathrm{H}}\mathbf{W}.$$

The degrees of freedom at baseband are reduced relative to M and K, tightening numerical conditioning and amplifying the role of prior information in equalization.

A delay–Doppler representation suited to high mobility constructs a 2D circular convolution

$$Z[\nu,\tau] = \sum_{\nu',\tau'} S[\nu-\nu',\tau-\tau']\, X[\nu',\tau'] + W[\nu,\tau],$$

where S is the sampled spreading function concentrating energy on a sparse set. Equalization leverages the near-circulant structure and the sparsity pattern of S, which suggests using structured transforms and iterative sparse recovery.

3 Linear Equalization: Conditioning, Optimality, and Regularization

Table 4. Linear Equalization Methods and Characteristics

Equalizer	Objective / Formula	Key Property
Zero-Forcing (ZF)	$\mathbf{W}_{\mathrm{ZF}} = (\mathbf{H}^{\mathrm{H}}\mathbf{H})^{-1}\mathbf{H}^{\mathrm{H}}$	Cancels ISI perfectly; sensitive to noise and ill-conditioning
LMMSE	$\mathbf{W}_{\mathrm{LMMSE}} = \sigma_x^2\mathbf{H}^{\mathrm{H}}(\sigma_x^2\mathbf{H}\mathbf{H}^{\mathrm{H}} + \sigma_w^2\mathbf{I})^{-1}$	Balances interference suppression and noise amplification
Frequency-Domain (Per-Subcarrier)	$\hat{X}[k] = \frac{H^*[k]}{\lvert H[k]\rvert^2+\alpha}\tilde{Y}[k]$	α regulates trade-off; efficient via FFT
Regularized / Structured	$\min_{\mathbf{w}} \frac{1}{2}\lVert\mathbf{A}\mathbf{w}-\mathbf{b}\rVert_2^2+\lambda\lVert\mathbf{G}\mathbf{w}\rVert_1$	Enforces smoothness or sparsity via ℓ_1 penalties

Linear equalizers address the estimation of $\mathbf{x}$ from $\mathbf{y}$ by minimizing a quadratic criterion. The zero-forcing filter seeks $\mathbf{W}_{\mathrm{ZF}}$ satisfying

$$\mathbf{W}_{\mathrm{ZF}} = \arg\min_{\mathbf{W}} \lVert\mathbf{W}\mathbf{H} - \mathbf{I}\rVert_F^2,$$

whose solution for full column rank $\mathbf{H}$ is

$$\mathbf{W}_{\mathrm{ZF}} = \left(\mathbf{H}^{\mathrm{H}}\mathbf{H}\right)^{-1}\mathbf{H}^{\mathrm{H}}.$$

Noise enhancement is controlled by the smallest singular value of $\mathbf{H}$, motivating regularization. The linear MMSE equalizer minimizes expected squared error under Gaussian assumptions,

$$\mathbf{W}_{\mathrm{LMMSE}} = \arg\min_{\mathbf{W}} \mathbb{E}\left[\lVert\mathbf{x} - \mathbf{W}\mathbf{y}\rVert_2^2\right],$$

Table 5. Conditioning, Regularization, and Optimization Aspects

Aspect	Description / Implication
Condition Number $\kappa_2(\mathbf{H})$	Ratio $\sigma_{\max}/\sigma_{\min}$ indicates sensitivity; poor conditioning increases noise amplification
Regularization Parameter α	Derived as σ_w^2/σ_x^2 in LMMSE; stabilizes inversion
Structured Regularizers	Derivative-based (smoothness) or group-sparse (banded) priors improve interpretability and robustness
First-Order Solvers	Proximal / soft-thresholding iterations feasible for fixed-point hardware

with closed-form

$$\mathbf{W}_{\text{LMMSE}} = \mathbf{R_{xx}}\mathbf{H}^{\text{H}}\left(\mathbf{H}\mathbf{R_{xx}}\mathbf{H}^{\text{H}} + \mathbf{R_{ww}}\right)^{-1}.$$

Assuming independent symbols with variance σ_x^2 and white noise $\sigma_w^2\mathbf{I}$, this reduces to

$$\mathbf{W}_{\text{LMMSE}} = \sigma_x^2\mathbf{H}^{\text{H}}\left(\sigma_x^2\mathbf{H}\mathbf{H}^{\text{H}} + \sigma_w^2\mathbf{I}\right)^{-1}.$$

In frequency domain, per-subcarrier equalization employs

$$\hat{X}[k] = \frac{H^*[k]}{|H[k]|^2 + \alpha}\,\tilde{Y}[k],$$

with $\alpha = \sigma_w^2/\sigma_x^2$ for LMMSE and $\alpha = 0$ for ZF. When doubly selective effects spill energy onto neighbors, a banded linear system arises with Toeplitz–circulant structure; efficient inversion uses conjugate gradients with FFT-based multiplications. Conditioning can be quantified via the spectral condition number

$$\kappa_2(\mathbf{H}) = \frac{\sigma_{\max}(\mathbf{H})}{\sigma_{\min}(\mathbf{H})},$$

and random-matrix results predict the empirical spectral distribution of $\mathbf{H}^{\text{H}}\mathbf{H}$ for i.i.d. channel gains, guiding the choice of α to stabilize inversion. [7]

Regularization that exploits structure beyond second-order moments includes Tikhonov terms aligned with derivative operators in time or frequency to penalize roughness of the equalizer response, and group-sparse penalties to encourage bandedness. A convex composite objective

$$\min_{\mathbf{w}} \ \frac{1}{2}\|\mathbf{A}\mathbf{w} - \mathbf{b}\|_2^2 + \lambda\|\mathbf{G}\mathbf{w}\|_1$$

captures these preferences, where $\mathbf{A}$ encodes convolutional mixing and $\mathbf{G}$ enforces locality. First-order proximal algorithms implement the corresponding soft-thresholding in fixed-point pipelines, with step sizes set by Lipschitz constants of $\mathbf{A}^{\text{H}}\mathbf{A}$.

For multiuser MIMO with M receive and K transmit dimensions, the post-equalization signal-to-interference-plus-noise ratio for user k under linear precoding–combining can be written as

$$\text{SINR}_k = \frac{|\mathbf{w}_k^{\text{H}}\mathbf{H}\mathbf{f}_k|^2}{\sum_{j\neq k}|\mathbf{w}_k^{\text{H}}\mathbf{H}\mathbf{f}_j|^2 + \sigma_w^2\|\mathbf{w}_k\|_2^2},$$

and optimization over $\{\mathbf{w}_k, \mathbf{f}_k\}$ under per-antenna or sum-power constraints aligns with weighted-MMSE iterations. The WMMSE fixed point satisfies stationarity conditions equivalent to KKT conditions of a sum-rate maximization with log terms; practical implementations embed equalization updates into this loop, amortizing matrix factorizations across users.

Table 6. Nonlinear Equalization: Algorithms and Robust Variants

Method	Formulation / Objective	Notable Features
Decision-Feedback (DFE)	$\hat{\mathbf{x}} = \mathrm{Q}(\mathbf{Fy} - \mathbf{B}\hat{\mathbf{x}})$	Cancels postcursor ISI; error propagation risk
Sequence Estimation (MAP/Viterbi)	$\arg\min_{\mathbf{x}} \frac{1}{2\sigma_w^2}\|\mathbf{y} - \mathbf{Hx}\|^2 - \log p(\mathbf{x})$	Optimal sequence detection; trellis complexity scales with memory
Channel Shortening	$\min_{\mathbf{g},\mathbf{c}} \|\mathbf{g}^{\mathsf{H}}\mathbf{H} - \mathbf{c}^{\mathsf{H}}\|^2 + \beta\|\mathbf{g}\|^2$	Reduces channel memory; enables lower-complexity detection
Distributionally Robust	$\min_{\mathbf{W}} \max_{\Delta\in\mathcal{U}} \mathbb{E}\|\mathbf{x} - \mathbf{W}(\mathbf{H} + \Delta)\mathbf{x}\|^2$	Mitigates model mismatch via uncertainty-aware penalties
Soft Interference Cancellation	$\mathbf{r} = \mathbf{y} - \mathbf{H}\mathbb{E}[\mathbf{x}]$, $\mathbf{C}_r = \mathbf{R}_{ww} + \mathbf{H}\mathrm{diag}(\mathrm{Var}[\mathbf{x}])\mathbf{H}^{\mathsf{H}}$	Combines LMMSE filtering with posterior-based symbol updates

4 Nonlinear Equalization: Decision Feedback, Sequence Estimation, and Robustness

Decision feedback equalization reduces postcursor interference by subtracting previously detected symbol contributions from the current decision statistic. In matrix form for a causal feedforward $\mathbf{F}$ and strictly lower-triangular feedback $\mathbf{B}$, the detector computes

$$\hat{\mathbf{x}} = \mathrm{Q}\left(\mathbf{Fy} - \mathbf{B}\hat{\mathbf{x}}\right),$$

where $\mathrm{Q}(\cdot)$ denotes a slicer. Designing $\mathbf{F}$ and $\mathbf{B}$ to minimize mean-squared error under unbiasedness constraints yields closed-form solutions via Cholesky factorizations of $\mathbf{H}^{\mathsf{H}}\mathbf{R}_{\mathbf{ww}}\mathbf{H}$, and Tomlinson–Harashima precoding implements a transmitter-side analog for pre-equalization.

Maximum a posteriori sequence estimation solves

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}\in\mathcal{X}^N} \frac{1}{2\sigma_w^2}\|\mathbf{y} - \mathbf{Hx}\|_2^2 - \log p(\mathbf{x}),$$

where $\mathcal{X}$ is the constellation. For memoryless priors and linear convolution, the metric decomposes into a trellis with branch metrics depending on tap vectors [8]. The Viterbi algorithm yields ML sequence estimates for small memory, while the BCJR algorithm delivers symbol-wise posteriors to feed iterative decoders. When channels are long or time varying, channel-shortening filters condense memory to a target length, reducing trellis complexity. The channel-shortening design solves

$$\min_{\mathbf{g},\,\mathbf{c}} \|\mathbf{g}^{\mathsf{H}}\mathbf{H} - \mathbf{c}^{\mathsf{H}}\|_2^2 + \beta\|\mathbf{g}\|_2^2,$$

subject to $\mathbf{c}$ being nonzero on a short support, where $\mathbf{g}$ is the prefilter and $\mathbf{c}$ the target impulse response.

Robustness to model mismatch is critical when front-end impairments or calibration drift generate structured errors. A distributionally robust formulation inflates the observation covariance within an uncertainty set $\mathcal{U}$:

$$\min_{\mathbf{W}} \max_{\Delta\in\mathcal{U}} \mathbb{E}\|\mathbf{x} - \mathbf{W}(\mathbf{H} + \mathbf{\Delta})\mathbf{x} - \mathbf{Ww}\|_2^2,$$

leading to regularized solutions where the penalty level depends on the radius of $\mathcal{U}$. In decision feedback, error propagation is ameliorated by soft interference cancellation using posterior means and variances rather than hard decisions, bridging toward iterative detection. Soft cancellation computes

$$\mathbf{r} = \mathbf{y} - \mathbf{H}\mathbb{E}[\mathbf{x}|\mathrm{prior}], \quad \mathbf{C}_r = \mathbf{R}_{\mathbf{ww}} + \mathbf{H}\mathrm{diag}(\mathrm{Var}[\mathbf{x}])\mathbf{H}^{\mathsf{H}},$$

and applies an LMMSE filter on $\mathbf{r}$, updating symbol beliefs.

5 Iterative Equalization via Factor Graphs and Message Passing

A factor-graph representation separates the linear mixing constraint from symbol priors and coding constraints. For the linear Gaussian channel, loopy belief propagation approximates marginal posteriors through messages that are Gaussian in continuous variables and discrete on constellation nodes [9]. Approximate message passing collapses these updates into a pair of vector operations with Onsager correction terms; for complex-valued systems with i.i.d. sub-Gaussian $\mathbf{H}$, the algorithm writes

$$\mathbf{x}^{t+1} = \eta_t\Big(\mathbf{x}^t + \mathbf{H}^{\mathsf{H}}\mathbf{z}^t\Big), \quad \mathbf{z}^t = \mathbf{y} - \mathbf{H}\mathbf{x}^t + \frac{1}{N}\mathbf{z}^{t-1}\sum_{i=1}^{N}\eta_t'(\cdot),$$

where η_t is a denoiser matched to the symbol prior and $\eta_t'(\cdot)$ denotes its divergence. State evolution predicts the effective noise variance τ_t^2 entering η_t via a scalar recursion

$$\tau_{t+1}^2 = \sigma_w^2 + \delta\,\mathbb{E}\left[|X - \eta_t(X + \tau_t Z)|^2\right], \quad \delta = \frac{K}{M},$$

with $Z \sim \mathcal{CN}(0,1)$ and X distributed as the prior. This recursion guides damping choices and stopping criteria, tying macroscopic performance to microscopic denoiser properties.

Orthogonal AMP variants replace independence assumptions with rotationally invariant $\mathbf{H}$, yielding stable convergence under broader ensembles through a decorrelated residual update. Vector AMP and expectation propagation generalize to colored noise and correlated channels by tracking covariance matrices or by whitening transformations. In doubly selective multicarrier systems, message passing can operate on a banded interference graph where each tone node exchanges beliefs with near neighbors; the sparsity pattern influences computational load and convergence speed.

Turbo equalization integrates a soft-in/soft-out equalizer with a channel decoder through extrinsic information exchange. Let L^{A} denote a priori log-likelihood ratios from the decoder and L^{E} the extrinsic LLRs returned by the equalizer. A Gaussian approximation to symbol priors produces equalizer means and variances that yield [10]

$$L^{\mathsf{E}} \approx \frac{2\,\Re\{g\,r\}}{\sigma_{\mathrm{eff}}^2} - L^{\mathsf{A}},$$

where r is a matched-filter output, g an effective gain, and σ_{eff}^2 an effective noise variance capturing residual interference after linear MMSE filtering with prior variance embedding. The exchange proceeds until a halting criterion is reached or a latency budget expires. Analytical tools based on extrinsic transfer curves predict tunnel openings that correlate with error floors, assisting in the selection of equalizer bandwidth and decoder strength.

6 Equalization for Doubly Selective Channels and Delay–Doppler Waveforms

In scenarios where the channel varies within an OFDM symbol, the intercarrier interference matrix Γ acquires banded structure with bandwidth proportional to the normalized Doppler. Windowed time-domain processing reduces leakage by tapering symbol edges, effectively convolving subcarrier responses. A joint time–frequency equalizer solves

$$\min_{\tilde{\mathbf{x}}}\ \|\tilde{\mathbf{y}} - \Gamma\tilde{\mathbf{x}}\|_2^2 + \alpha\|\mathbf{D}\tilde{\mathbf{x}}\|_2^2,$$

where $\mathbf{D}$ penalizes roughness across frequency or time indices. Efficient solutions rely on conjugate gradients with circulant preconditioners derived from the diagonal of $\Gamma^{\mathsf{H}}\Gamma$.

Delay–Doppler modulation maps symbols onto a lattice where the channel acts as a 2D circular convolution. Equalization operates in the symplectic finite Fourier transform domain using

$$\mathbf{Z} = \mathbf{F}_\nu\,\mathbf{Y}\,\mathbf{F}_\tau^{\mathsf{H}}, \quad \mathbf{X} = \mathbf{F}_\nu\,\mathbf{S}\,\mathbf{F}_\tau^{\mathsf{H}},$$

and the mixing becomes $\mathrm{vec}(\mathbf{Z}) = \mathbf{S}_{\mathrm{DD}}\,\mathrm{vec}(\mathbf{X}) + \mathrm{vec}(\mathbf{W})$ with $\mathbf{S}_{\mathrm{DD}}$ nearly block-circulant. A 2D MMSE equalizer adopts [11]

$$\hat{\mathbf{X}} = \left(\mathbf{S}_{\mathrm{DD}}^{\mathrm{H}}\mathbf{S}_{\mathrm{DD}} + \alpha\mathbf{I}\right)^{-1}\mathbf{S}_{\mathrm{DD}}^{\mathrm{H}}\mathrm{vec}(\mathbf{Z}),$$

computed through iterative methods exploiting FFT-based block multiplications. When the scattering function is sparse, a sparse Bayesian equalizer places a Laplacian or Bernoulli–Gaussian prior on significant taps and uses coordinate descent or expectation–maximization to refine support and amplitudes. The associated denoiser in an AMP framework becomes a soft-thresholding operator with parameters tied to Doppler and delay spreads.

Time-varying channels with partial reciprocity or pilot contamination motivate Kalman filtering for tracking effective coefficients. For a state-space model $\mathbf{h}_{n+1} = \mathbf{A}\mathbf{h}_n + \mathbf{q}_n$, $\mathbf{y}_n = \mathbf{X}_n\mathbf{h}_n + \mathbf{w}_n$, the equalizer relies on filtered and smoothed estimates $\hat{\mathbf{h}}_{n|n}$ and covariances $\mathbf{P}_{n|n}$ to construct per-symbol MMSE filters with uncertainty inflation. Stability hinges on spectral radius $\rho(\mathbf{A})$ and on observability of $(\mathbf{A}, \mathbf{X}_n)$. In high mobility, augmented-state models include Doppler slopes, and square-root implementations preserve numerical conditioning under fixed-point.

7 Massive and Cell-Free MIMO Equalization with Hardware Constraints

Massive MIMO aggregates many antennas to concentrate energy spatially, enabling simple linear equalizers to achieve near-optimal performance in favorable propagation. With $M \gg K$, matched filtering approximates the inverse of $\mathbf{H}^{\mathrm{H}}\mathbf{H}$ due to channel hardening. The post-combining interference behaves approximately Gaussian with variance predicted by large-system analysis. A deterministic equivalent for user k under LMMSE combining reads [12]

$$\gamma_k \approx \frac{p_k m_k^2}{\sum_{j\neq k} p_j \theta_{k,j} + \sigma_w^2 \xi_k},$$

where p_k is the power allocation, m_k the effective channel gain, $\theta_{k,j}$ cross terms, and ξ_k captures residual noise enhancement. These quantities satisfy fixed-point equations involving traces of resolvent matrices of $\mathbf{H}\mathbf{H}^{\mathrm{H}}$, allowing fast evaluation during resource allocation.

Hybrid beamforming constrains baseband dimension by analog networks with quantized phase shifters. The effective channel becomes $\mathbf{H}_{\mathrm{eff}} = \mathbf{W}_{\mathrm{RF}}^{\mathrm{H}}\mathbf{H}\mathbf{F}_{\mathrm{RF}}$, and equalizer design must respect constant-modulus constraints. Alternating optimization updates $\mathbf{F}_{\mathrm{RF}}$, $\mathbf{W}_{\mathrm{RF}}$, and baseband equalizers by solving

$$\min_{\mathbf{F}_{\mathrm{BB}},\mathbf{W}_{\mathrm{BB}}} \sum_k \|\mathbf{s}_k - \mathbf{W}_{\mathrm{BB}}^{\mathrm{H}}\mathbf{H}_{\mathrm{eff}}\mathbf{F}_{\mathrm{BB}}\mathbf{s}_k\|_2^2 + \alpha\|\mathbf{F}_{\mathrm{BB}}\|_F^2 + \beta\|\mathbf{W}_{\mathrm{BB}}\|_F^2,$$

subject to $|[\mathbf{F}_{\mathrm{RF}}]_{i,j}| = 1/\sqrt{N_{\mathrm{RF}}}$. Projected-gradient steps enforce constant-modulus constraints, while baseband updates use closed-form ridge regression.

Low-resolution analog-to-digital converters reduce power but complicate equalization. A Bussgang linearization expresses the quantized observation as

$$\mathbf{z} = \mathbf{G}\mathbf{y} + \mathbf{e}, \quad \mathbb{E}[\mathbf{y}\mathbf{e}^{\mathrm{H}}] = 0,$$

with $\mathbf{G}$ a gain matrix dependent on input covariance. Equalization proceeds with an effective linear model using $\mathbf{G}\mathbf{H}$ and augmented noise covariance $\mathbf{R}_{\mathbf{ee}} + \mathbf{G}\mathbf{R}_{\mathbf{ww}}\mathbf{G}^{\mathrm{H}}$. Iterative refinement incorporates symbol-dependent covariance updates, improving accuracy in moderate-SNR regimes. One-bit conversion introduces severe nonlinearity; generalized approximate message passing tailored to quantized likelihoods uses output-channel denoisers

$$\hat{y}_i = \mathbb{E}[Y_i|Z_i, q_i], \quad \mathrm{Var}[Y_i|Z_i, q_i],$$

with q_i the quantizer bin and Z_i the linear estimate.

Cell-free architectures distribute many access points over a wide area, each with local combining that is fused centrally through fronthaul links of limited capacity [13]. Equalization can be organized as local LMMSE with compression followed by centralized fusion,

$$\hat{\mathbf{s}} = \left(\sum_m \mathbf{G}_m^{\mathrm{H}} \mathbf{R}_m^{-1} \mathbf{G}_m \right)^{-1} \sum_m \mathbf{G}_m^{\mathrm{H}} \mathbf{R}_m^{-1} \mathbf{r}_m,$$

where $\mathbf{G}_m$ and $\mathbf{r}_m$ are local effective channels and observations, and $\mathbf{R}_m$ their covariances including quantization and thermal noise. Rate–distortion arguments guide fronthaul compression levels, and prewhitening at the edge simplifies fusion. Synchronization uncertainty across access points is mitigated by phase-tracking loops embedded within the equalization pipeline.

8 Learning-Augmented Equalization and Algorithm Unrolling

Learning-augmented equalizers exploit data to tune algorithm parameters or to replace heuristic steps with trainable modules while retaining structural priors. A canonical approach unrolls T iterations of a proximal gradient method for a composite objective

$$\min_{\mathbf{x}} \ \frac{1}{2} \|\mathbf{y} - \mathbf{H}\mathbf{x}\|_2^2 + \lambda \sum_i \phi(x_i),$$

where ϕ is a sparsity- or constellation-aware penalty. The iteration

$$\mathbf{x}^{t+1} = \mathcal{P}_{\theta_t} \Big(\mathbf{x}^t - \mu_t \mathbf{H}^{\mathrm{H}} (\mathbf{H}\mathbf{x}^t - \mathbf{y}) \Big)$$

uses step sizes μ_t and proximal parameters θ_t learned from data, with $\mathcal{P}_\theta$ implementing a parametric shrinkage or constellation-aware denoiser. Training employs synthetic channels or over-the-air captures and backpropagates through the unrolled graph with truncated backprop-through-time to control memory. [14]

Message-passing networks replace denoisers with small neural modules that map sufficient statistics to posterior means and variances. Stability is improved by spectral normalization and by tying parameters across iterations to limit overfitting. When channel state information is uncertain, a joint network accepts both $\mathbf{y}$ and pilot-based $\hat{\mathbf{H}}$ along with quality indicators to modulate equalization aggressiveness. Data augmentation with phase and frequency offsets improves robustness to synchronization errors.

A hybrid design treats front-end compensation as a differentiable layer. A phase-tracking module updates a latent $\hat{\theta}_n$ via gradient steps on a surrogate loss that measures residual intercarrier interference energy. The equalizer consumes the de-rotated observations and emits soft symbols. End-to-end training balances bias from the compensation layer with variance in the equalizer and decoder, and includes regularizers on phase increments to reflect oscillator physics.

Energy and latency constraints favor compact networks. Knowledge distillation transfers performance from a large teacher—potentially a high-iteration unrolled solver—to a small student by matching soft outputs and intermediate feature statistics [15]. Quantization-aware training ensures compatibility with fixed-point accelerators, and structured pruning aligns with systolic array tiling. The training objective includes a term penalizing memory traffic proxies, such as the number of off-chip accesses, steering architectures toward local reuse.

9 Bayesian and Probabilistic Optimization of Feed-Forward and Decision-Feedback Equalizers

A probabilistic perspective on equalizer tuning reframes the selection of filter lengths, step sizes, regularization weights, and architectural toggles as a sequential experimental design problem on a black-box objective that is expensive, noisy, and nonstationary. Instead of fixing hyperparameters offline or relying solely on deterministic gradient-based adaptation derived from simplified models, the receiver allocates on-air or emulated trials to probe the response surface of performance proxies such as post-equalization mutual information, generalized mutual information, error vector magnitude, or decoder-aided block-error probability under realistic waveform, channel, and hardware conditions. The outcome of each probe is inherently stochastic because it integrates thermal noise, residual interference, channel estimation error, and possibly quantization artifacts. A Bayesian optimizer compresses these observations into a posterior over the objective, balancing exploration of uncertain hyperparameter regions against exploitation of promising settings, and embracing constraints that arise from latency budgets, memory limits, and peak-to-average power restrictions.

Let $\theta \in \mathbb{R}^d$ collect continuous knobs, including feed-forward and feedback tap weights projected to a low-dimensional manifold, equalizer bandwidths, damping factors in iterative updates, and tunables tied to phase-noise tracking, while discrete switches choose decision-feedback enablement, slicer granularity, or partial-response targets. The performance metric $f(\theta)$ is not directly known and is observed as $y = f(\theta) + \varepsilon$, where ε models measurement noise with variance $\sigma^2(\theta)$ that may depend on the number of symbols used to evaluate f and on instantaneous channel variability. A Gaussian process prior with mean $m(\theta)$ and kernel $k(\theta, \theta')$ offers a flexible surrogate, [16]

$$f(\cdot) \sim \mathcal{GP}(m(\cdot),\ k(\cdot,\cdot)),$$

so that, given observations $\mathcal{D}_t = \{(\theta_i, y_i)\}_{i=1}^t$, the posterior mean and variance at a candidate θ become

$$\mu_t(\theta) = m(\theta) + \mathbf{k}_t(\theta)^\top \left(\mathbf{K}_t + \sigma^2\mathbf{I}\right)^{-1} (\mathbf{y}_t - \mathbf{m}_t),$$

$$\sigma_t^2(\theta) = k(\theta,\theta) - \mathbf{k}_t(\theta)^\top \left(\mathbf{K}_t + \sigma^2\mathbf{I}\right)^{-1} \mathbf{k}_t(\theta),$$

where $[\mathbf{K}_t]_{ij} = k(\theta_i, \theta_j)$, $\mathbf{k}_t(\theta) = [k(\theta_i, \theta)]_{i=1}^t$, and $\mathbf{m}_t = [m(\theta_i)]_{i=1}^t$. Acquisition rules such as expected improvement prioritize evaluation points that jointly consider predicted gain and uncertainty. For maximization with incumbent $f_t^\star$,

$$\mathrm{EI}_t(\theta) = \mathbb{E}\left[(f(\theta) - f_t^\star)_+\right] = (\mu_t(\theta) - f_t^\star)\Phi\left(\frac{\mu_t(\theta) - f_t^\star}{\sigma_t(\theta)}\right) + \sigma_t(\theta)\phi\left(\frac{\mu_t(\theta) - f_t^\star}{\sigma_t(\theta)}\right),$$

where Φ and ϕ denote the standard normal CDF and PDF. In resource-constrained receivers, the acquisition must also encode the cost of each probe, which depends on the number of symbols, pilots, and decoder iterations used to estimate f, motivating cost-aware variants that maximize improvement per unit time or energy.

Equalizer hyperparameters influence performance through nonconvex couplings that are painful to differentiate end-to-end under realistic impairments. Consider a decision-feedback structure with feed-forward filter $\mathbf{f} \in \mathbb{C}^{L_f}$ and strictly causal feedback filter $\mathbf{b} \in \mathbb{C}^{L_b}$, together with a linear MMSE prewhitener parameterized by a ridge coefficient $\alpha > 0$. An information-centric metric that avoids constellation-specific discontinuities is the mismatched achievable rate via the generalized mutual information,

$$\mathrm{GMI}(\theta) = \mathbb{E}\left[\log \frac{q(Y|X;\theta)}{\sum_{x' \in \mathcal{X}} \pi(x')\, q(Y|x';\theta)}\right],$$

with q the post-equalization Gaussian surrogate likelihood and π an assumed prior on symbols. The expectation is taken over true channel, noise, and residual interference [17]. While GMI correlates with decoder performance, evaluating it at scale still incurs nontrivial cost, reinforcing the role of surrogates. A practical surrogate is the posterior mean of symbol reliability aggregated over tones and time,

$$J(\theta) = \frac{1}{N} \sum_{n=1}^{N} \frac{|g_n(\theta) r_n(\theta)|^2}{\sigma_{\mathrm{eff},n}^2(\theta)},$$

with $g_n(\theta)$ an effective gain and $\sigma_{\mathrm{eff},n}^2(\theta)$ the residual interference-plus-noise variance implied by the chosen $(\mathbf{f}, \mathbf{b}, \alpha)$. Bayesian optimization treats $J(\theta)$ as the objective or as a correlated auxiliary to accelerate the search for θ that improves GMI.

Decision feedback introduces error propagation that depends on symbol reliability, burst structure of interference, and slicer operating region. A probabilistic optimizer can temper aggressive feedback by choosing L_b and feedback scaling to minimize a risk functional that penalizes tails of the reliability distribution. A coherent formulation introduces a conditional value-at-risk objective on per-symbol log-likelihood ratios $\{L_n(\theta)\}$,

$$\mathrm{CVaR}_\beta(\theta) = \inf_{\tau\in\mathbb{R}} \left\{ \tau + \frac{1}{\beta}\, \mathbb{E}\left[(\tau - L_n(\theta))_+ \right] \right\},$$

so that θ is tuned to improve the worst β-fraction of symbols, which frequently dominate block error outcomes after decoding. Since direct optimization is infeasible, the Bayesian surrogate learns the mapping $\theta \mapsto \mathrm{CVaR}_\beta(\theta)$ from batched reliability samples and proposes trials where the acquisition reflects tail-risk reduction.

In time-varying channels, the objective drifts with Doppler, oscillator phase noise, and interference geometry. A stationary kernel can underfit; nonstationary kernels address this by incorporating channel descriptors ξ such as delay spread, Doppler spread, and phase-noise variance into the input, yielding an augmented $k\big((\theta,\xi),(\theta',\xi')\big)$ that learns how optimal θ adapts across environments. Alternatively, a forgetting factor implements streaming posterior updates with exponential decay on old data, which is equivalent to inflating observation noise for stale points. The optimizer thus changes equalizer aggressiveness over time, e.g., shrinking feedback when mobility raises uncertainty in pre-cursor estimates or when low-resolution quantization amplifies nonlinearity. [18]

Hardware constraints embed directly into the search as inequality and budget constraints. Safe Bayesian optimization recognizes an unknown feasible set $\mathcal{F} = \{\theta : c_j(\theta) \le 0,\, j = 1,\ldots,J\}$, where constraints c_j reflect decoder-latency budgets, maximum allowed memory traffic, and fixed-point overflow margins. Each c_j is modeled with a GP and queried jointly with the objective. The acquisition selects candidates that satisfy feasibility with high posterior probability while maximizing predicted improvement. For example, a latency model grounded in measured cycle counts as a function of tap lengths, FFT sizes, and unrolling depth yields a constraint

$$c_{\mathrm{lat}}(\theta) = \mathbb{E}\left[T(\theta)\right] - T_{\max},$$

with $T_{\max}$ a hard deadline. The constrained optimizer explores only those θ for which $\mathbb{P}\{c_{\mathrm{lat}}(\theta) \le 0\}$ exceeds a safety threshold.

When the parameter dimension is large, structural priors curb sample complexity. Feed-forward and feedback taps lie near low-dimensional subspaces determined by dominant channel modes; a linear dictionary $\mathbf{U}$ with $r \ll L_f + L_b$ columns parameterizes $(\mathbf{f},\mathbf{b}) = \mathbf{U}\eta$. The optimizer searches over $\eta \in \mathbb{R}^r$ while $\mathbf{U}$ is either fixed from prior analysis or adapted slowly using principal subspace estimates of the channel convolution operator. Additive kernels exploit near-separability between groups of parameters,

$$k(\theta,\theta') = \sum_{g=1}^{G} k_g(\theta_g,\theta'_g),$$

yielding decomposable posteriors and acquisitions that scale with G rather than d. For mixed discrete–continuous designs, a Hamming-distance kernel covers combinatorial toggles such as enabling Tomlinson–Harashima precoding, while continuous kernels manage step sizes and ridge levels. [19]

To tame heteroscedasticity from variable-length probes, multi-fidelity modeling binds fast but biased metrics to slow but accurate ones. Let $f_0(\theta)$ be a low-cost proxy based on short frames or on decoder-free reliability summaries, and $f_1(\theta)$ the high-fidelity target measured with long frames and full decoding. A coregionalized GP captures cross-covariances,

$$\begin{bmatrix} f_0(\theta) \\ f_1(\theta) \end{bmatrix} \sim \mathcal{GP}\left(\mathbf{m}(\theta),\ \begin{bmatrix} k_{00}(\theta,\theta') & k_{01}(\theta,\theta') \\ k_{10}(\theta,\theta') & k_{11}(\theta,\theta') \end{bmatrix} \right),$$

so that frequent cheap queries of f_0 guide sparse expensive queries of f_1. Acquisition functions such as the multi-fidelity knowledge gradient select $(\theta, \text{fidelity})$ pairs that maximize expected value of information per unit cost, accelerating convergence under tight energy budgets.

Integration with iterative detection–decoding loops requires that the optimizer respect extrinsic-information dynamics. A coarse but effective bridge leverages state-evolution predictors of equalizer outputs to approximate how hyperparameters migrate the operating point on decoder transfer curves [20]. If $\tau^2(\theta)$ denotes the effective noise variance at the equalizer output predicted by large-system analysis, and T_{dec} is a decoder transfer function mapping input variance to output variance proxy, fixed-point relations

$$\tau_{t+1}^2(\theta) = \sigma_w^2 + \Psi(\theta, \tau_t^2), \qquad \upsilon_{t+1}^2 = T_{\text{dec}}\big(\tau_{t+1}^2(\theta)\big)$$

allow the optimizer to target θ that drives the pair (τ^2, υ^2) toward desirable regions without performing full decoding at every probe. The Bayesian surrogate can be trained to emulate $\Psi(\theta, \cdot)$ from limited pilot runs, trading exactness for speed while maintaining fidelity in the regime that determines error floors.

Probabilistic design extends beyond choosing scalar hyperparameters to shaping the feedback topology itself. Partial decision feedback limits propagation by selecting a subset of past decisions to subtract, guided by a posterior over their reliabilities. Represent the inclusion mask as $s \in \{0,1\}^{L_b}$ with a prior that favors sparsity. The performance surface $f(\eta, s)$ is combinatorial; a Thompson-sampling strategy samples f from the posterior surrogate and greedily constructs s by conditional maximization over entries, akin to Bayesian matching pursuit. Under block-fading, stability improves when the mask is frozen within a coherence block and updated only when posterior uncertainty exceeds a threshold that signals a likely topology change in interference paths.

In strongly hardware-impaired receivers, the observation model is quantized or phase-noisy, and gradients through the nonlinearity are unreliable. Bayesian optimization sidesteps this by learning effective hyperparameters for Bussgang-based linearizations and likelihood-aware message passing without differentiating through hard quantizers [21]. A robust objective that maintains performance across anticipated impairment distributions emerges by integrating the surrogate over a prior on impairment parameters ϕ,

$$\bar{f}(\theta) = \int f(\theta; \phi)\, p(\phi)\, d\phi \approx \frac{1}{S} \sum_{s=1}^{S} f(\theta; \phi_s),$$

with $\{\phi_s\}$ drawn from calibration-informed priors over phase-noise linewidths, gain/offset mismatch, and ADC thresholds. The posterior and acquisition are then built on $\bar{f}$, producing hyperparameters that generalize across day-to-day drift.

Co-design with beamforming exposes yet another layer of coupled decisions. Hybrid analog–digital front ends restrict the baseband dimension, shrinking equalizer degrees of freedom. A bilevel scheme looks attractive: the outer loop uses Bayesian optimization to choose analog beam codebook entries and baseband dimensionality, while the inner loop tunes equalizer parameters conditional on the chosen front end. Direct bilevel search is expensive; instead, a joint surrogate takes as input both beam indices and equalizer knobs, with a kernel that respects invariances such as global phase and permuted RF chains. Acquisition in this joint domain is tempered by a constraint on reconfiguration overhead, encoded as a cost term that reflects the transient penalty of retuning phase shifters and reinitializing channel estimates after beam changes.

Bandwidth- and latency-aware acquisition policies are vital for embedded deployments. If a probe at θ consumes time $\tau(\theta)$ and energy $E(\theta)$, the utility of an evaluation is scaled accordingly [22]. A normalized expected improvement,

$$\text{NEI}_t(\theta) = \frac{\text{EI}_t(\theta)}{a\,\tau(\theta) + b\,E(\theta)},$$

with design weights $a, b > 0$, biases the search toward evaluations that produce useful information at low operational cost. When real-time constraints are strict, batched Bayesian optimization dispatches B candidates in parallel, for example across subbands or user groups whose trials do not interfere, using fantasized posteriors to account for pending evaluations. Batched policies must incorporate interference coupling so that simultaneous probes do not corrupt each other's metrics; a conservative approach enforces spatial or spectral separation based on current channel covariance estimates.

A full-stack receiver can expose not only raw performance outcomes but also intermediate diagnostics as auxiliary observations that sharpen the surrogate. These include distributions of slicer distances, empirical residual spectra at equalizer outputs, and decoder syndrome weights. Multi-output Gaussian processes then model a vector-valued function whose components co-vary, enabling the acquisition to prioritize θ that reduces uncertainty in the most informative diagnostic, which in turn reduces uncertainty in the target. If the diagnostics are high-dimensional, a random-feature representation or a learned linear embedding compresses them into a few summary statistics that preserve sensitivity to deleterious phenomena such as intercarrier leakage or phase-noise induced common-phase errors. [23]

Reliability under out-of-distribution shifts benefits from priors that encode known invariances and smoothness. Kernels built from physically meaningful distances, such as geodesic distances between filters modulo linear-phase and scale, enforce that equivalent equalizers are nearby in input space. For feed-forward/feedback pairs $(\mathbf{f}, \mathbf{b})$ with unavoidable scaling ambiguities due to slicer normalization, define equivalence classes by normalizing $\|\mathbf{f}\|_2 = 1$ and embedding into a sphere, then use a heat kernel on the sphere to respect rotational symmetries. This improves sample efficiency and prevents the optimizer from wasting probes on redundant parameterizations that differ by trivial symmetries.

Although the Bayesian outer loop is non-intrusive, it must harmonize with inner-loop adaptive algorithms such as stochastic gradient updates on tap weights. A timescale separation ensures that the inner loop nearly equilibrates before the outer loop perturbs hyperparameters. This is captured by a two-timescale stochastic approximation model where the hyperparameters evolve according to

$$\theta_{t+1} = \theta_t + \gamma_t \, u_t, \qquad \gamma_t \ll \eta_t,$$

with η_t the inner-loop step size and u_t chosen by the acquisition. Stability requires $\sum_t \gamma_t = \infty$ and $\sum_t \gamma_t^2 < \infty$, while η_t decreases more slowly, reflecting faster convergence of tap adaptation relative to hyperparameter exploration [24]. In practice, this implies updating θ once per several coherence intervals, and freezing it within intervals to avoid confounding measurements with transient adaptation dynamics.

When training data from emulation or lab captures is available, a warm-start surrogate accelerates field convergence. Offline, one populates $\mathcal{D}_0$ with evaluations over synthetic channels whose statistics bracket expected deployments, optionally augmented by importance weighting toward regions of higher operational likelihood. The online phase then refines the surrogate using over-the-air data, with a robustification layer that downweights points whose residuals exceed what the kernel predicts under its noise model, a sign of distribution shift. Mathematically, a Student-t likelihood replaces the Gaussian noise model to accommodate occasional heavy-tailed deviations,

$$p(y|\theta) \propto \left(1 + \frac{(y - \mu(\theta))^2}{\nu \, \sigma^2(\theta)}\right)^{-(\nu+1)/2},$$

with degrees of freedom ν tuned to observed tail behavior.

Scalable implementation relies on sparse or low-rank kernel approximations. Inducing-point methods reduce cubic complexity $O(t^3)$ to $O(tm^2)$ with m pseudo-inputs, while random Fourier features linearize the surrogate in an expanded feature space. Both integrate naturally with batched parallelism on accelerators and map to fixed-point arithmetic by pre-quantizing kernel features. Numerical stability hinges on regularizing Gram matrices and monitoring condition numbers; jitter terms added to the diagonal prevent ill-conditioning when observations cluster [25].

Acquisition optimization itself is nontrivial; gradient-based searches over θ require differentiable kernels and cost models, while gradient-free global search strategies such as covariance matrix adaptation or direct space exploration provide robustness when θ contains discrete elements.

Practical objective choices determine whether the optimizer prioritizes throughput, reliability, or energy. For instance, a decoder-in-the-loop frame-error proxy with dynamic framing time $T(\theta)$ calls for a normalized objective $f(\theta) = -\mathrm{FER}(\theta)/T(\theta)$ that captures errors avoided per unit time. Alternatively, a constrained throughput maximization

$$\max_{\theta} \ \mathrm{SE}(\theta) \quad \text{s.t.} \quad \mathrm{BLER}(\theta) \leq \epsilon, \ \ T(\theta) \leq T_{\max},$$

is operated via a Lagrangian that the surrogate learns implicitly. The optimizer then tunes feed-forward bandwidth and feedback aggressiveness to trade slight increases in residual interference against reductions in iteration count or frame duration that unlock higher spectral efficiency at fixed block-error targets.

Within this broad landscape, specific studies illustrate the feasibility of probabilistic search for joint feed-forward and decision-feedback tuning. Methods that select hyperparameters by maximizing eye-related metrics like eye height or eye opening are convenient when the receiver exposes intuitive oscilloscopic diagnostics and when objective evaluation without the decoder is necessary to tighten the exploration budget. A Bayesian approach that places a surrogate over such metrics and queries them under uncertainty exemplifies the general template; for instance, Dikhaminjia et al (2021) [26]. report a probabilistic strategy that leverages an eye-height objective to drive joint equalizer configuration using sequential optimization, providing one concrete instantiation among many compatible designs.

10 Optimization Methods and Numerical Stability

Equalization algorithms often reduce to solving structured linear systems or composite convex programs. Preconditioned conjugate gradients solve normal equations $(\mathbf{H}^{\mathrm{H}}\mathbf{H} + \alpha\mathbf{I})\mathbf{x} = \mathbf{H}^{\mathrm{H}}\mathbf{y}$ with a circulant preconditioner $\mathbf{C}$ approximating $\mathbf{H}^{\mathrm{H}}\mathbf{H}$. The iteration uses

$$\mathbf{r}_0 = \mathbf{b} - \mathbf{A}\mathbf{x}_0, \quad \mathbf{z}_0 = \mathbf{C}^{-1}\mathbf{r}_0, \quad \mathbf{p}_0 = \mathbf{z}_0,$$

and updates

$$\alpha_t = \frac{\mathbf{r}_t^{\mathrm{H}}\mathbf{z}_t}{\mathbf{p}_t^{\mathrm{H}}\mathbf{A}\mathbf{p}_t}, \quad \mathbf{x}_{t+1} = \mathbf{x}_t + \alpha_t\mathbf{p}_t, \quad \mathbf{r}_{t+1} = \mathbf{r}_t - \alpha_t\mathbf{A}\mathbf{p}_t,$$

$$\mathbf{z}_{t+1} = \mathbf{C}^{-1}\mathbf{r}_{t+1}, \quad \beta_t = \frac{\mathbf{r}_{t+1}^{\mathrm{H}}\mathbf{z}_{t+1}}{\mathbf{r}_t^{\mathrm{H}}\mathbf{z}_t}, \quad \mathbf{p}_{t+1} = \mathbf{z}_{t+1} + \beta_t\mathbf{p}_t.$$

FFT-based multiplications reduce per-iteration cost to quasi-linear in block length, and residual-based stopping ensures numerical robustness with fixed iteration caps for latency predictability.

Alternating direction method of multipliers handles constraints such as constant-modulus, sparsity in transform domains, or bounded peak-to-average power. For

$$\min_{\mathbf{x}} \ f(\mathbf{x}) + g(\mathbf{K}\mathbf{x}),$$

the scaled ADMM iterations

$$\mathbf{x}^{t+1} = \arg\min_{\mathbf{x}} \ f(\mathbf{x}) + \frac{\rho}{2}\|\mathbf{K}\mathbf{x} - \mathbf{z}^t + \mathbf{u}^t\|_2^2,$$

$$\mathbf{z}^{t+1} = \mathrm{prox}_{g/\rho}\left(\mathbf{K}\mathbf{x}^{t+1} + \mathbf{u}^t\right), \quad \mathbf{u}^{t+1} = \mathbf{u}^t + \mathbf{K}\mathbf{x}^{t+1} - \mathbf{z}^{t+1}$$

decompose computation into a structured least-squares subproblem and a simple proximal map [27]. Parameter ρ is tuned via residual balancing rules, and over-relaxation improves convergence on ill-conditioned problems.

Fixed-point deployment requires dynamic range analysis. Given accumulator word length B and unit roundoff $u = 2^{-B}$, forward error bounds for a triangular solve with condition number κ yield relative error scaling approximately as

$$\frac{\|\hat{\mathbf{x}} - \mathbf{x}\|_2}{\|\mathbf{x}\|_2} \lesssim \kappa\, c\, u + O(u^2),$$

for constant c determined by algorithmic details. Mixed-precision approaches perform multiplications in low precision and accumulations in higher precision, preserving accuracy while saving energy. Stochastic rounding reduces bias in iterative solvers and stabilizes convergence in deep unrolling.

11 Complexity, Memory Traffic, and Energy Considerations

Throughput and energy per bit depend not only on arithmetic counts but also on memory traffic. For block length N and banded interference of half-bandwidth B, a banded conjugate-gradient equalizer has per-iteration complexity $O(NB)$ and memory proportional to $O(N)$, with a constant sensitive to caching. In massive MIMO, factorization-based linear solvers rely on QR or Cholesky updates whose complexity scales with $O(MK^2)$ or $O(K^3)$ depending on reuse [28]. Iterative solvers amortize cost across coherence intervals by warm-starting from previous solutions, reducing required iterations by a factor that correlates with channel temporal correlation ρ.

Energy models parameterize cost by picojoules per operation and per byte moved off-chip. For a given technology node, off-chip accesses may consume orders of magnitude more energy than arithmetic. Equalizers designed to maximize data locality, such as polyphase filters and blocked FFTs, lower energy by reusing intermediate results on-chip. Pipelining and systolic array scheduling ensure high utilization under strict latency constraints, and stream buffers exploit alignment between symbol boundaries and processing tiles.

Latency constraints in control-plane signaling motivate bounded-iteration algorithms with predictable worst-case execution. Techniques include early termination on reliable symbols, where symbols exceeding a reliability threshold bypass further iterations, and partial updates that focus computational effort on subcarriers or users with poor conditioning. In hybrid beamforming, precomputation of analog-domain transforms reduces online cost, while adaptive baseband dimensioning tailors digital equalization effort to instantaneous channel rank. [29]

12 Simulation Methodology and Performance Characterization

Performance evaluation proceeds by generating channels according to specified scattering functions, sampling Doppler processes consistent with mobility and oscillator behavior, and adding thermal noise at prescribed SNRs. For single-carrier frequency-selective channels, taps follow complex Gaussian distributions with exponential power delay profiles. For OFDM, phase noise follows a discrete-time Wiener process with linewidths that map to normalized phase variance across the symbol. Delay–Doppler channels adopt sparse support patterns with clusters of reflectors and micro-Doppler components.

Equalizer configurations include linear MMSE with banded interference models, decision feedback with soft cancellation, turbo equalization with iterative decoding, AMP-based detectors with matched denoisers, and unrolled proximal networks trained on synthetic datasets. Metrics include block error rate, bit error rate, and achievable rates under Gaussian approximation of residual interference. Complexity is tracked via operation counts, iteration counts, and memory traffic proxies, while latency is measured per frame based on cycle-accurate models.

Representative outcomes reveal operating regions where each equalizer class is advantageous [30]. Linear MMSE performs well when interference is moderate and channel estimates are accurate. Decision feedback gains manifest in channels with dominant postcursor energy and moderate error propagation risk. Turbo equalization closes much of the gap to optimal detection when coding is strong and latency budgets permit multiple exchanges. AMP-based detectors excel in large, moderately loaded systems with near-i.i.d. mixing, while unrolled methods provide improved convergence in ill-conditioned regimes. Under one-bit quantization, likelihood-aware message passing significantly improves error floors relative to linearization-based approaches.

13 Design Guidelines and Open Technical Questions

Several practical design guidelines emerge from the mathematical analysis and empirical characterization of adaptive equalizers in contemporary communication systems. Conditioning analysis, often expressed through singular value spectra or random-matrix equivalents, provides crucial insight into the stability and convergence behavior of iterative solvers [31]. When the effective channel matrix exhibits ill-conditioning, small perturbations in noise or estimation errors can produce disproportionately large fluctuations in the equalized output. Regularization thus becomes essential to mitigate noise amplification, and the optimal level can be inferred from the distribution of singular values or their large-system approximations. Similarly, understanding the condition number dynamics informs the appropriate number of iterations in conjugate-gradient or message-passing algorithms, balancing convergence accuracy against computational latency. In large-scale systems, random-matrix theory offers asymptotically accurate surrogates for these conditioning metrics, allowing regularization schedules to be set adaptively without exhaustive per-channel computation.

In architectures employing hybrid analog–digital beamforming, where the analog front end performs a dimensionality reduction before digital baseband processing, the equalizer must operate in a compressed domain. This compression can obscure channel features and introduce ambiguities if the analog beam patterns do not adequately span the relevant signal subspaces. To counteract this, the use of auxiliary pilot symbols aligned with analog beams proves beneficial. These pilots illuminate the effective channel within the reduced-dimensional space, improving identifiability and enhancing the conditioning of subsequent equalization [32]. Careful design of pilot allocation—accounting for analog beam directions, gain variations, and phase quantization constraints—ensures that the hybrid equalizer can recover sufficient information to compensate for analog-domain distortions without incurring excessive pilot overhead.

In systems characterized by delay–Doppler spreading, such as those employing orthogonal time–frequency space (OTFS) modulation, channel responses become sparse in a joint delay–Doppler representation. Exploiting this sparsity through structured priors or regularization promotes resilience against rapid time variation and frequency selectivity. Bayesian or compressed-sensing-inspired equalizers that enforce sparsity in the delay–Doppler domain can isolate dominant propagation paths and suppress diffuse interference, resulting in robust performance even under high mobility. These structured models naturally lend themselves to online updates through iterative thresholding or message-passing algorithms that adaptively track the sparse support as the channel evolves.

Despite these advances, several open research directions remain. One challenge lies in developing unified channel and impairment models that simultaneously capture oscillator phase noise, quantization effects, and nonlinear power amplifier memory, all within a tractable inferential framework. Each of these impairments introduces distinct forms of nonlinearity and correlation, and their combined effect complicates both analysis and algorithm design [33]. Achieving a model that is sufficiently expressive yet still admits low-complexity inference would enable more accurate equalization in practical transceivers operating under stringent power and hardware constraints. Another promising direction involves adaptive algorithms that dynamically adjust their iteration depth and operating mode based on instantaneous quality metrics, such as estimated post-equalization signal-to-interference-plus-noise ratio (SINR). Such adaptive control could conserve computational resources during favorable channel conditions while allocating more iterations when interference or distortion intensifies, achieving graceful complexity scaling with environmental difficulty.

The advent of learning-augmented equalizers introduces further possibilities for optimization beyond traditional arithmetic efficiency. Hardware-aware training that explicitly minimizes memory access and data movement—often the dominant energy consumers in modern digital signal processors—can yield substantial power savings. Incorporating memory-traffic cost functions into training objectives aligns algorithmic learning with hardware realities, promoting implementations that are not only accurate but also energy proportional. This co-design philosophy bridges the gap between theoretical performance metrics and practical deployment constraints, enabling sustainable operation in battery-powered or thermally limited platforms.

A deeper theoretical and architectural question concerns the interplay between equalization and medium-access control in emerging cell-free and distributed MIMO systems [34]. In these networks, coordination among distributed access points alters the interference landscape, while fronthaul compression and scheduling decisions affect the structure and dimensionality of the received signals. Equalization strategies must therefore be aware of such upper-layer dynamics, potentially adapting their model assumptions and prior distributions to reflect varying degrees of cooperation, latency, and data fidelity. Joint design of equalization and resource allocation could exploit these dependencies, leading to more coherent multiuser interference mitigation and improved spectral efficiency across distributed topologies.

From an optimization standpoint, the convergence behavior of unrolled or learned iterative equalizers remains an area of active inquiry. While these architectures achieve impressive empirical performance by mapping iterative inference steps onto differentiable network layers, their theoretical properties under realistic channel ensembles are not yet well understood. Establishing global convergence guarantees—or at least identifying conditions under which convergence to desirable fixed points is ensured—would significantly strengthen their reliability. Moreover, the development of state-evolution-like analytical tools for correlated, rotationally invariant channel models would provide predictive performance benchmarks akin to those available for idealized i.i.d. scenarios. Such predictors could guide hyperparameter tuning, regularization schedules, and iteration control with reduced reliance on empirical validation. [35]

14 Conclusion

References

[1] Q. Wei and B. Yang, "Adaptable vehicle detection and speed estimation for changeable urban traffic with anisotropic magnetoresistive sensors," *IEEE Sensors Journal*, vol. 17, pp. 2021–2028, 4 2017.

[2] A. S. Tomar and P. Jayaswal, "A hybrid fault diagnosis method using translation invariant wavelet denoising, hierarchical entropy, and support vector machine with pso algorithm," *Traitement du Signal*, vol. 39, pp. 2041–2053, 12 2022.

[3] S. Abdel-Hafeez and A. Gordon-Ross, "Reconfigurable fifo memory circuit for synchronous and asynchronous communication," *International Journal of Circuit Theory and Applications*, vol. 49, pp. 938–952, 1 2021.

[4] "in brief," *Electronics Letters*, vol. 53, pp. 442–442, 3 2017.

[5] T.-H. Liu, Y. Chen, M.-J. Wu, and B.-C. Dai, "Adaptive controller for an mtpa ipmsm drive system without using a high-frequency sinusoidal generator," *The Journal of Engineering*, vol. 2017, pp. 13–25, 2 2017.

[6] J. Li, Y. Chen, P. Tang, Z. Zhang, H. Wang, and H. Huang, "A low-power low-distortion 20-gs/s flash analog-to-digital converter for coherent optical receiver in 0.13-m sige bicmos," *Journal of Circuits, Systems and Computers*, vol. 28, pp. 1950167–, 9 2019.

[7] A. N. Z. Rashed and M. S. F. Tabbour, "Best candidate integrated technology for low-noise, high-speed, and wide bandwidth-based transimpedance amplifiers in optical computing systems and optical fiber applications," *International Journal of Communication Systems*, vol. 31, 9 2018.

[8] A. M. Yusof, Z. Abdul-Malek, and C. L. Wooi, "Lightning generated electric field over land and sea at northern region of peninsular malaysia: Measuring setup," *Indonesian Journal of Electrical Engineering and Informatics (IJEEI)*, vol. 8, pp. 457–464, 9 2020.

[9] I. S. Amiri, A. N. Z. Rashed, and P. P. Yupapin, "High-speed transmission circuits signaling in optical communication systems," *Journal of Optical Communications*, vol. 45, pp. s69–s75, 11 2019.

[10] L. Xiaoli and A. S. Mohd, "Implementation of embedded unspecific continuous english speech recognition based on hmm," *Recent Advances in Electrical & Electronic Engineering (Formerly Recent Patents on Electrical & Electronic Engineering)*, vol. 14, pp. 649–659, 11 2021.

[11] J. Kim, Y. G. Kim, J. Lee, and S. Kim, "Intended gait speed prediction via swing limb kinematics applicable for patients with slow gait," *Electronics Letters*, vol. 56, pp. 911–914, 9 2020.

[12] J. Ma, Z. Song, Y. Zhang, Y. Shen, and J. S. Thorp, "Model order reduction analysis of dfig integration on the power system small-signal stability considering the virtual inertia control," *IET Generation, Transmission & Distribution*, vol. 11, pp. 4087–4095, 10 2017.

[13] H. B. Valiveti and B. A. Kumar, "Handoff strategies between wireless fidelity to light fidelity systems for improving video streaming in high-speed vehicular networks," *International Journal of Communication Systems*, vol. 34, 1 2020.

[14] S. K. Kondoju and V. V. Mani, "Outage and ber analysis of dual-carrier modulation over frequency-selective nakagami-m fading channels," *COMPEL - The international journal for computation and mathematics in electrical and electronic engineering*, vol. 36, pp. 90–107, 1 2017.

[15] R. I. Sabitu, N. Khan, and A. Malekmohammadi, "40 gb/s high-speed mode-division multiplexing transmission employing nrz modulation format," *Journal of Optical Communications*, vol. 44, pp. s1157–s1165, 12 2020.

[16] M. Libin, K.-J. Xu, L. Bo, T. Lei, and L.-P. Liang, "Echo signal envelope fitting based signal processing methods for ultrasonic gas flow-meter.," *ISA transactions*, vol. 89, pp. 233–244, 12 2018.

[17] S. Wang, X. Xu, G. Zhou, Y. Chen, W. He, Y. Wenjun, X. Su, and Y. Hu, "Effects of microstrip line fabrication and design on high-speed signal integrity transmission of pcb manufacturing process," *Circuit World*, vol. 44, pp. 53–59, 5 2018.

[18] N. Kulchitsky, M. Automatics, A. Naumov, V. Startsev, and L. A. D. B. Jsc, "Development trends of infrared focal plane array detectors," *Nano- i Mikrosistemnaya Tehnika*, vol. 22, pp. 500–510, 12 2020.

[19] W. Zhang, W. Zhang, and D. Hao, "Adjustable coupled optoelectronic oscillator based on double frequency output," *Journal of Nanoelectronics and Optoelectronics*, vol. 15, pp. 769–776, 6 2020.

[20] B.-F. Qian, S.-J. Gao, Q.-F. Li, Q. Zhang, and Y. Wang, "A fast adaptive beamforming algorithm based on gram-schmidt orthogonalization," *Journal of Nanoelectronics and Optoelectronics*, vol. 16, pp. 642–650, 4 2021.

[21] J.-H. Park, D.-Y. Kong, C.-Y. Lee, C.-S. Kim, and J.-M. Kim, "Controller design considering operation characteristics of high-speed train hatch system," *The transactions of The Korean Institute of Electrical Engineers*, vol. 71, pp. 1923–1928, 12 2022.

[22] R. Katuri and S. Gorantla, "Modelling and analysis of a hybrid controller applied to the ultracapacitor based solar powered electric vehicle," *Modelling, Measurement and Control A*, vol. 91, pp. 114–122, 9 2018.

[23] E. Khoudry, A. Belfqih, T. Ouaderhman, J. Boukherouaa, and F. Elmariami, "A real-time fault diagnosis system for high-speed power system protection based on machine learning algorithms," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 10, pp. 6122–6138, 12 2020.

[24] S. M. Asad, P. V. Klaine, R. N. B. Rais, S. Hussain, Q. H. Abbasi, and M. A. Imran, "Blockchain-enabled handover skipping for high mobility train passengers," *Electronics Letters*, vol. 58, pp. 978–981, 10 2022.

[25] M. U. Hadi, N. Aslam, and H. Jung, "Performance appraisal of sigma delta modulated radio over fiber system," *Journal of Optical Communications*, vol. 43, pp. 273–279, 1 2019.

[26] N. Dikhaminjia, G. Tsintsadze, Z. Kiguradze, J. He, M. Tsiklauri, J. Drewniak, A. Chada, and B. Mutnury, "Optimization of joint equalization of high-speed signals using bayesian machine learning," in *2021 IEEE International Joint EMC/SI/PI and EMC Europe Symposium*, pp. 48–52, IEEE, 2021.

[27] X. Tang, Y. Xu, X. Sun, Y. Liu, Y. Jia, F. Gu, and A. D. Ball, "Intelligent fault diagnosis of helical gearboxes with compressive sensing based non-contact measurements.," *ISA transactions*, vol. 133, pp. 559–574, 7 2022.

[28] Y. Liu, F. Cao, X. Xiong, J. Huang, and W. Deng, "Parameter extraction for small-signal model of<scp>gan hemts</scp>on<scp>sic</scp>substrates based on modified firefly algorithm," *International Journal of RF and Microwave Computer-Aided Engineering*, vol. 32, 10 2022.

[29] A. R. Danesh and M. Habibi, "A signed pulse-train based image processor-array for parallel kernel convolution in vision sensors," *Sensor Review*, vol. 40, pp. 521–528, 6 2020.

[30] Y. Shi, Y. Yang, Y. Ruan, Y. Liao, and B. Zhang, "Track image acquisition based on the signal device of the photoelectric encoder," *Journal of Nanoelectronics and Optoelectronics*, vol. 18, pp. 184–192, 2 2023.

[31] X. Gao, H. Zhang, and T. Dang, "Radar detection and fast motion parameter estimation for complex manoeuvering targets at high speed and acceleration," *IET Radar, Sonar & Navigation*, vol. 16, pp. 1977–1996, 8 2022.

[32] L. Singkang, K. A. H. Ping, and P. R. P. Hoole, "Electric discharges localization for substation fault monitoring using two elements sensor," *Journal of Computational and Theoretical Nanoscience*, vol. 17, pp. 1009–1013, 2 2020.

[33] M. Toulabi, S. Bahrami, and A. M. Ranjbar, "Optimal supplementary frequency controller design using the wind farm frequency model and controller parameters stability region.," *ISA transactions*, vol. 74, pp. 175–184, 3 2018.

[34] J. Zhang, J. Li, H. Ma, and Y. Wang, "High-speed moving target detection and tracking method based on speed estimation," *The Journal of Engineering*, vol. 2019, pp. 6637–6641, 9 2019.

[35] R. Eswaramoorthi, M. L. Moses, J. B. S. Hameed, and B. Ghanti, "Deep graph neural network optimized with fertile field algorithm based detection model for uplink multiuser massive multiple-input and multiple-output system," *Transactions on Emerging Telecommunications Technologies*, vol. 33, 7 2022.