Fourierstudies Transactions on Artificial Intelligence, Machine Learning, and Cognitive Systems

ARTICLE

Integration of Bayesian Inference Principles in the Development of Modern Adaptive Equalization Algorithms

Rohit Deshmukh[1] **and Sandeep Varma**[2]

[1] Visvesvaraya Institute of Technology and Communication, Department of Electronics and Communication Engineering, 21 MG Road, Solapur, Maharashtra, India

[2] Rayalaseema University of Engineering, Department of Electrical and Electronics Engineering, 74 Srinivasa Nagar Main Road, Kurnool, Andhra Pradesh, India

Abstract

Adaptive equalization has long been a central mechanism for mitigating intersymbol interference and time variability in communication channels, with classical approaches relying on stochastic gradient and recursive least-squares recursions. Over the last decades, Bayesian inference has provided a complementary perspective that treats equalizer coefficients, latent channel states, and nuisance parameters as random variables with structured prior distributions. This perspective enables uncertainty quantification, principled regularization, and data-driven adaptation without ad hoc tuning. The present study examines the integration of Bayesian principles into modern adaptive equalization design under practical constraints of nonstationary propagation, non-Gaussian disturbances, finite-precision arithmetic, and stringent latency budgets. The discussion develops a modeling pipeline that connects parametric and nonparametric priors to state-space channel descriptions, explores online posterior inference via conjugate updates, Kalman-type filters, particle methods, and variational approximations, and evaluates robustness through heavy-tailed likelihood models and scale-mixture priors. Emphasis is placed on sparse and structured representations that align with wideband and millimeter-wave propagation, on hyperparameter learning through online evidence maximization, and on computational architectures that exploit matrix identities, low-rank structure, and streaming operators. The study also sketches links to coded and multicarrier systems, multiantenna processing, and differentiable implementations that borrow from probabilistic deep learning while preserving posterior interpretability. The treatment aims to remain neutral regarding algorithmic preferences, highlighting trade-offs among statistical fidelity, stability, and complexity, and it underscores regimes in which Bayesian equalizers can complement or subsume classical rules through calibrated uncertainty and regularized adaptation.

1 Introduction

Adaptive equalization represents a cornerstone technique in modern digital communication systems, serving as a dynamic countermeasure to channel-induced distortions that manifest due to multipath propagation, Doppler shifts, and other time-varying impairments [1]. The essential goal is to recover the transmitted symbol sequence from the received signal, which has been convolved with the channel's impulse response and contaminated by noise. Because practical channels exhibit both frequency selectivity and temporal variability, a fixed equalizer cannot suffice; instead, an adaptive mechanism is required that can continually refine its parameters to track the evolving channel characteristics. This process hinges on the estimation of an operator that effectively inverts or whitens the channel, thereby mitigating intersymbol interference (ISI). The equalizer must perform this estimation in real time, using streaming data that may come from known training sequences or from decision-directed operation based on previously detected symbols.

Classical Adaptive Equalizers

Algorithm	Principle	Remarks
LMS	Minimizes MSE via stochastic gradient descent	Simple, low cost; sensitive to step size
RLS	Minimizes exponentially weighted error sum	Fast convergence; higher complexity
[2] Kalman Filter	Bayesian recursive estimation (Gaussian case)	Optimal linear MMSE; high computational demand

Bayesian Equalization Components

Component	Description
Prior	Encodes beliefs on coefficients (e.g., AR(1), sparsity)
Likelihood	Relates received data to transmitted symbols and noise
Posterior	Updated beliefs combining prior and new evidence

Posterior Inference Methods

Method	Key Features
Variational Inference	Approximates posterior by a simpler family; efficient
[3] Particle Filtering	Captures nonlinearity via sampling
Moment Matching / EP	Low-rank, tractable approximations for real-time use

Traditional approaches to adaptive equalization, such as those employing stochastic gradient descent (SGD) or recursive least squares (RLS) updates, embody specific optimization principles and implicit statistical assumptions. The LMS (least mean squares) algorithm, for instance, seeks to minimize the mean squared error between the equalized output and the desired symbol, updating its weights along the negative gradient of this error surface. Its simplicity and low computational cost make it appealing for many applications, but its performance depends critically on the choice of step size, which trades off convergence speed against stability. The RLS algorithm, on the other hand, minimizes a weighted sum of squared errors over past data, yielding faster convergence at the expense of higher complexity. These classical methods can be viewed as particular instances of recursive estimation under certain probabilistic assumptions, even if such assumptions are not explicitly stated in their derivation. [4]

By recasting adaptive equalization within a Bayesian framework, one gains a unifying and principled perspective on these algorithms and their limitations. In the Bayesian view, all unknown quantities—including channel coefficients, equalizer taps, noise levels, and possibly even the transmitted symbols themselves—are treated as random variables with associated prior distributions. The measurement process is described probabilistically by a likelihood function that connects the observed received samples to these latent variables. The objective then becomes to infer the posterior distribution of the unknowns given the observed data. This posterior encapsulates all available information and uncertainty about the parameters, allowing not only point estimates (such as the mean or mode) but also credible intervals and predictive distributions that quantify confidence in the equalizer's decisions.

The Bayesian formalism clarifies several aspects of adaptive equalization that are often implicit in heuristic algorithms. For instance, the step-size parameter in LMS can be interpreted as a proxy for the prior variance on the coefficients or as a measure of how rapidly one expects the channel to vary [5]. Similarly, the exponential forgetting factor in RLS corresponds to a prior belief about the temporal correlation of the coefficients, controlling how much weight recent observations carry relative to older ones. Through this lens, both LMS and RLS can be understood as performing approximate posterior updates under Gaussian assumptions, with fixed or adaptive covariance structures. The Bayesian approach also enables a natural incorporation of non-Gaussian noise models, such as heavy-tailed or impulsive noise, through appropriate likelihood choices.

The structure of a Bayesian adaptive equalizer typically consists of three components: the prior, the likelihood, and the posterior inference mechanism. The prior specifies statistical beliefs about the unknown quantities before observing data—for example, that channel coefficients evolve according to a first-order autoregressive process with known variance, or that equalizer weights have sparse structure due to limited channel support. The likelihood encodes the measurement model, relating received symbols to transmitted ones through the channel and additive noise. The posterior, obtained via Bayes' rule, combines these two sources of information to yield updated beliefs after each observation. In an online context, where symbols arrive sequentially, posterior inference must be performed recursively to remain computationally tractable. [6]

Several algorithmic families can be derived from this probabilistic foundation. When both the prior and likelihood are Gaussian and linear relationships hold, the Bayesian recursion reduces to a Kalman filter, whose update equations provide optimal linear minimum mean-square estimates. In more complex cases, such as nonlinear equalization or non-Gaussian noise, approximate methods are necessary. Variational inference can be used to project the intractable posterior onto a simpler family of distributions, yielding update rules that generalize traditional gradient methods. Alternatively, sequential Monte Carlo (particle filtering) approaches can approximate the posterior using a weighted ensemble of samples, which naturally captures multimodality and nonlinearity at the cost of increased computation. The choice among these methods depends on latency, memory, and arithmetic constraints imposed by the communication system.

The advantages of a Bayesian approach are not limited to estimation accuracy [7]. It provides a coherent framework for uncertainty quantification, enabling adaptive control of decision thresholds and confidence-based modulation or coding adaptation. In decision-directed equalization, where detected symbols are fed back to guide future updates, explicit modeling of uncertainty can mitigate error propagation by discounting unreliable decisions. Bayesian reasoning also facilitates principled model selection and hyperparameter tuning, as priors and likelihoods can be calibrated to empirical data or physical channel knowledge rather than adjusted through ad hoc parameter sweeps.

Moreover, the Bayesian interpretation highlights connections between adaptive equalization and broader classes of machine learning and signal processing methods. For example, stochastic gradient updates correspond to stochastic variational inference under particular factorization assumptions, while RLS bears resemblance to natural gradient descent in information geometry. Regularization techniques such as ℓ_1 or ℓ_2 penalties emerge as log-prior terms in the posterior objective, thereby linking optimization-based and probabilistic formulations. This perspective fosters the design of hybrid algorithms that combine the interpretability and efficiency of classical filters with the flexibility of probabilistic inference. [8]

In practical implementations, the translation of Bayesian principles into real-time adaptive equalizers requires careful attention to computational feasibility. Exact inference may be intractable for high-dimensional systems or when symbol rates are extremely high. Therefore, approximate filtering recursions—such as those based on moment matching, expectation propagation, or low-rank covariance updates—are employed to balance fidelity and efficiency. The resulting algorithms retain the interpretive advantages of Bayesian reasoning while operating within stringent hardware limits. In wireless receivers, for instance, simplified Bayesian equalizers can be implemented using recursive updates that approximate posterior means and variances, achieving performance gains over purely deterministic schemes with only modest added complexity.

The stability and robustness of Bayesian adaptive equalizers also benefit from the explicit modeling of uncertainty. Because posterior updates depend on both prior beliefs and observed evidence, the system can gracefully handle transient degradations or abrupt channel changes. If the data suddenly deviate from prior expectations, the posterior variance inflates, signaling reduced confidence and prompting the equalizer to adapt more cautiously. Such self-calibrating behavior is especially valuable in mobile and multipath environments where channel statistics evolve unpredictably. Furthermore, under model mismatch or finite data conditions, Bayesian estimators tend to exhibit superior generalization, as the prior regularization prevents overfitting to noise or transient phenomena.

The following sections construct a modeling path from channel representations and priors to posterior updates and algorithmic forms. The discussion considers sparse and clustered responses, low-rank multiantenna couplings, and wideband operators, with emphasis on uncertainty-aware adaptation. Robustness is developed through heavy-tailed models and scale-mixture constructions that temper the influence of outliers and symbol decision errors. Hyperparameter learning is framed as online evidence maximization with safeguards for overfitting and drift [9]. Computational architectures exploit conjugacy, matrix identities, and structured operators to maintain constant or slowly growing per-sample cost while accommodating large tap counts, multicarrier transforms, and pilot-design constraints. Throughout, the analysis avoids exaggerated claims of dominance and instead develops conditions under which Bayesian formulations offer clear interpretability and well-calibrated updates that can complement classical approaches in practical receivers.

2 Bayesian Modeling Principles for Equalization

A Bayesian equalizer begins by defining random elements for channel responses, equalizer taps, and noise terms. Let the transmitted symbol sequence be represented as a complex process with alphabet-induced moment constraints, and consider a linear time-varying operator that maps input samples to received observations. The probabilistic structure is completed by a likelihood that encodes observation noise, quantization effects if present, and decision-directed uncertainty during blind or semi-blind adaptation. Priors reflect beliefs about sparsity, smoothness, or low-dimensional structure, and they may be hierarchical to enable data-driven regularization.

The discrete-time baseband observation model with symbol-period sampling can be summarized as a convolutional regression whose dimensionality can be controlled through windowing and tap truncation. A representative expression for a segment of observations is given by [10]

$$\mathbf{y}_t = \mathbf{X}_t \mathbf{h}_t + \mathbf{n}_t,$$

where $\mathbf{y}_t$ stacks m consecutive received samples at time t, $\mathbf{X}_t$ collects shifted input symbols (known during pilot intervals and random during decision-directed operation), $\mathbf{h}_t$ is a vector of effective channel taps or a composite representing the cascade of channel and front-end filtering, and $\mathbf{n}_t$ is a disturbance term. If the channel is considered slowly time varying, a state transition captures drift:

$$\mathbf{h}_t = \mathbf{F}_t \mathbf{h}_{t-1} + \mathbf{w}_t,$$

with process noise $\mathbf{w}_t$ modeling Doppler-induced evolution or unmodeled reception effects. Priors on $\mathbf{h}_t$ can be Gaussian to induce quadratic penalties, Laplace or Student constructions to encourage sparsity, or Gaussian process models to encode smoothness across tap index or across time. A conjugate Gaussian prior with covariance $\boldsymbol{\Sigma}_0$ yields closed-form posteriors under Gaussian noise. A scale-mixture representation transforms heavy-tailed priors and likelihoods into conditionally Gaussian forms, preserving tractability through augmented variables.

Designing a Bayesian equalizer involves choosing an estimator from the posterior such as a mean for minimum mean-square error, a mode for maximum a posteriori, or a predictive distribution for symbol decisions. The predictive distribution is central to uncertainty calibration, as it quantifies the dispersion of equalized outputs under posterior uncertainty in $\mathbf{h}_t$ and noise variance. Calibration matters when decision feedback introduces bias in training data; the Bayesian formulation accommodates this by propagating uncertainty from decisions to subsequent updates through likelihood tempering or latent-variable augmentation that represents unknown or soft-labeled symbols during decision-directed stages.

3 Channel Representation and Prior Structures

The channel representation determines the structural assumptions that priors should exploit. In rich multipath environments, the effective impulse response exhibits clusters of significant taps separated by near-zero regions. A sparse prior accommodates this with independent or group-structured shrinkage [11]. A widely used construction is a hierarchical Gaussian scale mixture in which each tap $h_{t,k}$ has a zero-mean Gaussian prior whose variance is itself random, yielding automatic relevance determination across tap indices. The model reads

$$h_{t,k} \,|\, \alpha_{t,k} \sim \mathcal{N}(0, \alpha_{t,k}^{-1}), \quad \alpha_{t,k} \sim \mathrm{Gamma}(a_0, b_0),$$

so that integrating out $\alpha_{t,k}$ induces heavy-tailed marginal behavior that favors many small coefficients and a few large ones. Grouped structures can be introduced by assigning a common scale to blocks of taps or by coupling adjacent scales through Markov or

Gaussian process priors across delay index. When the channel varies over time, a state-space prior captures correlation across t. For smooth drift, an autoregressive prior of order one with small process variance is natural, while for bursty changes a switching model with a latent regime index allows abrupt transitions.

Nonparametric priors provide flexibility for modeling long-delay dispersion or frequency selectivity without fixing the number of taps. A Gaussian process prior across delay index with kernel $k(\ell, \ell')$ allows control of smoothness and decay, and when combined with a state evolution across time it yields a separable spatiotemporal covariance. A Kronecker factorization of the covariance across delay and time indices, [12]

$$\mathrm{cov}[\mathrm{vec}(\mathbf{H})] = \mathbf{K}_\tau \otimes \mathbf{K}_t,$$

enables scalable inference by leveraging eigen decompositions or iterative solvers. For multicarrier systems, a frequency-domain representation can be preferable. Priors that are diagonal or banded in the discrete Fourier basis produce computational advantages through fast transforms and suggest evidence-driven regularization that varies across subcarriers to reflect differing signal-to-noise ratios. In multiantenna settings, joint priors across spatial channels encode low-rank coupling arising from common scattering clusters. A matrix-variate Gaussian or a factor-analytic prior

$$\mathbf{H}_t = \mathbf{A}\mathbf{B}_t^\top + \mathbf{E}_t$$

with small latent dimension captures shared structure while allowing per-link deviations through $\mathbf{E}_t$. These structures directly influence the conditioning of posterior updates, dictate the complexity of linear algebra operations, and shape the uncertainty that propagates to decision metrics.

4 Posterior Inference Mechanisms for Online Adaptation

Posterior Inference for Online Adaptation

Model Type	**Mechanism / Key Idea**
Gaussian Linear	Kalman recursion with $(\mathbf{F}_t, \mathbf{Q}_t, \mathbf{R}_t)$; RLS equivalence when $\mathbf{F}_t = \mathbf{I}$
Heavy-tailed / Sparse	Conditioned Gaussian with inverse-gamma mixing (Student-t) or EM / variational updates
Nonlinear / Non-Gaussian	Particle filters with resampling; Rao–Blackwellized hybrids for low latency
Variational	Online ELBO optimization with natural gradients and annealing for stability

Posterior inference must operate under streaming constraints while remaining stable under decision-directed noise and model mismatch. Conjugate Gaussian models with linear dynamics yield Kalman-type recursions for the channel state. With observation matrix $\mathbf{X}_t$ and process and noise covariances $\mathbf{Q}_t$ and $\mathbf{R}_t$, the filtering steps are

$$\widehat{\mathbf{h}}_{t|t-1} = \mathbf{F}_t \widehat{\mathbf{h}}_{t-1|t-1}, \quad \mathbf{P}_{t|t-1} = \mathbf{F}_t \mathbf{P}_{t-1|t-1} \mathbf{F}_t^\top + \mathbf{Q}_t,$$

$$\mathbf{S}_t = \mathbf{X}_t \mathbf{P}_{t|t-1} \mathbf{X}_t^\top + \mathbf{R}_t, \quad \mathbf{K}_t = \mathbf{P}_{t|t-1} \mathbf{X}_t^\top \mathbf{S}_t^{-1},$$

$$\widehat{\mathbf{h}}_{t|t} = \widehat{\mathbf{h}}_{t|t-1} + \mathbf{K}_t \left(\mathbf{y}_t - \mathbf{X}_t \widehat{\mathbf{h}}_{t|t-1} \right), \quad \mathbf{P}_{t|t} = (\mathbf{I} - \mathbf{K}_t \mathbf{X}_t) \mathbf{P}_{t|t-1}.$$

When $\mathbf{F}_t = \mathbf{I}$ and $\mathbf{Q}_t = \lambda^{-1}\sigma^2\mathbf{I}$, the recursion recovers a Bayesian interpretation of recursive least squares with forgetting factor λ upon appropriate identification of covariances. For heavy-tailed noise or sparse priors, conditionally Gaussian augmentations enable expectation–maximization or variational updates. A Student likelihood for residuals can be expressed as a Gaussian with an inverse-gamma mixing variable, turning robust inference into alternating updates for effective weights on residuals and standard Kalman-like steps with reweighted covariances.

Particle methods support nonlinear or non-Gaussian structures and preserve multimodality under severe decision uncertainty [13]. A bootstrap filter propagates weighted particles for $\mathbf{h}_t$ with importance weights derived from the likelihood, and resampling combats degeneracy. However, latency and memory constraints suggest combining small particle sets with Rao–Blackwellization when linear substructures exist. When the equalizer acts on soft decisions from a demodulator, a factor-graph perspective reveals messages between symbol variables and channel states. Expectation propagation and assumed-density filtering provide moment-matching projections onto tractable families, yielding updates that generalize Kalman filtering to non-Gaussian settings while controlling complexity through low-order moments and structured covariances.

Online variational inference approximates the posterior with a factorized or structured Gaussian and optimizes an evidence lower bound on a per-sample or mini-batch basis. For a factorization $q(\mathbf{h}_t, \boldsymbol{\theta}) = q(\mathbf{h}_t)q(\boldsymbol{\theta})$ where $\boldsymbol{\theta}$ denotes hyperparameters such as noise scales and shrinkage variables, coordinate updates follow from expected sufficient statistics under the current variational marginals. Stochastic natural-gradient steps stabilize adaptation by respecting the geometry of exponential families, and annealing schedules limit overconfident updates during early decision-directed iterations.

5 Sparse and Structured Bayesian Equalization

Sparse and Structured Bayesian Equalization

Structure	**Description / Update Rule**
Sparse Taps	Shrinkage via $\alpha_{t,k}^{\text{new}} = \frac{1-\alpha_{t,k}\Sigma_{t,kk}}{\lvert w_{t,k}\rvert^2}$
Block / Group Sparse	Shared precision per group; promotes cluster activation or suppression
[14] Low-rank MIMO	Factor model $\mathbf{W}_t = \mathbf{U}\mathbf{V}_t^\top + \mathbf{E}_t$ with alternating posterior updates

Sparse propagation suggests equalizers with many negligible taps, and Bayesian shrinkage offers a mechanism to control complexity and improve stability. A linear equalizer with tap vector $\mathbf{w}_t$ that minimizes mean-squared error under posterior uncertainty in $\mathbf{h}_t$ can be synthesized by targeting the posterior predictive covariance of the equalized error. Let $\mathbf{R}_{xx,t}$ denote the input covariance matrix, possibly estimated from pilots or imposed by modulation structure, and let $\mathbf{C}_{hy,t}$ denote the cross-covariance between channel taps and observations under the current posterior. The equalizer that minimizes posterior expected squared error satisfies

$$\mathbf{w}_t^\star = \arg\min_{\mathbf{w}} \mathbb{E}\left[\lvert d_t - \mathbf{w}^{\mathrm{H}}\mathbf{y}_t\rvert^2 \,\middle|\, \mathcal{D}_t \right],$$

where d_t is the desired symbol and $\mathcal{D}_t$ denotes observations and soft decisions up to time t. Under linear Gaussian modeling,

$$\mathbf{w}_t^\star = \left(\mathbf{R}_{yy,t}\right)^{-1}\mathbf{r}_{yd,t},$$

with $\mathbf{R}_{yy,t}$ and $\mathbf{r}_{yd,t}$ computed as posterior expectations, thereby embedding uncertainty calibration into the filter. To impose sparsity on $\mathbf{w}_t$, a hierarchical prior yields coordinate-wise shrinkage factors that are updated online. A canonical sparse Bayesian learning update for tap precisions $\boldsymbol{\alpha}_t$ resembles

$$\alpha_{t,k}^{\text{new}} = \frac{1 - \alpha_{t,k}\Sigma_{t,kk}}{\lvert w_{t,k}\rvert^2},$$

where $\Sigma_{t,kk}$ is the kth diagonal element of the posterior covariance of $\mathbf{w}_t$ under a linear-Gaussian model with Gaussian prior of precision $\alpha_{t,k}$. This rule adapts shrinkage based on the ratio of squared mean to variance, suppressing taps with low evidence while preserving those with stable support across time.

Group and block structures arise in equalizers that operate on polyphase components, subcarrier groups, or antenna clusters. A block-sparse prior assigns a common scale to a group of taps, encouraging whole-group activation or deactivation consistent with cluster scattering. The posterior update adopts blockwise statistics, replacing scalar $\alpha_{t,k}$ with group precisions and propagating group posterior covariances. In multiantenna equalizers, low-rank couplings across spatial streams can be enforced through a factorized prior on the equalizer matrix [15]. With $\mathbf{W}_t$ denoting a matrix of linear-combining coefficients across antennas and delays, a factor prior $\mathbf{W}_t = \mathbf{U}\mathbf{V}_t^\top + \mathbf{E}_t$ with small latent dimension constrains solutions to a subspace that evolves under smooth dynamics. Online updates for $\mathbf{U}$ and $\mathbf{V}_t$ follow from alternating conditional posteriors, while $\mathbf{E}_t$ captures deviations under model mismatch.

6 Non-Gaussian Disturbances and Robustness

Robustness under Non-Gaussian Disturbances

Approach	Effect / Interpretation
Student-t Likelihood	Inverse-gamma latent u_t reweights residuals, reducing outlier impact
Soft Decision Likelihood	Mixture-of-Gaussians over symbol hypotheses; approximated via EP
Temporal Priors	Fused-Laplacian or total-variation structures for smooth yet adaptive evolution
Colored Noise Models	Toeplitz / circulant covariances handled by fast transforms for efficiency

Communication receivers encounter impulsive noise and heavy-tailed residuals driven by interference, oscillator glitches, and decision errors. A robust Bayesian equalizer treats the disturbance as a mixture or scale mixture that allocates greater uncertainty to outliers [16]. A Student-t likelihood with degrees of freedom ν can be expressed as a Gaussian with an inverse-gamma weight for each sample, leading to reweighted innovations in Kalman-style updates. The per-sample auxiliary weight u_t modifies the innovation covariance:

$$p(\mathbf{y}_t \mid \mathbf{h}_t, u_t) = \mathcal{N}\left(\mathbf{X}_t\mathbf{h}_t,\, \frac{\mathbf{R}_t}{u_t}\right), \quad u_t \sim \mathrm{Gamma}\left(\frac{\nu}{2}, \frac{\nu}{2}\right).$$

Taking expectations or maximizing with respect to u_t yields effective downweighting of large residuals. When decision-directed likelihoods are used, symbol uncertainty can be represented as a mixture over constellation points with soft probabilities, which in turn induces a mixture-of-Gaussians likelihood for the observation. Expectation propagation projects this mixture onto a Gaussian by matching mean and covariance, maintaining tractable updates while retaining information from soft symbol beliefs.

Robustness can also be enforced through priors that limit overreaction to transient artifacts. A fused-Laplacian prior across time penalizes abrupt changes in adjacent tap values but allows occasional jumps when evidence is strong [17]. The posterior mode then resembles a total-variation regularized estimate, while a fully Bayesian treatment computes posterior means via proximal-operator-aware approximations that integrate over change points. In frequency-selective environments, spectral shaping of noise through colored-likelihood models introduces Toeplitz or circulant covariance structures; inference exploits fast transforms to avoid quadratic cost. The combination of heavy-tailed noise, colored disturbance models, and soft-decision likelihoods promotes resilience in the presence of interference and nonlinear front-end distortions.

7 Hyperparameter Learning and Uncertainty Calibration

Hyperparameters govern process noise levels, observation noise scales, shrinkage strengths, kernel parameters, and degrees of freedom of heavy-tailed models. An empirical Bayes approach updates these quantities by maximizing the marginal likelihood conditioned on data up to time t, with safeguards to avoid aggressive adaptation that destabilizes filters. For linear-Gaussian structures, the marginal likelihood is available in closed form via innovations:

$$\log p(\mathbf{y}_{1:t} \mid \phi) = -\frac{1}{2}\sum_{i=1}^{t}\left[\log\det \mathbf{S}_i + \boldsymbol{\epsilon}_i^\top \mathbf{S}_i^{-1}\boldsymbol{\epsilon}_i + c\right],$$

where ϕ collects hyperparameters, $\boldsymbol{\epsilon}_i = \mathbf{y}_i - \mathbf{X}_i\widehat{\mathbf{h}}_{i|i-1}$ are innovations, and $\mathbf{S}_i$ are their covariances. Gradient-based updates follow from matrix calculus identities, with natural-gradient or Fisher scoring steps improving stability. For hierarchical sparsity, gamma prior parameters (a_0, b_0) and their group analogues can be tuned by maximizing expected complete-data likelihoods under current posterior moments, effectively implementing online expectation–maximization [18]. Heavy-tailed likelihoods introduce additional expectations with respect to auxiliary weights, which admit closed-form updates for conjugate choices.

Uncertainty calibration is assessed by comparing nominal predictive intervals to empirical coverage across time. Let $\widehat{d}_t$ denote the equalized output and σ_t^2 its posterior predictive variance conditioned on $\mathcal{D}_t$. A nominal $q\%$ prediction band is

$$\widehat{d}_t \pm z_q\,\sigma_t,$$

with z_q determined by the approximate predictive distribution. Calibration quality is monitored by indicator averages over windows. If undercoverage is detected, tempering parameters scale process noise or likelihood precision to regularize confidence. In decision-directed regimes, it is often useful to inflate uncertainty during early iterations to prevent feedback bias [19]. A practical scheme introduces a tempering factor $\tau_t \in (0,1]$ that multiplies the information matrix contribution of decision-directed samples:

$$\mathbf{R}_{\text{eff},t}^{-1} = \tau_t\,\mathbf{R}_t^{-1},$$

with τ_t adapted based on recent residual statistics or symbol posterior entropy. Such mechanisms keep updates conservative when evidence is uncertain and tighten them as decisions stabilize.

8 Computational Architectures and Complexity Analysis

The computational viability of Bayesian adaptive equalizers depends on algebraic structure and streaming updates. The Woodbury identity and matrix determinant lemma are central to reducing per-sample costs when the observation dimension is small relative to the state dimension. Suppose the prior covariance is $\mathbf{P}_{t|t-1}$ and the observation model has the form $\mathbf{X}_t\mathbf{h}_t$ with $\mathbf{X}_t$ of size $m \times L$ where L is the tap count. The innovation covariance inversion reduces to

$$\mathbf{S}_t^{-1} = \mathbf{R}_t^{-1} - \mathbf{R}_t^{-1}\mathbf{X}_t\left(\mathbf{P}_{t|t-1}^{-1} + \mathbf{X}_t^\top\mathbf{R}_t^{-1}\mathbf{X}_t\right)^{-1}\mathbf{X}_t^\top\mathbf{R}_t^{-1},$$

which avoids forming dense $L \times L$ inverses when m is small [20]. In block processing, circulant embedding and fast Fourier transforms diagonalize convolution operators, transforming updates into elementwise operations across frequencies. A multicarrier receiver benefits from treating each subcarrier as a low-dimensional observation with cross-subcarrier coupling introduced only through priors or regularization that encourage smoothness. The resulting updates operate per-subcarrier with occasional consensus steps that share hyperparameters or low-rank factors.

Low-rank models for multiantenna coupling reduce the state dimension by projecting into latent subspaces. If $\mathbf{h}_t$ admits a rank-r factorization with $r \ll L$, state transitions and posterior covariances operate in the latent space with cost that scales with r^3 rather than L^3. When sparsity is prominent, iterative solvers with conjugate gradients and diagonal preconditioners exploit the structure of $\mathbf{X}_t^\top\mathbf{R}_t^{-1}\mathbf{X}_t + \mathbf{P}_{t|t-1}^{-1}$ to compute search directions without forming dense matrices. Sliding-window approximations truncate temporal correlations to maintain bounded memory, while exponential forgetting implements an implicit window with smoother weighting. Quantization and fixed-point constraints are addressed by propagating covariance approximations that account for arithmetic noise in $\mathbf{S}_t$ and for lookup-table nonlinearities in front ends. Stability is monitored through spectral radii of linearized update operators and through boundedness of posterior covariances under stationary excitation.

Hardware mapping organizes updates into pipelined stages: innovation computation, gain construction, state update, and covariance downdate [21]. Approximate diagonal or banded representations of $\mathbf{P}_{t|t}$ maintain locality and reduce memory traffic. When heavy-tailed models are employed, auxiliary weights can be updated in vectorized form and fused with innovation calculations, adding only a small overhead compared with Gaussian updates. The aggregate complexity can thus be controlled to nearly linear in the number of active taps, with constant factors determined by transform sizes, group structures, and the precision of iterative solves.

9 Bayesian Equalization in High-Speed Wireline Links and PAM-4 Signaling

High-speed wireline links in backplanes, chip-to-chip interconnects, and short-reach optical modules confront severe intersymbol interference generated by frequency-dependent loss, dielectric dispersion, skin effect, connector discontinuities, and reflections from imperfect impedance control. Equalizers in this regime are typically realized as analog or mixed-signal front-end filters and digital post-processors that include continuous-time linear equalization in the analog domain, feed-forward equalizers in the digital domain, and decision feedback equalizers that cancel postcursor energy by subtracting scaled past decisions. The prevailing adaptation strategies for these blocks rely on stochastic-gradient updates, sign-error rules, and limited-memory second-order schemes that are carefully engineered to converge under slicer noise, finite precision, and tight link training budgets. A Bayesian perspective reframes these mechanisms by introducing explicit priors on tap coefficients and nuisance parameters, likelihoods that account for quantization, clock/data recovery-induced timing jitter, and symbol uncertainty, and posteriors that quantify residual ambiguity in the equalizer state and its predictive consequences for bit error metrics. This perspective is compatible with the architectural constraints of serializer–deserializer receivers and can be specialized to non-return-to-zero signaling and to multi-level pulse-amplitude modulation where slicer thresholds and decision feedback become tightly coupled to symbol posterior distributions. [22]

Feed-forward equalizers in wireline receivers operate as finite-impulse-response filters that pre-compensate precursor and partially address postcursor energy prior to the slicer. In a Bayesian formulation, the tap vector is treated as a random variable whose prior encodes smoothness across delay, group sparsity aligned with pulse response clusters, or low-complexity constraints motivated by hardware resource limits. The observation model includes not only additive thermal noise but also colored interference from near-end and far-end crosstalk, quantization noise from analog-to-digital conversion if present, and effective noise due to timing jitter that appears as multiplicative phase perturbations when the sampling instant drifts. The Bayesian equalizer update uses posterior predictive quantities in place of point estimates, thereby controlling adaptation reactivity when soft evidence is ambiguous. In systems that rely on short training sequences or pseudo-random binary sequences for link calibration, the posterior sharpens quickly during the training window and then transitions to a tempered decision-directed phase in which likelihood precision is downscaled to reflect the nontrivial probability of symbol errors. This tempering can be implemented by scaling the information content of decision-directed samples, limiting aggressive tap changes in periods when the slicer confidence fluctuates because of transient temperature shifts, supply noise, or bursty crosstalk.

Decision feedback equalization in the wireline context cancels dominant postcursors by subtracting a weighted sum of past detected symbols. The challenge with decision feedback is error propagation: a single slicer error injects a structured disturbance into subsequent outputs through the feedback path [23]. Classical cures include conservative step sizes, leakage in the feedback tap adaptation, and pattern-based error monitors that freeze adaptation when eye opening is small. A Bayesian approach augments these heuristics by representing past decisions as random variables with nonzero error probabilities and propagating this uncertainty through the feedback convolution. The effective likelihood for the observation given taps becomes a mixture over symbol hypotheses whose weights are determined by soft information from the slicer or from a downstream forward-error-correcting decoder. Moment-matching projections of this mixture onto a Gaussian approximate likelihood yield tractable updates that naturally downweight epochs with high uncertainty. The posterior over feedback taps retains larger variance in ambiguous regimes, which in turn reduces the magnitude of feedback corrections and curtails error bursts without hand-tuned freeze logic. When the feedback memory is long, block-structured priors that tie groups of taps to delayed clusters are advantageous, because dispersion in copper channels is often localized in a few delay neighborhoods corresponding to major reflection points or frequency notches.

Non-return-to-zero signaling presents a two-level alphabet with relatively large eye openings at moderate rates but faces severe high-frequency loss at tens of gigabits per second on standard materials [24]. Pulse-amplitude modulation with four levels increases spectral efficiency by encoding two bits per symbol but narrows voltage margins at the slicer and raises sensitivity to both vertical and horizontal eye closure. Bayesian equalization for PAM-4 therefore benefits from linking equalizer state updates to multi-threshold slicer confidence metrics or soft decisions derived from offset-cancelled comparators. In this setting, the decision device produces, for each symbol, soft evidence over the four levels captured by log-likelihood ratios or equivalent metrics whose dispersion depends on equalized noise, jitter-induced timing error, and residual intersymbol interference. By treating these soft metrics as part of the data likelihood, the equalizer update reflects the fact that misclassifying adjacent levels is far more probable than confusing the outermost levels, which affects how predictive error variances propagate into subsequent decisions. In a Bayesian formulation, the predictive distribution of the equalized voltage at the slicer thresholds guides both the tap update and the dynamic placement of decision thresholds, with the latter modeled as latent parameters subject to slow drift priors to accommodate comparator offsets and temperature dependencies.

The practical wireline environment is dominated by stringent implementation budgets measured in power per bit and in area allocation within SerDes macros. Adaptation loops must run at line rate or at a decimated rate with minimal buffering, and they frequently execute in fixed-point logic with limited multipliers and table-lookup nonlinearities [25]. Bayesian equalization adapts to these constraints by exploiting conjugacy where possible and by using conditionally Gaussian augmentations that keep updates in terms of means and covariances that map cleanly to shift–accumulate operations. For example, the reweighting implicit in heavy-tailed likelihoods for impulsive noise can be realized as per-sample scalar multipliers applied to innovation terms, with the multipliers updated via simple rational functions of residual magnitudes. When full covariance updates are infeasible, diagonal or banded approximations retain stability benefits while respecting hardware limits, and Kronecker-separable structures become relevant when the equalizer operates jointly across tributaries or when lane bonding induces mild coupling that can be captured with a small number of shared parameters. The Bayesian interpretation also clarifies the role of leakage terms commonly used in fixed-point implementations: they correspond to priors that prevent unbounded drift and keep posterior variances finite when symbol excitation is insufficiently rich.

Training sequences and link negotiation protocols in standards such as Ethernet and PCI Express expose a valuable opportunity for empirical Bayes within the receiver. During the training phase the transmitter sweeps preset equalization configurations and pre-emphasis levels, while the receiver evaluates link quality metrics and feeds back requests. A Bayesian equalizer can use the structured training data to estimate hyperparameters for process noise, observation noise, and shrinkage intensities by marginal-likelihood maximization computed from innovation statistics, constrained so that the resulting parameters do not induce instability when the system transitions to decision-directed mode [26]. Hyperparameters estimated under benign training conditions must be tempered when operating conditions shift, for instance when the board temperature increases and dielectric losses deepen, which is reflected in a mismatch between nominal predictive dispersion and measured error rates. Online calibration methods that compare empirical error indicators with nominal prediction bands adjust likelihood precision multipliers or process noise intensities to restore coverage without resorting to manual retuning. Such calibration reduces the incidence of oscillatory behavior in tap trajectories that would otherwise emerge from a static step-size configuration that is misaligned with the current jitter–noise–loss regime.

Clock and data recovery loops interact with equalization through the timing of samples, and their dynamics strongly influence the residual error seen at the slicer. In practice, the sampling phase is updated via phase detectors and loop filters whose effective noise is colored and often non-Gaussian, especially in the presence of bursty supply noise and spread-spectrum clocking. A Bayesian state-space model that couples equalizer taps and timing offset evolves both sets of latent variables under physically motivated priors. Joint filtering in this model propagates uncertainty from timing into equalizer predictions and vice versa, moderating tap updates when timing is poorly localized and reducing the tendency to attribute timing errors to channel variation. When the phase detector exhibits nonlinearities and dead zones, conditionally Gaussian approximations become inaccurate, and particle-based updates with small particle counts and Rao–Blackwellization over the linear substructure maintain robust performance under latency constraints [27]. The outcome is a more coherent division of labor between timing and equalization, where the equalizer no longer compensates for phase errors that the clock loop should correct, thereby improving stability margins and reducing flutter in the feedback path.

Error correction influences equalization both through its soft outputs and through the temporal correlation it induces in the apparent symbol error process. With strong forward-error correction, the raw symbol error rate at the equalizer output can be relatively high while the post-decoding error rate remains acceptable, provided that errors are sufficiently random. This dichotomy complicates decision-directed adaptation, since equalizer updates based solely on pre-decoder symbol errors may react to patterns that the decoder will easily correct, or under-react to errors that overwhelm the decoder. By embedding decoder-derived soft information into the Bayesian likelihood and by smoothing the impact of individual symbol decisions via predictive variance, the adaptation loop can be tuned to respect the decoder's capability. A practical design introduces a schedule that scales the influence of decoder extrinsics on the equalizer updates as a function of symbol posterior entropy and of a rolling window of parity-check satisfaction indicators, diffusing the effect of occasional unreliable decoder messages.

The progression from two-level to four-level modulation intensifies the role of nonideal front-end components [28]. Comparator offset drift alters optimal threshold placements, differential-to-single-ended conversion imbalances tilt the constellation, and slice-to-slice gain mismatches distort symbol prior probabilities. Bayesian equalization reacts to these realities by modeling threshold and gain parameters as latent variables with smooth-drift priors and by exposing their posterior uncertainty to the decision logic. The feedback structure in a PAM-4 DFE must also contend with the asymmetry of error impacts: misclassifying a middle level has a different consequence for feedback than misclassifying an outer level, because the resulting feedback waveform differs in amplitude and phase alignment with subsequent symbols. A likelihood that marginalizes over discrete symbol transitions weighted by level-dependent priors captures this asymmetry in a principled manner, making the adaptation less sensitive to imbalanced error patterns that arise from skewed thermal noise or asymmetric crosstalk.

Quantization plays an outsized role in wireline receivers that prioritize energy efficiency by relying on low-resolution analog-to-digital conversion or even on analog-only decision devices without full-rate digitization. The Bayesian formulation remains applicable by replacing continuous-valued likelihoods with quantized or censored likelihoods that condition on threshold crossings rather than exact voltages. Under a one-bit front end, for example, the observation informs the sign of the equalized sample relative to a threshold, and the update proceeds by computing the posterior over taps that best aligns sign statistics with their predictive distribution [29]. This change of likelihood yields non-Gaussian posteriors, but variational approximations that retain only first and second moments suffice to produce stable and hardware-friendly updates. When a small number of intermediate comparators are available, as in multi-bit flash ADCs used at a decimated rate, the likelihood incorporates the interval within which the sample falls, further improving information content without abandoning a low-power front end. The resulting hybrid schemes align with contemporary SerDes design practice where a limited digital backend assists an analog front end rather than replacing it completely.

Co-optimization of transmitter pre-emphasis with receiver equalization is a central element of modern link training. Adopting a Bayesian view at the transmitter side amounts to treating pre-emphasis coefficients as latent parameters tuned to maximize a marginal likelihood of training observations at the receiver, as communicated through a sparse feedback channel. Although the feedback bandwidth is small, informative aggregates such as innovations-based scores or compressed sufficient statistics suffice for the transmitter to run a stochastic search over a constrained coefficient space. Joint Bayesian treatment at both ends regularizes the search to avoid overfitting to transient noise realizations during short training intervals and to produce settings that remain robust as the channel warms up or cools down [30]. When multiple lanes share a board region or cable bundle, a hierarchical prior that couples pre-emphasis and equalizer parameters across lanes captures shared loss profiles while permitting per-lane adjustments based on local discontinuities and connector variability.

Link diagnostics and lifecycle management benefit from the uncertainty summaries delivered by Bayesian equalizers. Rather than reporting only point-valued tap coefficients and a single eye opening estimate, the receiver can maintain rolling distributions over tap magnitudes, predicted eye margins at target bit error rates, and the expected impact of further tap changes on these margins. These distributions support operational decisions such as when to trigger retraining, when to relax power settings during low-activity periods, and how to prioritize maintenance for channels exhibiting narrowing predictive margins. Because predictive variance naturally inflates when environmental shifts push the system into regimes not seen during initial calibration, operators receive early warnings before error rates spike, allowing proactive mitigation such as small transmitter coefficient nudges that restore margins without full retraining.

In optical short-reach links where chromatic dispersion and modal dispersion contribute to frequency-selective fading, the same Bayesian equalization principles apply with moderate modifications to account for laser phase noise and nonlinearities in direct-detection schemes. The state-space evolution for taps includes faster dynamics due to temperature-induced wavelength drift, and the likelihood might incorporate signal-dependent noise [31]. Bayesian shrinkage priors are particularly helpful when dispersion compensation produces long though sparse impulse responses, keeping tap counts manageable without sacrificing the ability to represent long-delay echoes. In coherent links, frequency-offset and phase tracking are already cast as probabilistic filtering problems; integrating equalizer tap inference into the same probabilistic framework reduces the burden on ad hoc control logic that coordinates loops for carrier recovery, timing, and equalization.

The literature reflects a gradual shift toward the uncertainty-aware and structure-exploiting design sketched above, covering both classical copper backplanes and optical modules, with reports spanning robust variants of recursive least squares viewed through a probabilistic lens, decision-directed equalizers stabilized by tempering schedules derived from predictive variance, and link training algorithms that interpret aggregate error metrics as innovations in an evidence maximization loop. Within this broader transition, there are specific demonstrations of Bayesian machine learning applied to the co-optimization of feed-forward and decision feedback equalization for two-level and four-level signaling in practical laboratory settings that align with the conceptual framework described here, such as the application to optimizing both feed-forward and decision feedback equalizers for NRZ and PAM-4 signals reported by Dikhaminjia et al. (2021) [32].

10 Connections to Coded, Multicarrier, and Multiantenna Systems

Equalization rarely operates in isolation; it is coupled to channel coding, interleaving, and multicarrier modulation. In coded systems, soft information exchange between the equalizer and decoder closes a loop that can be framed as iterative inference on a joint probabilistic model. Symbol variables, parity-check constraints, and channel states form a factor graph, and messages propagate among them [33]. A Bayesian equalizer furnishes predictive means and variances for symbol likelihoods conditioned on posterior uncertainty in $\mathbf{h}_t$, while the decoder produces extrinsic probabilities that refine the symbol prior. Expectation propagation or loopy belief propagation with moment matching allows the equalizer to incorporate decoder feedback without double counting information. The resulting schedule aligns with turbo equalization but inherits uncertainty calibration from the Bayesian construction.

Multicarrier modulation transforms convolution into per-subcarrier multiplication, but time variability introduces intercarrier interference that couples subcarriers. A Bayesian state-space model across frequency bins captures Doppler leakage through banded coupling in the frequency domain. Let $\mathbf{z}_t$ denote subcarrier-wise channel coefficients and let a banded matrix $\mathbf{G}$ express leakage to neighboring bins; the evolution

$$\mathbf{z}_t = \mathbf{G}\mathbf{z}_{t-1} + \boldsymbol{\eta}_t$$

with small-bandwidth $\mathbf{G}$ yields updates whose cost per subcarrier depends only on a limited neighborhood. Priors that vary across frequency reflect power delay profiles and subband-selective sparsity. In the presence of pilot patterns, evidence accumulation can be concentrated on pilot-bearing tones while decision-directed information supplies soft constraints on data-bearing tones; tempering and calibration prevent overconfident cross-subcarrier propagation.

In multiantenna receivers, spatial multiplexing and diversity create a matrix-valued channel whose dimension grows with antenna counts [34]. A Bayesian equalizer exploits Kronecker or factorized priors to limit effective dimension. Suppose $\mathbf{H}_t \in \mathbb{C}^{N_r \times N_t}$ evolves with a low-rank model. The vectorized state obeys

$$\mathrm{vec}(\mathbf{H}_t) = (\mathbf{I} \otimes \mathbf{F}_t)\mathrm{vec}(\mathbf{H}_{t-1}) + \mathrm{vec}(\mathbf{W}_t),$$

with structured $\mathbf{F}_t$ and noise covariance that factorizes as $\mathbf{K}_s \otimes \mathbf{K}_t$. The equalizer combines spatial filtering with temporal updates, and predictive symbol variances drive soft detection and decoding. Joint priors across streams mitigate error propagation when decision-directed feedback is used for higher-order constellations. Calibration in this context requires monitoring of per-stream predictive intervals and cross-covariances, as underestimation leads to brittle decisions on weaker streams.

11 Differentiable Implementations and Hybrid Bayesian–Learning Approaches

Modern receivers often include learned components that are trained offline and adapted online. Bayesian principles interact with such components through uncertainty-aware layers, amortized inference networks, and priors on weights that control overfitting in limited-pilot regimes [35]. A differentiable Bayesian equalizer exposes its update equations as computational graphs, enabling gradient-based tuning of hyperparameters or meta-parameters on representative data while maintaining closed-form filtering during deployment. Consider a parameter vector ϕ collecting process and observation noise scales, shrinkage targets, and tempering schedules. A meta-objective averages symbol error proxies over datasets, and gradients flow through the filtering recursions via implicit differentiation:

$$\frac{\partial \mathcal{L}}{\partial \phi} = \sum_t \frac{\partial \mathcal{L}}{\partial \widehat{\mathbf{h}}_{t|t}} \frac{\partial \widehat{\mathbf{h}}_{t|t}}{\partial \phi} + \sum_t \frac{\partial \mathcal{L}}{\partial \mathbf{P}_{t|t}} : \frac{\partial \mathbf{P}_{t|t}}{\partial \phi},$$

where the colon denotes the Frobenius inner product. To ensure stability, gradients are regularized and truncated in time, and hyperparameters are constrained to physically meaningful ranges by reparameterization such as softplus for variances and sigmoid for tempering factors.

Amortized inference replaces parts of the analytic update with neural proposals that map local features of $\mathbf{y}_t$ and soft decisions to approximate posterior moments. The proposals serve as initializations for analytic corrections or as full surrogates when computation is constrained. A Bayesian calibration layer rescales proposal variances based on recent innovation statistics, maintaining coverage while leveraging expressive features. Hybrid equalizers also impose priors directly on network weights, treating them as random variables with variational posteriors [36]. The predictive distribution of equalized symbols then marginalizes both channel and network uncertainty, yielding outputs with calibrated dispersion. Training uses stochastic variational objectives aggregated over time windows, and deployment employs low-rank approximations of weight covariances to preserve throughput.

12 Conclusion

The integration of Bayesian inference principles into adaptive equalization provides a structured lens through which classical and modern algorithms can be unified, generalized, and interpreted. Within this perspective, the equalizer no longer appears merely as a deterministic filter trained to minimize instantaneous error, but rather as an inferential mechanism that continually updates beliefs about the channel, transmitted symbols, and model uncertainties. This viewpoint naturally accommodates streaming constraints and limited hardware resources by framing each update as an approximate Bayesian inference step—often realized through recursive moment updates or factorized posterior approximations. The resulting balance between computational feasibility and statistical rigor defines the practicality of Bayesian adaptive equalization in real-world systems, where symbol rates, memory, and arithmetic throughput impose strict bounds on algorithmic complexity.

A key aspect of this formulation lies in the specification of priors that encode structural assumptions about the communication channel and equalizer [37]. Sparse priors, for instance, reflect the fact that many physical channels exhibit only a few dominant multipath components. Low-rank priors, by contrast, arise naturally in multiantenna or multicarrier configurations where the channel responses across subcarriers or antennas share latent correlations. Hierarchical shrinkage models, such as those employing Student-t or horseshoe distributions, offer a flexible compromise between sparsity and adaptivity, enabling automatic relevance determination without imposing hard thresholds. These priors directly influence the equalizer's response to new observations, determining how aggressively coefficients are updated and how uncertainty propagates through the estimation process.

The likelihood model forms the second cornerstone of the Bayesian equalization framework. Gaussian likelihoods remain the default choice when thermal noise dominates and linear assumptions hold, leading to analytically tractable updates and efficient implementations. However, real channels often encounter impulsive or non-Gaussian interference, motivating the adoption of heavy-tailed likelihoods such as Laplace or mixture models that downweight outliers. In such cases, conditionally Gaussian reweighting schemes can approximate robust inference while retaining the convenience of quadratic update equations [38]. These reweighted updates introduce adaptive variance scaling that moderates the influence of corrupted observations, thereby stabilizing decision-directed operation and mitigating error bursts.

Temporal variability of the channel is naturally modeled within a state-space formulation, where the evolution of the channel or equalizer coefficients follows a stochastic dynamical process. Smooth drift can be represented by autoregressive Gaussian transitions, while abrupt changes—such as those caused by handoffs or deep fades—can be accommodated through switching or jump-diffusion dynamics. Bayesian filtering techniques, including Kalman-type updates for linear-Gaussian models and particle filtering for nonlinear or multimodal scenarios, provide recursive mechanisms to track these dynamics. When dimensionality or latency constraints preclude full posterior maintenance, variational projection methods can compress belief states into low-rank or diagonal approximations, ensuring that per-sample complexity remains bounded while preserving the essential uncertainty structure needed for reliable symbol decisions.

From an algorithmic standpoint, the Bayesian approach encompasses a rich spectrum of update mechanisms. Kalman filters and their variants yield optimal linear estimators under Gaussian assumptions, offering a well-understood baseline [39]. When the noise or dynamics deviate from linearity, particle filters approximate the posterior through weighted ensembles of samples, capturing nonlinear effects and multimodal uncertainty at the cost of increased computational demand. Variational inference, meanwhile, projects the true posterior onto a tractable family—often Gaussian with factorized covariance—yielding scalable recursions that can be executed in high-dimensional systems such as massive MIMO or wideband OFDM receivers. These variational updates can be interpreted as structured stochastic gradient steps in an evidence lower bound (ELBO) optimization, connecting probabilistic inference with modern optimization theory.

To make these methods feasible for streaming applications, computational strategies exploit matrix identities, transform-domain diagonalization, and sliding-window approximations. The use of the matrix inversion lemma, for example, enables efficient recursive updates of covariance estimates without explicit matrix inversion. Transform-domain equalizers leverage the approximate diagonalization of convolution operators in the frequency domain, allowing per-subcarrier updates that are computationally decoupled yet statistically coherent through shared priors. Sliding-window schemes restrict inference to recent data while maintaining an approximate sufficient statistic for the posterior, striking a balance between adaptivity and computational economy [40]. Such strategies ensure that each incoming symbol can be processed within constant time, preserving the throughput requirements of high-rate systems.

The Bayesian equalization framework also extends naturally to coded communication systems, multicarrier modulation, and multiantenna architectures. In turbo equalization and related iterative detection schemes, the equalizer exchanges soft information with a channel decoder, using posterior variances to quantify symbol reliability. Bayesian modeling provides a consistent way to generate and interpret this soft information, with predictive distributions guiding the log-likelihood ratios passed to the decoder. In multiantenna settings, joint inference across spatial channels benefits from low-rank priors that capture inter-antenna correlation, enabling both improved estimation accuracy and reduced complexity. Multicarrier systems, such as OFDM, gain from hierarchical priors that couple subcarriers through shared latent variables, reflecting the coherence bandwidth of the channel and stabilizing equalization under frequency-selective fading.

Recent developments also explore hybrid differentiable implementations in which Bayesian equalization modules are embedded within neural architectures [41]. These designs retain interpretable update equations derived from probabilistic reasoning while introducing data-driven components that learn hyperparameters, noise models, or proposal mechanisms. By making the inference pipeline differentiable, such systems can be trained end-to-end using gradient-based optimization while preserving the uncertainty quantification and calibration benefits of the Bayesian formulation. The predictive variances produced by these models can be used to regulate confidence, schedule feedback adaptation, or allocate computational effort dynamically according to uncertainty levels.

Throughout this exploration, the focus remains on maintaining a conservative and balanced assessment of performance gains. While Bayesian equalization provides conceptual elegance and often improved robustness, its benefits depend critically on the match between prior assumptions, computational budgets, and channel behavior. In regimes where structural priors align with the true propagation characteristics and where hardware allows modest additional complexity, uncertainty-aware adaptation can significantly enhance reliability and convergence stability. Conversely, in highly resource-constrained environments or when prior mis-specification dominates, simpler heuristic methods may remain competitive. The value of the Bayesian framework thus lies not in universal superiority but in providing a coherent set of tools for reasoning about these trade-offs. [42]

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgement

This work was supported without any funding.

References

[1] C.-K. Yu, D.-B. Lin, and H.-P. Lin, "Wideband common-mode suppression filter using defected corrugated reference plane structures," *Microwave and Optical Technology Letters*, vol. 65, no. 6, pp. 1615–1621, Feb. 16, 2023. DOI: 10.1002/mop.33630

[2] M. Sahani, S. K. Rout, and P. K. Dash, "Epileptic seizure recognition using reduced deep convolutional stack autoencoder and improved kernel rvfln from eeg signals," *IEEE transactions on biomedical circuits and systems*, vol. 15, no. 3, pp. 595–605, Aug. 12, 2021. DOI: 10.1109/tbcas.2021.3090995

[3] I. Chinaeke-Ogbuka et al., "A robust high-speed sliding mode control of permanent magnet synchronous motor based on simplified hysteresis current comparison," *International Journal of Power Electronics and Drive Systems (IJPEDS)*, vol. 12, no. 1, pp. 1–9, Mar. 1, 2021. DOI: 10.11591/ijpeds.v12.i1.pp1-9

[4] F. Dimc, M. Bažec, D. Borio, C. Gioia, G. Baldini, and M. Basso, "An experimental evaluation of low-cost gnss jamming sensors," *Navigation*, vol. 64, no. 1, pp. 93–109, May 7, 2017. DOI: 10.1002/navi.184

[5] H. Kim and J.-K. Kang, "A novel pwam signaling scheme for high-speed serial interface," *JOURNAL OF SEMICONDUCTOR TECHNOLOGY AND SCIENCE*, vol. 22, no. 5, pp. 326–340, Oct. 31, 2022. DOI: 10.5573/jsts.2022.22.5.326

[6] T. Sarmiento, L. Zhao, P. Moser, T. Li, Y. Huo, and J. S. Harris, "Continuous-wave operation of gaas-based 1.5- μ m gainnassb vcsels," *IEEE Photonics Technology Letters*, vol. 31, no. 20, pp. 1607–1610, Oct. 15, 2019. DOI: 10.1109/lpt.2019.2938177

[7] G. V. Kavitha, S. S. Sindhu, and J. Zacharias, "Dual band radio-over-fibre millimetre–wave system utilizing optical frequency combs," *Journal of Optical Communications*, vol. 44, no. s1, s1433–s1438, Oct. 23, 2020. DOI: 10.1515/joc-2020-0177

[8] H. Shishido et al., "Neutron detection using the superconducting nb-based current-biased kinetic inductance detector," *Superconductor Science and Technology*, vol. 30, no. 9, pp. 094 003–, Aug. 14, 2017. DOI: 10.1088/1361-6668/aa7a3d

[9] L. Tian and Q. Yu, "Wide-speed-range sensorless control scheme for ipmsms based on super-twisting observer and hf signal injection," *International Journal of Applied Electromagnetics and Mechanics*, vol. 54, no. 3, pp. 367–387, Jul. 1, 2017. DOI: 10.3233/jae-160150

[10] V. I. Lovchakov and O. A. Shibyakin, "The solution of a problem of speed of response on output coordinate for linear dynamic systems," *Mekhatronika, Avtomatizatsiya, Upravlenie*, vol. 20, no. 9, pp. 532–541, Sep. 5, 2019. DOI: 10.17587/mau.20.532-541

[11] X.-T. Zhang, J.-H. Zhang, Y. Hu, and C.-L. Dou, "Elevator speed control method based on wireless sensor network," *The Journal of Engineering*, vol. 2019, no. 22, pp. 8400–8403, Nov. 25, 2019. DOI: 10.1049/joe.2019.1090

[12] J. Schaude, A. Gröschl, and T. Hausotte, "Stitched open-loop measurements with a focal-distance-modulated confocal sensor," *tm - Technisches Messen*, vol. 88, no. 9, pp. 544–555, May 5, 2021. DOI: 10.1515/teme-2021-0036

[13] G. Dushaq, B. Paredes, and M. Rasras, "Strong enhancement of direct transition photoluminescence at room temperature for highly tensile-strained ge decorated using 5 nm gold nanoparticles.," *Nanotechnology*, vol. 31, no. 31, pp. 315 201–, Apr. 17, 2020. DOI: 10.1088/1361-6528/ab8a8d

[14] M. Amraee, E. Farshidi, and A. Kosarian, "Improved turn-on speed of low-power loads in pulsed power supply scheme and high-energy efficiency," *Journal of Circuits, Systems and Computers*, vol. 32, no. 8, May 12, 2023. DOI: 10.1142/s0218126623500937

[15] Y.-K. Yoon, S.-J. Kim, K. Ko, C.-O. Park, and Y.-K. Kim, "A study on the analysis and recommendations for ground current in the signal machine room of high-speed rail," *The transactions of The Korean Institute of Electrical Engineers*, vol. 70, no. 10, pp. 1517–1525, Oct. 31, 2021. DOI: 10.5370/kiee.2021.70.10.1517

[16] M. K. Matters-Kammerer, D. van Goor, and L. Tripodi, "Broadband sub-thz spectroscopy modules integrated in 65-nm cmos technology," *International Journal of Microwave and Wireless Technologies*, vol. 9, no. 6, pp. 1211–1218, Jun. 9, 2017. DOI: 10 . 1017 / s1759078717000599

[17] H. He, T. Chen, M. Chen, D. Li, and P. Cheng, "Remote sensing image super-resolution using deep–shallow cascaded convolutional neural networks," *Sensor Review*, vol. 39, no. 5, pp. 629–635, Sep. 16, 2019. DOI: 10.1108/sr-11-2018-0301

[18] L. G. S. S. Harsha, B. R. Jammu, V. R. Samoju, S. Veeramachaneni, and N. M. S, "A low-error, memory-based fast binary antilogarithmic converter," *International Journal of Circuit Theory and Applications*, vol. 49, no. 7, pp. 2214–2226, Feb. 28, 2021. DOI: 10.1002/cta.2981

[19] M. L. Psiaki, "Ionosphere ray tracing of radio-frequency signals and solution sensitivities to model parameters," *Radio Science*, vol. 54, no. 8, pp. 738–757, Aug. 15, 2019. DOI: 10.1029/2019rs006792

[20] J. Zeng, J. Wan, D. Chen, X. Yang, and Z. Yu, "An analog-to-digital converter calibration algorithm with clock jitter compensation based on bidirectional long-short-time-memory," *Electronics Letters*, vol. 58, no. 20, pp. 753–755, Aug. 11, 2022. DOI: 10.1049/ell2.12598

[21] D. Xu, S. Xu, X. Li, and J. Pu, "A 10-bit 110 mhz sar adc with asynchronous trimming in 65-nm cmos*," *Journal of Semiconductors*, vol. 38, no. 4, pp. 045 003–, Apr. 20, 2017. DOI: 10.1088/1674-4926/38/4/045003

[22] C. Wu, S. Jiang, and C. Bian, "Online parameter identification of spmsm based on improved artificial bee colony algorithm," *Archives of Electrical Engineering*, pp. 777–790, Nov. 30, 2021. DOI: 10.24425/aee.2021.138260

[23] K. Mariammal, M. H. Banu, J. B. Pari, and V. Dhandapani, "Decisive structures for multirate fir filter incorporating retiming and pipelining schemes," *Circuit World*, vol. 47, no. 4, pp. 427–444, Sep. 28, 2020. DOI: 10.1108/cw-05-2020-0094

[24] S. Rzepka, A. Otto, D. Vogel, and R. Dudek, "Application-driven reliability research of next generation for automotive electronics: Challenges and approaches," *Journal of Electronic Packaging*, vol. 140, no. 1, pp. 010 903–, Mar. 1, 2018. DOI: 10 . 1115/1.4039333

[25] I. S. Amiri, A. N. Z. Rashed, and P. P. Yupapin, "High-speed light sources in high-speed optical passive local area communication networks," *Journal of Optical Communications*, vol. 44, no. 1, pp. 61–67, Apr. 20, 2019. DOI: 10.1515/joc-2019-0070

[26] S. Ye, "Design and performance analysis of an iterative flux sliding-mode observer for the sensorless control of pmsm drives," *ISA transactions*, vol. 94, pp. 255–264, Apr. 25, 2019. DOI: 10.1016/j.isatra.2019.04.009

[27] R. Wang, J. Tang, Z. Deng, and Y. Kang, "Motion induced eddy current based testing method for the detection of circumferential defects under circumferential magnetization," *International Journal of Applied Electromagnetics and Mechanics*, vol. 64, no. 1-4, pp. 501–508, Dec. 10, 2020. DOI: 10.3233/jae-209357

[28] M. Habibi and A. R. Danesh, "A digital arbitrary size kernel convolution smart image sensor based on in-pixel pulse width processors," *Sensor Review*, vol. 37, no. 4, pp. 468–477, Sep. 18, 2017. DOI: 10.1108/sr-03-2017-0035

[29] Y. Shimoda, K. Ikuta, K. Hayashi, D. Ito, and M. Nakamura, "Ctle compensation design methodology for pam4 signals," *IEEJ Transactions on Electronics, Information and Systems*, vol. 142, no. 1, pp. 1–5, Jan. 1, 2022. DOI: 10.1541/ieejeiss.142.1

[30] S. Eladl, A. Nasr, and A. Sharaf, "Analysis of a vertical cavity surface emitting laser excited by a rectangular pulse," *JOURNAL OF SEMICONDUCTOR TECHNOLOGY AND SCIENCE*, vol. 22, no. 1, pp. 17–23, Feb. 28, 2022. DOI: 10.5573/jsts.2022.22.1.17

[31] A. Sharma and P. Chauhan, "High speed radio over fiber system for wireless local area networks by incorporating alternate mark inversion scheme," *Journal of Optical Communications*, vol. 42, no. 2, pp. 273–277, Aug. 8, 2018. DOI: 10.1515/joc-2018-0084

[32] N. Dikhaminjia et al., "Optimization of joint equalization of high-speed signals using bayesian machine learning," in *2021 IEEE International Joint EMC/SI/PI and EMC Europe Symposium*, IEEE, 2021, pp. 48–52.

[33] ., and., "" *Radioelectronic and Computer Systems*, no. 4, pp. 50–56, Mar. 26, 2019. DOI: 10.32620/reks.2017.4.05

[34] A. Abuelhaija, S. Salama, and M. El-Absi, "Multi-tuned radiofrequency coil using microfluidically tunable capacitor for magnetic resonance imaging/spectroscopy at 7-tesla," *International Journal on Communications Antenna and Propagation (IRECAP)*, vol. 9, no. 6, pp. 419–427, Dec. 31, 2019. DOI: 10.15866/irecap.v9i6.17834

[35] A. A. Komarskiy, S. R. Korzhenevskiy, A. V. Ponomarev, and N. A. Komarov, "Pulsed x-ray source with the pulse duration of 50 ns and the peak power of 70 mw for capturing moving objects.," *Journal of X-ray science and technology*, vol. 29, no. 4, pp. 567–576, May 8, 2021. DOI: 10.3233/xst-210873

[36] C. A. Parmar, B. Ramanadham, and A. D. Darji, "Fpga implementation of hardware efficient adaptive filter robust to impulsive noise," *IET Computers & Digital Techniques*, vol. 11, no. 3, pp. 107–116, Feb. 21, 2017. DOI: 10.1049/iet-cdt.2016.0067

[37] S. M. R. Rasid, A. Michael, and H. R. Pota, "On-chip self-sensing piezoelectric micro-lens actuator with feedback control," *IEEE Sensors Journal*, vol. 23, no. 10, pp. 10 275–10 284, May 15, 2023. DOI: 10.1109/jsen.2023.3261965

[38] S. Ye, "Fuzzy sliding mode observer with dual sogi-fll in sensorless control of pmsm drives.," *ISA transactions*, vol. 85, pp. 161–176, Oct. 16, 2018. DOI: 10.1016/j.isatra.2018.10.004

[39] S. Kaur, M. L. Singh, null Priyanka, and M. Singh, "Performance comparison of all-optical logic gates using electro-optic effect in mzi-based waveguide switch at 1.46 µm," *Journal of Optical Communications*, vol. 44, no. s1, s231–s243, Nov. 27, 2020. DOI: 10.1515/joc-2020-0125

[40] H. B. Eldeeb, H. A. I. Selmy, H. M. ElSayed, and R. I. Badr, "Interference mitigation and capacity enhancement using constraint field of view adr in downlink vlc channel," *IET Communications*, vol. 12, no. 16, pp. 1968–1978, Jul. 31, 2018. DOI: 10.1049/iet-com.2017.1174

[41] V. Kober, "Recursive algorithms for computing sliding dct with arbitrary step," *IEEE Sensors Journal*, vol. 21, no. 10, pp. 11 507–11 513, May 15, 2021. DOI: 10.1109/jsen.2020.3023892

[42] M. H. H. Kani, M. J. Yazdanpanah, and A. H. Markazi, "Stability analysis of a class of uncertain switched time-delay systems with sliding modes," *International Journal of Robust and Nonlinear Control*, vol. 29, no. 1, pp. 19–42, Oct. 26, 2018. DOI: 10.1002/rnc.4369