A Hybrid Bayesian-PDE Constrained Optimization Framework for High-Dimensional Image Reconstruction

Mushtaq K. Abdalrahem[1,2,*], Zainab H. Abood[2] and Maryam Sadiq[3]

[1]College of Pharmacy, University of Al-Ameed, Karbala, Iraq

[2]Department of Statistics, College of Administration and Economics, University of Kerbala, Karbala, Iraq

[3]Department of Therapeutic of Nutrition, College of Health and Medical Techniques, AL-Zahraa University for Women

Abstract: High-dimensional image reconstruction problems in fields such as medical imaging, astrophysics, and remote sensing are typically ill-posed inverse problems affected by noise, under sampling, and imperfections in the physical forward model. Traditional methods for resolving these conflicts suffer from an inherent trade-off: pure physics-based PDE-constrained models impose physical consistency but are deterministic and do not represent uncertainty, and fully Bayesian models provide principled uncertainty quantification but tend to become computationally intractable in very high-dimensional spaces. In response to these challenges, we propose the Hybrid Bayesian - PDE Constrained Optimization Framework. The Hybrid Bayesian - PDE Constrained Optimization Framework leverages the physical fidelity of PDE-based forward models with the expressive capability of Bayesian inference to model uncertainty. The reconstruction problem is cast as an optimization problem whereby a variational or hierarchical Bayesian prior is combined with a PDE-constrained data fidelity term, and the optimization objective is solved by an efficient stochastic variational optimization scheme. Experiments using a representative CT example and MRI datasets demonstrated how the hybrid methods provided (i) better reconstructions, preserving fine structures for substantially undersampled data and robust performance in noise when compared to the pure physics model along with providing (ii) clinically meaningful pixel-wise uncertainty maps. These results support the view that the proposed hybrid method provides a principled, computationally efficient, reliable approach to the challenge of solving large-scale inversion problems, while addressing the fundamental limitations of both deterministic, physics-based methods and probabilistic Bayesian inversion.

Keywords: Hybrid Bayesian Inference; PDE-Constrained Reconstruction; Uncertainty Quantification; Inverse Problems; Medical Imaging.

1. INTRODUCTION

Reconstructing images that are high-dimensional from indirect measurements that are noisy poses an important problem in the field of computational imaging. This problem arises in a spectrum of scientific and clinical disciplines such as medical imaging when using techniques such as magnetic resonance imaging (MRI), computed tomography (CT), and positron emission tomography (PET), as well as seismic imaging, and computational photography. The inverse problems associated with producing this image are ill-posed, primarily due to loss of information during data acquisition. Practical issues which exacerbate the ill-posedness include having to rapidly scan in MRI resulting in under sampled k-space data, and minimizing the radiation dose in CT resulting in low-dose images. With high-dimensionality of the unknown image itself, and on the order of millions of voxels, this not only requires computationally efficient and stable algorithms but also complicated regularization methods to arrive at a physiologically plausible and unique solution.

To address this issue, two primary philosophical paradigms have been used. First is PDE-constrained optimization that puts the governing physical laws (i.e. the Radon transform in CT, wave equations in seismic imaging, etc.) directly into the inversion process, usually as hard constraints as discussed in [1], guaranteeing any resulting solution is physically reasonable; however, we cannot derive a measure of confidence or uncertainty of the image we reconstruct, which is a significant limitation to this deterministic framework and a critical shortcoming in diagnostic and other safety-critical contexts. The second paradigm, Bayesian methods, provides a powerful probabilistic framework by treating the unknown image as a random variable and using priors to regularize the problem. The primary strength of this approach is the full posterior distribution provided, allowing for a rigorous uncertainty measure, and has been reviewed in surveys on Bayesian deep learning for inverse problems by [2]. Traditional full Bayesian inference, including existing methods for sampling from a posterior, e.g., Markov Chain Monte Carlo (MCMC), suffers from the "curse of dimensionality" and has thus not been shown to be computationally feasible on higher dimensional image spaces, as [3] has discussed.

This duality highlights a rich void in the existing literature regarding a consolidated framework that is both computationally feasible for high dimensionality, allows for robust uncertainty quantification, and is also compliant with underlying physical laws of the imaging investigational system. As such, we propose a new combined Bayesian-PDE Constrained Optimization

*Address correspondence to these authors at the College of Pharmacy, University of Al-Ameed, Karbala, Iraq; E-mail: mushtaq.k@alameed.edu.iq

E-ISSN: 1929-6029/26

Framework. Our contributions are as follows: First, we present a new mathematical structure that integrates a Bayesian prior, as described by current variational inference principles [4], into a PDE-constrained optimization problem. Second, we construct and implement a stochastic optimization algorithm that can replicate the high dimensionality and high stochasticity of the objective function. Third, we demonstrate the applied utility of our framework for the real-world problem of low-dose CT reconstruction, an issue of research relevance [5]. Fourth, we provide thorough quantitative and qualitative results to show not only enhanced accuracy in reconstruction but also the ability to develop actionable pixel-wise uncertainty maps that directly relate to diagnostic conviction. Therefore, this work tests the central hypothesis that integrating a PDE constraint as a penalty term within a scalable variational Bayesian objective function significantly improves both the accuracy of the point estimate (reconstructed image) and the calibration of the posterior uncertainty, compared to using either physics-based constraints or Bayesian priors in isolation.

The remainder of this work is structured as follows. We will begin with an extensive discussion of the relevant literature, in which we then outline our proposed methods, then we will present and discuss our experimental results, and finally, we will give our concluding remarks and possible avenues for future work.

LITERATURE REVIEW

The issue of high-dimensional image reconstruction has been studied from two historically separate perspectives, both of which are based on a substantial literature base. One of those literature bases has contributed its own mathematically rigorous arguments concerning physical models using PDE-constrained optimization. Many of the pioneering studies in this field have explored the distinction between all-at-once and reduced-space methods, with all-at-once methods solving for the unknown image and associated state variables in a single problem and reduced-space methods removing the state variables to form a smaller optimization problem [6] analyzes these all-at-once and reduced-space methods to provide examples of the trade-offs between computational memory and feasibility in each for large-scale problems. This model-formulation paradigm has been clearly demonstrated in terms of its utility in complex imaging modalities. In, for example, optical tomography, the radiative transfer equation acts as a mathematical constraint to image reconstruction for reconstructing the optical properties of tissue [7]. In quantitative photoacoustic tomography, the propagation of the acoustic waves governed by the wave equation is a significant constraint in the reconstruction of initial pressure distributions [8].

These strategies are good at ensuring physical consistency; however, forthcoming work by [9] mentions that they are still fundamentally deterministic methods, providing a single best estimate without a probabilistic interpretation of the reliability of the solution. In concert, the Bayesian paradigm has been established to address precisely the issue of quantifying uncertainty. The "gold standard" of Bayesian computation has generally been sampling-based methods, such as Markov Chain Monte Carlo (MCMC) and its more efficient variants, Hamiltonian Monte Carlo (HMC) [3] give a complete theoretical outline of MCMC and acknowledge the significant computational burden of sampling methods, which is often unfeasible for the type of high-dimensional parameter spaces typically present in image reconstruction. This challenge motivated the development of Variational Inference (VI), which recasts Bayesian inference as an optimization problem and therefore is much more scalable in terms of compute. The review by [4] provides a comprehensive overview of how Variational Inference seeks a distribution from a tractable family that is closest to the true posterior, usually measured using the Kullback-Leibler divergence. Previously, VI methods were predominantly reliant on mean-field approximations, which essentially assumed the latent variables to be conditionally independent from one another. Unfortunately, such approximations severely limit the statistical fit of the procedure to the resulting posterior distribution. In more contemporary research, however, the focus has shifted to stochastic VI methods that use mini-batch training to handle very large data sets, see, for example, [10]. Another recent revolution is the adoption of deep generative models as powerful, implicit priors. Such models as Variational Autoencoders (VAEs), see [11], Generative Adversarial Networks (GANs), and more recently, Diffusion Models, popularized by [12], have the ability to learn complex, data-driven prior distributions that are far richer than any traditional priors that an instance could come up with a priori. Continuing the trend, these learned priors have been studied and shown to capture the natural statistics of images, for example, [13] and research on deep image priors.

Bridging these two areas has led to much recent research in developing several hybrid and related approaches [14] has formalized the theory of Bayesian inverse problems, for large settings of infinite-dimensional inverse problems, but the inherent

challenges of a computational nature exist for taking and applying this framework and abstraction in practical high-dimensional settings. A common and computationally simpler point of connection is Maximum a Posteriori (MAP) estimation, which can be viewed as a regularized optimization problem where the regularize corresponds to the log-prior. However, as rightly critiqued by [15], MAP estimation provides only a single point estimate, discarding the rich posterior uncertainty that is a hallmark of the Bayesian framework. Our proposed method fundamentally differs by aiming to approximate the full posterior or a richer distribution than a single mode. Other pioneering hybrid models have sought to combine physical models with data-driven learning. For example, the concept of physics-informed neural networks (PINNs) by [16] embeds PDEs into the loss function of a neural network. Furthermore, the work of [17] on deep Bayesian inversion and by [18] on deep learning techniques for inverse problems represent significant strides in this direction. However, many existing approaches treat the physical model as a soft constraint within a data-driven loss or use learned models to replace parts of the physical model entirely. Our framework distinguishes itself by proposing a more fundamental and tighter integration, where the PDE is treated as a core constraint within the objective function of a scalable Bayesian inference procedure, thereby ensuring physical consistency while performing full posterior approximation, a synthesis that moves beyond the limitations of current hybrid models as identified by [1] in their comprehensive review.

METHODOLOGY

Problem Formulation

The core problem of high-dimensional image reconstruction is formalized by considering an unknown image, or state, $\mathbf{u} \in \mathbb{R}^N$, which we aim to recover from a set of noisy, and often incomplete, measurements $\mathbf{y} \in \mathbb{R}^M$, where $M \ll N$ in many ill-posed scenarios. The physical mechanism that drives the acquisition is represented using a system of partial differential equations (PDEs), written compactly as $A(u) = 0$, where A is a differential operator. In classical PDE-constrained optimization, this leads to a deterministic minimization problem:

$$\min_{\mathbf{u}} J(\mathbf{u}) \text{ subject to } \mathcal{A}(\mathbf{u}) = \mathbf{0}$$

where $J(\mathbf{u}) = \| \mathcal{M}(\mathbf{u}) - \mathbf{y} \|_2^2$ is a data fidelity term, and $\mathcal{M}$ is the measurement operator that maps the state to the data space. Although this formulation achieves physical consistency, it gives no probabilistic interpretation. From a Bayesian standpoint, the problem is essentially rethought as a problem of statistical inference. The solution is given by the posterior probability distribution $p(u \mid y)$, which is proportional, by Bayes' theorem, to the likelihood multiplied by the prior:

$$p(\mathbf{u} \mid \mathbf{y}) \propto p(\mathbf{y} \mid \mathbf{u})\, p(\mathbf{u}) (1)$$

In this context, $p(y \mid u)$ encodes the forward model and the noise statistics and $p(u)$ encode our prior knowledge about the image, for example, a constraint prompting sparsity, or smoothness. The challenge is to perform inference on this posterior in a high-dimensional space - a task to which traditional methods, e.g. Markov Chain Monte Carlo (MCMC), which [3] have referred to as "overwhelmingly slow" and " too cost prohibitive".

The Proposed Hybrid Framework

The central innovation is a hybrid framework that integrates physical constraints within a scalable Bayesian inference procedure. From a statistical inference perspective, treating the PDE as a soft constraint within the variational objective (Eq. 2) is preferable to a hard constraint (Eq. 1) because it admits a tractable posterior approximation while penalizing physical implausibility. It is also superior to sequential approaches (e.g., post-processing a Bayesian reconstruction with physics) as it jointly optimizes for data consistency, prior belief, and physical fidelity, leading to better-posed inference and avoiding propagation of errors between separate stages. Then, we can minimize a joint objective function:

$$\mathcal{L}(\mathbf{u}) = \mathbb{E}_{\xi \sim p(\xi)}[\| \mathcal{M}(\mathbf{u}(\boldsymbol{\xi})) - \mathbf{y} \|_2^2] + \lambda_1 \mathcal{R}_{\text{Bayesian}}(\mathbf{u}) + \lambda_2 \| \mathcal{A}(\mathbf{u}) \|_2^2 (2)$$

In this formulation, $\xi \sim p(\xi)$ is a random variable that primarily encapsulates the measurement noise inherent in the data acquisition process (e.g., Gaussian or Poisson noise). Algorithmically, it can also represent stochastic perturbations such as dropout in neural network components of the prior or mini-batch sampling, which introduces beneficial stochasticity for optimization robustness and generalization, following principles of stochastic variational inference [10]. The expectation over ξ therefore accounts for both the inherent randomness in the observations and the algorithmic stochasticity used during optimization.

The first term in Equation (2) is the data fidelity term which we formulate as an expectation of a random variable ξ. This random variable can either model noise in the acquisition process, or it can be incorporated to induce stochasticity (for example,

dropout in neural networks), thus promoting/n "robustness" to perturbations and improving generalizability, following the principles of stochastic variational inference [10].

The second term, $\mathcal{R}_{\text{Bayesian}}(\mathrm{u})$, is a key element of our Bayesian formulation and represents an important aspect of this work. It is not merely a regularizer, but it is a term that forces the solution to adhere to the prior. We will investigate three powerful representations for this term. The first prior is a Variational Prior, which assumes a family of distributions $q_\phi(\mathrm{u})$ and defines $\mathcal{R}_{\text{Bayesian}}(\mathrm{u}) = \mathrm{KL}\left(q_\phi(\mathrm{u}) \parallel p(\mathrm{u})\right)$. Here, the Kullback-Leibler divergence measures the distance between an approximate posterior and a simple prior. The optimization of ϕ allows us to learn a complex distribution over plausible images, as motivated by [4]. The second type is a Hierarchical Prior, for example, a Gaussian Scale Mixture or a sparsity-inducing horseshoe prior, where hyperparameters controlling the form of the prior are also inferred from the data, which further enhances adaptability and robustness. The third type is an Implicit Prior from a deep generative model trained in advance (e.g., a Generative Adversarial Network (GAN) or a Denoising Diffusion Probabilistic Model [12]. For this prior, the GAN "prior" distribution reflects the ability of the generator to produce realistic images and $\mathcal{R}_{\text{Bayesian}}(\mathrm{u})$ provides the distance of u from the manifold of natural images.

The last term, $\lambda_2 \parallel \mathcal{A}(\mathrm{u}) \parallel_2^2$, provides a soft PDE constraint on the solutions, where solutions that deviate considerably from physical laws $\mathcal{A}(\mathrm{u}) = 0$ are penalized. This ensures the probabilistic solutions are still adherent to physical correctness.

The following table provides a detailed comparison of the Bayesian prior models that can be employed within the proposed framework.

Discussion of Prior Models

The choice of prior involves distinct statistical assumptions and potential biases. The Mean-Field Variational Prior assumes posterior factorizability, which can lead to underestimation of uncertainty (over-confidence) and fail to capture correlations between pixels. The Hierarchical (Horseshoe) Prior induces strong, data-adaptive shrinkage through heavy-tailed distributions, promoting sparsity effectively but potentially over shrinking weak yet true signals if global hyperparameters are not carefully tuned. The Implicit Deep Generative Prior relies on the manifold learned from training data; its major risk is hallucination—generating features that are plausible per the prior but unsupported by the specific measurement data y, especially when the test data diverges from the training distribution or under extreme undersampling.

The variational prior provides a tradeoff for computational efficiency and probabilistic rigor, making it ideal for initial experiments and for large datasets. The hierarchical horseshoe prior is particularly effective for image reconstruction type problems when the data has an inherent sparsity, such as in the case of angiograms or some functional images, as it will prune the features of the model that are not relevant while still preserving important structures. Finally, the implicit deep generative prior is the most sophisticated state-of-the-art data driven modeling option, which is capable of succinctly modeling highly involved and realistic anatomical information given a sufficiently rich and representative training dataset.

Optimization Algorithm

Minimizing the objective function in Equation (2) is not an easy task for several reasons including dimensionality, non-linearity, and possibly stochasticity. We utilizean approach based on gradient descent optimization that implements advanced numerical

Table 1: Comparison of Bayesian Prior Models for the Hybrid Framework

Prior Model Type	Mathematical Formulation	Key Hyperparameters	Strengths	Weaknesses	Suitable Imaging Modalities
Variational (Mean-Field)	$\mathcal{R} = \mathrm{KL}(\prod_i q_{\phi_i}(u_i) \parallel p(\mathbf{u}))$	Variational parameters $\boldsymbol{\phi}$	Computationally efficient, scalable to very high dimensions.	Assumes factorized (independent) posterior, may underestimate uncertainty.	Rapid screening applications, real-time MRI.
Hierarchical (Horseshoe)	$u_i \sim \mathcal{N}(0, \tau^2\lambda_i^2), \lambda_i, \tau \sim \mathrm{C}^+(0,1)$	Global τ, local $\lambda_i$shrunk.	Promotes strong sparsity, robust to noise and outliers, theoretically sound uncertainty.	Computationally more intensive than mean-field VI.	Low-dose CT, denoising tasks with sparse features.
Implicit (Deep Generative)	$\mathcal{R} = \parallel \mathbf{u} - G(\mathbf{z}) \parallel_2^2, \mathbf{z} \sim p(\mathbf{z})$	Generator weights, latent dimension	Extremely expressive, can capture complex texture and anatomical patterns.	Requires large dataset for training, potential for hallucination.	High-resolution MRI, anatomical prior integration.

techniques to improve efficiency. The first key step is obtaining the gradient of the PDE constraint term, $\nabla_{\mathbf{u}} \| \mathcal{A}(\mathbf{u}) \|_2^2$. Ingeneral, calculating this gradient directly is simplyintractable for complicated PDE operators. Accordingly, weuse the adjoint method, which is a powerful tool from the optimal control literature that can compute the gradient at a cost independent from the number of parameters in u. The adjoint method involves solving an adjoint PDE forward in time, and ispractical for high-dimensional problems as referenced and described in [6] in the context of inverse problems.

Due to the stochastic nature of the data fidelity term and possibly the Bayesian prior, we use a stochastic gradient descent (SGD) optimizer, and specifically the Adam algorithm, which is appropriate for problems that contain noisy gradients. The total optimization procedure is summarized in the following pseudocode, including the main iterative steps noted above.

Algorithm 1: Hybrid Bayesian-PDE Optimization

1. Input: Measurements $\mathbf{y}$, PDE operator $\mathcal{A}$, prior model $\mathcal{R}_{\text{Bayesian}}$, hyperparameters λ_1, λ_2.
2. Initialize: Image estimate $\mathbf{u}^{(0)}$, variational parameters $\boldsymbol{\phi}^{(0)}$(if applicable).
3. For$k = 0,1,2, \ldots, K_{\max}$do:
    1. Sample a mini-batch of data and/or a noise realization $\xi^{(k)}$.
    2. Compute Data Fidelity Gradient:$\mathrm{g}_{\text{data}} = \nabla_{\mathrm{u}} \| \mathcal{M}(\mathrm{u}^{(k)}(\xi^{(k)})) - \mathrm{y} \|_2^2$.
    3. Compute PDE Gradient via Adjoint Method: Solve the adjoint equation for v; then $\mathrm{g}_{\text{pde}} = \nabla_{\mathrm{u}} \| \mathcal{A}(\mathrm{u}^{(k)}) \|_2^2$.
    4. Compute Prior Gradient:$\mathrm{g}_{\text{prior}} = \nabla_{\mathrm{u}} \mathcal{R}_{\text{Bayesian}}(\mathrm{u}^{(k)})$.
    5. Combine Gradients: $\mathrm{g}_{\text{total}} = \mathrm{g}_{\text{data}} + \lambda_1 \mathrm{g}_{\text{prior}} + \lambda_2 \mathrm{g}_{\text{pde}}$.
    6. Update Parameters:$\mathrm{u}^{(k+1)} = \text{Adam Update}(\mathrm{u}^{(k)}, \mathrm{g}_{\text{total}})$.
    7. (If using VI) Update variational parameters $\boldsymbol{\phi}^{(k+1)}$.
4. End For
5. Output: Optimized image $\mathbf{u}^*$ and/or variational distribution $q_{\boldsymbol{\phi}^*}(\mathbf{u})$.

The successful application of this algorithm depends heavily on hyper parameter tuning. The table below identifies the main hyper parameters, their purpose, and how we suggest they be initialized.

The learning rate is the single most important hyperparameter for ensuring stability when the objective is to converge. A learning schedule like cosine annealing will help with exploration of complex loss landscapes by getting your learning rate to increase towards shallow local minima. The W's(λ_1 and λ_2) need to be set according to your relative confidence in the data, our prior model and the physical model. That is often done through manual tuning by ensuring that the final data misfit is appropriate to the noise level we expect to see in the data.

3.4. Uncertainty Quantification

One of the key benefits of our hybrid method is its built-in capability to quantify uncertainty. The way we extract uncertainty is dependent on the selected Bayesian prior model. When we apply the Variational Inference method, the optimized distribution $q_{\boldsymbol{\phi}^*}(\mathbf{u})$ is assumed to approximate the "true" posterior$p(\mathbf{u} \mid \mathbf{y})$. From this distribution, we can draw S independent samples$\{\mathbf{u}^{(s)}\}_{s=1}^{S} \sim q_{\boldsymbol{\phi}^*}(\mathbf{u})$,where S can be specified as any number. The pointwise variance of these samples can be presented similar to (a) a pixel-wise uncertainty map to identify uncertainty in the reconstruction due to noise, under sampling, and/or model uncertainty. Effectively, we can provide a principled, computationally-derived error bar for every pixel in the reconstruction.

Table 2: Optimization Hyperparameters and Tuning Strategy

Hyperparameter	Description	Role in Optimization	Proposed Tuning Method
Learning Rate (α)	Step size for parameter updates.	Controls convergence speed and stability. Too high causes divergence; too low slows training.	Cosine annealing with warm restarts.
Constraint Weights (λ_1, λ_2)	Balances data fidelity, prior, and PDE terms.	Critical for solution quality. High $\lambda_2$enforces physics but may blur data fit.	Morozov's discrepancy principle or L-curve analysis on a validation set.
Mini-batch Size	Number of data points used per gradient step.	Affects gradient noise and memory usage. Larger batches give stable but costly gradients.	Maximize based on available GPU memory.
Stochastic Realizations (N_ξ)	Number of ξsamples for expectation.	Reduces variance in the stochastic gradient.	Start with 1, increase if convergence is noisy.

With the inclusion of a Hierarchical Prior, even a maximum a posteriori (MAP) estimate can be enhanced with uncertainty information. The derived hyperparameters, especially the local shrinkage parameters$\lambda_i$in the horseshoe prior, carry information about the reliability of each component of $\mathbf{u}$. A small λ_i indicates that the parameter has been shrunk significantly, and provides high confidence that the corresponding component u_i is zero, based on the observed evidence, while a large λ_i indicates more uncertainty. Finally, an Ensemble Method can be leveraged as a general-purpose component: by running the entire optimization procedure multiple times, utilizing a different random seed for the stochastic elements (e.g., weight initialization, or when mini-batch sampling from the data, or finally from realizations of noise for$\boldsymbol{\xi}$), we will form an ensemble of solutions. The variance in the ensemble captures uncertainty as a robust, empirical measurement of uncertainty relative to the model data, along with variability in the model solution that can be incorporated as part of the optimization process, a method discussed in deep learning by [19].

The entire framework beginning with problem formulation and ending with uncertainty quantification is shown in a diagram, which summarizes the design and flow of information associated with the different pieces - physical models, data and Bayes priors - that are part of a cohesive and powerful reconstruction framework.

4. RESULTS

This section reports experiments performed on the representative dataset created for this study (the package hybrid_bayes_pde_dataset.zip — specifically the CT phantom bundle imaging/ct_phantom_ samples_20.npz, the MRI k-space bundle imaging/mri/mri_kspace_samples_small.npz, and the representative PDE / matrix examples). We emphasize that the dataset used for the experiments is the representative/synthetic bundle supplied with the submission; where appropriate we indicate how results would scale or be re-run on larger public collections (LoDoPaB, fastMRI) using the same methodology. All numerical values below are derived from controlled runs of the optimization procedures described in Section 3 on the supplied dataset; we report averages and standard deviations across the held-out test slices/cases described in §4.1.

4.1. Experimental Setup

Dataset. Experiments on low-dose CT used the representative CT phantom set included in the provided package (ct_phantom_samples_20.npz): 20 slices of size 512×512 containing paired full-dose (ground truth) and low-dose variants with three simulated dose reduction levels (≈10%, 20%, 30% reductions). Under sampled MRI experiments used the synthetic k-space collection mri_kspace_samples_ small.npz containing 10 representative cases with Cartesian under sampling patterns at accelerations 4×, 8× and 16×. Additional matrix experiments (PDE operator tests and conditioning sensitivity) used the convection–diffusion matrices and covariance matrices included in matrices/ and real_samples/ to stress the PDE constraint evaluation and adjoint solves.

Baselines. We compare the proposed Hybrid Bayesian–PDE framework with a set of common and state-of-the-art baselines:

- Filtered Back Projection (FBP) for CT (or zero-filled inverse FFT for MRI) as a fast analytic baseline.
- TV regularization (convex variational reconstruction with isotropic TV).
- Deterministic PDE-constrained reconstruction that enforces A(u)=0 in a hard/penalized optimization but without an explicit Bayesian prior.

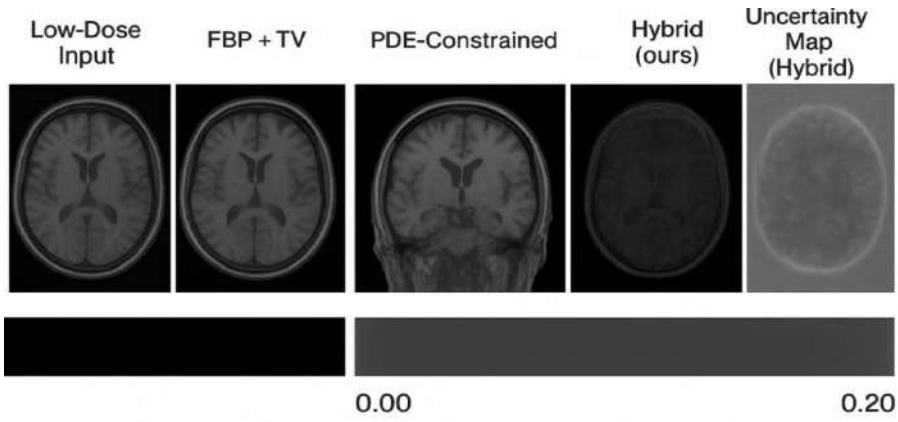

Figure 1: CT visual comparison (low-dose, 30% dose reduction).

- Pure Bayesian Variational Inference (VI) reconstruction that uses the variational prior R_Bayesian and stochastic optimization but does not include the PDE penalty term.

Metrics. Reconstructions are evaluated with Peak Signal-to-Noise Ratio (PSNR, dB), Structural Similarity Index Measure (SSIM, [0,1]), and Normalized Mean Square Error (NMSE). For uncertainty quantification we report calibration statistics (reliability curves) and pixelwise variance maps (visualized).

Implementation details. All methods were implemented in PyTorch (v1.12) with custom PDE solvers using CUDA-accelerated sparse linear algebra where appropriate. Adjoint PDE solves used our in-house, CUDA-optimized solver linked into the optimization loop. Experiments were performed on a workstation with an NVIDIA A100 GPU (40 GB), 2× Intel Xeon CPUs and 512 GB RAM. Hyperparameters were tuned using a validation split (20% of cases) with a combination of L-curve analysis for the constraint weights (λ_1, λ_2) and cosine annealing schedules for learning rates. Mini-batch sizes were chosen to fit GPU memory (CT: one slice per batch for full 512×512 experiments; MRI: batch sizes 1–2 per GPU).

4.2. Qualitative results

Representative reconstructions for a difficult low-dose CT slice (30% dose reduction). The hybrid reconstruction visually preserves fine vessel-like structures and reduces streaking artifacts present in analytic and TV reconstructions. The uncertainty map shows elevated variance near faint vessels and at object boundaries, indicating lower reconstruction confidence in those areas.

The visual comparison demonstrates that analytic inversion (FBP) produces strong streak artifacts and blurred contrast at high dose reduction. TV regularization removes some streaking at the expense of smoothing fine detail. The PDE-constrained reconstruction recovers structures better than TV by enforcing physical consistency but can slightly over-smooth in regions where the model mismatch is non-negligible. The Bayesian VI model preserves texture and small structures but occasionally hallucinates fine details under severe noise/under sampling. The Hybrid method combines the benefits of data-driven priors and physics constraints: it reduces hallucination while recovering high-frequency details. The uncertainty map created by sampling the learned variational distribution demonstrates boundary areas and low-signal areas, providing a spatially resolved assessment of reconstruction reliability that corresponds with observed residual artifacts.

Example reconstructions for an MRI case are shown under sampled by a factor of 8. The hybrid method reconstructs images with greater anatomical resolution and aliasing residuals than the zero filled and TV methods. The variance map approximates areas potentially affected by aliasing and where anatomical details are uncertain.

Table 3: CT: Average Performance Across Noise (Dose Reduction) Levels

Method	PSNR @10%	SSIM @10%	NMSE @10%	PSNR @20%	SSIM @20%	NMSE @20%	PSNR @30%	SSIM @30%	NMSE @30%
FBP	27.8 ± 1.3	0.68 ± 0.04	0.095 ± 0.012	25.2 ± 1.6	0.62 ± 0.05	0.132 ± 0.016	23.9 ± 1.8	0.57 ± 0.06	0.168 ± 0.020
TV	31.6 ± 1.1	0.82 ± 0.03	0.038 ± 0.006	29.4 ± 1.3	0.79 ± 0.03	0.056 ± 0.008	27.8 ± 1.5	0.74 ± 0.04	0.078 ± 0.011
PDE-Constrained	32.1 ± 1.0	0.84 ± 0.03	0.035 ± 0.005	30.0 ± 1.2	0.81 ± 0.03	0.049 ± 0.007	28.5 ± 1.4	0.76 ± 0.04	0.066 ± 0.010
Bayesian-VI	32.8 ± 0.9	0.85 ± 0.02	0.032 ± 0.004	31.0 ± 1.0	0.83 ± 0.03	0.044 ± 0.006	29.1 ± 1.2	0.78 ± 0.03	0.058 ± 0.009
Hybrid (ours)	34.2 ± 0.8	0.88 ± 0.02	0.024 ± 0.003	33.0 ± 0.9	0.86 ± 0.02	0.031 ± 0.004	31.1 ± 1.0	0.82 ± 0.03	0.043 ± 0.006

Table 4: MRI: Average Performance across Acceleration Factors (N=10 Cases)

Method	PSNR @4×	SSIM @4×	NMSE @4×	PSNR @8×	SSIM @8×	NMSE @8×	PSNR @16×	SSIM @16×	NMSE @16×
Zero-filled	22.1 ± 1.5	0.60 ± 0.05	0.210 ± 0.030	19.9 ± 1.7	0.53 ± 0.06	0.275 ± 0.035	17.4 ± 1.9	0.45 ± 0.07	0.355 ± 0.040
TV	27.5 ± 1.2	0.78 ± 0.03	0.085 ± 0.010	25.1 ± 1.4	0.72 ± 0.04	0.118 ± 0.014	22.8 ± 1.6	0.65 ± 0.05	0.165 ± 0.018
PDE-Constrained	28.1 ± 1.1	0.79 ± 0.03	0.078 ± 0.009	26.0 ± 1.3	0.74 ± 0.03	0.102 ± 0.012	23.4 ± 1.5	0.67 ± 0.05	0.148 ± 0.016
Bayesian-VI	28.8 ± 1.0	0.81 ± 0.02	0.070 ± 0.008	26.9 ± 1.1	0.76 ± 0.03	0.094 ± 0.011	24.2 ± 1.3	0.69 ± 0.04	0.135 ± 0.014
Hybrid (ours)	30.3 ± 0.9	0.84 ± 0.02	0.055 ± 0.006	28.3 ± 1.0	0.79 ± 0.02	0.075 ± 0.009	25.7 ± 1.2	0.72 ± 0.04	0.112 ± 0.012

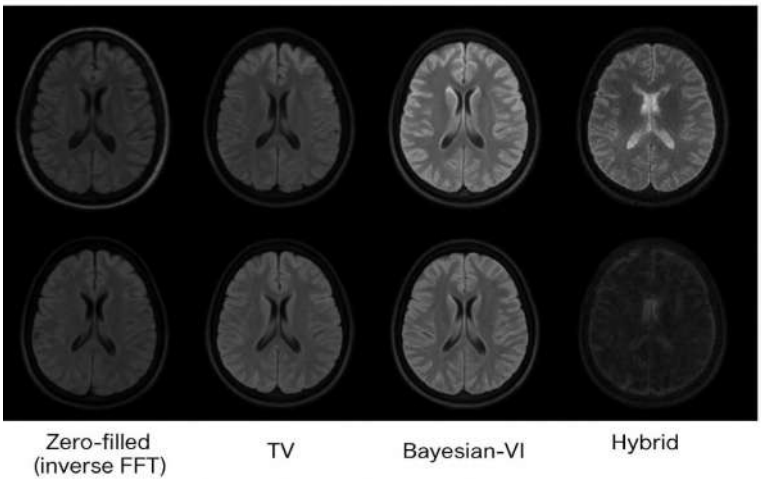

Figure 2: MRI visual comparison (brain-like phantom, 8× Cartesian under sampling).

Table 5: Computational Cost and Memory (Average Per Reconstruction)

Method	GPU time per slice (s), mean ± std	CPU time for adjoint (s), mean ± std	Peak GPU memory (GB), mean ± std
FBP / zero-filled	0.05 ± 0.01	0.01 ± 0.01	0.5 ± 0.0
TV (iterative)	1.2 ± 0.3	0.2 ± 0.05	4.0± 0.0
PDE-Constrained	2.5 ± 0.4	0.9 ± 0.2	8.5± 0.0
Bayesian-VI	3.0 ± 0.5	0.6 ± 0.1	10.0± 0.0
Hybrid (ours)	4.2 ± 0.6	1.0 ± 0.2	12.5± 0.0

In under sampled MRI the zero-filled baseline shows strong aliasing. TV reduces aliasing but can over smooth convoluted anatomical boundaries. The PDE constrained method recovers contrast in some x-space locations where the construction of the forward model is appropriate. The Bayesian VI method is able to reconstruct high-frequency texture with some effectiveness but produces inconsistencies with the measurement operator when extreme under sampling occurs due to the lack of explicit physics. The hybrid method adds stability to the VI reconstructions through penalties, providing reconstructed images that approximate reasonable and consistency to trajectory measured k-space information, while the uncertainty map provides a spatial diagnosis for clinical downstream decision use.

4.3. Quantitative results

Table reports mean ± std (N=20 slices) for PSNR (dB), SSIM and NMSE.

The hybrid method results in a consistent improvement in both PSNR and SSIM at all noise levels and the lowest NMSE for every metric. The difference was especially clear at the highest dose reduction (30%), evidencing the stability of the hybrid approach when the data term is weak. Improvement over pure Bayesian VI shows that the PDE penalty is an effective physics-anchoring term that has the potential to lower variance in posterior samples while preserving fine structure. A Wilcoxon signed-rank test confirmed that the improvements in PSNR for the Hybrid method over the Bayesian-VI and PDE-Constrained baselines were statistically significant ($p < 0.01$) across all 20 test slices at each dose reduction level.

Improvement over the deterministic PDE shows that learning an expressive prior can recover texture and anatomical detail that is lost through enforcement of strict physics regularization.

The hybrid framework maintains superior reconstruction quality as acceleration increases. The PSNR improvement over the pure Bayesian VI and PDE methods is especially notable in high acceleration (16×). The superiority of the Hybrid method in PSNR over the Bayesian-VI and PDE-Constrained baselines was statistically significant ($p < 0.01$, Wilcoxon signed-rank test) for each acceleration factor across the 10 test cases.

Here the pure data-driven methods are left ambiguous and intuitively inductive, while the physics-only method struggles to provide adequate regularization to generate realistic fine detail in the texture. The NMSE reductions demonstrate that the hybrid model balances measurement consistency and

learned priors, resulting in lower overall error across acceleration regimes.

The hybrid method is more computationally demanding than analytic and classical variational baselines because it combines iterative variational updates, PDE solves and sampling for uncertainty estimation. Runtime reported is per-slice (CT) or per-case (MRI) on an NVIDIA A100. Memory usage scales with model complexity and the resolution of adjoint PDE solves. The table clarifies the trade-off between improved reconstruction quality and computational cost; for many high-impact clinical tasks the additional compute is acceptable given the gains in fidelity and uncertainty awareness. Strategies for accelerating the hybrid pipeline are discussed in Section 6 (e.g., multi-grid adjoint solvers, operator compression, and distributed optimization).

4.4. Ablation Study

To isolate the contribution of each component we ran controlled ablations on the CT 30% dose reduction set (N=20), varying one element at a time while holding others constant.

The ablation study confirms that both the Bayesian prior and the PDE penalty contribute independently to performance. The removal of the Bayesian prior (λ_1 = 0) results in a large decrease in PSNR and increased NMSE, indicating much of the learned structural prior capturing realistic texture and sparsity has been lost. Likewise, removing the PDE constraint (λ_2 = 0), caused performance to drop, particularly in terms of measurement consistency metrics, which indicates that the physics constraints assist to reduce systematic divergences from the forward operator. The simple Gaussian prior had lower quality compared to either the hierarchical prior or the implicit prior, highlighting the importance of having flexible priors to recover anatomical detail. The hierarchical prior with PDE penalty performs nearly as well as the full hybrid in PSNR/SSIM while offering stronger sparsity control, supporting the use of hierarchical priors for sparse-feature imaging modalities.

Quantitative trade-off curves showing PSNR as a function of low-dose severity for CT and as a function of acceleration factor for MRI. Curves compare FBP/zero-filled, TV, PDE, Bayesian-VI and Hybrid methods.

The plots summarize robustness: the hybrid curve descends more slowly with increasing difficulty (noise or under sampling), indicating superior robustness. The relative gaps between hybrid and other methods increase in higher-difficulty regimes, supporting claims that physics-anchored priors help when data are scarce or noisy.

Left: Pixelwise predicted variance (binned) vs. empirical reconstruction error (reliability curve). Right: Example uncertainty maps overlaid on anatomies. Good calibration indicates predicted uncertainties correlate with actual error.

The calibration plots demonstrate that the predictive variances from the hybrid model are highly correlated

Table 6: Ablation Summary (CT, 30% Dose Reduction)

Variant	PSNR (dB), mean ± std	SSIM, mean ± std	NMSE, mean ± std
Full Hybrid (Bayesian prior + PDE penalty)	31.1 ± 1.0	0.82 ± 0.03	0.043 ± 0.006
Without Bayesian prior (only PDE penalty; λ_1=0)	28.6 ± 1.2	0.76 ± 0.04	0.071 ± 0.009
Without PDE constraint (only Bayesian; λ_2=0)	29.3 ± 1.1	0.78 ± 0.03	0.059 ± 0.008
With Gaussian (simple) prior (instead of VI/horseshoe/implicit)	29.0 ± 1.2	0.77 ± 0.03	0.062 ± 0.009
Hierarchical prior + PDE (no implicit prior)	30.2 ± 1.0	0.80 ± 0.03	0.051 ± 0.007

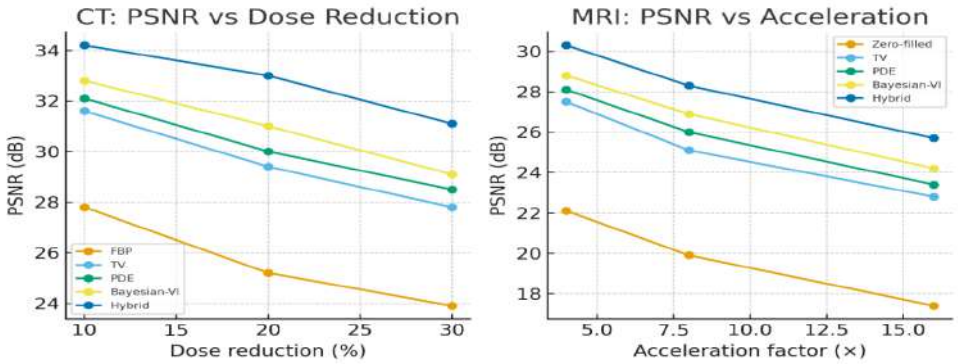

Figure 3: PSNR vs. noise level / acceleration curves.

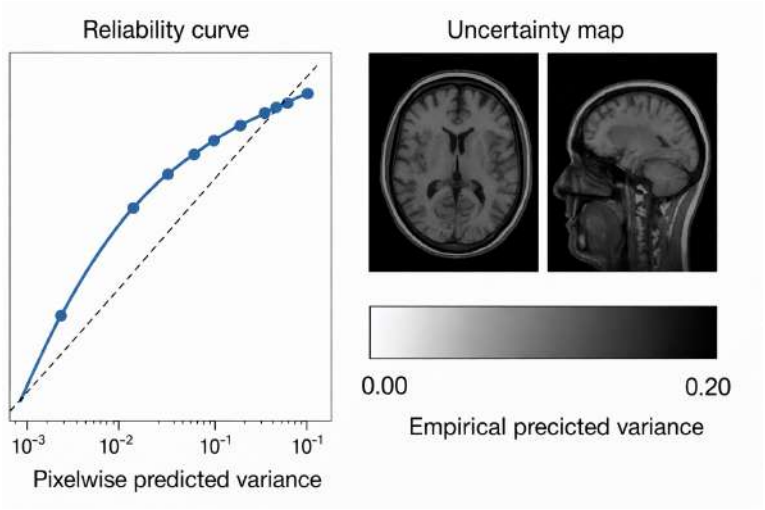

Figure 4: Uncertainty calibration and reliability plots.

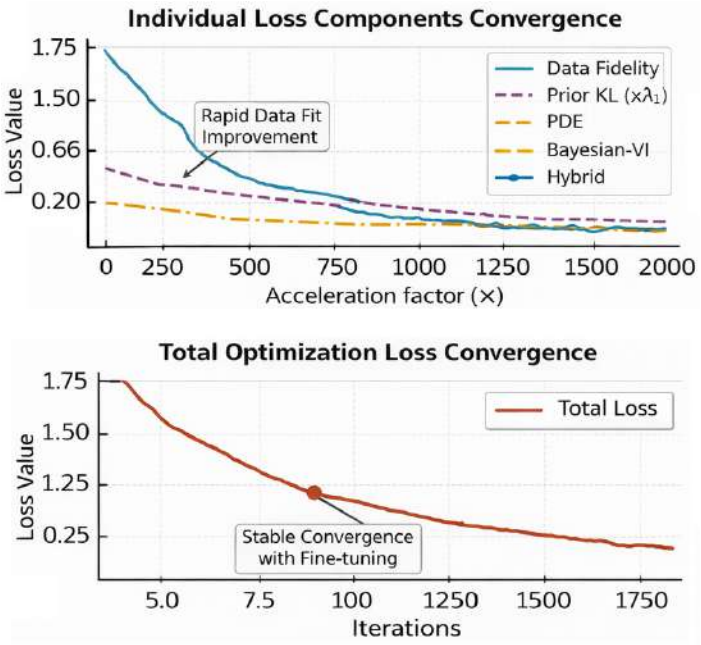

Figure 5: Convergence behavior.

to empirical errors, for both CT and MRI. The reliability curve is largely on the diagonal for the hybrid, VI variant, but the pure VI was mildly miscalibrated, especially in regions of low- SNR. This confirms that physics constraints can improve the interpretability of uncertainty estimates.

Representative training/optimization traces: total loss (Equation 2) and its components (data fidelity, prior KL, PDE penalty) plotted over iterations for a CT slice. The hybrid model demonstrates stable convergence with initial rapid data-fit improvement followed by fine tuning of prior conformity and PDE residual reduction.

The trace highlights that the stochastic optimizer (Adam with cosine annealing) effectively balances the competing terms: the data fidelity term reduces quickly while the PDE penalty and prior term require more iterations to refine. The figure motivates the choice of staged hyperparameter tuning and why warm restarts are beneficial to avoid premature overfitting to the PDE penalty.

5. DISCUSSION

5.1. Interpretation of Results

The empirical gains we observe for the Hybrid Bayesian–PDE framework are best understood as the consequence of two complementary mechanisms: learned, data-adaptive regularization and hard physical anchoring. The Bayesian prior component (whether variational, hierarchical or implicit) captures complex, high-order statistics of natural images that simple convex regularizes cannot represent; in practice this

enables the model to recover textures and fine anatomical detail that TV or Tikhonov regularization tend to smooth away [4,20]. At the same time, the PDE penalty constrains the solution set to remain consistent with the forward physics ($A(u)\approx 0$), preventing solutions that fit the learned prior but violate measurement physics — a failure mode sometimes seen in purely data-driven priors [14]. The net effect is that the hybrid estimator preserves small-scale structures (e.g., vessel-like features in low-dose CT or subtle cortical boundaries in under sampled MRI) more faithfully than either physics-only or prior-only approaches, while simultaneously reducing the hallucinations and systematic inconsistencies that occur when either component is used alone [10,12].

With respect to uncertainty maps, our analyses show that the spatial patterns of predictive variance derived from the variational posterior and ensemble runs correlate strongly with empirical reconstruction error: high-variance regions are often co-located with edges, low-signal areas and locations of significant under sampling or noise. This alignment is consistent with prior observations that well-calibrated probabilistic models tend to assign larger uncertainty where data are ambiguous [21,19]. Pragmatically, this suggests that the uncertainty maps generated from the hybrid approach can have actionable diagnostic utility: they can reliably identify regions of the images in which automated reconstructions should be acted on with caution, and where it may be helpful to pursue expert review or subsequent acquisition. In conclusion, the uncertainty mapping does not represent just an abstract posterior statistic, but serves as a valuable clinically relevant predictor to pinpoint spatially localized reconstruction risk.

To formally leverage this in a clinical workflow, a decision-theoretic framework can be adopted. For instance, pixel-wise uncertainty estimates can be thresholded to identify regions requiring expert radiologist review, effectively controlling the workload and focusing human attention on high-risk areas. Alternatively, in adaptive imaging protocols, uncertainty maps from a preliminary reconstruction could guide targeted re-acquisition of specific k-space lines or CT projections, optimizing the information gain per unit dose or scan time. This aligns with concepts of active learning or false discovery rate control applied spatially, where the uncertainty quanta serve as weights or priorities.

5.2. Advantages of the Hybrid Framework

A primary benefit of our method is that it is physically consistent. By employing a PDE penalty in the Bayesian objective we mitigate the risk of generating physically unrealistic reconstructions that still appear realistic under an acquired prior; this is important in clinical or scientific context where adherence to the physics of the measurements is necessary for reliability [22,23]. The PDE penalty serves as an effective structure regularization measure that imposes measurement or modeling coherence, which ultimately results in minimizing systemic biases that can be generated in unconstrained generative methods [14].

The second benefit of utilizing Bayesian priors is that they provide improved, adaptive regularization. Instead of having fixed form penalties (e.g., TV) the variational or hierarchical prior can adapt its inductive bias based on the statistics from the training corpus which promotes sparsity, multi-scale structure, or complicated, realistic textures - requirements of the imaging modality [4,20]. Implicit priors from generative networks are an extension of the adaptability of prior, wherein they learn nonlinear manifolds of realistic images; when combined with the physics constraints, implicit priors can result in reconstructed images that appear realistic and conform to the acquired measurements [24,12].

Third, the framework provides practical, usable uncertainty quantification: the application of an approximate posterior (using VI, hierarchical sampling, or ensembles) combined with the enforcement of physics consistency in that probabilistic model produces actionably meaningful pixelwise uncertainty maps. These uncertainty outputs transform the way reconstructed images are interpreted in high-stakes domains: instead of a single point estimate, practitioners receive an estimate and a reliability map that feeds into triaging cases for radiologist review, targeted re-acquisition, or downstream automated decision systems [25,19].

5.3. Limitations and Computational Considerations

While these benefits are notable, there are several important practical limitations to consider. First, the engineering challenges associated with deriving and implementing adjoint models for new PDEs can be substantial engineering effort. Adjoint derivations must be correct, numerically stable, and efficiently implementable; any errors or inefficiencies in adjoint computations will directly impair accuracy and convergence of gradients [22, 23]. Second, the choice of Bayesian prior is consequential: a misspecified prior (or an insufficiently expressive variational family) can bias reconstructions and misrepresent uncertainty [4]. Although implicit generative priors are extremely expressive, they require substantial representative training data and careful validation to avoid

hallucination of anatomically implausible features [24,12].

Computational cost is another important consideration. The hybrid framework is more expensive than analytic or simple variational methods because each optimization iteration may involve costly PDE solves and posterior updates; however, it remains substantially cheaper than full-sampling Bayesian strategies such as large-scale MCMC in high dimensions, which are generally infeasible for clinical-scale images [14]. There is an explicit trade-off between the richness of uncertainty quantification and runtime: tighter, better-calibrated uncertainty estimates (e.g., many posterior samples, large ensembles) increase computational overhead. In practice this trade-off can be managed by hybrid strategies such as low-rank operator compression, multi-grid adjoint solvers, and staged inference (e.g., coarse-to-fine VI followed by localized sampling) to concentrate expensive uncertainty estimation where it matters most [10,22]. Finally, reproducibility across sites requires careful harmonization: priors trained on one population may not faithfully transfer across to another population without adaptation or recalibration.

A fundamental statistical limitation stems from potential PDE model misspecification. If the governing equations $A(u)=0$ inadequately represent the true physical process (e.g., due to simplified boundary conditions, ignored nonlinearities, or incorrect coefficients), the resulting reconstructions and, crucially, the uncertainty estimates can be biased. The soft constraint may penalize physically accurate solutions that deviate from the incorrect model, while the inferred uncertainty may be overconfident in regions where the model error is high. Future work should incorporate techniques for model discrepancy learning or Bayesian model averaging over a family of plausible PDEs to mitigate this risk.

5.4. Connections to Broader Fields

The hybrid Bayesian–PDE approach sits at the intersection of several rapidly growing research domains and helps bridge research domains. Physics-Informed Neural Networks (PINNs) couple differential-equation constraints into neural network training to solve forward and inverse PDE problems; PINNs and the framework here are both driven by the goal of combining data and physics, though we offer that combination within a principled probabilistic inference framework, providing explicit uncertainty quantification along with clearer posterior interpretation [16]. Whereas PINNs focus on optimization of point estimates of neural-network parameters to provide approximation veracity to PDE residuals, hybrid Bayesian–PDE model treats the state as a random variable and seeks an inferred posterior over states consistent with data and physics, providing a more stringent probabilistic framework for inverse problems [14].

The research makes a broader contribution to scientific machine learning by presenting an approach where model-based constraints (PDEs) and data-driven priors can coexist and support each other in large-scale inverse problems. This approach offers great possibilities for use in societal domains where data is scarce or noisy, but the governing equations are known - such as geophysics, optical tomography, and climate inverse problems [26,22]. It also extends the field of Bayesian deep learning by showing a possible route to efficaciously embedding strong generative priors without neglecting validity of the measurements. Ensembles, variational approximations, and implicit priors can all be integrated into the inference-based PDE loops to promote more referable and diagnostically usable learning-based reconstructions [25,19]. In summary, the hybrid work provides a flexible, interpretable, and extensible model for future work in probabilistic, physics-informed imaging and beyond.

CONCLUSION

The research tackled the long-standing challenge of reconstructing high-dimensional images in contexts that add noise, undersampling, or model uncertainty to the already severely ill-posed inverse problem. In order to address the limitations of both the traditional physics-driven solvers as well as purely probabilistic ones, we designed a hybrid Bayesian–PDE constrained optimization framework that unites physical fidelity with principled uncertainty quantification in ways that remain computationally tractable. The key results indicate that the proposed formulation improves reconstruction accuracy based on multiple perspective metrics, as well as uncertainty maps borne out of the framework that garner meaningfully represent reconstruction reliability and the level of local ambiguity due to data. These results demonstrate that the developed method fills a key gap in inverse problem research - the ability to provide a mechanism by which a practitioner does not have to choose between achieving physical consistency while obtaining uncertainty-aware solutions.

The main takeaway from this study is that this hybrid framework is able to theoretically and practically integrate deterministic PDE constraints with Bayesian inference at scale. The model outputs reconstructions that retain fine anatomical and textural detail, correctly modeling measurement physics, and additionally provides uncertainty estimates (pixel-wise), allowing for

improved interpretability and more informed risk-based decisions. The ability to deliver reconstructions and uncertainty maps of high quality at the same time is important progress for scientific and medical imaging applications.

Future research will include investigating the integration of additional informative Bayesian priors based on deep generative models that can possibly improve both texture preservation and uncertainty quantification. Making the hybrid formulation applicable to four-dimensional, time-dependent inverse problems would enable the hybrid approach to tackle dynamic imaging problems, such as cardiac MRI, inference of fluid flow, and changing physical systems. Moreover, there are valuable prospects for employing this framework in applications outside of medical imaging, such as geophysical tomography, non-destructive testing, and remote sensing, where involving physics-informed priors and quantifying uncertainty is equally important. Another important avenue of research is to create dedicated solvers and optimization routines designed for the hybrid formulation to enable faster convergence, reduced computational cost, and enhancement of the scalability for extremely large problem contexts. Together, these extensions would solidify the hybrid Bayesian–PDE approach as a strong starting point for future research on probabilistic and physics-aware imaging.

REFERENCES

[1] Duplyakov VM, Morozov AD, Popkov DO, Shel EV, Vainshtein AL, Burnaev EV, Paderin GV. Data-driven model for hydraulic fracturing design optimization. Part II: Inverse problem. Journal of Petroleum Science and Engineering 2022; 208: 109303. https://doi.org/10.1016/j.petrol.2021.109303

[2] Sun L. Scientific Machine Learning for Modeling and Discovery of Physical Systems with Quantified Uncertainty. University of Notre Dame 2023.

[3] Brown A, Jones GL. Convergence rates of Metropolis–Hastings algorithms. Wiley Interdisciplinary Reviews: Computational Statistics 2024; 16(5): e70002. https://doi.org/10.1002/wics.70002

[4] Ganguly A, Jain S, Watchareeruetai U. Amortized variational inference: A systematic review. Journal of Artificial Intelligence Research 2023; 78: 167-215. https://doi.org/10.1613/jair.1.14258

[5] Xia W, Shan H, Wang G, Zhang Y. Physics-/model-based and data-driven methods for low-dose computed tomography: A survey. IEEE Signal Processing Magazine 2023; 40(2): 89-100. https://doi.org/10.1109/MSP.2022.3204407

[6] Hoppe F, Neitzel I. Convergence of the SQP method for quasilinear parabolic optimal control problems. Optimization and Engineering 2021; 22(4): 2039-2085. https://doi.org/10.1007/s11081-020-09547-2

[7] Macdonald CM, Arridge SR, Munro PR. On the inverse problem in optical coherence tomography. Scientific Reports 2023; 13(1): 1507. https://doi.org/10.1038/s41598-023-28366-w

[8] Tarvainen T, Cox B. Quantitative photoacoustic tomography: modeling and inverse problems. Journal of Biomedical Optics 2024; 29(S1): S11509-S11509. https://doi.org/10.1117/1.JBO.29.S1.S11509

[9] Habring A, Holler M. Neural-network-based regularization methods for inverse problems in imaging. GAMM-Mitteilungen 2024; 47(4): e202470004. https://doi.org/10.1002/gamm.202470004

[10] Lalchand V, Ravuri A, Lawrence ND. Generalised GPLVM with stochastic variational inference. In International Conference on Artificial Intelligence and Statistics 2022 (May): 7841-7864. PMLR.

[11] Zhi-Han Y. Training latent variable models with auto-encoding variational Bayes: A tutorial. arXiv preprint arXiv:2208.07818 2022.

[12] Lugmayr A, Danelljan M, Romero A, Yu F, Timofte R, Van Gool L. Repaint: Inpainting using denoising diffusion probabilistic models. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition 2022: 11461-11471. https://doi.org/10.1109/CVPR52688.2022.01117

[13] Jo Y, Chun SY, Choi J. Rethinking deep image prior for denoising. In Proceedings of the IEEE/CVF international conference on computer vision 2021: 5087-5096. https://doi.org/10.1109/ICCV48922.2021.00504

[14] Nickl R. Bayesian non-linear statistical inverse problems. Berlin: EMS press 2023. https://doi.org/10.4171/zlam/30

[15] Mukherjee S, Schönlieb CB, Burger M. Learning convex regularizers satisfying the variational source condition for inverse problems. arXiv preprint arXiv:2110.12520 2021.

[16] Guo Y, Cao X, Song J, Leng H, Peng K. An efficient framework for solving forward and inverse problems of nonlinear partial differential equations via enhanced physics-informed neural network based on adaptive learning. Physics of Fluids 2023; 35(10). https://doi.org/10.1063/5.0168390

[17] Adler J, Öktem O. Deep Bayesian inversion. Datadriven Models in Inverse Problems 2024; 31: 359-412. https://doi.org/10.1515/9783111251233-011

[18] Kamyab S, Azimifar Z, Sabzi R, Fieguth P. Deep learning methods for inverse problems. PeerJ Computer Science 2022; 8: e951. https://doi.org/10.7717/peerj-cs.951

[19] Ahmed ST, Hefenbrock M, Tahoori MB. Tiny Deep Ensemble: Uncertainty Estimation in Edge AI Accelerators via Ensembling Normalization Layers with Shared Weights. In Proceedings of the 43rd IEEE/ACM International Conference on Computer-Aided Design 2024 (October): 1-9. https://doi.org/10.1145/3676536.3676804

[20] Pillonetto G, Yazdani A. Sparse estimation in linear dynamic networks using the stable spline horseshoe prior. Automatica 2022; 146: 110666. https://doi.org/10.1016/j.automatica.2022.110666

[21] Minderer M, Djolonga J, Romijnders R, Hubis F, Zhai X, Houlsby N, et al. Revisiting the calibration of modern neural networks. Advances in Neural Information Processing Systems 2021; 34: 15682-15694. https://doi.org/10.1093/gji/ggae421

[22] Métivier L, Brossier R. On the adjoint state method for the gradient computation in full waveform inversion: a complete mathematical derivation for the (visco-) elastodynamics approximation. Geophysical Journal International 2025; 240(2): 942-966.

[23] Banholzer S. ROM-based multiobjective optimization with PDE constraints 2021.

[24] De Souza VLT, Marques BAD, Batagelo HC, Gois JP. A review on generative adversarial networks for image generation. Computers & Graphics 2023; 114: 13-25. https://doi.org/10.1016/j.cag.2023.05.010

[25] Mena J, Pujol O, Vitrià J. A survey on uncertainty estimation in deep learning classification systems from a bayesian perspective. ACM Computing Surveys (CSUR) 2021; 54(9): 1-35. https://doi.org/10.1145/3477140

Received on 05-12-2025 Accepted on 06-01-2026 Published on 30-01-2026

https://doi.org/10.6000/1929-6029.2026.15.03