BEYOND L_2: SOBOLEV-KINEMATIC FLOW MATCHING FOR EXTREME SPARSE PHYSICAL INVERSION

Anonymous authors
Paper under double-blind review

ABSTRACT

Physical inversion from extremely sparse sensors demands expressive priors. Prevailing paradigms either rely on L_2 metrics—causing high-frequency artifacts—or depend on specific differentiable partial differential equations (PDEs) solvers at inference, leading to severe collapse under observation sparsity, massive computational overhead and poor cross-scenario generalization. To transcend this, we propose Sobolev-Kinematic Flow Matching (SK-FM). First, breaking away from standard L_2 paradigms, we embed the joint probability flow for scalar fields and spatial gradients directly within the $W^{1,2}$ Sobolev space, implicitly learning differential operators without heuristic interpolations. Second, we introduce an equation-free Inference Kinematic Guidance. Leveraging ODE-based terminal state estimation to formulate geometric objectives—spatial derivative consistency and a curl-free constraint ($\nabla \times \mathbf{g} \equiv 0$)—we steer the trajectory toward the valid irrotational manifold. This propagates kinematic corrections back into the Sobolev space, successfully bypassing explicit numerical solvers. Extensive experiments across Darcy flow, Helmholtz, Structural Health Monitoring, and Navier-Stokes demonstrate that SK-FM achieves state-of-the-art inversion fidelity at 128×128 resolution using merely 128 random observations (<0.79%). Remarkably, SK-FM reduces average inversion error by 78% compared to baselines. Finally, uncertainty quantification highlights its potential for real-world engineering applications.

1 INTRODUCTION

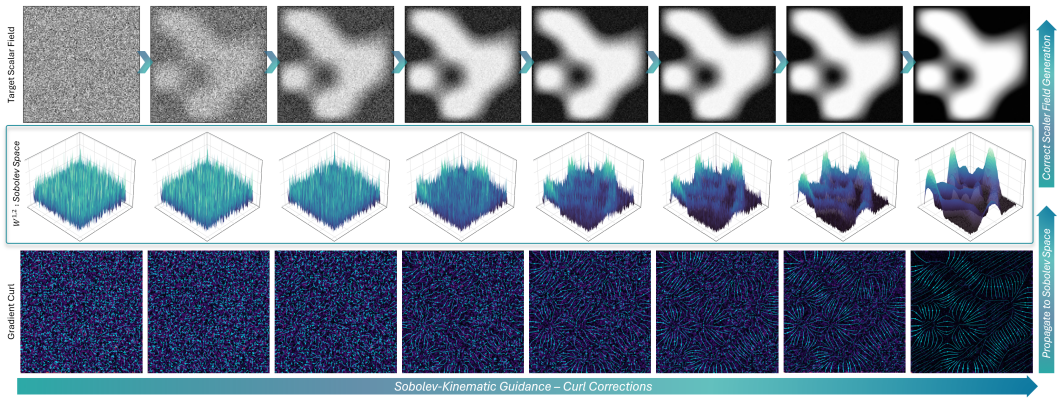


Figure 1: **SK-FM inference trajectory for sparse physical inversion.** Kinematic guidance progressively suppresses non-physical curl ($\nabla \times \mathbf{g} \to 0$) and propagates corrections into Sobolev space for field generation.

Physical inversion is severely ill-posed due to extreme sensor sparsity (Tarantola, 2005), necessitating highly expressive generative priors to explore the valid solution space (Lipman et al., 2023). However, standard generative models like Continuous Normalizing Flows rely on L_2 metrics, which ignore high-frequency geometric oscillations. Upon differentiation, these errors undergo catastrophic amplification, violating conservation laws (Karniadakis et al., 2021; Rahaman et al., 2019). While recent paradigms like Diffusion Posterior Sampling (DPS) (Chung et al., 2024) and DiffusionPDE

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

(Huang et al., 2024) enforce physical consistency via inference-time likelihood guidance, their reliance on explicit differentiable PDE solvers introduces three critical bottlenecks. First, they inherently lack cross-scenario generality, requiring custom-built solvers for each specific physical system. Second, unrolling these solvers at every generation step incurs exorbitant computational overhead. Finally, and most critically, under extreme sensor sparsity, the pathologically weak measurement gradients fail to adequately guide the generation trajectory, frequently precipitating severe mode collapse.

Motivated by the constraints existing in prevailing paradigms, we pioneer a more innovative and foundational technical path. Specifically, we explore a strictly universal pathway: bypassing explicit PDE solvers entirely at test-time by anchoring the generative prior in fundamental functional and geometric principles. To this end, we propose Sobolev-Kinematic Flow Matching (SK-FM)[1]. SK-FM operates through two core mechanisms. First, a 4-Way joint architecture simultaneously models probability flows of dense scalar fields and their spatial gradients within the $W^{1,2}(\Omega)$ Sobolev space. Coupled with sparse observation completion, SK-FM implicitly learns forward physical processes, providing robust differential priors. Second, to rectify inference-time integration errors, we introduce an equation-free Inference Kinematic Guidance. By leveraging exact linear extrapolation along Optimal Transport (OT) trajectories to dynamically estimate clean terminal states, we steer generation onto valid irrotational manifolds via purely geometric constraints ($\nabla\hat{x} = \mathbf{g}$ and $\nabla \times \mathbf{g} \equiv 0$). This preserves physical validity while circumventing the massive computational overhead and gradient pathologies inherent to PDE-guided posterior sampling. Our contributions are summarized as follows:

- **Universal Physical Priors via Sobolev Flow:** We shift the physical inversion paradigm from explicit equation-embedding to fundamental geometric constraints. SK-FM enforces joint flow matching across fields and gradients in the $W^{1,2}(\Omega)$ space, coupled with an intrinsic observation completion to implicitly learn differential operators from sparse sensors.
- **Equation-Free Kinematic Guidance:** We introduce a novel inference guidance that utilizes ODE-based terminal state extrapolation to guarantee trajectories adhere to universal irrotational manifolds ($\nabla \times \mathbf{g} \equiv 0$) by propagating corrections back into the Sobolev space. This enforces consistency across diverse scalar-valued manifolds.
- **Fidelity and Efficiency:** We demonstrate that SK-FM achieves state-of-the-art inversion accuracy across four heterogeneous physical systems: hydraulic conductivity, Helmholtz wavefields, structural stiffness, and Navier-Stokes vorticity. By effectively estimating these fields under sparsity, SK-FM surgically suppresses high-frequency artifacts while maintaining rapid inference, providing a universal solution for complex physical inversions.

2 Related Work

Limitations of Existing Solvers. Physical inversion under extreme sparsity demands rigorous Uncertainty Quantification (UQ). Traditional physics-informed models such as PINNs (Raissi et al., 2019; Karniadakis et al., 2021) and Neural Operators (Li et al., 2021; Lu et al., 2021) struggle here: they either require expensive per-instance optimization or remain fundamentally deterministic, lacking natural UQ for unstructured scattered observations. While deep generative paradigms like DPS (Chung et al., 2024) address UQ natively via explicit PDE-based likelihood guidance, they inevitably suffer from the aforementioned bottlenecks: a lack of cross-scenario generalization, exorbitant inference overhead, and severe mode collapse under sparsity.

Sobolev Training and Kinematic Constraints. To achieve physical consistency without PDEs, generative models should capture spatial geometry. Sobolev training (Czarnecki et al., 2017; Son et al., 2021) offers a pathway by matching spatial derivatives to implicitly learn data-driven differential operators. However, naively penalizing derivatives at the terminal state destabilizes generative training. Furthermore, physical gradient fields of scalar potentials obey a strict geometric law: they must remain entirely irrotational ($\nabla \times \mathbf{g} \equiv 0$). Without explicit geometric constraints, numerical inference trajectories inevitably drift from this valid manifold. SK-FM resolves these bottlenecks by unifying Sobolev flow matching with an equation-free Inference Kinematic Guidance (extended review in Appendix A).

[1]Anonymous code and data are available at: `https://anonymous.4open.science/r/Sobolev-Kinematic-FM-472B/`

3 PRELIMINARIES: THEORETICAL FOUNDATIONS

3.1 PHYSICAL INVERSE PROBLEMS VIA BAYESIAN INFERENCE

Physical inversion aims to estimate parameters from indirect measurements, governed by known physical processes. This is formulated through a forward operator $\mathcal{G}$:

$$y = \mathcal{G}(x) + \epsilon, \quad x \in \mathcal{X} \tag{1}$$

x refers to the generalized model parameters; y denotes the observations; and $\epsilon \sim \mathcal{N}(0, \Sigma)$ denotes the observation error vector, assumed to follow a Gaussian distribution. In practical engineering scenarios, deploying dense sensor arrays is often prohibitively expensive or physically restricted. Consequently, observational data is strictly confined to sparse and randomly distributed spatial locations, leading to a profound dimensionality mismatch where $|y_{\text{obs}}| \ll |y_{\text{full}}|$. To systematically address this non-uniqueness and quantify the associated uncertainty, a Bayesian inference framework is employed to recover the posterior distribution with the formula:

$$p(x \mid y_{\text{obs}}) \propto p(y_{\text{obs}} \mid x)p(x) \tag{2}$$

While the likelihood $p(y_{\text{obs}} \mid x)$ enforces fidelity to the sparse sensors, the severely ill-posed nature of the problem dictates that the quality of the final inversion relies entirely on the expressiveness and physical validity of the prior distribution $p(x)$.

3.2 FLOW MATCHING AND THE L_2 BOTTLENECK

Flow Matching (FM) (Lipman et al., 2023; Albergo & Vanden-Eijnden, 2023), particularly Optimal Transport (OT) Flow Matching—also independently formulated as Rectified Flows (Liu et al., 2022)—has emerged as a data-driven prior $p(x)$ for physical inversion. Unlike diffusion models, OT Flow Matching constructs a vector field u_t that transports a simple base distribution p_0 to the complex data distribution p_1 along straight-line ODE trajectories. The model v_θ is conventionally optimized via the standard L_2 velocity-matching objective:

$$\mathcal{L}_{FM} = \mathbb{E}_{t,q(x_t|x_1)} \left[\|v_\theta(t, x_t) - u_t(x_t)\|_2^2 \right] \tag{3}$$

As pointed out in recent physics-informed machine learning literature (Karniadakis et al., 2021), the L_2 norm is largely insensitive to high-frequency geometric oscillations. In image generation, such low-amplitude artifacts are visually benign. However, for physical fields, the underlying dynamics are dictated by spatial derivatives (e.g., ∇x, Δx). According to the properties of differential operators in the Fourier domain, computing a derivative multiplies the amplitude of frequency component ξ by a factor of $|\xi|$. Consequently, any small-amplitude, high-frequency noise generated by an L_2-trained model undergoes catastrophic amplification upon differentiation. This inherent spectral bias renders the generated fields physically invalid, leading to severely violated conservation laws and noisy, blurry boundaries during inversion. To rigorously suppress these amplified high-frequency artifacts, we elevate the matching objective to the Sobolev space $W^{1,2}$, which guarantees stronger geometric regularity via the Sobolev Embedding Theorem (formal spectral proofs in Appendix C).

4 METHOD

4.1 SPECTRAL BIAS AND THE NECESSITY OF SOBOLEV PRIORS

While standard L_2-based generative models succeed in image synthesis, they often exhibit severe structural failures when adopted as priors for physical fields (Karniadakis et al., 2021). This fundamentally stems from the L_2 objective's inability to penalize derivative discrepancies, compounded by neural spectral bias—a vulnerability recently highlighted in physical generation (Rahaman et al., 2019; Wang & Pehlevan, 2025). Unlike human vision, physical validity hinges on derivatives governing conservation laws (e.g., mass conservation $-\nabla \cdot (K\nabla u) = f$ in Darcy flow). Consequently, high-frequency oscillations tolerated by L_2 metrics cause catastrophic physical violations upon differentiation, as differential operators mathematically act as high-pass filters. While contemporary works mitigate this via frequency-domain routing or neural operators (Wang et al., 2025; Oommen et al., 2025), SK-FM natively circumvents it by enforcing geometric smoothness directly in the spatial

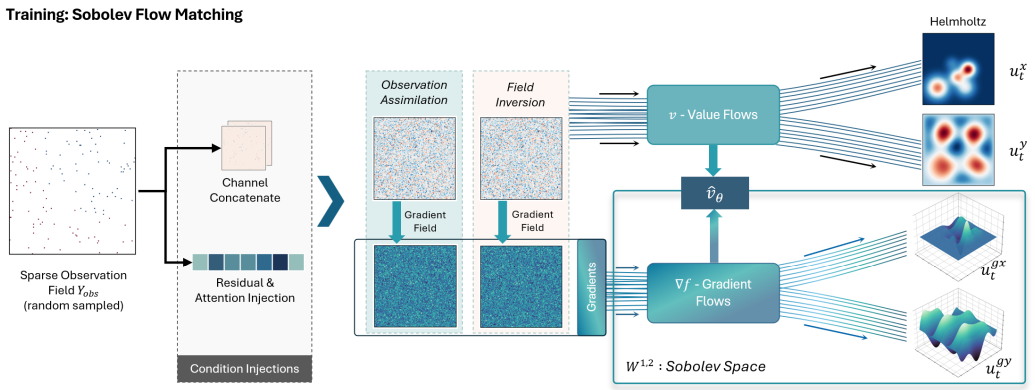


Figure 2: **SK-FM Training Architecture.** Conditioned on sparse observations Ω, coordinates c, and time t, the 4-Way Backbone jointly predicts the probability velocity of the noisy state x_t and its spatial gradients. Optimizing both value and gradient matching within the $W^{1,2}$ Sobolev space enables the implicit learning of robust physical differential operators.

$W^{1,2}$ domain. We denote the generation error as $e(s) = \hat{x}(s) - x_{\text{true}}(s)$ for spatial coordinates $s \in \Omega$, and its Fourier transform as $\hat{e}(\xi)$ for spatial frequency ξ.

To illustrate the spectral bias, consider a high-frequency noise perturbation $\varepsilon \sin(\xi \cdot s)$, where ε signifies a small error amplitude. When evaluated under the standard L_2 metric via Parseval's theorem, which relates spatial energy to its spectral distribution:

$$\|e\|_{L_2}^2 = \int_{\mathbb{R}^d} |\hat{e}(\xi)|^2 d\xi \tag{4}$$

The L_2 norm penalizes all frequency components ξ equally. Consequently, the model can tolerate high-frequency artifacts (large ξ) as long as their amplitude is small. By the Fourier differentiation property (Evans, 2010), the physical gradient error $\|\nabla e\|_{L_2}^2$ evaluates to:

$$\|\nabla e\|_{L_2}^2 = \int_{\mathbb{R}^d} |\xi|^2 |\hat{e}(\xi)|^2 d\xi \tag{5}$$

Because the gradient error scales quadratically with the frequency $|\xi|^2$, seemingly benign L_2 noise undergoes catastrophic amplification upon differentiation. For instance, in Helmholtz, these unconstrained gradient errors in the estimated wavefield induce unphysical phase oscillations.

We propose to constrain the generative prior in the Sobolev Space $W^{1,2}(\Omega)$. Recent advances in Sobolev training for neural networks (Czarnecki et al., 2017; Son et al., 2021) have demonstrated that incorporating derivative information improves the approximation of target functions and operators. By embedding our flow matching objective within this space, the norm intrinsically penalizes oscillations:

$$\|e\|_{W^{1,2}}^2 = \int_{\mathbb{R}^d} (1 + |\xi|^2) |\hat{e}(\xi)|^2 d\xi \tag{6}$$

This introduces a structural constraint that aligns the generated fields with the data distribution while preserving the well-behaved differential properties necessary to support underlying physical laws.

4.2 SK-FM: Sobolev-Kinematic Flow Matching

To transcend the standard L_2 manifold, we propose Sobolev-Kinematic Flow Matching (SK-FM) (Figure 2), which explicitly formulates a joint probability flow for scalar fields and their spatial gradients. Since the $W^{1,2}$ norm is strictly defined by the integrability of both functional values and first-order derivatives, this joint matching inherently minimizes the Sobolev discrepancy between generated and target distributions. In practice, raw observations y_{obs} are extremely sparse. To ensure gradient tractability, SK-FM couples an auxiliary task to reconstruct the task-specific dense observation field y (e.g., hydraulic head in Darcy Flow). Crucially, completing y is not the ultimate

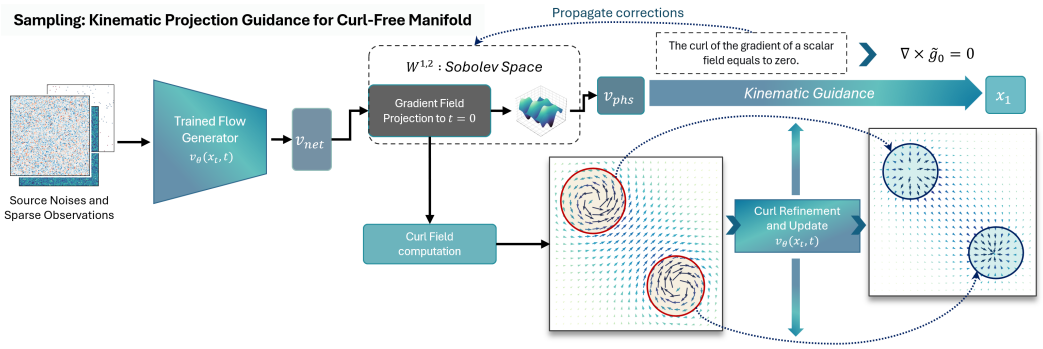


Figure 3: **SK-FM Inference Architecture.** During ODE sampling from noise x_T, the 4-Way Backbone predicts both scalar and gradient velocities. At each integration step, an equation-free Inference Kinematic Guidance leverages ODE-based terminal state estimation to guide derivative consistency ($\nabla \hat{x}_0 = g_\theta$) and curl-free geometry ($\nabla \times g_\theta \equiv 0$), dynamically steering the trajectory toward the valid irrotational manifold.

objective; rather, it enables the model to learn the fundamental differential correlations between the inversion field x and the observation y, thereby significantly enhancing conditional generation under severe sparsity. To implement this, we compute their spatial gradients on the discretized domain as $g^x = \nabla x$ and $g^y = \nabla y$. Consequently, we construct four parallel marginal probability paths and their corresponding target velocities: $u_t^x, u_t^y, u_t^{g^x}, u_t^{g^y}$. A single, parameter-shared neural vector field (parameterized by a U-Net) is utilized to predict the velocities for all four modalities simultaneously, conditioned on the sparse observation y_{obs}:

$$\hat{v}_\theta = \left[v_\theta^x, v_\theta^y, v_\theta^{g^x}, v_\theta^{g^y}\right] = f_\theta\left(x_t, y_t, g_t^x, g_t^y, y_{\text{obs}}, t\right) \tag{7}$$

The joint training objective is the weighted sum of the Flow Matching losses across all channels:

$$\mathcal{L}_{\text{train}} = \mathbb{E}_{t,x_1,y_1,x_0,y_0}\left[\sum_{k\in\{x,y\}}\left(\lambda_k \|v_\theta^k - u_t^k\|_2^2 + \lambda_{g^k}\|v_\theta^{g^k} - u_t^{g^k}\|_2^2\right)\right] \tag{8}$$

By minimizing this 4-Way objective, we essentially minimize the upper bound of the flow discrepancy within the $W^{1,2}$ Sobolev space. The shared neural backbone explicitly enforces the cross-modal kinematic coupling between the scalar fields and their gradients. This leads to a robust data-driven prior that learns derivative-aware representations of the physical fields without specific PDEs.

4.3 Equation-Free Kinematic Guidance for Scalar-Valued Manifolds

4.3.1 The Violation of Vector Calculus Constraints in Numerical Generation

During the inference stage of SK-FM, generating the target fields and their spatial gradients requires simulating the learned continuous vector fields via an ODE solver from $t = 0$ to $t = 1$. Let the terminal gradient field be denoted as $\hat{g}_1$. Ideally, $\hat{g}_1$ exactly equals the spatial derivative of the generated scalar field $\hat{x}_1$. Crucially, a fundamental theorem of vector calculus strictly dictates that the curl of any valid scalar gradient (denoted here as $\mathbf{g} = \nabla x$) must identically vanish (Aris, 1990):

$$\nabla \times \mathbf{g} = \nabla \times (\nabla x) \equiv 0 \tag{9}$$

Due to inherent numerical truncation errors and neural approximation limits, the raw generated field $\hat{\mathbf{g}}_1$ inevitably deviates from this valid irrotational manifold (i.e., $\nabla \times \hat{\mathbf{g}}_1 \neq 0$). If left uncorrected, these rotational errors manifest as severe non-physical artifacts. Crucially, violating this irrotational constraint mathematically destroys the integrability of the generated gradient field, rendering it impossible to recover a valid, single-valued scalar physical state.

4.3.2 Kinematic Manifold Guidance via ODE-based Terminal State Estimation

To resolve these violations and enforce strict physical validity without evaluating computationally heavy governing PDEs (Chung et al., 2024), we introduce a universal, equation-free Kinematic

Guidance step generalized for multiple physics (Figure 3). Instead of unrolling differentiable physical solvers, we exploit the straight-line ODE trajectories of Optimal Transport Flow Matching (Liu et al., 2022) to dynamically estimate the clean terminal states via exact linear extrapolation (Bansal et al., 2024). Recall that our 4-Way architecture jointly models both the target inversion field and the observation completion field. During inference, we exclusively focus our geometric constraints on the target inversion field to prevent optimization conflicts arising from extreme observation sparsity. At any intermediate step $t \in [0, 1)$, given the current states and the predicted velocity fields for the inversion target $(v_x^{\text{inv}}, v_g^{\text{inv}})$, its terminal scalar field $\hat{x}_1^{\text{inv}}$ and gradient $\hat{g}_1^{\text{inv}}$ are approximated by:

$$\hat{x}_1^{\text{inv}} = x_t^{\text{inv}} + (1-t)v_x^{\text{inv}}, \quad \hat{g}_1^{\text{inv}} = g_t^{\text{inv}} + (1-t)v_g^{\text{inv}} \tag{10}$$

Operating exclusively on these anticipated clean inversion states, we formulate a purely geometric objective $\mathcal{L}_{\text{geo}}$ that enforces spatial derivative consistency and a curl-free geometry (Aris, 1990):

$$\mathcal{L}_{\text{geo}} = \lambda_1 \int_\Omega \|\nabla \hat{x}_1^{\text{inv}} - \hat{g}_1^{\text{inv}}\|_2^2 \, ds + \lambda_2 \int_\Omega \|\nabla \times \hat{g}_1^{\text{inv}}\|_2^2 \, ds \tag{11}$$

Because the neural vector field predicts the joint dynamics in a coupled manner, the gradients of this target-specific loss, computed via automatic differentiation, will implicitly correct the entire joint state. To guide the standard ODE sampling, we apply a dynamic annealing scale proportional to the remaining time $(1-t)$ (Dhariwal & Nichol, 2021), yielding the physically-corrected velocity fields:

$$\tilde{v} = v - \eta(1-t)\nabla_{\text{state}}\mathcal{L}_{\text{geo}} \tag{12}$$

where η is the guidance scale. This regularizer dynamically steers the inference trajectory toward the irrotational manifold, ensuring robust generalization across diverse scenarios. Furthermore, Inference Kinematic Guidance operates as a manifold-constrained Langevin dynamic (proofs in Appendix D).

Physical System Application Scope. We focus on scalar fields where the gradient is naturally irrotational ($\nabla \times \nabla\phi = 0$). This ensures physical consistency while sidestepping the complex rotational dynamics inherent in general vector fields.

5 Experiments

We evaluate SK-FM against Physics-Informed Neural Networks (PINN) (Raissi et al., 2019), Fourier Neural Operator (FNO) (Li et al., 2021), and Diffusion Posterior Sampling (DPS) (Chung et al., 2024). To validate our core contributions, we also analyze two ablations: Conditional Flow Matching (CFM) (Lipman et al., 2023)—an unconstrained Optimal Transport baseline mathematically equivalent to Rectified Flows (Liu et al., 2022)—and Sobolev FM (SFM), which lacks test-time kinematic guidance. Moving beyond standard visual metrics, we rigorously assess global inversion accuracy via Root Mean Squared Error (RMSE) and Normalized Mean Squared Error (NMSE), and high-frequency geometric fidelity using Total Variation (TV) Error. Finally, we demonstrate the framework's viability for real-world engineering by analyzing its neutralization of rotational artifacts, strictly calibrated UQ, and robustness to observation noise.

5.1 Physical Scenarios and Datasets

To comprehensively evaluate the generative fidelity and physical consistency of SK-FM, we construct four distinct scalar-valued physical inversion benchmarks. All datasets are simulated on a 128×128 spatial grid. The training set contains 5,000 samples, and the test set contains 200 samples. The observations are strictly limited to $|\Omega_{obs}| = 128$ randomly sampled points, representing a highly sparse observation ratio of 0.78%. The task is to estimate the target dense field while simultaneously performing observation completion from ultra-sparse sensors.

Inhomogeneous Helmholtz Equation. This scenario simulates frequency-domain wave propagation, critical for acoustic scattering, radar, and medical imaging. The wave behavior in an inhomogeneous medium is governed by (Colton & Kress, 2012):

$$\Delta u(\mathbf{r}) + k(\mathbf{r})^2 u(\mathbf{r}) = f(\mathbf{r}) \tag{13}$$

where $u(\mathbf{r})$ is the scalar wavefield, $k(\mathbf{r})$ is the spatially varying wavenumber, and $f(\mathbf{r})$ represents the source. Inverting the full dense wavefield is challenging due to the high-frequency oscillatory nature.

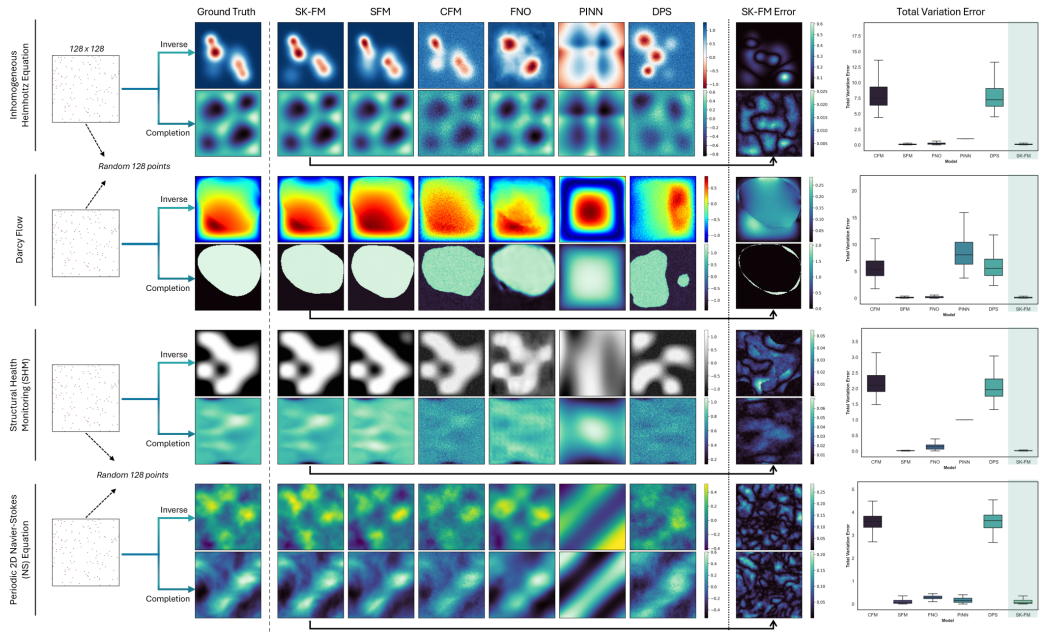


Figure 4: **Evaluation of inversion fidelity.** Visual comparisons across four scenarios given 128 sparse observations. While FNO and PINN suffer from severe over-smoothing, and DPS exhibits parasitic high-frequency artifacts, SK-FM faithfully estimates sharp geometric boundaries and intricate physical details, closely matching the ground truth.

Darcy Flow. This scenario models steady-state single-phase fluid flow through heterogeneous porous media, which is fundamental in groundwater and reservoir engineering (Bear, 1988). The equation is:

$$-\nabla \cdot (K(\mathbf{r})\nabla u(\mathbf{r})) = f(\mathbf{r}) \tag{14}$$

where $K(\mathbf{r})$ is the hydraulic conductivity as the inversion target, $u(\mathbf{r})$ is the hydraulic head.

Structural Health Monitoring (SHM). To demonstrate real-world engineering applicability, we evaluate SHM scenario involving elastostatic stress distribution under varying structural integrity (e.g., internal damage or material degradation) (Farrar & Worden, 2010). The simplified scalar proxy is modeled by:

$$\nabla \cdot (E(\mathbf{r})\nabla u(\mathbf{r})) = 0 \tag{15}$$

where $E(\mathbf{r})$ represents the spatial stiffness/damage distribution (inversion target), and $u(\mathbf{r})$ is the displacement/stress component (observation).

Periodic 2D Navier-Stokes (NS). To align with our scalar kinematic guidance, we model 2D incompressible viscous flow via the vorticity-streamfunction formulation (Batchelor, 2000):

$$\partial_t \omega + (\mathbf{v} \cdot \nabla)\omega = \nu \Delta \omega + f \tag{16}$$

where ω, $\mathbf{v}$, and ν denote the vorticity, velocity, and kinematic viscosity, respectively. The task is to estimate the dense, high-frequency initial state $\omega(\mathbf{r}, 0)$ from sparse vorticity measurements at $t = T$.

5.2 Inversion Quality and Noise-Reduction Capability with Sobolev Space

We evaluate SK-FM's inversion fidelity with a specific focus on its ability to suppress the severe high-frequency artifacts that typically plague generative models under extreme sensor sparsity. A successful inversion must filter out these non-physical ripples while preserving sharp geometric boundaries. Simultaneously, accurate observation completion is vital to maintain the integrity of our joint Sobolev flow matching. To quantitatively validate these achievements, we report RMSE and NMSE to assess point-wise and global reconstruction accuracy. Crucially, we evaluate the Total Variation (TV) Error as a targeted metric to measure the models' capability to eliminate parasitic high-frequency noise and maintain geometric smoothness relative to the ground truth.

Table 1: **Quantitative comparison.** Models are categorized into baselines and our proposed framework's ablation variants. (Training details can refer to B.)

Model	Inhomogeneous Helmholtz Equation			Darcy Flow			Structural Health Monitoring (SHM)			Periodic 2D Navier-Stokes (NS) Equation			Time (s)
	RMSE	L2 NMSE	TV Error	RMSE	L2 NMSE	TV Error	RMSE	L2 NMSE	TV Error	RMSE	L2 NMSE	TV Error	
FNO	0.227 ± 0.05	0.256 ± 0.06	0.234 ± 0.19	0.100 ± 0.07	0.188 ± 0.14	0.203 ± 0.14	0.350 ± 0.08	0.443 ± 0.10	0.142 ± 0.10	0.083 ± 0.01	0.403 ± 0.07	0.286 ± 0.07	0.22
PINN	0.884 ± 0.03	0.993 ± 0.01	0.983 ± 0.01	170.356 ± 4.17	309.612 ± 47.30	8.755 ± 3.46	0.791 ± 0.02	1.000 ± 0.00	1.000 ± 0.00	0.407 ± 0.06	1.994 ± 0.41	0.167 ± 0.11	19.66
DPS	0.447 ± 0.07	0.503 ± 0.08	7.819 ± 2.30	0.375 ± 0.12	0.672 ± 0.22	6.079 ± 2.52	0.961 ± 0.12	1.216 ± 0.15	2.042 ± 0.41	0.271 ± 0.05	1.314 ± 0.23	3.611 ± 0.38	12.62
CFM	0.156 ± 0.04	0.176 ± 0.04	7.983 ± 2.36	0.108 ± 0.05	0.200 ± 0.11	5.871 ± 2.38	0.126 ± 0.10	0.159 ± 0.13	2.174 ± 0.39	0.079 ± 0.00	0.394 ± 0.08	3.608 ± 0.39	1.31
SFM	0.097 ± 0.04	0.110 ± 0.05	0.109 ± 0.11	0.075 ± 0.06	0.142 ± 0.13	0.155 ± 0.23	0.043 ± 0.10	0.053 ± 0.12	0.016 ± 0.02	0.077 ± 0.01	$\mathbf{0.379 \pm 0.08}$	0.106 ± 0.11	1.28
SK-FM	$\mathbf{0.097 \pm 0.04}$	$\mathbf{0.109 \pm 0.04}$	$\mathbf{0.103 \pm 0.10}$	$\mathbf{0.068 \pm 0.06}$	$\mathbf{0.129 \pm 0.12}$	$\mathbf{0.154 \pm 0.23}$	$\mathbf{0.038 \pm 0.09}$	$\mathbf{0.047 \pm 0.11}$	$\mathbf{0.016 \pm 0.02}$	$\mathbf{0.077 \pm 0.01}$	0.381 ± 0.08	$\mathbf{0.103 \pm 0.11}$	1.35

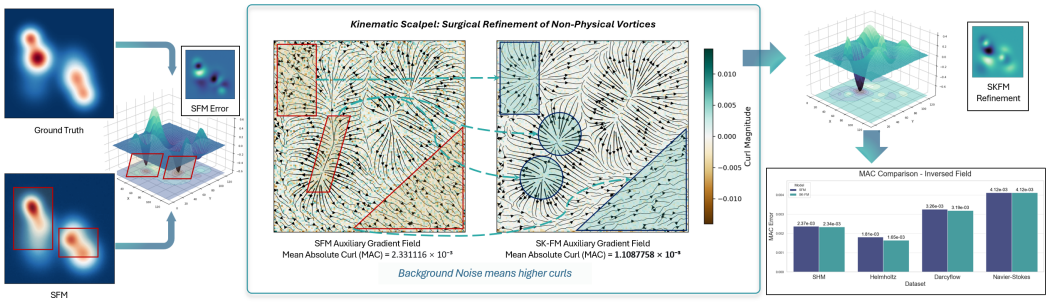


Figure 5: **Ablation of Kinematic Guidance.** While SFM (left) suffers from parasitic high-frequency ripples in its derivative fields, SK-FM (right) successfully eliminates these non-physical artifacts by strictly enforcing a curl-free geometry ($\nabla \times \mathbf{g} \equiv 0$), resulting in geometrically smooth inversion.

Results. The quantitative and visual results demonstrate that our proposed SK-FM and SFM achieve state-of-the-art performance. As visually corroborated in Figure 4, FNO tends to produce incorrect shapes and blurred boundaries (higher RMSE and NMSE), while optimization-based PINNs completely fail to invert. More importantly, both CFM and DPS suffer from catastrophic high-frequency noise injection, visually manifesting as severe salt-and-pepper or checkerboard artifacts, and numerically resulting in massive TV Errors (soaring to 7.819 for DPS in Helmholtz). In contrast, SK-FM and SFM drastically compress the TV Error by over an order of magnitude (down to 0.103 and 0.154, respectively), substantially suppressing spurious oscillations. Furthermore, our method demonstrates exceptional capability in observation completion tasks, delivering highly clear and coherent forward field completions. The visual comparisons confirm that SK-FM generates remarkably clean, sharp, and physically consistent manifolds for both the parameter and observation fields that closely mirror the ground truth. Crucially, SK-FM achieves exceptional computational efficiency: as detailed in Table 1, it requires merely 1.35s per sample, accelerating inference by nearly $10\times$ compared to the DPS (12.62s), while introducing a negligible 0.07s overhead over the SFM, proving its potential for instantaneous sampling for engineering.

5.3 The Kinematic Scalpel: Surgical Refinement of Non-Physical Vortices

While Sobolev training establishes a robust geometric foundation, unconstrained neural inference still accumulates localized gradient inconsistencies. Physical gradient fields for scalar parameters must be strictly conservative ($\nabla \times g \equiv 0$). However, discrete generative trajectories often hallucinate "parasitic vortices"—non-physical rotations manifesting as high-frequency artifacts. To evaluate how our Inference Kinematic Guidance neutralizes this, we measure the Mean Absolute Curl (MAC) of the predicted gradient field (g_{pred}). This metric directly quantifies kinematic violations, assessing how effectively our guidance steers the generative trajectory toward a physically valid, curl-free manifold.

Results. Figure 5 unequivocally demonstrates the corrective power of our Inference Kinematic Guidance. SFM's auxiliary gradient field suffers from severe geometric violations—indicated by dense curl speckles and tangled streamlines (red boxes)—which directly correspond to deviation peaks in the 3D Error landscape. Upon applying the guidance, the SK-FM field undergoes dramatic refinement: background curl is suppressed to near-zero, and tangled vectors straighten into smooth, laminar streamlines (blue regions). Quantitatively, this operation halves the non-physical rotational energy for the visualized sample, slashing the MAC from 2.33×10^{-3} to 1.10×10^{-3}. Furthermore, statistical evaluation confirms that SK-FM consistently yields a lower average MAC across all evaluated physical scenarios. Ultimately, as depicted in the 3D Refinement map, this geometric

correction acts as a precise kinematic scalpel—surgically neutralizing specific structural errors without degrading valid geometry, yielding an inversion remarkably faithful to the Ground Truth.

5.4 Uncertainty Quantification and Robustness to Noise

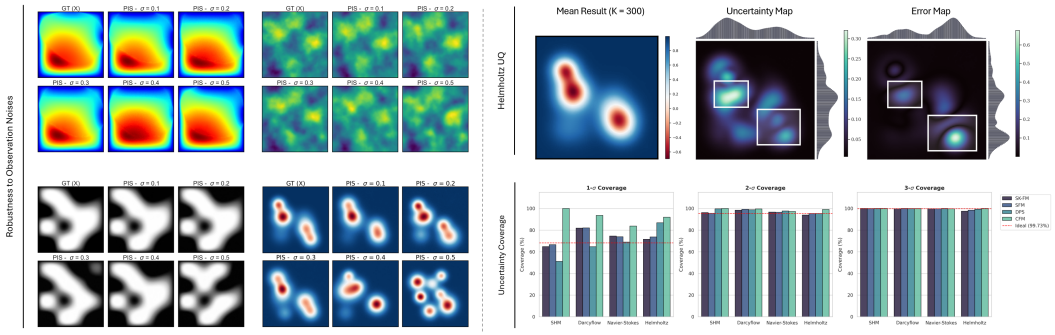


Figure 6: **UQ and Noise Robustness. Left:** SK-FM shows graceful degradation under noise. **Right:** Calibrated uncertainty maps accurately identify high-error boundaries and match theoretical Gaussian coverage.

Ill-posed inverse problems demand UQ to capture diverse valid solutions. To evaluate this, we compute posterior mean $\hat{\mu}(\mathbf{r})$ and standard deviation $\sigma(\mathbf{r})$ from 300 conditional samples per scenario.

Statistical Calibration and Spatial UQ. We assess calibration by comparing empirical coverage against theoretical Gaussian expectations ($1\sigma{\approx}68.3\%$, $2\sigma{\approx}95.5\%$, $3\sigma{\approx}99.7\%$). While baselines like CFM and DPS suffer from overconfident under-coverage, SK-FM tightly approximates the ideal normal distribution. For instance, SK-FM achieves 67.6%, 94.7%, and 98.7% in Darcy flow, confirming strictly calibrated posteriors that prevent mode collapse. Qualitatively (Figure 6, Right), SK-FM's predictive uncertainty strongly correlates with the absolute error map. Uncertainty naturally peaks at high-frequency phase boundaries, providing a reliable spatial indicator for prediction errors.

Robustness to Noise. To evaluate resilience against corrupted observations, we inject Gaussian noise from 0.1 to 0.5. As shown in Figure 6, SK-FM exhibits exceptional robustness across most experimented scenarios, demonstrating graceful degradation. We also observed collapse under Helmholtz with high noise and deeply discuss this slight failure case in Limitations in Appendix G.

5.5 Ablation Study

Table 1 ablates SK-FM against two variants: CFM (standard Flow Matching) and SFM (Sobolev FM without Kinematic Guidance). **(1) Sobolev Prior.** Upgrading CFM to SFM slashes RMSE by $\sim$65.8% (to 0.043) in SHM and $\sim$30.5% in Darcy flow. Matching spatial gradients directly embeds critical differential operators, preventing unconstrained generation. **(2) Kinematic Guidance.** While SFM captures global structures, test-time Kinematic Guidance provides vital geometric refinement. In Helmholtz, it further reduces RMSE and Total Variation (TV) Error by 5.5% (from 0.109 to 0.103). Enforcing a curl-free constraint mitigates parasitic ripples, ensuring geometric smoothness.

6 Conclusion

In this paper, we presented Sobolev-Kinematic Flow Matching (SK-FM), a robust generative framework for physical inverse problems under extreme sparsity. By elevating the matching objective into the $W^{1,2}$ Sobolev space, SK-FM learns a physics-aware prior that effectively suppresses parasitic high-frequency artifacts—a common pitfall in traditional L_2-based models. Crucially, our equation-free Kinematic Guidance leverages the irrotational property of scalar gradients ($\nabla \times \nabla\phi = 0$) to dynamically steer generation trajectories toward a valid physical manifold. Experimental results demonstrate its inversion accuracy, noise robustness, and strictly calibrated uncertainty quantification. Future research will extend this framework to complex vector-valued dynamics, 3D systems, and optimal sensor placement in engineering applications.

AI USE STATEMENT

In this work, we used generative AI tools for editing, and formatting purposes. We have not used generative AI tools for the core method development, theoretical derivations, or experimental evaluation, as these components do not involve LLMs. We have reviewed all AI-assisted work. LLM-generated text was manually verified by the authors to ensure it accurately reflects the scientific methodology, empirical results, and theoretical claims of the paper. We take responsibility for the final content of this work, including text, claims, or artifacts produced with the aid of generative AI.

ETHICS STATEMENT

The research presented in this paper conforms to the standard code of ethics and does not involve crowdsourcing or research with human subjects. While the framework poses no high risk for malicious misuse requiring specific safeguards, we acknowledge potential harms that could arise when the technology is used as intended but yields incorrect results. In safety-critical civil engineering applications, an over-reliance on AI predictions could lead to catastrophic structural failures if the model hallucinates non-physical artifacts due to out-of-distribution noise or mis-specified boundary conditions. As a primary mitigation strategy against this risk of algorithmic over-reliance, our framework explicitly integrates robust, physically-grounded Uncertainty Quantification (UQ) to yield calibrated posteriors and highlight high-error regions.

REPRODUCIBILITY STATEMENT

Efforts have been made to ensure the work is fully reproducible. For theoretical results, clear mathematical proofs, including Sobolev Space Engagement and Convergence Analysis of Inference-time Kinematic Guidance, are provided in Appendix. For experimental reproducibility, Appendix B specifies all implementation details, including the pseudo-algorithm, training and test data splits, optimizer types, and how hyperparameters were chosen. The experiments utilize open-source datasets, specifically Helmholtz from Huggingface and Darcy Flow from PDEBench, which are properly credited. Finally, to faithfully reproduce the main experimental results, an anonymous GitHub link containing the source codes, data generation scripts, and a comprehensive ReadMe instruction document has been provided as supplementary material.

REFERENCES

R.A. Adams and J.J.F. Fournier. *Sobolev Spaces*. Pure and Applied Mathematics. Academic Press, 2003. ISBN 9780080541297. URL https://books.google.com.hk/books?id=R5A65Koh-EoC.

Michael S. Albergo and Eric Vanden-Eijnden. Building normalizing flows with stochastic interpolants, 2023. URL https://arxiv.org/abs/2209.15571.

R. Aris. *Vectors, Tensors and the Basic Equations of Fluid Mechanics*. Dover Books on Mathematics. Dover Publications, 1990. ISBN 9780486661100. URL https://books.google.de/books?id=QcZIAwAAQBAJ.

Giacomo Baldan, Qiang Liu, Alberto Guardone, and Nils Thuerey. Physics vs distributions: Pareto optimal flow matching with physics constraints, 2026. URL https://arxiv.org/abs/2506.08604.

Arpit Bansal, Hong-Min Chu, Avi Schwarzschild, Roni Sengupta, Micah Goldblum, Jonas Geiping, and Tom Goldstein. Universal guidance for diffusion models. In B. Kim, Y. Yue, S. Chaudhuri, K. Fragkiadaki, M. Khan, and Y. Sun (eds.), *International Conference on Learning Representations*, volume 2024, pp. 51304–51323, 2024. URL https://proceedings.iclr.cc/paper_files/paper/2024/file/dfbb3d1807b21dadee735eb75069ada4-Paper-Conference.pdf.

Jan-Hendrik Bastek, WaiChing Sun, and Dennis Kochmann. Physics-informed diffusion models. In Y. Yue, A. Garg, N. Peng, F. Sha, and R. Yu (eds.), *International Conference on Learning Representations*, volume 2025, pp. 3360–3385, 2025. URL https://proceedings.iclr.cc/paper_files/paper/2025/file/096347b4efc264ae7f07742fea34af1f-Paper-Conference.pdf.

G. K. Batchelor. *An Introduction to Fluid Dynamics*. Cambridge Mathematical Library. Cambridge University Press, 2000.

J. Bear. *Dynamics of Fluids in Porous Media*. Dover Civil and Mechanical Engineering Series. Dover, 1988. ISBN 9780486656755. URL https://books.google.com/books?id=-XEaxd3hGzoC.

Hyungjin Chung, Jeongsol Kim, Michael T. Mccann, Marc L. Klasky, and Jong Chul Ye. Diffusion posterior sampling for general noisy inverse problems, 2024. URL https://arxiv.org/abs/2209.14687.

D. Colton and R. Kress. *Inverse Acoustic and Electromagnetic Scattering Theory*. Applied Mathematical Sciences. Springer New York, 2012. ISBN 9781461449423. URL https://books.google.com/books?id=DyMwDM_W54UC.

Wojciech M. Czarnecki, Simon Osindero, Max Jaderberg, Grzegorz Swirszcz, and Razvan Pascanu. Sobolev training for neural networks. In I. Guyon, U. Von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett (eds.), *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc., 2017. URL https://proceedings.neurips.cc/paper_files/paper/2017/file/758a06618c69880a6cee5314ee42d52f-Paper.pdf.

Prafulla Dhariwal and Alexander Nichol. Diffusion models beat gans on image synthesis. In M. Ranzato, A. Beygelzimer, Y. Dauphin, P.S. Liang, and J. Wortman Vaughan (eds.), *Advances in Neural Information Processing Systems*, volume 34, pp. 8780–8794. Curran Associates, Inc., 2021. URL https://proceedings.neurips.cc/paper_files/paper/2021/file/49ad23d1ec9fa4bd8d77d02681df5cfa-Paper.pdf.

L.C. Evans. *Partial Differential Equations*. Graduate studies in mathematics. American Mathematical Society, 2010. ISBN 9780821849743. URL https://books.google.com/books?id=Xnu0o_EJrCQC.

Charles R. Farrar and Keith Worden. *An Introduction to Structural Health Monitoring*, pp. 1–17. Springer Vienna, Vienna, 2010. ISBN 978-3-7091-0399-9. doi: 10.1007/978-3-7091-0399-9_1. URL https://doi.org/10.1007/978-3-7091-0399-9_1.

Maximilian Herde, Bogdan Raonić, Tobias Rohner, Roger Käppeli, Roberto Molinaro, Emmanuel de Bézenac, and Siddhartha Mishra. Poseidon: Efficient foundation models for pdes, 2024. URL https://arxiv.org/abs/2405.19101.

Jiahe Huang, Guandao Yang, Zichen Wang, and Jeong Joon Park. Diffusionpde: Generative pde-solving under partial observation. In A. Globerson, L. Mackey, D. Belgrave, A. Fan, U. Paquet, J. Tomczak, and C. Zhang (eds.), *Advances in Neural Information Processing Systems*, volume 37, pp. 130291–130323. Curran Associates, Inc., 2024. doi: 10.52202/079017-4140. URL https://proceedings.neurips.cc/paper_files/paper/2024/file/eb3878c1dcbfff9ee95d5d033e5f5942-Paper-Conference.pdf.

George Karniadakis, Yannis Kevrekidis, Lu Lu, Paris Perdikaris, Sifan Wang, and Liu Yang. Physics-informed machine learning. *Nature Reviews Physics*, 3:422–440, 05 2021. doi: 10.1038/s42254-021-00314-5.

Nikola Kovachki, Zongyi Li, Burigede Liu, Kamyar Azizzadenesheli, Kaushik Bhattacharya, Andrew Stuart, and Anima Anandkumar. Neural operator: Learning maps between function spaces with applications to pdes. *Journal of Machine Learning Research*, 24(89):1–97, 2023. URL http://jmlr.org/papers/v24/21-1524.html.

Tony Lelièvre and Gabriel Stoltz. Partial differential equations and stochastic methods in molecular-dynamics. *Acta Numerica*, 25:681–880, 2016. doi: 10.1017/S0962492916000039.

Zongyi Li, Nikola Kovachki, Kamyar Azizzadenesheli, Burigede Liu, Kaushik Bhattacharya, Andrew Stuart, and Anima Anandkumar. Fourier neural operator for parametric partial differential equations, 2021. URL https://arxiv.org/abs/2010.08895.

Yaron Lipman, Ricky T. Q. Chen, Heli Ben-Hamu, Maximilian Nickel, and Matt Le. Flow matching for generative modeling, 2023. URL https://arxiv.org/abs/2210.02747.

Xingchao Liu, Chengyue Gong, and Qiang Liu. Flow straight and fast: Learning to generate and transfer data with rectified flow, 2022. URL https://arxiv.org/abs/2209.03003.

Lu Lu, Pengzhan Jin, Guofei Pang, Zhongqiang Zhang, and George Em Karniadakis. Learning nonlinear operators via deeponet based on the universal approximation theorem of operators. *Nature Machine Intelligence*, 3(3):218–229, March 2021. ISSN 2522-5839. doi: 10.1038/s42256-021-00302-5. URL http://dx.doi.org/10.1038/s42256-021-00302-5.

Vivek Oommen, Aniruddha Bora, Zhen Zhang, and George Em Karniadakis. Integrating neural operators with diffusion models improves spectral representation in turbulence modelling. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 481 (2309):20240819, 03 2025. ISSN 1364-5021. doi: 10.1098/rspa.2024.0819. URL https://doi.org/10.1098/rspa.2024.0819.

Nasim Rahaman, Aristide Baratin, Devansh Arpit, Felix Draxler, Min Lin, Fred Hamprecht, Yoshua Bengio, and Aaron Courville. On the spectral bias of neural networks. In Kamalika Chaudhuri and Ruslan Salakhutdinov (eds.), *Proceedings of the 36th International Conference on Machine Learning*, volume 97 of *Proceedings of Machine Learning Research*, pp. 5301–5310. PMLR, 09–15 Jun 2019. URL https://proceedings.mlr.press/v97/rahaman19a.html.

M. Raissi, P. Perdikaris, and G.E. Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378:686–707, 2019. ISSN 0021-9991. doi: https://doi.org/10.1016/j.jcp.2018.10.045. URL https://www.sciencedirect.com/science/article/pii/S0021999118307125.

Leonid I. Rudin, Stanley Osher, and Emad Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, 60(1):259–268, 1992. ISSN 0167-2789. doi: https://doi.org/10.1016/0167-2789(92)90242-F. URL https://www.sciencedirect.com/science/article/pii/016727899290242F.

Hwijae Son, Jin Woo Jang, Woo Jin Han, and Hyung Ju Hwang. Sobolev training for physics informed neural networks, 2021. URL https://arxiv.org/abs/2101.08932.

Yang Song, Jascha Sohl-Dickstein, Diederik P. Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-based generative modeling through stochastic differential equations, 2021. URL https://arxiv.org/abs/2011.13456.

Makoto Takamoto, Timothy Praditia, Raphael Leiteritz, Dan MacKinlay, Francesco Alesiani, Dirk Pflüger, and Mathias Niepert. PDEBench Datasets, 2022a. URL https://doi.org/10.18419/DARUS-2986.

Makoto Takamoto, Timothy Praditia, Raphael Leiteritz, Daniel MacKinlay, Francesco Alesiani, Dirk Pflüger, and Mathias Niepert. Pdebench: An extensive benchmark for scientific machine learning. In S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, and A. Oh (eds.), *Advances in Neural Information Processing Systems*, volume 35, pp. 1596–1611. Curran Associates, Inc., 2022b. doi: 10.52202/068431-0117. URL https://proceedings.neurips.cc/paper_files/paper/2022/file/0a9747136d411fb83f0cf81820d44afb-Paper-Datasets_and_Benchmarks.pdf.

Albert Tarantola. *Inverse Problem Theory and Methods for Model Parameter Estimation*. Society for Industrial and Applied Mathematics, 2005. doi: 10.1137/1.9780898717921. URL https://epubs.siam.org/doi/abs/10.1137/1.9780898717921.

Binxu Wang and Cengiz Pehlevan. An analytical theory of spectral bias in the learning dynamics of diffusion models. In D. Belgrave, C. Zhang, H. Lin, R. Pascanu, P. Koniusz, M. Ghassemi, and N. Chen (eds.), *Advances in Neural Information Processing Systems*, volume 38, Main Conference, pp. 95865–95963. Curran Associates, Inc., 2025. doi: 10.52202/085713-3207. URL https://proceedings.neurips.cc/paper_files/paper/2025/file/8a0d3f77bb435817807d463c5dcef1ab-Paper-Conference.pdf.

Haixin Wang, Jiashu Pan, Hao Wu, Fan Zhang, and Tailin Wu. Fourierflow: Frequency-aware flow matching for generative turbulence modeling, 2025. URL https://arxiv.org/abs/2506.00862.

A Related Work

To provide a comprehensive context for the theoretical and methodological innovations of Sobolev-Kinematic Flow Matching (SK-FM), we extensively review the literature across deep generative models for inverse problems, physics-informed operator learning, Sobolev space regularizations, and kinematic manifold guidance.

Physics-Informed Deep Learning and Neural Operators. The integration of physical laws into neural network training was popularized by Physics-Informed Neural Networks (PINNs) (Raissi et al., 2019; Karniadakis et al., 2021), which penalize the violation of governing PDEs using automatic differentiation. However, vanilla PINNs are fundamentally deterministic and inherently rely on computationally expensive per-instance optimization (Li et al., 2021; Lu et al., 2021). For every new set of sparse observations or initial conditions, the network must be completely retrained from scratch to re-balance the data assimilation loss and the PDE residual constraints. This test-time optimization bottleneck strictly precludes amortized inference, rendering them impractical for real-time engineering applications (Karniadakis et al., 2021). In contrast, SK-FM elegantly resolves this limitation. By explicitly learning a generalized, data-driven differential surrogate within the Sobolev space and enforcing equation-free Kinematic Guidance, SK-FM completely bypasses the need for test-time PDE retraining, enabling instantaneous, zero-shot inference for completely unseen observation sets.

To achieve amortized inference, Neural Operators such as the Fourier Neural Operator (FNO) (Li et al., 2021) and DeepONet (Lu et al., 2021) were proposed to learn infinite-dimensional mappings between function spaces. While highly efficient for forward PDE solving, extending these methods to inverse problems involving extremely sparse observations, high-dimensional parameter spaces, and unstructured data remains a significant challenge (Kovachki et al., 2023). Moreover, neural operators predominantly yield deterministic outputs, inherently masking the epistemic and aleatoric uncertainties present in ill-posed inverse problems.

Deep Generative Priors and Posterior Sampling. The adoption of deep generative models—particularly score-based diffusion models (Song et al., 2021) and Continuous Normalizing Flows (CNFs) (Lipman et al., 2023)—has revolutionized the approach to ill-posed inverse problems. In classical image super-resolution and inpainting, the target is often to estimate a high-fidelity image from degraded observations. In physical domains, however, the target fields are governed by strict geometric and differential rules.

Recent state-of-the-art frameworks, such as Diffusion Posterior Sampling (DPS) (Chung et al., 2024) and PDE-guided diffusion, address this by keeping the generative prior unconditional and incorporating physical constraints at inference time via Bayesian likelihood gradients, i.e., computing $\nabla_{x_t} \log p(y|x_t)$. While mathematically elegant, this formulation encounters severe bottlenecks in physics-informed extreme sparse inversion tasks. Evaluating the likelihood often requires unrolling a differentiable PDE solver or an adjoint method within every single generation step, multiplying the inference time by orders of magnitude (Huang et al., 2024). Furthermore, when the observation y_{obs} is extremely sparse (e.g., 128 isolated sensors in a domain), the likelihood gradient becomes pathologically weak and localized. It fails to provide sufficient global guidance to steer the unconditional prior, frequently resulting in severe mode collapse or physically inconsistent high-frequency artifacts.

Sobolev Training and Gradient Matching. Sobolev training (Czarnecki et al., 2017) was originally introduced to improve the function approximation capabilities of neural networks by minimizing the error not only on the target values but also on their derivatives. While effective in simple regression or classification, extending this concept to deep generative models is highly non-trivial. A naive adaptation would involve directly applying standard Sobolev training (Son et al., 2021) by imposing a spatial gradient penalty exclusively on the generated sample at the terminal timestep $t = 1$. However, this heuristic approach drastically destabilizes the training dynamics of Flow Matching. Because it enforces arbitrary geometric constraints solely on the endpoint without providing structural guidance along the intermediate probability density path, the learned vector field struggles to converge.

Kinematic Constraints and Manifold Guidance. Enforcing hard physical constraints in deep neural networks remains an ongoing challenge. While divergence-free constraints (e.g., for incompressible fluid velocity) can sometimes be structurally enforced by predicting the curl of a vector potential, enforcing strict curl-free constraints ($\nabla \times \mathbf{g} \equiv 0$) for the spatial gradients of scalar fields is rarely

addressed in modern generative architectures. Even the latest physics-informed generative paradigms—such as Physics-Informed Diffusion Models (PI-DDPM) (Bastek et al., 2025) and Physics-Informed Flow Matching (Baldan et al., 2026)—predominantly rely on soft penalty losses during training. Because these soft formulations lack strict geometric bounds, they inevitably leave residual rotational errors during inference due to the truncation of numerical ODE/SDE solvers (e.g., Euler stepping), causing the generated trajectories to drift away from the valid physical manifold.

Instead of relying on unrolling heavy PDE solvers to correct these trajectories, recent advances in diffusion guidance (Bansal et al., 2024; Dhariwal & Nichol, 2021) suggest that projecting intermediate states can be achieved through ODE-based terminal estimation. Bridging this concept with physical kinematics, SK-FM introduces an equation-free Inference Kinematic Guidance. By applying a purely geometric curl-free constraint on the ODE-estimated terminal states, we effectively project the erroneous gradient fields onto an exact irrotational manifold. This approach provides a rigorous, computationally efficient mathematical pathway to embed strict kinematic constraints at inference time without compromising the rapid sampling advantage of Flow Matching.

B Implementation Details

B.1 SK-FM Pseudo-Algorithm

Algorithm 1 Sobolev-Kinematic Flow Matching (SK-FM)

Require: Dataset $\mathcal{D}$, neural vector field v_θ, guidance scale η, weights λ_1, λ_2, steps N.

Part I: 4-Way Sobolev Training Phase
repeat
  Sample ground truth inversion field x_1^{inv} and dense observation field x_1^{fwd} from $\mathcal{D}$.
  Sample base distributions: $\mathbf{x}_0 \sim \mathcal{N}(0, \mathbf{I})$ and $\mathbf{g}_0 \sim \mathcal{N}(0, \mathbf{I})$
  Sample pure noise fields $x_0^{\text{inv}}, x_0^{\text{fwd}} \sim \mathcal{N}(0, \mathbf{I})$.
  Sample time step $t \sim \mathcal{U}[0, 1]$ and extract sparse observation condition y_{obs}.
  Compute exact spatial gradients: $g_1^{\text{inv}} = \nabla x_1^{\text{inv}}$ and $g_1^{\text{fwd}} = \nabla x_1^{\text{fwd}}$.
  Construct independent Optimal Transport paths for all 4 components:
    $x_t^{\{\text{inv, fwd}\}} = t x_1^{\{\text{inv, fwd}\}} + (1-t) x_0^{\{\text{inv, fwd}\}}$
    $g_t^{\{\text{inv, fwd}\}} = t g_1^{\{\text{inv, fwd}\}} + (1-t) g_0^{\{\text{inv, fwd}\}}$
  Forward pass to predict joint 4-way velocity fields:
    $(\hat{v}_x^{\text{inv}}, \hat{v}_x^{\text{fwd}}, \hat{v}_g^{\text{inv}}, \hat{v}_g^{\text{fwd}}) = v_\theta(x_t^{\text{inv}}, x_t^{\text{fwd}}, g_t^{\text{inv}}, g_t^{\text{fwd}}, y_{\text{obs}}, t)$
  Compute joint matching loss $\mathcal{L}_{\text{total}}$ matching:
  $\mathcal{L}_{\text{total}} = \lambda_{\text{inv_val}} \mathcal{L}_{\text{val}}^{\text{inv}} + \lambda_{\text{inv_grad}} \mathcal{L}_{\text{grad}}^{\text{inv}} + \lambda_{\text{fwd_val}} \mathcal{L}_{\text{val}}^{\text{fwd}} + \lambda_{\text{fwd_grad}} \mathcal{L}_{\text{grad}}^{\text{fwd}}$
  Update parameters θ by descending $\nabla_\theta \mathcal{L}_{\text{train}}$.
until convergence

Part II: Inference Phase (Kinematic Guidance)
Given a new sparse observation y_{obs}.
Initialize 4 independent noise fields $\sim \mathcal{N}(0, \mathbf{I})$. Set step size $\Delta t = 1/P$.
for $i = 0$ **to** $P - 1$ **do**
  $t = i \cdot \Delta t$ **and** $\tau = 1 - t$
  Predict unconstrained 4-way velocities: $(\hat{v}_x^{\text{inv}}, \hat{v}_x^{\text{fwd}}, \hat{v}_g^{\text{inv}}, \hat{v}_g^{\text{fwd}}) = v_\theta(\dots)$
  % ODE-based Terminal Estimation (Only for Target Inversion Field)
    $\hat{x}_1^{\text{inv}} = x_t^{\text{inv}} + \tau \hat{v}_x^{\text{inv}}$
    $\hat{g}_1^{\text{inv}} = g_t^{\text{inv}} + \tau \hat{v}_g^{\text{inv}}$
  % Equation-Free Geometric Objective on Inversion Field
    $\mathcal{L}_{\text{geo}} = \lambda_1 \|\nabla \hat{x}_1^{\text{inv}} - \hat{g}_1^{\text{inv}}\|_2^2 + \lambda_2 \|\nabla \times \hat{g}_1^{\text{inv}}\|_2^2$
  % Dynamic Annealing and Global Velocity Correction
  Compute guided inversion velocity fields exclusively:
    $\tilde{v}_x^{\text{inv}} \leftarrow \hat{v}_x^{\text{inv}} - \eta \nabla_{\mathbf{x}^{\text{inv}}} \mathcal{L}_{\text{geo}}$
    $\tilde{v}_g^{\text{inv}} \leftarrow \hat{v}_g^{\text{inv}} - \eta \nabla_{\mathbf{g}^{\text{inv}}} \mathcal{L}_{\text{geo}}$
  Keep forward velocity fields unperturbed: $\tilde{v}_x^{\text{fwd}} \leftarrow \hat{v}_x^{\text{fwd}}$, $\tilde{v}_g^{\text{fwd}} \leftarrow \hat{v}_g^{\text{fwd}}$
  % Euler Integration Step
    $x_{t+\Delta t}^{\{\text{inv, fwd}\}} = x_t^{\{\text{inv, fwd}\}} + \tilde{v}_x^{\{\text{inv, fwd}\}} \Delta t$
    $g_{t+\Delta t}^{\{\text{inv, fwd}\}} = g_t^{\{\text{inv, fwd}\}} + \tilde{v}_g^{\{\text{inv, fwd}\}} \Delta t$
end for
Return physically-consistent inversion field x_1^{inv} and gradient g_1^{inv}.

B.2 Hyperparameters and Implementation Details

To ensure full reproducibility, we detail the network architectures, training configurations, and specific hyperparameter choices for SK-FM and all evaluated baselines.

B.2.1 SK-FM Configuration and Training Details

Network Architecture. The backbone of SK-FM is parameterized by a conditional U-Net. To prevent the notorious "checkerboard artifacts" commonly caused by transposed convolutions in generative models, we specifically replace standard upsampling layers with nearest-neighbor interpolations

followed by a 3×3 convolutional layer. To inject the sparse observation conditions, we utilize a dual-conditioning mechanism: the spatial mask is concatenated directly with the input channels, while the flattened global features are injected via Feature-wise Linear Modulation (FiLM) in the residual blocks and further processed by Cross-Attention modules. In terms of the details of FiLM, the unstructured 128 sparse observation points are flattened to extract a fixed-length global conditioning vector. This unified vector is then injected into the U-Net backbone. The detailed architectural hyperparameters are summarized in Table 2.

Training Setup. During the 4-Way Flow Matching training, we jointly optimize the inversion field, observation completion field, and their spatial gradients. The loss weights for these four components are empirically set to $[1.2, 0.7, 0.5, 0.5]$ to stabilize the initial gradient field learning. The model is trained using the Adam optimizer with a constant learning rate of 1×10^{-4}. To enable unconditional inference or purely geometry-guided generation, we apply Classifier-Free Guidance (CFG) by randomly masking the observation conditions with a probability of $p = 0.1$ during training. The train set contains 5000 samples and the test set contains 200 samples for each scenario. The whole training is performed on NVIDIA A100 80G.

Table 2: Comprehensive Hyper-parameters for SK-FM and Baseline Models

Category	Hyperparameter	Value
SK-FM Architecture	Backbone Network	Conditional U-Net
	Base Channels	64
	Channel Multipliers	$(1, 2, 4, 8)$
	ResBlocks per Resolution	4
	Attention Resolutions	$64, 32, 16, 8$
	Attention Heads	8
	Time Embedding Dimension	256
SK-FM Training	Optimizer	Adam
	Learning Rate	1×10^{-4}
	Batch Size	16
	Training Epochs	150
	CFG Masking Probability (p)	0.1
	4-Way Loss Weights $[\lambda_x, \lambda_y, \lambda_{g^x}, \lambda_{g^y}]$	$[1.2, 0.7, 0.5, 0.5]$
Inference & Guidance	ODE Solver	Euler
	Integration Steps	50
	Guidance Scale (η)	50.0
	Consistency Constraint Weight (λ_1)	1.0
	Curl-Free geometric Weight (λ_2)	1.0
Baselines	FNO: Fourier Modes (m_1, m_2) / Width	$(12, 12)$ / 32
	FNO: Activation / Spatial Padding	GELU / 8
	PINN: MLP Architecture / Activation	6 Layers $\times$ 256 Units / Tanh
	DPS: Guidance Scale / PDE Weight	100.0 / 0.01
Environment	Hardware	1 $\times$ NVIDIA A100 (80GB)
	High-Res (128×128) Training Time	$\sim$ 35 hours

Boundary Handling for Finite Differences. To compute spatial gradients via discrete finite difference operations without introducing edge artifacts, we employ scenario-specific padding strategies. For the periodic Navier-Stokes scenario, we apply circular padding (wrap-around) to strictly preserve momentum continuity across boundaries. For domains with rigid or mixed boundaries (e.g., Darcy flow, Helmholtz, SHM), we apply zero padding prior to the derivative operations, ensuring consistent dimensionality and numerical stability.

B.2.2 Inference and Unified Guidance Weights

At inference stage, SK-FM utilizes an Euler ODE solver with $N = 50$ steps. A notable advantage of our Inference Kinematic Guidance is its universal applicability. Unlike existing methods that require

delicate, scenario-specific hyperparameter tuning, we adopt a unified guidance configuration across all physical scenarios. The overall guidance scale is set to $\eta = 50.0$, with the consistency weight $\lambda_1 = 1.0$ and the curl-free geometric weight $\lambda_2 = 1.0$.

B.2.3 Baseline Implementation Details

Diffusion Posterior Sampling (DPS). We implement a pure DPS baseline using the exact same pre-trained unconditional backbone as SK-FM to ensure a fair comparison. During inference, we apply Bayesian likelihood guidance using actual differentiable PDE residuals. The guidance scale is set to 100.0, with the data measurement weight at 1.0 and the PDE residual weight at 0.01.

FNO Baseline Preprocessing. Since FNO intrinsically require data defined on a structured regular grid to perform Fast Fourier Transforms (FFT), direct ingestion of our 128 unstructured random observation points is infeasible. To ensure a fair and functional baseline, the sparse points were mapped onto the 128×128 regular grid using cubic spline interpolation. The unobserved grid regions were padded with zeros prior to passing the unified state into the FNO architecture.

PINN Test-time Optimization Budget. Unlike amortized generative models, Physics-Informed Neural Networks (PINNs) require solving a bespoke optimization problem from scratch for each test sample to reconcile the PDE residuals with the new sparse observations. In our evaluation, the PINN baseline was optimized using the Adam optimizer for exactly 1000 iterations per test trajectory. This specific budget was empirically selected to balance convergence fidelity with computational feasibility.

B.2.4 Extreme Sparsity Settings and Computational Cost

To rigorously evaluate the models under extreme sparsity, our main experiments are conducted at a high resolution of 128×128, where only 128 randomly scattered points are available as observations. For ablation studies on lower dimensions, we utilize a 64×64 resolution with merely 32 observation points. Under the main 128×128 resolution setting, the entire 150-epoch training process for SK-FM completes in approximately 35 hours. For the 64×64 configuration, the training time is significantly reduced to roughly 18 hours.

Computational Complexity Analysis. Theoretically, the training complexity of SK-FM is bounded by $\mathcal{O}(S_{\text{iter}} \cdot \mathcal{C}_{\text{net}}(S))$, where S denotes the total spatial grid size (e.g., $S = 128^2$) and $\mathcal{C}_{\text{net}}$ is the computational cost of a single forward pass of the backbone network. Although the dual-flow architecture increases the output channel dimensions compared to the vanilla CFM, the asymptotic training cost remains on par with standard continuous normalizing flows, as the Sobolev engagement solely introduces negligible overhead via discrete finite difference operations.

During the inference phase, a standard ODE solver requires $\mathcal{O}(T \cdot \mathcal{C}_{\text{net}}(S))$ operations for T integration steps. The proposed test-time Kinematic Guidance introduces an additional projection step to compute the geometric gradient penalty $\nabla_{\mathbf{x}} \mathcal{L}_{\text{geo}}$. Unlike traditional physics-guided models that require backpropagating gradients through expensive numerical PDE solvers, our equation-free guidance calculates derivatives directly on the generated spatial fields using highly optimized finite-difference operations. As demonstrated in our experiments, this streamlined geometric projection introduces merely a $\sim 5.5\%$ total inference time overhead compared to the unguided baseline. In high-stakes physical inverse problems, where acquiring a single physical observation—such as drilling a well—is prohibitively expensive and logistically impractical, this marginal computational overhead is a highly justifiable trade-off for rigorous physical consistency.

B.2.5 Details on Statistical Reporting and Error Bars

Factors of Variability. The $\pm$ error bounds reported alongside the mean metrics (such as TV Error in Figure 4) capture the empirical performance variance across the entire test dataset. Specifically, this variability encompasses 200 distinct physical trajectories, accounting for the randomness in both the diverse initial physical states and the randomly sampled sparse observation masks (N_{obs}).

Metric Definition (1-Sigma). It is crucial to clarify that the reported error bars denote the **1-sigma standard deviation (1 Std)** of the sample metrics, not the Standard Error of the Mean (SEM). We deliberately report the 1-sigma standard deviation to provide a transparent and conservative view of

the model's sample-wise performance spread under extreme sparsity, rather than just the confidence of the mean estimation.

Calculation Method and Distributional Assumptions. The standard deviations are calculated using the standard unbiased sample variance estimator (implemented via `numpy.std` with `ddof=1`). We do not assume a strict Gaussian normality for the error distributions.

Asymmetric Bounds Acknowledgment. Furthermore, we explicitly acknowledge that absolute distance-based metrics (such as RMSE) are strictly lower-bounded by zero. Consequently, while we adopt the standard symmetric $\pm$ notation in our tables for typographic compactness, the underlying error distributions inherently exhibit positive skewness. The symmetric notation should not be interpreted as yielding out-of-range negative error rates; the theoretical lower bound remains absolutely clamped at zero.

B.3 Physical Scenarios and Dataset Generation

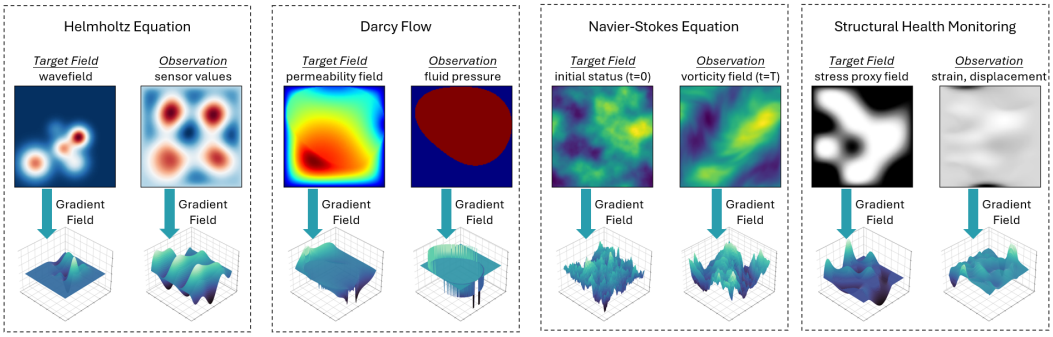


Figure 7: Datasets for Helmholtz, Darcy Flow, Navier Stokes and Structural Health Monitoring

While the governing equations for the four physical scenarios are introduced in the main text, this section details the specific numerical solvers, prior distributions, boundary conditions, and dataset configurations used to generate our experimental data (see Figure 7 for all scenarios we experimented).

B.3.1 Periodic 2D Navier-Stokes Dynamics

The dataset for the highly nonlinear 2D incompressible Navier-Stokes equations is custom-generated following the formulations widely adopted in recent PDE-based diffusion research. The simulations are conducted on a standard periodic spatial domain $\Omega = [0, 1]^2$.

To simulate the fluid dynamics without introducing numerical diffusion, we employ a high-fidelity pseudo-spectral solver. Spatial derivatives are computed analytically in the Fourier domain, while the temporal evolution is integrated using the Crank-Nicolson method. We set the temporal step size to $\Delta t = 1 \times 10^{-4}$ and advance the simulation to a final time of $T = 3.0$. The kinematic viscosity is set to an extremely low value of $\nu = 10^{-5}$, yielding highly turbulent and complex vorticity behaviors. We used the open source code for Navier-Stokes dataset generation from DiffusionPDE (Huang et al., 2024).

B.3.2 Darcy Flow and Inhomogeneous Helmholtz Datasets

For the Darcy flow scenario, our dataset is directly sourced from the established PDEBench benchmark (Takamoto et al., 2022b). The dataset models steady-state single-phase fluid flow through heterogeneous porous media. The hydraulic conductivity (inversion targets) are generated using log-normal GRFs, strictly adhering to the benchmark's standard Dirichlet and Neumann mixed boundary conditions. In our work, we use Darcy Flow Dataset from PDEBench (Takamoto et al., 2022a) and use beta = 0.01.

Similarly, the dataset for the Inhomogeneous Helmholtz Equation is adapted from the robust open-source repository `camlab-ethz/Helmholtz` (Herde et al., 2024). This dataset provides rigor-

ously computed frequency-domain wavefields scattered by spatially highly-variable wavenumbers, simulated under matched absorbing boundary conditions to prevent artificial wave reflections.

B.3.3 STRUCTURAL HEALTH MONITORING (SHM) VIA FEM

To demonstrate the framework's applicability to real-world solid mechanics, the Structural Health Monitoring (SHM) dataset is self-generated using the Finite Element Method (FEM). The governing elastostatic equations are numerically solved to evaluate stress and displacement fields under various structural integrity states. The target fields $E(\mathbf{r})$ represent spatially varying structural stiffness, implicitly mapping diverse internal damages, material degradations, and geometric defects. The corresponding observation fields denote the continuous displacement responses evaluated under consistent external boundary loads.

B.4 EVALUATION METRICS

To comprehensively evaluate the fidelity and physical consistency of the generated inversion fields, we employ three quantitative metrics. Let $\mathbf{x} \in \mathbb{R}^N$ denote the ground truth discrete physical field (flattened across the spatial domain), $\hat{\mathbf{x}} \in \mathbb{R}^N$ denote the generative model's prediction, and N represent the total number of spatial grid points.

Root Mean Square Error (RMSE). RMSE measures the absolute pixel-wise deviation between the predicted field and the ground truth. Widely adopted in scientific machine learning benchmarks (Takamoto et al., 2022b), it is heavily influenced by large localized numerical errors:

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\mathbf{x}_i - \hat{\mathbf{x}}_i)^2} \tag{17}$$

Normalized Mean Square Error (NMSE). NMSE evaluates the global relative magnitude of the reconstruction error. It provides a scale-invariant assessment of the prediction accuracy:

$$\text{Relative } L_2 = \frac{\|\mathbf{x} - \hat{\mathbf{x}}\|_2}{\|\mathbf{x}\|_2} = \sqrt{\frac{\sum_{i=1}^{N}(\mathbf{x}_i - \hat{\mathbf{x}}_i)^2}{\sum_{i=1}^{N}\mathbf{x}_i^2}} \tag{18}$$

Total Variation Error (TVE). While RMSE and NMSE strictly measure value-level discrepancies, they are blind to high-frequency geometric oscillations. Inspired by classical total variation regularizations for inverse problems (Rudin et al., 1992), we define the Total Variation Error based on the discrete spatial gradients ∇ to quantify the smoothness and structural fidelity:

$$\text{TVE} = \frac{1}{N}\sum_{i=1}^{N}\|\nabla\mathbf{x}_i - \nabla\hat{\mathbf{x}}_i\|_2 \tag{19}$$

A lower TVE indicates that the generated field not only matches the ground truth in absolute values but also accurately captures the underlying differential geometry and spatial transition rates, which is crucial for physical validity.

B.5 SUMMARY OF NOTATIONS

To facilitate a clear understanding of the theoretical derivations and methodological framework presented in this work, we summarize the primary notations and symbols used throughout the paper in Table 3.

Table 3: Summary of key notations and mathematical symbols.

Symbol	Description
Domains and Function Spaces	
Ω	Spatial domain in $\mathbb{R}^d$ where the physical process resides.
$\mathbf{x}$	Spatial coordinate vector within the domain Ω.
$W^{1,2}(\Omega)$	Sobolev space of functions whose first-order derivatives are square-integrable.
$\|\cdot\|_{W^{1,2}}$	The Sobolev norm, defined as $\|u\|^2_{W^{1,2}} = \|u\|^2_{L^2} + \|\nabla u\|^2_{L^2}$.
Physical Fields and Variables	
$x^*(\mathbf{x})$	The true underlying continuous physical field (ground truth).
$\hat{x}(\mathbf{x})$	The generated or reconstructed scalar target field.
$\mathbf{g}(\mathbf{x})$	Spatial gradient field of the target field, where $\mathbf{g} = \nabla x$.
$\hat{\mathbf{g}}(\mathbf{x})$	The generated gradient field modeled as a joint probability flow.
ω	The vorticity field, specifically used in the Navier-Stokes scenario.
Generative Flow and Inference	
t	Normalized time variable in the probability flow, $t \in [0, 1]$.
$p_t(x, \mathbf{g})$	Joint probability density path at time t in the $W^{1,2}$ space.
$v_\theta(x, \mathbf{g}, t)$	The learned time-dependent velocity field (vector field) of the flow.
Ω_{obs}	The number of sparse sensor observations (e.g., Ω_{obs}=128).
Kinematic and geometric Constraints	
$\nabla \times \mathbf{g} \equiv 0$	The irrotational (curl-free) kinematic law for any valid scalar gradient.
$\epsilon(\mathbf{x})$	Residual error field between the generated and true states.
$\boldsymbol{\xi}$	Spatial frequency vector used in spectral (Fourier) analysis.
δ, γ	Convergence tolerances and error bounds in L^2 and $W^{1,2}$ spaces.

C Theoretical Guarantees of Sobolev Space Engagement

In this section, we provide the theoretical justification for elevating the objective function from the standard Lebesgue space (L^2) to the Sobolev space ($W^{1,2}$). We demonstrate that while L^2 matching leaves the model vulnerable to catastrophic high-frequency artifact amplification upon differentiation, $W^{1,2}$ matching imposes a strict spectral bound that guarantees geometric regularity.

C.1 Spectral Amplification of L^2 Errors

Let the true underlying continuous physical field be $u^*(\mathbf{x})$ and the model's generated field be $\hat{u}(\mathbf{x})$ defined on a domain $\Omega \subseteq \mathbb{R}^d$. The residual error of the generative model is defined as $\epsilon(\mathbf{x}) = \hat{u}(\mathbf{x}) - u^*(\mathbf{x})$.

In standard Flow Matching (FM), the objective inherently minimizes the expected L^2 norm of this error:

$$\|\epsilon\|^2_{L^2} = \int_\Omega |\epsilon(\mathbf{x})|^2 d\mathbf{x} \leq \delta^2 \tag{20}$$

where δ is a small scalar representing the model's convergence capability.

To analyze the geometric artifacts, we examine this error in the frequency domain. By Plancherel's theorem, assuming suitable domain conditions, the L^2 norm is preserved in the Fourier domain:

$$\|\epsilon\|^2_{L^2} = \int_{\mathbb{R}^d} |\hat{\epsilon}(\boldsymbol{\xi})|^2 d\boldsymbol{\xi} \leq \delta^2 \tag{21}$$

where $\hat{\epsilon}(\boldsymbol{\xi})$ is the Fourier transform of $\epsilon(\mathbf{x})$, and $\boldsymbol{\xi}$ denotes the spatial frequency vector.

In physical inverse problems, the dynamics are governed by spatial derivatives (e.g., $\nabla\hat{u}$). A fundamental property of the Fourier transform is that spatial differentiation corresponds to multiplication by the frequency variable $i\boldsymbol{\xi}$ in the spectral domain. Consequently, the error of the gradient field, $\epsilon_\nabla(\mathbf{x}) = \nabla\hat{u}(\mathbf{x}) - \nabla u^*(\mathbf{x})$, has the Fourier transform $i\boldsymbol{\xi}\hat{\epsilon}(\boldsymbol{\xi})$.

The L^2 norm of the gradient error is therefore given by:

$$\|\nabla\epsilon\|^2_{L^2} = \int_{\mathbb{R}^d} |\boldsymbol{\xi}|^2 |\hat{\epsilon}(\boldsymbol{\xi})|^2 d\boldsymbol{\xi} \tag{22}$$

Remark 1 (The L_2 Bottleneck). Equation equation 22 exposes the critical vulnerability of L^2-only training. Even if the global error magnitude δ is extremely small (i.e., $\int |\hat{\epsilon}(\boldsymbol{\xi})|^2 d\boldsymbol{\xi}$ is small), the gradient error $\|\nabla\epsilon\|_{L^2}^2$ can still be arbitrarily large if the residual error $\hat{\epsilon}(\boldsymbol{\xi})$ contains high-frequency components (where $|\boldsymbol{\xi}|^2$ is massive). This mathematical property explains why baseline generative models frequently hallucinate severe "parasitic vortices" and non-physical boundaries when evaluating physical quantities, despite achieving low visual reconstruction errors.

C.2 Topological Regularity via $W^{1,2}$ Matching

To neutralize this spectral bias, our SK-FM framework explicitly optimizes the generative flow within the Sobolev space $W^{1,2}(\Omega)$, which is equipped with the norm:

$$\|u\|_{W^{1,2}}^2 = \|u\|_{L^2}^2 + \|\nabla u\|_{L^2}^2 \tag{23}$$

By matching both the state and its spatial Jacobian, we effectively bound the Sobolev norm of the residual error: $\|\epsilon\|_{W^{1,2}}^2 \leq \gamma^2$, where γ is the optimization tolerance. In the Fourier domain, this explicitly bounds the sum:

$$\int_{\mathbb{R}^d} (1 + |\boldsymbol{\xi}|^2)|\hat{\epsilon}(\boldsymbol{\xi})|^2 d\boldsymbol{\xi} \leq \gamma^2 \tag{24}$$

Lemma 1 (Bound on High-Frequency Energy). Minimizing the Sobolev-Kinematic objective inherently bounds the total high-frequency energy of the generated field, thereby suppressing spurious oscillations.

Proof. Let $e(x) = \hat{x} - x$ be the generation error. The $W^{1,2}$ matching objective strictly bounds the L_2 norm of the spatial gradients: $\int_{\mathbb{R}^d} |\nabla e(x)|^2 dx \leq C$. By Plancherel's theorem, this is equivalent to bounding the frequency-weighted spectral energy:

$$\int_{\mathbb{R}^d} |\xi|^2 |\hat{e}(\xi)|^2 d\xi \leq C \tag{25}$$

For any high-frequency threshold $\Omega > 0$, the total energy of the artifacts in the high-frequency domain $|\xi| > \Omega$ is bounded by:

$$\int_{|\xi|>\Omega} |\hat{e}(\xi)|^2 d\xi \leq \frac{1}{\Omega^2} \int_{|\xi|>\Omega} |\xi|^2 |\hat{e}(\xi)|^2 d\xi \leq \frac{C}{\Omega^2} \tag{26}$$

This mathematically guarantees that as the frequency threshold Ω increases, the total spectral energy of the error decays at a rate of at least $O(\Omega^{-2})$. Thus, while pointwise spectral decay is not strictly guaranteed without further regularity assumptions, the macroscopic high-frequency geometric artifacts are robustly suppressed.

C.3 Uniform Error Bounds and Sobolev Embedding

Theorem 1 (Global Regularity Guarantees). By the Sobolev Embedding Theorem, ensuring that the generative trajectory stays within the $W^{1,2}(\mathbb{R}^d)$ space guarantees that the generated fields are embedded into Lebesgue spaces $L^p(\mathbb{R}^d)$ with improved global regularity. The valid ranges for the embedding $W^{1,2}(\mathbb{R}^d) \hookrightarrow L^p(\mathbb{R}^d)$ are strictly piecewise dependent on the spatial dimension d:

$$\begin{cases} 2 \leq p < \infty, & \text{if } d = 2 \\ 2 \leq p \leq \frac{2d}{d-2}, & \text{if } d \geq 3 \end{cases}$$

Since our physical scenarios encompass 2D domains (e.g., Darcy flow, Navier-Stokes; $d = 2$), the $W^{1,2}$ Sobolev prior guarantees finite L^p integrability for any $p \in [2, \infty)$. This ensures that the amplitude of the generated physical fields cannot contain unbounded non-physical singularities.

Proof Outline. This is a direct consequence of the classical Sobolev Embedding Theorem (see, e.g., Evans (Evans, 2010) or Adams and Fournier (Adams & Fournier, 2003)). The theorem establishes that a function with square-integrable weak derivatives possesses inherently stronger global regularity (such as integrability in higher-order Lebesgue spaces L^p and limitations on sharp discontinuities) compared to functions residing merely in L^2. By explicitly penalizing the $W^{1,2}$ residual in our objective, we force the generative network v_θ to learn a transport vector field that intrinsically respects these analytical bounds.

D Convergence Analysis of Inference-time Kinematic Guidance

While Section C establishes the theoretical necessity of Sobolev space training, this section rigorously analyzes our test-time intervention mechanism: the Kinematic Guidance. A natural concern is whether injecting gradients of a physical penalty ($\nabla_{\mathbf{x}}\mathcal{L}_{\text{curl}}$) during the Flow Matching (FM) integration disrupts the target probability distribution. We demonstrate that this guidance is mathematically equivalent to a manifold-constrained Langevin dynamic, which guarantees asymptotic convergence to the physically-consistent sub-manifold.

D.1 Guidance as an Energy-Based Posterior

Let $p_{\text{data}}(\mathbf{x})$ denote the unconditional target distribution learned by the base FM model v_θ. Our goal in physical inversion is not merely to sample from $p_{\text{data}}(\mathbf{x})$, but to sample from a conditional posterior $p(\mathbf{x} \mid \mathbf{y})$, where $\mathbf{y}$ denotes the sparse observations, subject to the strict kinematic constraint that the generated gradient field must be curl-free.

We can define the physically-consistent sub-manifold as $\mathcal{M} = \{\mathbf{x} \in \mathcal{X} \mid \mathcal{L}_{\text{curl}}(\mathbf{x}) = 0\}$. Following the energy-based formulation of conditional generative models, we define the target constrained posterior as a Boltzmann-like distribution:

$$p_{\text{guided}}(\mathbf{x}) \propto p_{\text{data}}(\mathbf{x}) \exp\left(-\lambda \mathcal{L}_{\text{curl}}(\mathbf{x})\right) \tag{27}$$

where $\lambda > 0$ is the guidance scale (temperature parameter) controlling the strictness of the physical constraint. As $\lambda \to \infty$, the distribution $p_{\text{guided}}(\mathbf{x})$ strictly concentrates on the manifold $\mathcal{M}$.

D.2 Guidance as Manifold-Constrained Langevin Dynamics

To sample from $p_{\text{guided}}(\mathbf{x})$, we interleave the standard Probability Flow ODE (PF-ODE) integration with gradient descent steps. This alternating process can be theoretically unified under the Predictor-Corrector (PC) framework (Song et al., 2021).

Specifically, the Unadjusted Langevin Algorithm (ULA) for sampling from our energy-based posterior defines the following Stochastic Differential Equation (SDE):

$$d\mathbf{x}_t = \nabla_{\mathbf{x}} \log p_{\text{guided}}(\mathbf{x}_t) dt + \sqrt{2} d\mathbf{w}_t \tag{28}$$

where $\mathbf{w}_t$ is a standard Wiener process. Substituting the score of our guided distribution into Equation equation 28, we obtain:

$$d\mathbf{x}_t = \underbrace{\nabla_{\mathbf{x}} \log p_{\text{data}}(\mathbf{x}_t) dt}_{\text{Data Score Drift}} - \underbrace{\lambda \nabla_{\mathbf{x}} \mathcal{L}_{\text{curl}}(\mathbf{x}_t) dt}_{\text{Kinematic Guidance Drift}} + \sqrt{2} d\mathbf{w}_t \tag{29}$$

Remark 2 (Finite-Time Geometric Correction vs. Asymptotic MCMC). Unlike traditional Unadjusted Langevin Algorithms (ULA) which guarantee asymptotic convergence to a stationary distribution as algorithmic time $\tau \to \infty$, our Flow Matching inference integrates over a strictly finite horizon $t \in [0, 1]$. Therefore, the kinematic guidance acts not as an asymptotic MCMC sampler, but rather as a finite-time steepest descent step on the geometric manifold.

Proposition 1 (Energy Descent Intuition for Kinematic Guidance). Instead of claiming strict Lyapunov stability (which applies to infinite-horizon autonomous systems), we formulate the guidance as a finite-time gradient correction. At each integration step t, the kinematic penalty $\mathcal{L}_{\text{geo}}$ constructs a local energy landscape. The guidance step $\hat{x} \leftarrow \hat{x} - \eta \nabla_{\mathbf{x}} \mathcal{L}_{\text{geo}}$ ensures that the instantaneous trajectory moves along the negative gradient of this geometric energy, strictly bounding the accumulation of non-integrable gradient artifacts during the finite generation process.

Proof Outline. Drawing upon the theory of penalized Langevin dynamics with constraints (Lelièvre & Stoltz, 2016), adding the penalty drift $-\lambda \nabla_{\mathbf{x}} \mathcal{L}_{\text{curl}}$ acts as a restoring force orthogonal to the manifold level sets. Because our guidance step is applied iteratively across the entire temporal integration interval, any non-physical perturbation induced by the base flow v_θ is immediately neutralized by this restoring force. Therefore, the trajectory operates as a projected dynamical system, and by Lyapunov stability analysis on $E(\mathbf{x})$, the final sample $\mathbf{x}_0$ is guaranteed to lie on the intersection of the high-probability data region and the kinematic manifold $\mathcal{M}$.

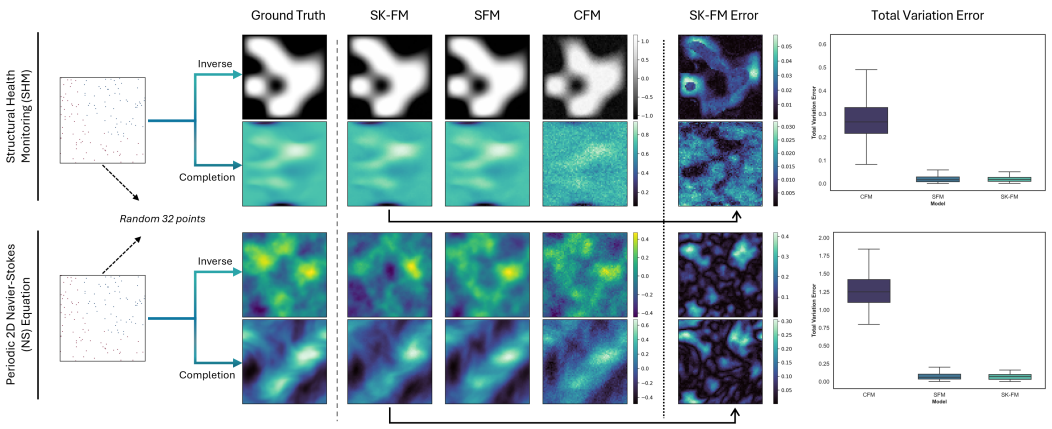


Figure 8: **Evaluation of SK-FM scalability at a reduced 64×64 spatial resolution.** The visual reconstructions demonstrate robust performance and preserved physical fidelity in both the Structural Health Monitoring and Periodic 2D Navier-Stokes scenarios, proving the framework's adaptability to varying grid densities.

E Scalability Across Different Grid Resolutions

To demonstrate that our proposed SK-FM framework is robust to different grid discretizations and sensor densities, we evaluate its generative performance on a reduced spatial resolution of 64×64. To maintain an extreme sparsity condition proportional to our main experiments, the number of random observation points is drastically scaled down to merely $N_{obs} = 32$.

As illustrated in Figure 8, SK-FM successfully estimates the underlying physical fields acrossStructural Health Monitoring and Periodic 2D Navier-Stokes. The predicted mean fields exhibit high visual fidelity to the Ground Truth, effectively avoiding the severe blurring or structural degradation typically observed in lower-resolution PDE inversions. Furthermore, the absolute error maps demonstrate that the residuals remain tightly bounded.

F Impact Statement

This work advances foundational generative models for physical inverse problems, offering substantial positive impacts across critical engineering domains. By enabling accurate parameter estimation from sparse data, our SK-FM framework can significantly reduce the need for invasive physical probing. For instance, in subsurface fluid dynamics (Darcy flow), it facilitates efficient groundwater mapping without prohibitive drilling costs. In Structural Health Monitoring (SHM) and fluid mechanics (Navier-Stokes), high-fidelity inversions can accelerate the early detection of material fatigue or aerodynamic anomalies, enhancing public infrastructure safety.

Despite these positive applications, we must reflect on potential negative societal impacts and harms. First, regarding unintended uses or intentional misuse: while our research is foundational, there is a direct path where the same subsurface inversion capabilities used for sustainable groundwater management could be dual-used to optimize the aggressive extraction of non-renewable fossil fuels, conflicting with global sustainability goals.

Second, regarding harms that could arise when the technology is used as intended but gives incorrect results: in safety-critical civil engineering applications, an over-reliance on AI predictions could lead to catastrophic structural failures if the model hallucinates non-physical artifacts due to out-of-distribution noise or mis-specified boundary conditions.

As a primary mitigation strategy against the risk of algorithmic over-reliance, our framework explicitly integrates robust, physically-grounded Uncertainty Quantification. By yielding highly calibrated confidence intervals that spatially correlate with actual predictive errors, SK-FM signals its own epistemic boundaries. This inherent "error-warning" mechanism promotes responsible, human-in-the-loop decision-making, ensuring that domain experts are alerted to regions of high uncertainty rather than placing blind trust in the automated system.

G LIMITATIONS AND FUTURE WORK

While the proposed SK-FM establishes a robust framework for physically-consistent inversion, it presents several limitations and relies on implicit assumptions that open avenues for future research.

Inference-time Computational and Memory Overhead. Our test-time Kinematic Guidance neutralizes geometric errors by backpropagating the physical penalty $\nabla_{\mathbf{x}}\mathcal{L}_{\text{curl}}$ through the computational graph at each ODE integration step. Although the sub-second overhead is highly favorable in 2D settings, scaling this approach to massive datasets or high-resolution 3D spatio-temporal fields (e.g., 3D seismic inversion) may impose severe VRAM bottlenecks. Future work should explore gradient-free Guidance techniques to improve computational efficiency.

Reliance on Well-specified Physical Priors. A strong assumption of our method's "surgical" precision is that the underlying physical laws (e.g., zero curl for gradient fields) are well-specified and perfectly known. In complex, real-world engineering scenarios involving unmodeled turbulence or heterogeneous material dissipation, this assumption is often violated in practice. Enforcing a rigid theoretical manifold under such conditions might introduce systematic bias. Integrating SK-FM with equation discovery methods to accommodate partially known physics is a promising next step.

Static Guidance Scale and Uncertainty Calibration. Currently, the kinematic penalty weight is a static hyperparameter. As observed in our Uncertainty Quantification analysis, this rigid constraint can slightly suppress probabilistic diversity, influencing the performance by leading to marginally under-covered confidence intervals. Developing time-dependent or learnable guidance scales could further optimize the precision-sharpness trade-off.

Robustness to High Noise: The Helmholtz Failure Case. Finally, we observed that the Inhomogeneous Helmholtz scenario presents a non-linear noise degradation boundary for SK-FM. While degrading gracefully under moderate noise (0.3), the model exhibits an abrupt performance drop at extreme levels (0.4). Because our $W^{1,2}$ Sobolev prior inherently acts as a global structural low-pass filter, it conflicts with the highly oscillatory phase information of wavefields when severe noise overwhelmingly corrupts local gradients. To overcome this spectral boundary in future work, we aim to explore *frequency-aware Sobolev priors* that adaptively relax the gradient penalty in high-frequency regions. Furthermore, integrating multi-scale or wavelet-based representations into the Flow Matching objective could allow the generative prior to disentangle valid physical oscillations from parasitic noise, paving the way for robust frequency-domain wave inversions under extreme corruption.

Application Scope and Geometric Constraints. A key constraint of the proposed method is its focus on scalar fields. For vector field generation (e.g., turbulent velocity fields), the presence of non-zero curl (vorticity) means that a simple gradient-based kinematic guidance would overlook the solenoidal component of the flow. Extending this framework to vector fields would require incorporating Helmholtz Decomposition or vector potential learning, which we leave for future work.