EquiReg: Equivariance Regularized Diffusion for Inverse Problems

Anonymous authors
Paper under double-blind review

Abstract

Diffusion models represent the state-of-the-art for solving inverse problems such as image restoration tasks. Diffusion-based inverse solvers incorporate a likelihood term to guide prior sampling, generating data consistent with the posterior distribution. However, due to the intractability of the likelihood, most methods rely on isotropic Gaussian approximations, which can push estimates off the data manifold and produce inconsistent, poor reconstructions. We propose *Equivariance Regularized* (EquiReg) diffusion, a general plug-and-play framework that improves posterior sampling by penalizing trajectories that deviate from the data manifold. EquiReg formalizes manifold-preferential equivariant functions that exhibit low equivariance error for on-manifold samples and high error for off-manifold ones, thereby guiding sampling toward symmetry-preserving regions of the solution space. We highlight that such functions naturally emerge when training non-equivariant models with augmentation or on data with symmetries. EquiReg is particularly effective under reduced sampling and measurement consistency steps, where many methods suffer severe quality degradation. By regularizing trajectories toward the manifold, EquiReg implicitly accelerates convergence and enables high-quality reconstructions. EquiReg consistently improves performance in linear and nonlinear image restoration tasks and solving partial differential equations.

1 Introduction

Inverse problems aim to recover an unknown signal $\boldsymbol{x}^* \in \mathbb{R}^d$ from undersampled noisy measurements:

$$\boldsymbol{y} = \mathcal{A}(\boldsymbol{x}^*) + \boldsymbol{\nu} \in \mathbb{R}^m, \tag{1}$$

where $\mathcal{A}$ is a known measurement operator, and $\boldsymbol{\nu}$ is an unknown noise (Groetsch, 1993). Inverse problems are widely studied in science and engineering, including imaging and astrophotography.

Inverse problems are ill-posed, i.e., the inversion process can have many solutions; hence, they require prior information about the desired solution (Kabanikhin, 2008). In the Bayesian formulation, the solution maximizes the posterior distribution $p(\boldsymbol{x}|\boldsymbol{y}) \propto p(\boldsymbol{y}|\boldsymbol{x})p(\boldsymbol{x})$, where $p(\boldsymbol{y}|\boldsymbol{x})$ is the likelihood of the measurements and $p(\boldsymbol{x})$ is a prior describing the signal structure (Stuart, 2010). Examples of handcrafted priors include sparsity (Donoho, 2006) and low-rankness (Candès et al., 2011).

This paper focuses on methods that leverage unconditionally pre-trained score-based generative diffusion models as learned priors (Ho et al., 2020; Song & Ermon, 2019) with applications in image restoration (Chung et al., 2023), medical imaging (Chung et al., 2022a), and solving partial differential equations (PDEs) (Huang et al., 2024; Yao et al., 2025). These methods define a sequential noising process $\boldsymbol{x}_0 \sim p_{\text{data}} \to \boldsymbol{x}_t \to \boldsymbol{x}_T \sim p_T(\boldsymbol{x}) \approx \mathcal{N}(\boldsymbol{0}, \boldsymbol{I})$ and a reverse denoising process parameterized by a neural network score $\nabla_{\boldsymbol{x}_t} \log p_t(\boldsymbol{x}_t)$ (Vincent, 2011). During sampling, these approaches incorporate gradient signals carrying likelihood information to solve inverse problems.

Solving inverse problems with diffusion (Zhang et al., 2025a; Alkhouri et al., 2025) requires computing the conditional score $\nabla_{\boldsymbol{x}_t} \log p_t(\boldsymbol{x}_t|\boldsymbol{y})$, decomposed into $\nabla_{\boldsymbol{x}_t} \log p_t(\boldsymbol{x}_t) + \nabla_{\boldsymbol{x}_t} \log p_t(\boldsymbol{y}|\boldsymbol{x}_t)$. This introduces challenges, as the likelihood score $\nabla_{\boldsymbol{x}_t} \log p_t(\boldsymbol{y}|\boldsymbol{x}_t) = \nabla_{\boldsymbol{x}_t} \log \int p(\boldsymbol{y}|\boldsymbol{x}_0)p_t(\boldsymbol{x}_0|\boldsymbol{x}_t)\mathrm{d}\boldsymbol{x}_0$ is only computationally

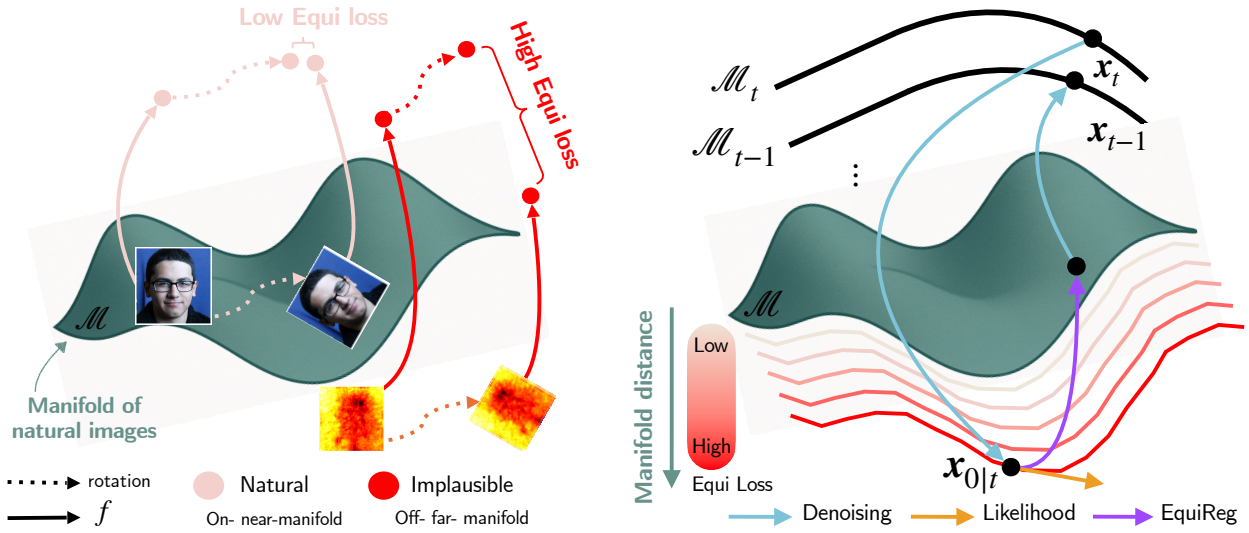


Figure 1: **Equivariance Regularized (EquiReg) diffusion for inverse problems.** (left) Manifold preferential equivariance (MPE) functions whose equivariance error is lower for on-manifold and higher for off-manifold data. (right) EquiReg regularizes the posterior sampling trajectory for improved performance. It penalizes off-manifold trajectories via MPE-based regularization.

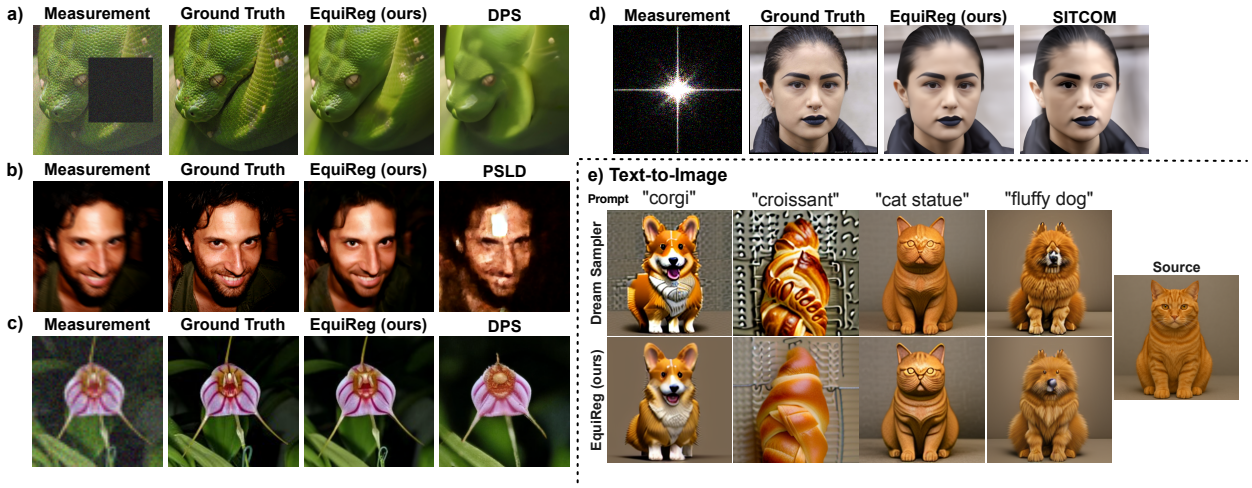


Figure 2: **EquiReg's broad applicability.** a-d) image restoration inverse problems and e) text-guided image generation, resulting in artifact reduction and more realistic generation. Here, EquiReg refers to our regularization being applied to the diffusion sampling method on the same row.

tractable when $t = 0$. To handle the likelihood for $t > 0$, many methods approximate the posterior $p_t(\boldsymbol{x}_0|\boldsymbol{x}_t)$ with the isotropic Gaussian distribution (Zhang et al., 2025a), where the distribution expectation is computed using the optimal denoising score (Robbins, 1956). The Gaussian approximation can be inaccurate for complex distributions (Figure 3), leading to errors in likelihood computation, especially with point estimations (Chung et al., 2023). Since the posterior expectation is a conditional expectation, a linear combination of all possible $\boldsymbol{x}_0$, it may lie off the data manifold even when individual samples remain on it. These issues are further amplified in latent diffusion models (LDMs), introducing artifacts (Rout et al., 2023).

Prior work has attempted to address this challenge via projection-based (He et al., 2024; Zirvi et al., 2025) or decoupled optimization strategies (Zhang et al., 2025a), aimed at reducing the propagation of measurement errors during sampling. However, they rely on the isotropic Gaussian assumption, which can lead to failures on difficult tasks or at reduced sampling steps. While higher-order statistics reduces errors (Boys et al., 2024), most approaches employ such approximation for its efficiency, scalability, and simplicity (Alkhouri et al.,

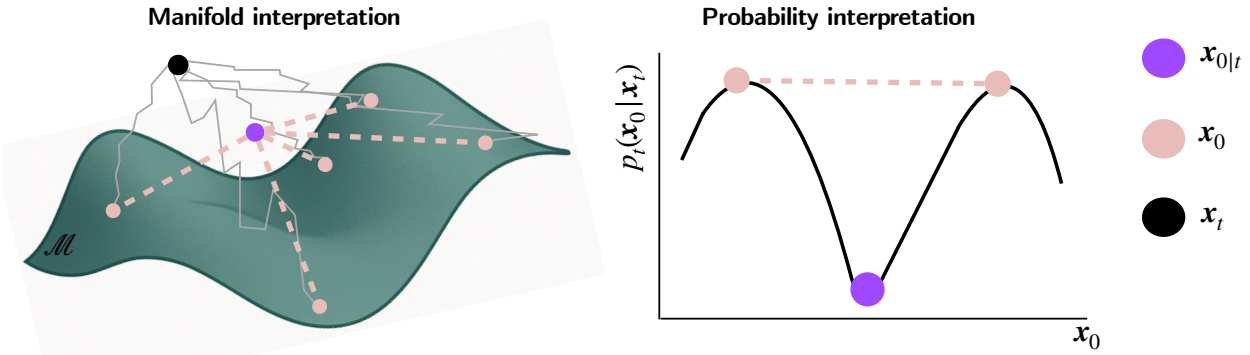


Figure 3: **Off-manifold posterior expectation.** This impacts the likelihood score $p_t(\boldsymbol{y}|\boldsymbol{x}_t) = \int p(\boldsymbol{y}|\boldsymbol{x}_0)p_t(\boldsymbol{x}_0|\boldsymbol{x}_t)\mathrm{d}\boldsymbol{x}_0$ computation achieved via isotropic Gaussian modelling of $p_t(\boldsymbol{x}_0|\boldsymbol{x}_t)$.

2025), often coupled with large-scale LDMs (Peebles & Xie, 2023). This raises a key question: how can we ensure the reliability and practicality of conditional diffusion models under this approximation?

Equivariance offers a natural mechanism to keep sampling trajectories close to the data manifold. We address this challenge with a regularization scheme that leverages equivariance to improve posterior sampling by guiding trajectories toward symmetry-preserving solution spaces. Prior work has enforced equivariance directly within generation or denoising processes (Chen et al., 2023a; Terris et al., 2024), with extensions to probabilistic symmetries (Bloem-Reddy et al., 2020) enabling sample efficiency (Wang et al., 2024).

Our approach differs as follows: rather than strictly enforcing equivariance within denoising architectures, which can hinder tasks requiring symmetry breaking (Lawrence et al., 2025), we employ equivariance as a plug-and-play regularizer to guide diffusion trajectories toward the data manifold.

Our contributions. We propose *Equivariance Regularized* (EquiReg) diffusion, an equivariance-based regularization framework for solving inverse problems with diffusion models (Figure 1). EquiReg leverages equivariance to *regularize* likelihood-induced errors during posterior sampling, guiding diffusion trajectories toward more consistent, on-manifold solutions. Crucially, it employs *Manifold-Preferential Equivariant* (MPE) functions, which discriminate on-manifold from off-manifold data by exhibiting low equivariance error in-distribution and higher error out-of-distribution. We formalize that an effective regularizer should capture such a global property, and MPE functions provide a principled way to direct sampling toward plausible solutions. This design makes EquiReg architecture-agnostic: the regularizer operates independently of the diffusion model itself. With a suitable MPE function, EquiReg improves performance across models, including those with equivariant scores, where likelihood guidance may otherwise push trajectories off the manifold.

We observe that many practical functions behave as MPEs: their equivariance error is small on the training or data manifold but grows off-manifold. This behavior arises in learned models trained with data augmentation, as well as in data with inherent symmetries such as those from physical systems. Rather than treating the degradation off-manifold as a limitation, we exploit it as a signal: equivariance error serves as a natural discriminator for identifying undesirable states during diffusion sampling. Building on this idea, we construct pre-trained MPEs as the foundation of our EquiReg loss. The choice of this function is independent of the denoiser in diffusion models and can be derived separately. For instance, if the diffusion architecture is itself equivariant to a set of group actions, it cannot be regularized via EquiReg using these same equivariances, as the MPE cannot discriminate between on- and off-manifold samples (the equivariance loss would remain low). In this case, a separate MPE with different equivariance properties can be trained. We explore several constructed and pre-trained MPE functions and show that EquiReg is widely robust to the choice of MPE.

Beyond architectural flexibility, we systematically analyze the diversity properties of EquiReg-guided posterior sampling. We demonstrate that EquiReg offers favourable fidelity–diversity trade-offs, and as inverse problems become more ill-posed, diversity increases approximately linearly, indicating that EquiReg naturally expands posterior exploration with growing uncertainty rather than collapsing to a single mode.

We further demonstrate that MPE functions are not rare or specialized constructs, but emerge broadly across practical neural networks. We empirically analyze multiple architectures, including the latent diffusion encoder (Rombach et al., 2022), CNN autoencoders trained with symmetry-based augmentations, pre-trained

ResNet-50 (He et al., 2016), and CLIP encoders (Radford et al., 2021), and show that their equivariance error increases systematically as Gaussian noise pushes samples off the data manifold. We further show that using these MPE functions for EquiReg guidance improves solutions to inverse problems. Overall, this study shows that MPEs are widespread, and EquiReg leverages this intrinsic property as a principled regularization signal.

We validate the efficacy of EquiReg through extensive experiments across diverse diffusion models, inverse problems, and datasets. We demonstrate that EquiReg improves perceptual image quality and remains effective in cases where baselines fail. We show that EquiReg improves the performance of SITCOM (Alkhouri et al., 2025) and DPS (Chung et al., 2023) when the number of measurement consistency and sampling steps are reduced, thus moving toward more efficient diffusion-based solvers. Our method is particularly useful when applied to LDMs. EquiReg reduces failure cases, and consistently improves PSLD (Rout et al., 2023), ReSample (Song et al., 2023a), and DPS (Chung et al., 2023) on linear and nonlinear image restoration tasks. For example, EquiReg significantly improves the LPIPS (Song et al., 2023a) of ReSample by 51% for motion deblur and the FID of DPS (Chung et al., 2023) by 59% on super-resolution. We also include diversity analyses, demonstrating that EquiReg maintains diversity without collapse of single mode reconstruction.

We extend EquiReg's applicability to function-space diffusion models and demonstrate its added benefit for solving PDEs. EquiReg achieves a 7.3% relative reduction in the ℓ_2 error of FunDPS (Mammadov et al., 2024a; Yao et al., 2025) on the Helmholtz equation and a 7.5% relative reduction on the Navier-Stokes equation. Lastly, we include preliminary experiments on EquiReg improving the realism and plausibility of text-guided image generation, emphasizing that the benefits of EquiReg extend beyond image restorations. Overall, the flexibility of EquiReg as a plug-and-play regularization framework suggests that its utility will extend well beyond the specific methods studied in this paper.

2 Preliminaries and Related Works

Diffusion models. Diffusion generative models (Ho et al., 2020; Song & Ermon, 2019; Sohl-Dickstein et al., 2015; Kadkhodaie & Simoncelli, 2021) are state-of-the-art in computer vision for image (Esser et al., 2024) and video generation (Brooks et al., 2024; Zhang et al., 2025b), with score-based methods (Song et al., 2021) being among the most widely used. Diffusion models generate data via a reverse noising process. The forward noising process transforms the data sample $\boldsymbol{x}_0 \sim p_{\text{data}}$ via a series of additive noise into an approximately Gaussian distribution ($p_{\text{data}} \to p_t \to \mathcal{N}(0, I)$ as $t \to \infty$), described by the stochastic differential equation (SDE) $\mathrm{d}\boldsymbol{x} = -\frac{\beta_t}{2}\boldsymbol{x}\mathrm{d}t + \sqrt{\beta_t}\mathrm{d}\boldsymbol{w}$, where $\boldsymbol{w}$ is a standard Wiener process, and the drift and diffusion coefficients are parameterized by a monotonically increasing noise scheduler $\beta_t \in (0, 1)$ in time t (Ho et al., 2020). Reversing the forward diffusion process is described by (Anderson, 1982)

$$\mathrm{d}\boldsymbol{x} = [-\tfrac{\beta_t}{2}\boldsymbol{x} - \beta_t \nabla_{\boldsymbol{x}_t} \log p_t(\boldsymbol{x}_t)]\ \mathrm{d}t + \sqrt{\beta_t}\mathrm{d}\bar{\boldsymbol{w}} \tag{2}$$

with $\mathrm{d}t$ moving backward in time or in discrete steps from T to 0. This reverse SDE is used to sample data from the distribution p_{data}, where the unknown gradient $\nabla_{\boldsymbol{x}_t} \log p_t(\boldsymbol{x}_t)$ is approximated by a scoring function $s_\theta(\boldsymbol{x}_t, t)$, parameterized by a neural network and learned via denoising score matching methods (Hyvärinen & Dayan, 2005; Vincent, 2011). Solving inverse problems is described as a conditional generation where the data is sampled from the posterior $p(\boldsymbol{x}|\boldsymbol{y})$:

$$\mathrm{d}\boldsymbol{x} = [-\tfrac{\beta_t}{2}\boldsymbol{x}\mathrm{d}t - \beta_t(\nabla_{\boldsymbol{x}_t} \log p_t(\boldsymbol{x}_t) + \nabla_{\boldsymbol{x}_t} \log p_t(\boldsymbol{y}|\boldsymbol{x}_t))]\mathrm{d}t + \sqrt{\beta_t}\mathrm{d}\bar{\boldsymbol{w}} \tag{3}$$

For solving general inverse problems where the diffusion is *pre-trained* unconditionally, the prior score $\nabla_{\boldsymbol{x}_t} \log p_t(\boldsymbol{x}_t)$ can be estimated using $s_\theta(\boldsymbol{x}_t, t)$. However, the likelihood score $\nabla_{\boldsymbol{x}_t} \log p_t(\boldsymbol{y}|\boldsymbol{x}_t)$ is only known at $t = 0$, otherwise it is computationally intractable.

Diffusion models for inverse problems. Solving inverse problems with pre-trained diffusion models requires approximating the intractable likelihood score $\nabla_{\boldsymbol{x}_t} \log p_t(\boldsymbol{y}|\boldsymbol{x}_t)$. Training-free solvers differ in how they approximate $p_t(\boldsymbol{y}|\boldsymbol{x}_t)$ and combine it with the prior $p_t(\boldsymbol{x}_t)$ (Peng et al., 2024). Since $p_t(\boldsymbol{y}|\boldsymbol{x}_t) = \int p(\boldsymbol{y}|\boldsymbol{x}_0)p_t(\boldsymbol{x}_0|\boldsymbol{x}_t)\mathrm{d}\boldsymbol{x}_0$, the common choice is to approximate $p_t(\boldsymbol{x}_0|\boldsymbol{x}_t)$ by an isotropic Gaussian $\mathcal{N}(\boldsymbol{x}_{0|t}, r_t^2\boldsymbol{I})$ (Chung et al., 2023; Song et al., 2023b; Zhu et al., 2023; Zhang et al., 2025a). With an optimal denoising score $s_\theta(\boldsymbol{x}_t, t)$, the posterior mean $\boldsymbol{x}_{0|t} \coloneqq \mathbb{E}[\boldsymbol{x}_0|\boldsymbol{x}_t]$ follows from Tweedie's formula (Robbins, 1956;

Miyasawa et al., 1961; Efron, 2011). While an MMSE estimate, $p_t(\boldsymbol{x}_0|\boldsymbol{x}_t)$ may not be concentrated around its mean for complex or multimodal distributions, leading to off-manifold solutions (see Figure 3).

Equivariance. Equivariance describes how functions transform predictably under group actions and provides a principled way to incorporate symmetry into deep learning (Bronstein et al., 2021). It has been applied to graphs (Satorras et al., 2021), convolutional networks (Cohen & Welling, 2016; Romero & Lohit, 2022), Lie groups for dynamical systems (Finzi et al., 2020), and diffusion models (Wang et al., 2024), with applications spanning molecular generation (Hoogeboom et al., 2022; Cornet et al., 2024), autonomous driving (Chen et al., 2023b), robotics (Brehmer et al., 2023), crystal structure prediction (Jiao et al., 2023), and audio inverse problems (Moliner et al., 2023). Equivariance guidance has also improved temporal consistency in video generation (Daras et al., 2024), and its role as a prior in inverse problems is theoretically supported in compressed sensing (Tachella et al., 2023).

Definition 2.1 (Equivariance)**.** *Let G act on $\mathcal{Z}$ via $T_g : \mathcal{Z} \to \mathcal{Z}$ and on $\mathcal{X}$ via $S_g : \mathcal{X} \to \mathcal{X}$. A function $f : \mathcal{Z} \to \mathcal{X}$ is equivariant if for all $g \in G$ and $\boldsymbol{z} \in \mathcal{Z}$, $f(T_g(\boldsymbol{z})) = S_g(f(\boldsymbol{z}))$.*

An equivariant function preserves structure under group transformations (Definition 2.1). While prior work leverages exact equivariance to encode symmetries directly into neural networks, recent studies investigate approximate equivariance to relax strict symmetry assumptions that may not hold in real-world data, aiming to improve performance (Wang et al., 2022). These works introduce a formal definition of approximate equivariance (Definition 2.2) and an equivariance error to quantify deviations from perfect symmetry.

Definition 2.2 (Approximate Equivariant Functions)**.** *Let G act on $\mathcal{Z}$ via $T_g : \mathcal{Z} \to \mathcal{Z}$ and on $\mathcal{X}$ via $S_g : \mathcal{X} \to \mathcal{X}$. A function $f : \mathcal{Z} \to \mathcal{X}$ is ϵ-approximate equivariant if for all $g \in G$ and $\boldsymbol{z} \in \mathcal{Z}$, $\|S_g(f(\boldsymbol{z})) - f(T_g(\boldsymbol{z}))\| \leq \epsilon$. The equivariance error of the function $f : \mathcal{Z} \to \mathcal{X}$ is defined as $\sup_{\boldsymbol{z},g} \|S_g(f(\boldsymbol{z})) - f(T_g(\boldsymbol{z}))\|$. Hence, f is ϵ-approximate equivariant iff its error $< \epsilon$.*

We use the term manifold which refers to the data manifold hypothesis (see Assumption H.1) (Cayton et al., 2005) that assumes data is sampled from a low-dimensional manifold embedded in a high-dimensional space. This hypothesis is popular in machine learning (Bordt et al., 2023) and diffusion-based solvers (He et al., 2024; Chung et al., 2022b; 2023), supported by empirical evidence for imaging (Weinberger & Saul, 2006).

3 EquiReg: Equivariance Regularized Diffusion

We begin by presenting a generalized regularization framework for improving diffusion-based inverse solvers. We focus on the property of *equivariance* and introduce a new class of functions whose equivariance errors are distribution-dependent (low for on- or near-manifold samples and high for off-manifold samples). We leverage these functions to regularize diffusion models, guiding sampling trajectories toward better inverse solutions.

Algorithm 1 Equi-DPS for Inverse Problems.

Require: $T, \boldsymbol{y}, \{\zeta_t\}_{t=1}^T, \{\tilde{\sigma}_t\}_{t=1}^T, \boldsymbol{s}_\theta$, $\mathcal{R}(\cdot)$, $\{\lambda_t\}_{t=1}^T$

$\boldsymbol{x}_T \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I})$

for $t = T-1$ **to** 0 **do**

  $\hat{\boldsymbol{s}} \leftarrow \boldsymbol{s}_\theta(\boldsymbol{x}_t, t)$

  $\boldsymbol{x}_{0|t} \leftarrow \frac{1}{\sqrt{\bar{\alpha}_t}}(\boldsymbol{x}_t + (1-\bar{\alpha}_t)\hat{\boldsymbol{s}})$

  $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I})$

  $\boldsymbol{x}'_{t-1} \leftarrow \frac{\sqrt{\alpha_t}(1-\bar{\alpha}_{t-1})}{1-\bar{\alpha}_t}\boldsymbol{x}_t + \frac{\sqrt{\bar{\alpha}_{t-1}}\beta_t}{1-\bar{\alpha}_t}\boldsymbol{x}_{0|t} + \tilde{\sigma}_t\boldsymbol{\epsilon}$

  $\boldsymbol{x}_{t-1} \leftarrow \boldsymbol{x}'_{t-1} - \zeta_t \nabla_{\boldsymbol{x}_t}\|\boldsymbol{y} - \mathcal{A}(\boldsymbol{x}_{0|t})\|_2^2$

  $\boldsymbol{x}_{t-1} \leftarrow \boldsymbol{x}_{t-1} - \lambda_t \nabla_{\boldsymbol{x}_t}\mathcal{R}(\boldsymbol{x}_t)$

end for

return $\boldsymbol{x}_0$

This paper addresses the propagation error introduced by the approximation of posterior $p_t(\boldsymbol{x}_0|\boldsymbol{x}_t)$ by incorporating an explicit regularization term. The proposed framework is general and can be applied as plug-in on a wide range of pixel and latent-space diffusion models. Given $p_t(\boldsymbol{y}|\boldsymbol{x}_t) = \int p(\boldsymbol{y}|\boldsymbol{x}_0)p_t(\boldsymbol{x}_0|\boldsymbol{x}_t)\mathrm{d}\boldsymbol{x}_0$, let $\tilde{p}_t(\boldsymbol{x}_0|\boldsymbol{x}_t)$ denote an approximation of the posterior to make the likelihood tractable. We formulate the regularized reverse diffusion dynamics as

$$\mathrm{d}\boldsymbol{x} = [-\tfrac{\beta_t}{2}\boldsymbol{x}\mathrm{d}t - \beta_t\nabla_{\boldsymbol{x}_t}(\log p_t(\boldsymbol{x}_t) + \log\textstyle\int p(\boldsymbol{y}|\boldsymbol{x}_0)\tilde{p}_t(\boldsymbol{x}_0|\boldsymbol{x}_t)\mathrm{d}\boldsymbol{x}_0 - \mathcal{R}(\boldsymbol{x}_t))]\mathrm{d}t + \sqrt{\beta_t}\mathrm{d}\bar{\boldsymbol{w}}, \tag{4}$$

where $\mathcal{R}(\boldsymbol{x}_t)$ is the regularizer. Applying this to DPS (Chung et al., 2023) takes the form in Algorithm 1). This formulation brings us to the key contribution: how to design the regularizer to be effective in improving posterior sampling of diffusion models. Regularization in optimization aims to penalize undesired solutions, moving the algorithm towards solutions with low regularizer value. In the context of diffusion models for

sampling from data distributions, we interpret an ideal regularizer as follows: an ideal regularizer should yield low values for on-manifold and high values for off-manifold samples, enabling accurate posterior sampling even when the likelihood score is approximated.

In terms of sampling dynamics, i.e., when applied at each reverse-diffusion step, the regularizer should effectively penalize trajectories leaving the data manifold and reinforce those aligned with high-probability regions. This motivates designing a regularizer that applies a global correction to the entire functional, in contrast to prior works that focus only on locally reducing likelihood error. The ideal property of a regularizer would be to produce high error on undesirable samples and low error on desirable samples. We use equivariance, a global property that enforces geometric symmetries to instantiate such a regularizer to guide the diffusion process toward the data manifold. To realize this idea, we seek functions that exhibit approximate equivariance and discriminate on- from off-manifold samples.

We propose to quantify the equivariance of a function relative to a data distribution. Specifically, while the literature has primarily studied the equivariance properties of functions for general inputs, we propose a new definition for functions in which their equivariance error is distribution-dependent and defined under the support of an input data distribution (Definition 3.1).

Definition 3.1 (Distribution-Dependent Equivariant Functions)**.** *Let G act on $\mathcal{Z}$ via $T_g : \mathcal{Z} \to \mathcal{Z}$ and on $\mathcal{X}$ via $S_g : \mathcal{X} \to \mathcal{X}$. The equivariance error of the function $f : \mathcal{Z} \to \mathcal{X}$ under the distribution p is defined as* $\sup_g \mathbb{E}_{\boldsymbol{z}\sim p}[\|S_g(f(\boldsymbol{z})) - f(T_g(\boldsymbol{z}))\|]$.

The above definition enables us to define functions whose equivariance error can differentiate on-manifold samples from off-manifold ones. We aim to find functions whose equivariance error is low for on-manifold data and high elsewhere. We also introduce a constrained version of equivariance error, where the input is implicitly regularized to lie on the manifold $\mathcal{M}$ in addition to minimizing the equivariance error (Definition 3.2). Both equivariance errors are non-local, defined at the distribution level. When used to regularize the reverse conditional diffusion process, they are computed via local evaluations over the sampled data.

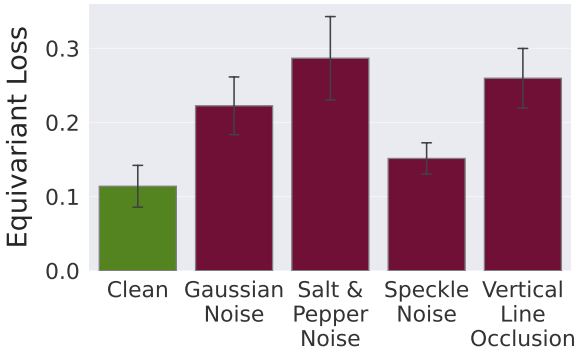
Definition 3.2 (Manifold-Constrained Distribution-Dependent Equivariant Functions)**.** *Let G act on $\mathcal{Z}$ via $T_g : \mathcal{Z} \to \mathcal{Z}$ and on $\mathcal{X}$ via $S_g : \mathcal{X} \to \mathcal{X}$. The manifold-constrained equivariance error of the function $f : \mathcal{Z} \to \mathcal{X}$ under the data distribution p is* $\sup_g \mathbb{E}_{\boldsymbol{z}\sim p}[\|\boldsymbol{z} - h(S_g^{-1}(f(T_g(\boldsymbol{z}))))\|]$ *where $h : \mathcal{X} \to \mathcal{Z}$, and the pair (f, h) forms a vanishing-error autoencoder (see Appendix I).*

To define our method, we term a class of *manifold-preferential equivariant (MPE)* functions, whose equivariance error is lower for samples on the data manifold than for off-manifold samples. EquiReg is a regularization framework, not a manifold projection method. EquiReg penalizes states that deviate from symmetry-preserving regions; when an MPE function is used, these regions align with the data manifold. In practice, MPE functions can emerge in different ways, which we illustrate with examples from augmented training and from data symmetries. MPE can emerge when functions are trained with symmetry-preserving mechanisms such as data augmentation. Prior work has studied equivariant properties of learned representations in deep networks (Lenc & Vedaldi, 2015), showing that data augmentations (Krizhevsky et al., 2012) and representation compression via reduced model capacity (Bruintjes et al., 2023) promote equivariant features even when equivariance is not explicitly built into the architecture. Importantly, the trained network is only approximately equivariant, and prior studies have noted that symmetry-preserving properties degrade for inputs deviating from in-distribution data (Azulay & Weiss, 2019). A few studies have leveraged this emergent MPE in trained networks for out-of-distribution detection (Zhou, 2022; Kaur et al., 2022; 2023).

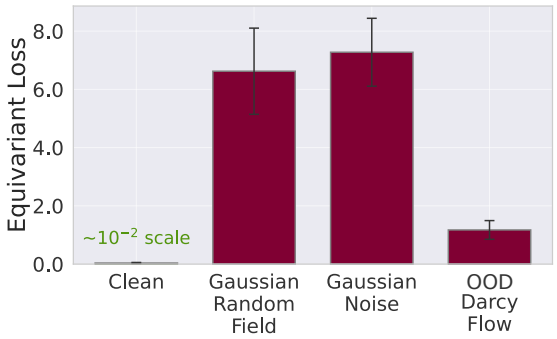
To demonstrate the widespread MPE property of learned mappings, we have considered additional pre-trained models and quantified their equivariance loss for several datasets, i.e., natural images and corrupted ones (see Section I of Appendix.) Figure 4a illustrates the MPE property, emergent via training with augmentations, of $\mathcal{E}$-$\mathcal{D}$ of a pre-trained autoencoder, currently used in LDMs. Specifically, it shows that the equivariance error is lower for natural images and increases when images deviate from the clean data distribution. Based on Definitions 3.1 and 3.2, we propose *Equi* loss using an MPE function f for diffusion-based inverse solvers:

$$\mathcal{R}_{f_{\text{MPE}}}(\boldsymbol{x}_t) = \|S_g(f_{\text{MPE}}(\boldsymbol{x}_{0|t})) - f_{\text{MPE}}(T_g(\boldsymbol{x}_{0|t}))\|_2^2 \tag{5}$$

where $\boldsymbol{x}_{0|t}$ and $\boldsymbol{z}_{0|t}$ are function of $\boldsymbol{x}_t$ and $\boldsymbol{z}_t$, respectively. MPE can also emerge due to symmetries present in the data itself during training. This often occurs in physics systems where coefficient functions, boundary



(a) Equivariance from augmentation, error computed with $f_{\text{MPE}} = \mathcal{E}$ from autoencoder pretrained on FFHQ-256 and evaluated on examples from FFHQ-256.



(b) Equivariance arising from data symmetries. Error computed with $f_{\text{MPE}} = \text{FNO}$ and evaluated on Navier-Stokes data.

Figure 4: **Equivariance error is consistently lower for clean vs perturbed examples.**

values, and solution functions of PDEs remain valid under invertible coordinate transformations. Formally, let $\mathcal{G}(a) \mapsto u$ be a PDE operator mapping initial condition a to solution u, and let T_g and S_g be invertible transformations that preserve PDE structure and boundary conditions. Then, $S_g(\mathcal{G}(a)) = \mathcal{G}(T_g(a))$. Neural operators (Kovachki et al., 2021), popular architectures for modelling physics, trained on PDEs with such inherent symmetries can learn equivariance properties. Figure 4b shows that an MPE function can be constructed using Fourier Neural Operators (FNOs (Li et al., 2021)) trained on non-augmented Navier–Stokes physics data. When a pre-trained FNO is used as the MPE function, the equivariance loss in equation 5 is lower for in-distribution data than for out-of-distribution data under reflection as the group action.

The key message from our MPE examples is that MPE behaviour naturally emerges when a function (e.g., a neural network) is trained with appropriate augmentations or when the data exhibit inherent symmetries. Our results extend this analysis and demonstrate that MPE functions are widely present and can be effectively integrated into the EquiReg framework to improve posterior sampling (Table 8). We leverage this property to distinguish on-manifold from off-manifold samples and to regularize the posterior sampling trajectory toward high-probability regions. Finally, we note that the choice of symmetry group may often be a challenge depending on application domain, a shared challenge in the broader equivariance literature. We provide guidelines on how to choose symmetry groups in Section H with reference on automatic symmetry discovery from data (Zhou et al., 2021; Quessard et al., 2020; Dehmamy et al., 2021; Mohapatra et al., 2025).

Finally, we complement our empirical results with a theoretical discussion of the role of the function $\mathcal{R}$ in regularizing conditional diffusion models and its impact on the sampling trajectory. In particular, we gain insight into the desirable properties of an optimal regularizer by reinterpreting the reverse conditional diffusion process as a time-inhomogeneous Wasserstein gradient flow (Ferreira & Valencia-Guevara, 2018) (see Propositions G.1 and G.2 in Appendix).

4 Results

This section provides experimental results evaluating the performance of EquiReg on inverse problems, including linear and nonlinear image restorations and PDE solving. To fairly assess the impact of EquiReg, we adopt a paired comparison setting (e.g., PSLD vs. Equi-PSLD) across experiments, keeping all other factors such as architecture, training, and sampling fixed. This design ensures that any observed improvements can be attributed specifically to EquiReg rather than to differences in the underlying model or inference procedure. For fair comparison, we set a fixed seed for the initial noise and evaluate the impact of EquiReg under reduced measurement consistency and sampling steps, providing a path toward faster diffusion-based inversion. Results emphasize EquiReg's usefulness when the baseline performance deteriorates. We analyze the diversity trade-offs, and demonstrate the broad emergence of MPEs and their usefulness within EquiReg. Lastly, we provide preliminary analysis on EquiReg improving the realism of text-guided image generation.

Image restoration tasks. We evaluate EquiReg when applied to: SITCOM (Alkhouri et al., 2025), PSLD (Rout et al., 2023), ReSample (Song et al., 2023a), and DPS (Chung et al., 2023). We compare against manifold-preserving or geometry-constraint approaches including MCG (Chung et al., 2022b), MPGD-AE (He

Table 1: **EquiReq Characterization under change of EquiReg period, λ_t, and DDIM steps.**

		Super Resolution				Gaussian Blur			
Method	**Period**	Runtime (s)	PSNR↑	LPIPS↓	FID↓	Runtime (s)	PSNR↑	LPIPS↓	FID↓
DPS	N/A	46.20	22.99 (1.93)	0.20 (0.05)	135.7	46.50	24.59 (2.25)	0.15 (0.03)	88.70
Equi-DPS	1	51.10	26.73 (1.99)	0.12 (0.03)	87.97	52.20	26.08 (2.25)	0.12 (0.03)	87.11
Equi-DPS	2	48.90	26.73 (1.99)	0.12 (0.03)	87.98	49.10	26.06 (2.24)	0.12 (0.03)	87.19
Equi-DPS	5	47.10	26.73 (1.99)	0.12 (0.03)	87.98	47.30	26.06 (2.24)	0.12 (0.03)	87.32
Equi-DPS	10	46.90	26.73 (1.99)	0.12 (0.03)	87.99	47.00	26.05 (2.24)	0.12 (0.03)	87.04

(a) Robustness and computational efficiency of applying EquiReg under various periods during sampling. EquiReg maintains performance when applied every $\{1, 2, 5, 10\}$ DDIM steps while incurring minimal computational overhead.

	DPS				PSLD		
λ_t^{DPS}	PSNR↑	SSIM↑	LPIPS↓	λ_t^{PSLD}	PSNR↑	SSIM↑	LPIPS↓
0.0	24.34 (1.03)	0.664 (0.061)	0.156 (0.051)	0.0	23.83 (2.61)	0.63 (0.12)	0.315 (0.07)
0.001	25.44 (1.22)	0.708 (0.057)	0.118 (0.038)	0.01	25.35 (2.24)	0.70 (0.09)	0.280 (0.07)
0.01	25.44 (1.22)	0.708 (0.057)	0.118 (0.038)	0.1	26.63 (1.68)	0.74 (0.08)	0.337 (0.06)
0.1	25.44 (1.22)	0.708 (0.057)	0.118 (0.038)	0.25	26.22 (1.57)	0.72 (0.08)	0.366 (0.05)
1.0	25.44 (1.22)	0.709 (0.057)	0.118 (0.038)	1.0	24.74 (1.28)	0.66 (0.07)	0.438 (0.05)

(b) Robustness of EquiReg to the choice of λ_t. Sensitivity analysis for DPS and PSLD.

	DPS				**Equi-DPS** (ours)			
Steps	PSNR↑	SSIM↑	LPIPS↓	FID↓	PSNR↑	SSIM↑	LPIPS↓	FID↓
500	13.89	0.0937	0.955	417.07	20.61	0.366	0.500	238.51
750	21.77	0.540	0.254	153.74	25.60	0.704	0.160	110.89
900	22.97	0.628	0.201	148.03	26.52	0.755	0.126	88.46
1000	22.99	0.649	0.201	135.71	26.73	0.767	0.120	88.00

(c) EquiReg improves performance under reduced DDIM steps. Pixel-based super-resolution on FFHQ 256×256.

et al., 2024), and DiffStateGrad (Zirvi et al., 2025). We measure performance via perceptual similarity (LPIPS), distribution alignment (FID), pixel-wise fidelity (PSNR), and structural consistency (SSIM).

We test EquiReg on a) the FFHQ 256×256 (Karras et al., 2021) and b) ImageNet 256×256 validation set (Deng et al., 2009). For pixel-based experiments, we use i) the pre-trained model from (Chung et al., 2023) on FFHQ, and ii) the pre-trained model from (Dhariwal & Nichol, 2021) on ImageNet. For latent diffusion experiments, we use i) the unconditional LDM-VQ-4 model (Rombach et al., 2022) on FFHQ, and ii) the Stable Diffusion v1.5 (Rombach et al., 2022) model on ImageNet.

We evaluate EquiReg on linear and nonlinear restoration tasks for natural images (see Section E for task details). We adopt the pre-trained encoder-decoder $\mathcal{E}$-$\mathcal{D}$ as our MPE function. For FFHQ, we use vertical reflection as the symmetry group, which preserves upright facial orientation. For ImageNet, we define a rotation group $G = \{0, \pi/2, \pi, 3\pi/2\}$, and uniformly at random select the group action for each sample. Finally, the loss functions given in Equation (5) are used to regularize. While our main experiment explore the reflection and rotation groups with small cardinality, EquiReg does not rely on full group coverage. Sampling even a sparse or randomly chosen subset of group actions is sufficient, as long as the function used for regularization exhibits the MPE property across the group (Table 10).

Table 2: **EquiReg improves SITCOM under reduced measurement consistency steps (K_{meas}).** We reduce K_{meas} and add an equal amount of EquiReg steps (K_{EquiReg}). Evaluated with 50 DDIM steps on motion deblur for FFHQ.

$K_{\text{meas.}}$	K_{EquiReg}	PSNR↑	SSIM↑	Runtime (s)
10	N/A	28.06	0.81	21.57
5	5	**29.26**	**0.83**	**11.09**
20	N/A	27.04	0.79	38.85
10	10	**28.93**	**0.82**	**20.92**
30	N/A	27.79	0.80	58.84
15	15	**29.63**	**0.84**	**30.19**
40	N/A	**30.40**	**0.85**	78.08
20	20	29.50	0.83	**41.02**
60	N/A	28.35	0.81	108.57
30	30	**31.36**	**0.87**	**59.38**

Table 3: **EquiReg for diffusion models on FFHQ**. 256×256 with $\sigma_{\boldsymbol{y}} = 0.05$.

Method	Gaussian deblur			Motion deblur			Super-resolution (×4)			Box inpainting			Random inpainting		
	LPIPS↓	FID↓	PSNR↑	LPIPS↓	FID↓	PSNR↑	LPIPS↓	FID↓	PSNR↑	LPIPS↓	FID↓	PSNR↑	LPIPS↓	FID↓	PSNR↑
PSLD	0.357	106.2	22.87	**0.322**	**84.62**	24.25	0.313	89.72	24.51	0.158	43.02	24.22	0.246	49.77	29.05
Equi-PSLD	0.344	94.09	**24.42**	0.338	99.14	24.83	0.289	90.88	**26.32**	0.098	**31.54**	24.19	**0.188**	41.61	**30.43**
EquiCon-PSLD	**0.320**	**83.18**	24.38	**0.322**	89.87	**25.14**	**0.277**	**79.39**	26.14	**0.092**	35.07	**24.26**	0.204	**40.75**	29.99

(a) Latent diffusion.

Method	Gaussian deblur			Motion deblur			Super-resolution (×4)			Box inpainting			Random inpainting		
	LPIPS↓	FID↓	PSNR↑	LPIPS↓	FID↓	PSNR↑	LPIPS↓	FID↓	PSNR↑	LPIPS↓	FID↓	PSNR↑	LPIPS↓	FID↓	PSNR↑
DPS	0.145	104.8	25.48	0.132	99.75	26.75	0.191	125.4	24.38	0.133	56.89	23.10	0.113	51.32	29.63
Equi-DPS (ours)	**0.114**	**48.76**	**26.32**	**0.094**	**41.71**	**28.23**	**0.120**	**51.00**	**27.15**	**0.099**	40.47	23.39	**0.068**	33.65	**32.16**
DiffStateGrad-DPS	0.128	52.73	26.29	0.118	50.14	27.61	0.186	73.02	24.65	0.114	47.53	**24.10**	0.107	49.42	30.15
MCG	0.340	101.2	6.72	0.702	310.5	6.72	0.520	87.64	20.05	0.309	**40.11**	19.97	0.286	**29.26**	21.57
MPGD-AE	0.150	114.9	24.42	0.120	104.5	25.72	0.168	137.7	24.01	0.138	248.7	21.59	0.172	339.0	25.22

(b) Pixel-based diffusion.

Table 4: **EquiReg for latent diffusion models on ImageNet**. 256×256 with $\sigma_{\boldsymbol{y}} = 0.05$.

Method	Gaussian deblur		Motion deblur		Super-resolution (x4)		Box inpainting		Random inpainting	
	FID↓	PSNR↑	FID↓	PSNR↑	FID↓	PSNR↑	FID↓	PSNR↑	FID↓	PSNR↑
PSLD	263.9	20.70	252.1	21.26	224.3	22.29	151.4	16.28	83.22	26.56
EquiCon-PSLD	**214.5**	**22.01**	**196.3**	**22.69**	**198.5**	**22.34**	**137.6**	**19.25**	**65.14**	**27.03**

We characterize EquiReg under change of its period, regularization parameter λ_t, and DDIM steps. EquiReg can be applied efficiently with longer period' it preserves performance when applied with lower frequency (Table 1a); and it is robust to the choice of its hyperparameter λ_t, demonstrated on DPS and PSLD (Table 1b). EquiReg maintains strong performance as the number of DDIM steps is reduced, whereas DPS suffers a significant drop. Importantly, Equi-DPS consistently outperforms DPS, with the performance gap widening at lower step counts (Table 1c, Figure 12). An important advantage of EquiReg is that it allows the user to reduce the number of measurement consistency optimization steps, thereby introducing an implicit acceleration in solving the inverse problem. Table 2 shows that EquiReg regularization consistently enables SITCOM to achieve superior performance with significantly reduced runtime by using fewer measurement consistency steps.

Table 3a, Table 4, and Table 5 highlights the benefits of EquiReg for latent models by consistently improving the performance of ReSample and PSLD across several tasks on FFHQ and ImageNet. We attribute this improvement in part to the reduction of failure cases (Figure 5d). EquiReg also significantly improves the performance of pixel-based methods (see Equi-DPS vs. DPS, Table 3d).

Table 5: **EquiReg for ReSample on linear and nonlinear tasks.** FFHQ 256×256 with $\sigma_{\boldsymbol{y}} = 0.01$.

Task	Method	LPIPS↓	FID↓	PSNR↑	SSIM↑
Linear					
Gaussian deblur	ReSample	0.253	55.65	27.78	0.757
	Equi-ReSample	0.197	64.86	**29.08**	**0.825**
	EquiCon-ReSample	**0.156**	**54.72**	28.18	0.777
Motion deblur	ReSample	0.160	40.14	30.55	0.854
	Equi-ReSample	0.120	46.28	**30.92**	**0.870**
	EquiCon-ReSample	**0.078**	**37.61**	30.73	0.860
Super-res. (×4)	ReSample	0.204	40.46	28.02	0.790
	Equi-ReSample	**0.098**	43.56	**29.74**	**0.849**
	EquiCon-ReSample	0.112	**40.38**	28.27	0.801
Box inpainting	ReSample	0.198	108.30	19.91	0.807
	Equi-ReSample	**0.150**	**59.69**	**22.56**	**0.832**
	EquiCon-ReSample	0.171	110.70	21.04	0.815
Random inpainting	ReSample	0.115	36.12	31.27	0.892
	Equi-ReSample	**0.047**	29.88	**31.47**	**0.908**
	EquiCon-ReSample	**0.047**	**28.81**	31.21	0.904
Nonlinear					
HDR	ReSample	0.190	**49.06**	**24.88**	**0.819**
	Equi-ReSample	**0.133**	49.52	24.71	0.815
	EquiCon-ReSample	0.135	49.98	24.67	0.817
Phase retrieval	ReSample	0.237	97.86	27.61	0.750
	Equi-ReSample	**0.155**	**85.22**	**28.16**	0.770
	EquiCon-ReSample	0.159	88.75	28.11	**0.774**
Nonlinear deblur	ReSample	0.188	56.06	29.54	0.842
	Equi-ReSample	0.128	55.09	29.45	0.840
	EquiCon-ReSample	**0.125**	**54.62**	**29.55**	**0.843**

We observe that EquiReg achieves its largest improvements on perceptual metrics (FID and LPIPS), suggesting it generates more realistic images that lie closer to the data manifold (see Appendix E for supporting qualitative results). EquiReg improves performance under high measurement noise (Figure 5b). This result aligns with Figure 5a, which shows the equivariance error is lower on clean images than noisy ones, indicating that EquiReg enforces an effective denoising. Lastly, we note that EquiReg is robust to regularizing hyperparameter λ_t (Figure 5c, see Section C for details). For qualitative performance of EquiReg, see Figure 15 and Figure 16.

Solving PDEs from sparse observations. EquiReg is evaluated on two important PDE problems: the Helmholtz and Navier-Stokes equations (see Section F). The objective is to solve both forward and inverse problems in sparse sensor settings. The forward problem involves predicting the solution function or the final state using measurements from only 3% of the coefficient field or the initial state. The inverse problem, conversely, aims to predict the input conditions from observations of 3% of the system's output. This task is challenging due to the nonlinearity of the equations, the complex structure of Gaussian random fields, and the sparsity of observations.

Table 6: **Solving PDEs from sparse observations.**

	Steps (N)	Helmholtz		Navier-Stokes	
		Forward	Inverse	Forward	Inverse
DiffusionPDE	2000	12.64%	19.07%	3.78%	9.63%
FunDPS	500	2.13%	17.16%	3.32%	8.48%
Equi-FunDPS (ours)	500	**2.12%**	**15.91%**	**3.06%**	**7.84%**

Recent studies (Huang et al., 2024; Mammadov et al., 2024a; Yao et al., 2025) have demonstrated the superiority of diffusion models over deterministic single-forward methods for solving PDEs. DiffusionPDE (Huang et al., 2024) decomposes the conditional log-likelihood into a learned diffusion prior and a measurement score. FunDPS (Yao et al., 2025) extends the sampling process to a more natural infinite-dimensional spaces, achieving better accuracy and speed via function space models.

We integrate EquiReg into the state-of-the-art FunDPS framework (Mammadov et al., 2024a; Yao et al., 2025), where the equivariance loss is computed using an FNO trained on the corresponding inverse problem. We employ reflection symmetry (i.e., flipping along the $y = x$ axis) and observe no significant performance differences when using alternative transformations such as rotations or alternating flips. Equi-FunDPS improves performance (Table 6), measured by relative ℓ_2 error, across multiple tasks, particularly in inverse problems where a strong data prior is essential. Importantly, these results highlight that EquiReg is not limited to conventional neural networks and is naturally applicable to neural operators, enabling regularization in both finite-dimensional and function-space diffusion models.

Diversity analysis. To study posterior sampling diversity of EquiReg, we generated $K = 10$ posterior samples for 20 test images across three inverse problems of box inpainting, Gaussian deblurring, $4\times$ super-resolution, and measured diversity using two complementary metrics: Intra-LPIPS for perceptual diversity and Pixel-Std for spatial diversity. Table 7 demonstrates that Equi-DPS achieves favorable fidelity-diversity trade-offs. We further investigated diversity scaling by varying box inpainting mask size from 128×128 to 192×192 pixels (Figure 17). Results show that diversity metrics increase linearly with task difficulty, demonstrating that Equi-DPS naturally expands sampling as problems become more ill-posed rather than artificially constraining solutions. This linear relationship indicates healthy, predictable posterior sampling behavior across the difficulty spectrum. Lastly, Figures 6 and 18 provides qualitative confirmation through visual examples showing four posterior samples per image. Observable variations in facial features, expressions, and eye gaze validate our quantitative measurements, confirming EquiReg can generate genuinely diverse reconstructions rather than collapsing to a single solution (for additional discussion and analysis, see Section D).

MPE functions are widespread and effective for EquiReg. We empirically evaluate whether MPE behaviour naturally emerges across commonly used neural networks. We examine i) the emergence of MPE properties in different functions (neural networks) and ii) the effect of using these functions within EquiReg on identical inverse problem settings. For FFHQ and ImageNet, we analyze four function classes: the LDM encoder (Rombach et al., 2022), a CNN autoencoder trained with symmetry augmentations (flip for FFHQ and rotation for ImageNet), pre-trained ResNet-50 (He et al., 2016), and CLIP (Radford et al., 2021). Across all architectures, equivariance error increases systematically as Gaussian noise pushes samples off the data manifold, confirming clear MPE behaviour. The strength of this effect varies: the CNN autoencoder exhibits the strongest MPE signal, while the LDM encoder shows the weakest; ResNet-50 and CLIP lie between

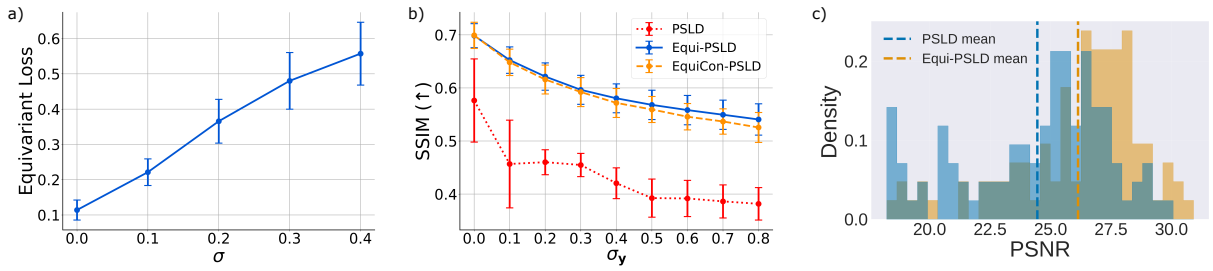


Figure 5: **EquiReg effectiveness and robustness across a range of measurement noise levels and regularization parameter** (a) Equivariance error computed over a pre-trained decoder on increasingly noisy inputs. (b) EquiReg performance computed over a range of measurement noise levels on the FFHQ dataset. (c) EquiReg reduces failure cases and enhances reconstruction fidelity for super-resolution on FFHQ.

Figure 6: **Qualitative diversity examples for super-resolution.** We show $K = 4$ posterior samples for two test images. Samples differ in facial features (i.e., teeth in the first test image, eye color and eyelashes in the second test image) while maintaining high fidelity to the ground truth, demonstrating EquiReg generates diverse plausible reconstructions rather than collapsing to a single mode.

these extremes (see Figures 14 and 19). Notably, the strongest MPE behaviour of the CNN autoencoder is in line with our systematic guidelines for constructing MPE functions; This is precisely the regime where the training distribution of the function matches the distribution of the inverse problem (i.e., training on ImageNet train and evaluating on ImageNet test).

We then integrate each function into EquiReg under identical inverse problem settings (FFHQ/ImageNet, DPS/SITCOM, super-resolution/motion deblurring). In all cases, EquiReg improves reconstruction quality over the corresponding baseline without regularization (Tables 8 and 13). These results demonstrate that i) MPE behaviour naturally arises in widely used pre-trained networks, and ii) EquiReg is robust to the choice of MPE function, even when the MPE signal is relatively weak. Notably, our main experiments use the weakest MPE (the LDM encoder), suggesting further gains are possible with stronger MPE constructions.

Text-to-image guidance. Given a "source" image, DreamSampler (Kim et al., 2024) transforms it according to a text prompt. When applying EquiReg to DreamSampler, perceptual quality and reduced artifacts are improved in the generated images. Figure 2 shows a source "cat" transformed into a "corgi": Equi-DreamSampler produces more realistic results and resolves anatomical inconsistencies (e.g., correcting a three-front-legged corgi to two front legs). We also observe an implicit acceleration effect when EquiReg is imposed (Figure 7). Equi-DreamSampler with 50 DDIM steps produces images comparable to DreamSampler with substantially more steps. Increasing the regularization strength λ_t at fixed 50 DDIM steps results in effects similar to increasing DDIM steps in DreamSampler (from 50 to 75 to 100), suggesting that EquiReg guides sampling trajectories closer to the data manifold. These experiments illustrate that EquiReg is readily

Table 7: **Fidelity and diversity comparison across inverse problems.** EquiReg improves fidelity while largely preserving or enhancing sampling diversity (20 test images with $K = 10$ samples per image.)

Task	Method	Fidelity Metrics		Diversity Metrics	
		LPIPS↓	FID↓	Intra-LPIPS↑	Pixel-Std↑
Box inpainting	DPS	0.140	70.89	0.112	9.286
	Equi-DPS (ours)	**0.112**	**59.70**	**0.118**	**10.59**
Gaussian deblur	DPS	0.150	76.71	**0.114**	**6.565**
	Equi-DPS (ours)	**0.120**	**63.02**	0.092	5.669
Super-resolution (×4)	DPS	0.683	99.11	0.134	7.956
	Equi-DPS (ours)	**0.703**	**87.52**	**0.187**	**23.52**

Table 8: **MPE behaviour emerges across diverse network architectures and consistently improves diffusion-based inverse problems with EquiReg.** SITCOM motion deblurring on FFHQ 256 with $\lambda = 0.05$. Results are reported as mean (standard deviation).

MPE function	PSNR	SSIM	LPIPS
None	27.670 (1.343)	0.790 (0.031)	0.221 (0.040)
LDM Encoder (FFHQ)	28.357 (1.379)	0.806 (0.031)	0.200 (0.036)
CNN Autoencoder (FFHQ)	**28.852 (1.376)**	**0.819 (0.044)**	**0.193 (0.033)**
Pretrained ResNet50	28.682 (1.388)	0.811 (0.036)	0.198 (0.036)

applicable beyond inverse problems, including text-to-image guidance. While not our primary focus, they demonstrate the broader utility and flexibility of EquiReg as a plug-and-play regularization framework.

5 Conclusion

We introduce *Equivariance Regularized* (EquiReg) diffusion for inverse problems. EquiReg regularizes sampling trajectories to stay closer to the data manifold, leveraging manifold-preferential equivariance (MPE): functions with low equivariance error on-manifold and high error off-manifold. Such functions arise naturally in trained networks and can serve as plug-and-play regularizers without modifying the diffusion denoiser. EquiReg is agnostic across pixel- and latent-space diffusion models and remains robust under reduced sampling, effectively accelerating convergence. Across diverse inverse problems, it consistently improves perceptual and reconstruction metrics while reducing failure cases, highlighting its generality and efficiency.

EquiReg operates as a plug-and-play regularization framework and therefore builds upon the quality of the underlying diffusion solver. It does not modify the diffusion architecture itself, but instead improves sampling by guiding trajectories toward more consistent, on-manifold solutions. As a regularization mechanism, its impact is most pronounced in challenging regimes where likelihood guidance alone may degrade or become unstable. Applying EquiReg requires selecting appropriate symmetry groups and constructing suitable MPE functions for the task at hand. While we provide systematic guidelines for imaging and PDE settings, extending these constructions to new domains remains an important direction for future work; we have provided reference to prior work on how symmetries for various application can be learned and set. Finally, although we formalize distribution-dependent equivariant and MPE functions, a deeper theoretical understanding of when MPE behaviour emerges in trained networks and how it may interact with joint training of diffusion models, remains an interesting avenue for further study.

Broader Impact Statement. On the positive side, high-fidelity image restoration can improve downstream tasks in medical imaging, remote-sensing and environmental monitoring (e.g., denoising satellite observations to track pollution or deforestation). Likewise, accelerated PDE-solving via learned diffusion priors may enable faster, more accurate simulations for climate modeling, fluid-dynamics research, and engineering design. On the other hand, robust reconstruction methods could be misappropriated for privacy-invasive surveillance or to create deceptive imagery. We emphasize that our method does not amplify these existing risks.

References

Ismail Alkhouri, Shijun Liang, Cheng-Han Huang, Jimmy Dai, Qing Qu, Saiprasad Ravishankar, and Rongrong Wang. Sitcom: Step-wise triple-consistent diffusion sampling for inverse problems. In *Forty-second International Conference on Machine Learning*, 2025.

Christopher Anders, Plamen Pasliev, Ann-Kathrin Dombrowski, Klaus-Robert Müller, and Pan Kessel. Fairwashing explanations with off-manifold detergent. In *International Conference on Machine Learning*, pp. 314–323. PMLR, 2020.

Brian DO Anderson. Reverse-time diffusion equation models. *Stochastic Processes and their Applications*, 12 (3):313–326, 1982.

Aharon Azulay and Yair Weiss. Why do deep convolutional networks generalize so poorly to small image transformations? *Journal of Machine Learning Research*, 20(184):1–25, 2019.

Lorenzo Baldassari, Ali Siahkoohi, Josselin Garnier, Knut Solna, and Maarten V de Hoop. Conditional score-based diffusion models for bayesian inference in infinite dimensions. *Advances in Neural Information Processing Systems*, 36:24262–24290, 2023.

Wanyu Bian, Albert Jang, Liping Zhang, Xiaonan Yang, Zachary Stewart, and Fang Liu. Diffusion modeling with domain-conditioned prior guidance for accelerated mri and qmri reconstruction. *IEEE Transactions on Medical Imaging*, 2024.

Benjamin Bloem-Reddy, Yee Whye, et al. Probabilistic symmetries and invariant neural networks. *Journal of Machine Learning Research*, 21(90):1–61, 2020.

Ashish Bora, Ajil Jalal, Eric Price, and Alexandros G Dimakis. Compressed Sensing using Generative Models. In *International conference on machine learning*, pp. 537–546. PMLR, 2017.

Sebastian Bordt, Uddeshya Upadhyay, Zeynep Akata, and Ulrike von Luxburg. The manifold hypothesis for gradient-based explanations. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 3697–3702, 2023.

Valentin De Bortoli. Convergence of denoising diffusion models under the manifold hypothesis. *Transactions on Machine Learning Research*, 2022. ISSN 2835-8856. URL https://openreview.net/forum?id=MhK5aXo3gB.

Benjamin Boys, Mark Girolami, Jakiw Pidstrigach, Sebastian Reich, Alan Mosca, and Omer Deniz Akyildiz. Tweedie moment projected diffusions for inverse problems. *Transactions on Machine Learning Research*, 2024. ISSN 2835-8856. URL https://openreview.net/forum?id=4unJi0qrTE. Featured Certification.

Johann Brehmer, Joey Bose, Pim De Haan, and Taco S Cohen. Edgi: Equivariant diffusion for planning with embodied agents. *Advances in Neural Information Processing Systems*, 36:63818–63834, 2023.

Michael M Bronstein, Joan Bruna, Taco Cohen, and Petar Veličković. Geometric deep learning: Grids, groups, graphs, geodesics, and gauges. *preprint arXiv:2104.13478*, 2021.

Tim Brooks, Bill Peebles, Connor Holmes, Will DePue, Yufei Guo, Li Jing, David Schnurr, Joe Taylor, Troy Luhman, Eric Luhman, et al. Video generation models as world simulators. *OpenAI Blog*, 1:8, 2024.

Robert-Jan Bruintjes, Tomasz Motyka, and Jan van Gemert. What affects learned equivariance in deep image recognition models? In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 4839–4847, 2023.

Emmanuel J Candès, Xiaodong Li, Yi Ma, and John Wright. Robust principal component analysis? *Journal of the ACM (JACM)*, 58(3):1–37, 2011.

Gabriel Cardoso, Yazid Janati el idrissi, Sylvain Le Corff, and Eric Moulines. Monte carlo guided denoising diffusion models for bayesian linear inverse problems. In *The Twelfth International Conference on Learning Representations*, 2024. URL https://openreview.net/forum?id=nHESwXvxWK.

Lawrence Cayton et al. *Algorithms for manifold learning*. Univ. of California at San Diego Tech. Rep, 2005.

Dongdong Chen, Mike Davies, Matthias J Ehrhardt, Carola-Bibiane Schönlieb, Ferdia Sherry, and Julián Tachella. Imaging with equivariant deep learning: From unrolled network design to fully unsupervised learning. *IEEE Signal Processing Magazine*, 40(1):134–147, 2023a.

Kehua Chen, Xianda Chen, Zihan Yu, Meixin Zhu, and Hai Yang. Equidiff: A conditional equivariant diffusion model for trajectory prediction. In *2023 IEEE 26th International Conference on Intelligent Transportation Systems (ITSC)*, pp. 746–751. IEEE, 2023b.

Hyungjin Chung and Jong Chul Ye. Score-Based Diffusion Models for Accelerated MRI. *Medical Image Analysis*, pp. 102479, 2022.

Hyungjin Chung, Eun Sun Lee, and Jong Chul Ye. MR Image Denoising and Super-Resolution Using Regularized Reverse Diffusion. *IEEE Transactions on Medical Imaging*, 42(4):922–934, 2022a.

Hyungjin Chung, Byeongsu Sim, Dohoon Ryu, and Jong Chul Ye. Improving Diffusion Models for Inverse Problems using Manifold Constraints. *Advances in Neural Information Processing Systems*, 35:25683–25696, 2022b.

Hyungjin Chung, Jeongsol Kim, Michael Thompson Mccann, Marc Louis Klasky, and Jong Chul Ye. Diffusion Posterior Sampling for General Noisy Inverse Problems. In *The Eleventh International Conference on Learning Representations*, 2023. URL https://openreview.net/forum?id=OnD9zGAGT0k.

Taco Cohen and Max Welling. Group equivariant convolutional networks. In *International conference on machine learning*, pp. 2990–2999. PMLR, 2016.

François Cornet, Grigory Bartosh, Mikkel Schmidt, and Christian Andersson Naesseth. Equivariant neural diffusion for molecule generation. *Advances in Neural Information Processing Systems*, 37:49429–49460, 2024.

Giannis Daras, Weili Nie, Karsten Kreis, Alex Dimakis, Morteza Mardani, Nikola Kovachki, and Arash Vahdat. Warped diffusion: Solving video inverse problems with image diffusion models. *Advances in Neural Information Processing Systems*, 37:101116–101143, 2024.

Nima Dehmamy, Robin Walters, Yanchen Liu, Dashun Wang, and Rose Yu. Automatic symmetry discovery with lie algebra convolutional network. *Advances in Neural Information Processing Systems*, 34:2503–2515, 2021.

Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. ImageNet: A Large-Scale Hierarchical Image Database. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 248–255. IEEE, 2009.

Prafulla Dhariwal and Alexander Nichol. Diffusion models beat gans on image synthesis. *Advances in neural information processing systems*, 34:8780–8794, 2021.

David L Donoho. Compressed sensing. *IEEE Transactions on information theory*, 52(4):1289–1306, 2006.

Zolnamar Dorjsembe, Hsing-Kuo Pao, Sodtavilan Odonchimed, and Furen Xiao. Conditional Diffusion Models for Semantic 3D Brain MRI Synthesis. *IEEE Journal of Biomedical and Health Informatics*, 2024.

Bradley Efron. Tweedie's formula and selection bias. *Journal of the American Statistical Association*, 106 (496):1602–1614, 2011.

Patrick Esser, Sumith Kulal, Andreas Blattmann, Rahim Entezari, Jonas Müller, Harry Saini, Yam Levi, Dominik Lorenz, Axel Sauer, Frederic Boesel, et al. Scaling rectified flow transformers for high-resolution image synthesis. In *Forty-first international conference on machine learning*, 2024.

Charles Fefferman, Sanjoy Mitter, and Hariharan Narayanan. Testing the manifold hypothesis. *Journal of the American Mathematical Society*, 29(4):983–1049, 2016.

Lucas CF Ferreira and Julio C Valencia-Guevara. Gradient flows of time-dependent functionals in metric spaces and applications to pdes. *Monatshefte für Mathematik*, 185(2):231–268, 2018.

Marc Finzi, Samuel Stanton, Pavel Izmailov, and Andrew Gordon Wilson. Generalizing convolutional neural networks for equivariance to lie groups on arbitrary continuous data. In *International Conference on Machine Learning*, pp. 3165–3176. PMLR, 2020.

Fabian Fuchs, Daniel Worrall, Volker Fischer, and Max Welling. Se (3)-transformers: 3d roto-translation equivariant attention networks. *Advances in neural information processing systems*, 33:1970–1981, 2020.

Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative Adversarial Nets. *Advances in neural information processing systems*, 27, 2014.

Charles W Groetsch. *Inverse problems in the mathematical sciences*, volume 52. Springer, 1993.

Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778, 2016.

Yutong He, Naoki Murata, Chieh-Hsin Lai, Yuhta Takida, Toshimitsu Uesaka, Dongjun Kim, Wei-Hsiang Liao, Yuki Mitsufuji, J Zico Kolter, Ruslan Salakhutdinov, and Stefano Ermon. Manifold preserving guided diffusion. In *International Conference on Learning Representations*, 2024.

Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. *Advances in neural information processing systems*, 33:6840–6851, 2020.

Emiel Hoogeboom, Vıctor Garcia Satorras, Clément Vignac, and Max Welling. Equivariant diffusion for molecule generation in 3d. In *International conference on machine learning*, pp. 8867–8887, 2022.

Jiahe Huang, Guandao Yang, Zichen Wang, and Jeong Joon Park. DiffusionPDE: Generative PDE-solving under partial observation. In *The Thirty-eighth Annual Conference on Neural Information Processing Systems*, 2024.

Alex Ling Yu Hung, Kai Zhao, Haoxin Zheng, Ran Yan, Steven S Raman, Demetri Terzopoulos, and Kyunghyun Sung. Med-CDiff: Conditional Medical Image Generation with Diffusion Models. *Bioengineering*, 10(11):1258, 2023.

Aapo Hyvärinen and Peter Dayan. Estimation of non-normalized statistical models by score matching. *Journal of Machine Learning Research*, 6(4), 2005.

Victor Isakov. *Inverse problems for Partial Differential Equations*, volume 127. Springer, 2006.

Ajil Jalal, Marius Arvinte, Giannis Daras, Eric Price, Alexandros G Dimakis, and Jon Tamir. Robust Compressed Sensing MRI with Deep Generative Priors. In *Advances in Neural Information Processing Systems*, volume 34, pp. 14938–14954. Curran Associates, Inc., 2021.

Rui Jiao, Wenbing Huang, Peijia Lin, Jiaqi Han, Pin Chen, Yutong Lu, and Yang Liu. Crystal structure prediction by joint equivariant diffusion. *Advances in Neural Information Processing Systems*, 36:17464–17497, 2023.

Richard Jordan, David Kinderlehrer, and Felix Otto. The variational formulation of the fokker–planck equation. *SIAM journal on mathematical analysis*, 29(1):1–17, 1998.

SI Kabanikhin. Definitions and examples of inverse and ill-posed problems. *Journal of Inverse and Ill-Posed Problems*, 16(4):317–357, 2008.

Zahra Kadkhodaie and Eero Simoncelli. Stochastic solutions for linear inverse problems using the prior implicit in a denoiser. *Advances in Neural Information Processing Systems*, 34:13242–13254, 2021.

Tero Karras, Samuli Laine, and Timo Aila. A Style-Based Generator Architecture for Generative Adversarial Networks. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 43(12):4217–4228, Dec 2021.

Ramneet Kaur, Susmit Jha, Anirban Roy, Sangdon Park, Edgar Dobriban, Oleg Sokolsky, and Insup Lee. idecode: In-distribution equivariance for conformal out-of-distribution detection. In *Proceedings of the AAAI conference on artificial intelligence*, volume 36, pp. 7104–7114, 2022.

Ramneet Kaur, Kaustubh Sridhar, Sangdon Park, Yahan Yang, Susmit Jha, Anirban Roy, Oleg Sokolsky, and Insup Lee. Codit: Conformal out-of-distribution detection in time-series data for cyber-physical systems. In *Proceedings of the ACM/IEEE 14th International Conference on Cyber-Physical Systems (with CPS-IoT Week 2023)*, pp. 120–131, 2023.

Bahjat Kawar, Michael Elad, Stefano Ermon, and Jiaming Song. Denoising diffusion restoration models. *Advances in Neural Information Processing Systems*, 35:23593–23606, 2022.

Amirhossein Kazerouni, Ehsan Khodapanah Aghdam, Moein Heidari, Reza Azad, Mohsen Fayyaz, Ilker Hacihaliloglu, and Dorit Merhof. Diffusion models in medical imaging: A comprehensive survey. *Medical image analysis*, 88:102846, 2023.

Jeongsol Kim, Geon Yeong Park, and Jong Chul Ye. Dreamsampler: Unifying diffusion sampling and score distillation for image manipulation. In *European Conference on Computer Vision*, pp. 398–414. Springer, 2024.

Diederik P Kingma. Auto-encoding Variational Bayes. *arXiv preprint arXiv:1312.6114*, 2013.

Nikola B. Kovachki, Zongyi Li, Burigede Liu, Kamyar Azizzadenesheli, Kaushik Bhattacharya, Andrew M. Stuart, and Anima Anandkumar. Neural operator: Learning maps between function spaces. *CoRR*, abs/2108.08481, 2021.

Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25, 2012.

Hannah Lawrence, Vasco Portilheiro, Yan Zhang, and Sékou-Oumar Kaba. Improving equivariant networks with probabilistic symmetry breaking. *International Conference on Learning Representations*, 2025.

Karel Lenc and Andrea Vedaldi. Understanding image representations by measuring their equivariance and equivalence. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 991–999, 2015.

Guangyuan Li, Chen Rao, Juncheng Mo, Zhanjie Zhang, Wei Xing, and Lei Zhao. Rethinking diffusion model for multi-contrast mri super-resolution. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 11365–11374, 2024.

Zijie Li, Anthony Zhou, and Amir Barati Farimani. Generative latent neural pde solver using flow matching. *arXiv preprint arXiv:2503.22600*, 2025.

Zongyi Li, Nikola Kovachki, Kamyar Azizzadenesheli, Burigede Liu, Kaushik Bhattacharya, Andrew Stuart, and Anima Anandkumar. Fourier neural operator for parametric partial differential equations. *arXiv preprint arXiv:2010.08895*, 2020.

Zongyi Li, Nikola Borislavov Kovachki, Kamyar Azizzadenesheli, Kaushik Bhattacharya, Andrew Stuart, Anima Anandkumar, et al. Fourier neural operator for parametric partial differential equations. In *International Conference on Learning Representations*, 2021.

Ziming Liu, Di Luo, Yilun Xu, Tommi Jaakkola, and Max Tegmark. Genphys: From physical processes to generative models. *arXiv preprint arXiv:2304.02637*, 2023.

Andreas Lugmayr, Martin Danelljan, Andres Romero, Fisher Yu, Radu Timofte, and Luc Van Gool. Repaint: Inpainting using denoising diffusion probabilistic models. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 11461–11471, 2022.

Yunqian Ma and Yun Fu. *Manifold learning theory and applications*, volume 434. CRC press Boca Raton, 2012.

Abbas Mammadov, Julius Berner, Kamyar Azizzadenesheli, Jong Chul Ye, and Anima Anandkumar. Diffusion-based inverse solver on function spaces with applications to pdes. *Machine Learning and the Physical Sciences Workshop at NeurIPS*, 2024a. URL https://ml4physicalsciences.github.io/2024/files/NeurIPS_ML4PS_2024_253.pdf.

Abbas Mammadov, Hyungjin Chung, and Jong Chul Ye. Amortized posterior sampling with diffusion prior distillation, 2024b. URL https://arxiv.org/abs/2407.17907.

Chris Metzler, Ali Mousavi, and Richard Baraniuk. Learned d-amp: Principled neural network based compressive image recovery. *Advances in neural information processing systems*, 30, 2017.

Christopher A Metzler, Arian Maleki, and Richard G Baraniuk. From denoising to compressed sensing. *IEEE Transactions on Information Theory*, 62(9):5117–5144, 2016.

Koichi Miyasawa et al. An empirical bayes estimator of the mean of a normal population. *Bull. Inst. Internat. Statist*, 38(181-188):1–2, 1961.

Jeet Mohapatra, Nima Dehmamy, Csaba Both, Subhro Das, and Tommi Jaakkola. Symmetry-driven discovery of dynamical variables in molecular simulations. In *Forty-second International Conference on Machine Learning*, 2025.

Eloi Moliner, Jaakko Lehtinen, and Vesa Välimäki. Solving audio inverse problems with a diffusion model. In *ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 1–5, 2023.

Hariharan Narayanan and Sanjoy Mitter. Sample complexity of testing the manifold hypothesis. *Advances in neural information processing systems*, 23, 2010.

Jung Yeon Park, Sujay Bhatt, Sihan Zeng, Lawson LS Wong, Alec Koppel, Sumitra Ganesh, and Robin Walters. Approximate equivariance in reinforcement learning. In *International Conference on Artificial Intelligence and Statistics*, pp. 4177–4185. PMLR, 2025.

William Peebles and Saining Xie. Scalable diffusion models with transformers. In *Proceedings of the IEEE/CVF international conference on computer vision*, pp. 4195–4205, 2023.

Xinyu Peng, Ziyang Zheng, Wenrui Dai, Nuoqian Xiao, Chenglin Li, Junni Zou, and Hongkai Xiong. Improving diffusion models for inverse problems using optimal posterior covariance. In *Forty-first International Conference on Machine Learning*, 2024.

Robin Quessard, Thomas Barrett, and William Clements. Learning disentangled representations and group structure of dynamical environments. *Advances in Neural Information Processing Systems*, 33:19727–19737, 2020.

Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, Gretchen Krueger, and Ilya Sutskever. Learning transferable visual models from natural language supervision. In *Proceedings of the 38th International Conference on Machine Learning (ICML)*, 2021.

Herbert Ellis Robbins. An empirical bayes approach to statistics. In *Proc. Third Berkley Symposium on Mathematcal Statistics*, pp. 157–163, 1956.

Yaniv Romano, Michael Elad, and Peyman Milanfar. The little engine that could: Regularization by denoising (red). *SIAM Journal on Imaging Sciences*, 10(4):1804–1844, 2017.

Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. High-resolution Image Synthesis with Latent Diffusion Models. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 10684–10695, 2022.

David W Romero and Suhas Lohit. Learning partial equivariances from data. *Advances in Neural Information Processing Systems*, 35:36466–36478, 2022.

Litu Rout, Negin Raoof, Giannis Daras, Constantine Caramanis, Alex Dimakis, and Sanjay Shakkottai. Solving Linear Inverse Problems Provably via Posterior Sampling with Latent Diffusion Models. In *Thirty-seventh Conference on Neural Information Processing Systems*, 2023. URL https://openreview.net/forum?id=XKBFdYwfRo.

Chitwan Saharia, Jonathan Ho, William Chan, Tim Salimans, David J Fleet, and Mohammad Norouzi. Image super-resolution via iterative refinement. *IEEE transactions on pattern analysis and machine intelligence*, 45(4):4713–4726, 2022.

Vıctor Garcia Satorras, Emiel Hoogeboom, and Max Welling. E (n) equivariant graph neural networks. In *International conference on machine learning*, pp. 9323–9332. PMLR, 2021.

Hang Shao, Abhishek Kumar, and P Thomas Fletcher. The riemannian geometry of deep generative models. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 315–323, 2018.

Aliaksandra Shysheya, Cristiana Diaconu, Federico Bergamin, Paris Perdikaris, José Miguel Hernández-Lobato, Richard Turner, and Emile Mathieu. On conditional diffusion models for pde simulations. *Advances in Neural Information Processing Systems*, 37:23246–23300, 2024.

Jascha Sohl-Dickstein, Eric Weiss, Niru Maheswaranathan, and Surya Ganguli. Deep unsupervised learning using nonequilibrium thermodynamics. In *International conference on machine learning*, pp. 2256–2265. pmlr, 2015.

Bowen Song, Soo Min Kwon, Zecheng Zhang, Xinyu Hu, Qing Qu, and Liyue Shen. Solving Inverse Problems with Latent Diffusion Models via Hard Data Consistency. In *Conference on Parsimony and Learning (Recent Spotlight Track)*, 2023a. URL https://openreview.net/forum?id=iHcarDCZLn.

Jiaming Song, Arash Vahdat, Morteza Mardani, and Jan Kautz. Pseudoinverse-guided diffusion models for inverse problems. In *International Conference on Learning Representations*, 2023b. URL https://openreview.net/forum?id=9_gsMA8MRKQ.

Yang Song and Stefano Ermon. Generative Modeling by Estimating Gradients of the Data Distribution. In *Advances in Neural Information Processing Systems*, volume 32, 2019.

Yang Song, Jascha Sohl-Dickstein, Diederik Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-based Generative Modeling through Stochastic Differential Equations. In *The International Conference on Learning Representations*, 2021. URL https://openreview.net/pdf/ef0eadbe07115b0853e964f17aa09d811cd490f1.pdf.

Yang Song, Liyue Shen, Lei Xing, and Stefano Ermon. Solving Inverse Problems in Medical Imaging with Score-Based Generative Models. In *International Conference on Learning Representations*, 2022.

Andrew M Stuart. Inverse problems: a bayesian perspective. *Acta numerica*, 19:451–559, 2010.

Julián Tachella, Dongdong Chen, and Mike Davies. Sensing theorems for unsupervised learning in linear inverse problems. *Journal of Machine Learning Research*, 24(39):1–45, 2023.

Matthieu Terris, Thomas Moreau, Nelly Pustelnik, and Julian Tachella. Equivariant plug-and-play image reconstruction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 25255–25264, 2024.

Nathaniel Thomas, Tess Smidt, Steven Kearnes, Lusann Yang, Li Li, Kai Kohlhoff, and Patrick Riley. Tensor field networks: Rotation-and translation-equivariant neural networks for 3d point clouds. *arXiv preprint arXiv:1802.08219*, 2018.

Phong Tran, Anh Tran, Quynh Phung, and Minh Hoai. Explore Image Deblurring via Encoded Blur Kernel Space. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, 2021.

Pascal Vincent. A Connection Between Score Matching and Denoising Autoencoders. *Neural Computation*, 23(7):1661–1674, 2011. doi: 10.1162/NECO_a_00142.

Dian Wang, Stephen Hart, David Surovik, Tarik Kelestemur, Haojie Huang, Haibo Zhao, Mark Yeatman, Jiuguang Wang, Robin Walters, and Robert Platt. Equivariant diffusion policy. In *8th Annual Conference on Robot Learning*, 2024. URL https://openreview.net/forum?id=wD2kUVLT1g.

Rui Wang, Robin Walters, and Rose Yu. Approximately equivariant networks for imperfectly symmetric dynamics. In *International Conference on Machine Learning*, pp. 23078–23091. PMLR, 2022.

Kilian Q Weinberger and Lawrence K Saul. Unsupervised learning of image manifolds by semidefinite programming. *International journal of computer vision*, 70:77–90, 2006.

Minkai Xu, Jiaqi Han, Aaron Lou, Jean Kossaifi, Arvind Ramanathan, Kamyar Azizzadenesheli, Jure Leskovec, Stefano Ermon, and Anima Anandkumar. Equivariant graph neural operator for modeling 3d dynamics. In *International Conference on Machine Learning*, pp. 55015–55032. PMLR, 2024.

Jiachen Yao, Abbas Mammadov, Julius Berner, Gavin Kerrigan, Jong Chul Ye, Kamyar Azizzadenesheli, and Anima Anandkumar. Guided diffusion sampling on function spaces with applications to pdes, 2025.

Bingliang Zhang, Wenda Chu, Julius Berner, Chenlin Meng, Anima Anandkumar, and Yang Song. Improving diffusion inverse problem solving with decoupled noise annealing. In *Proceedings of the Computer Vision and Pattern Recognition Conference*, pp. 20895–20905, 2025a.

Bingliang Zhang, Zihui Wu, Berthy T Feng, Yang Song, Yisong Yue, and Katherine L Bouman. Step: A general and scalable framework for solving video inverse problems with spatiotemporal diffusion priors. *preprint arXiv:2504.07549*, 2025b.

Kai Zhang, Wangmeng Zuo, Shuhang Gu, and Lei Zhang. Learning deep cnn denoiser prior for image restoration. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 3929–3938, 2017.

Hongkai Zheng, Wenda Chu, Austin Wang, Nikola Borislavov Kovachki, Ricardo Baptista, and Yisong Yue. Ensemble kalman diffusion guidance: A derivative-free method for inverse problems. *Transactions on Machine Learning Research*, 2025.

Allan Zhou, Tom Knowles, and Chelsea Finn. Meta-learning symmetries by reparameterization. In *International Conference on Learning Representations*, 2021.

Yibo Zhou. Rethinking reconstruction autoencoder-based out-of-distribution detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 7379–7387, 2022.

Yuanzhi Zhu, Kai Zhang, Jingyun Liang, Jiezhang Cao, Bihan Wen, Radu Timofte, and Luc Van Gool. Denoising diffusion models for plug-and-play image restoration. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 1219–1229, 2023.

Rayhan Zirvi, Bahareh Tolooshams, and Anima Anandkumar. Diffusion state-guided projected gradient for inverse problems. In *The Thirteenth International Conference on Learning Representations*, 2025. URL https://openreview.net/forum?id=kRBQwlkFSP.

Appendices for "EquiReg: Equivariance Regularized Diffusion for Inverse Problems"

We provide our source code when EquiReg. We will provide a publicly available source code upon acceptance. This supplementary materials contain the following:

- Section A includes additional experiments on text-to-image guidance. We regularize DreamSampler (Kim et al., 2024) with EquiReg for an improved performance (see Figures 7 to 11).
- Section B includes additional experiments on robustness including robustness to λ_t, reduced number of DDIM steps, and reduced number of measurent consistency steps.
- Section C includes qualitative analysis on the performance of methods with and without EquiReg. Results show a reduction of artifacts and an improved perceptual quality of the solution. This section also includes the equivariance error of a pre-trained encoder used in EquiReg (Figure 14a) and a histogram of Equi's improvement on DPS (Figure 13).
- Section D includes diversity experiments. Results show that EquiReg achieves favorable fidelity-diversity tradeoffs (Table 7, Figure 17, and Figure 18).
- Section E demonstrates EquiReg experimental setup and implementation for PSLD, ReSample, and DPS (Algorithms 2 to 6). It also contains information about the EquiReg hyperparameters for image restoration tasks.
- Section F contains information on the PDE reconstruction experiment. It discusses the equations along with implementation details and hyperparameters.
- Section G provides theoretical proofs of Propositions G.1 and G.2.
- Section H contains additional background information on solving inverse problems, vanishing-error autoencoders, and equivariance.
- Section I provides additional figures on MPE functions.
- Section J discloses computing resources used to conduct the experiments.
- Section K credits code assets used for our experiments.
- Section L concludes the appendix with a "responsible release" statement.

The authors acknowledge the usage of LLMs on proofreading and improving the coherency of the manuscript. The authors have not used LLMs for content generation.

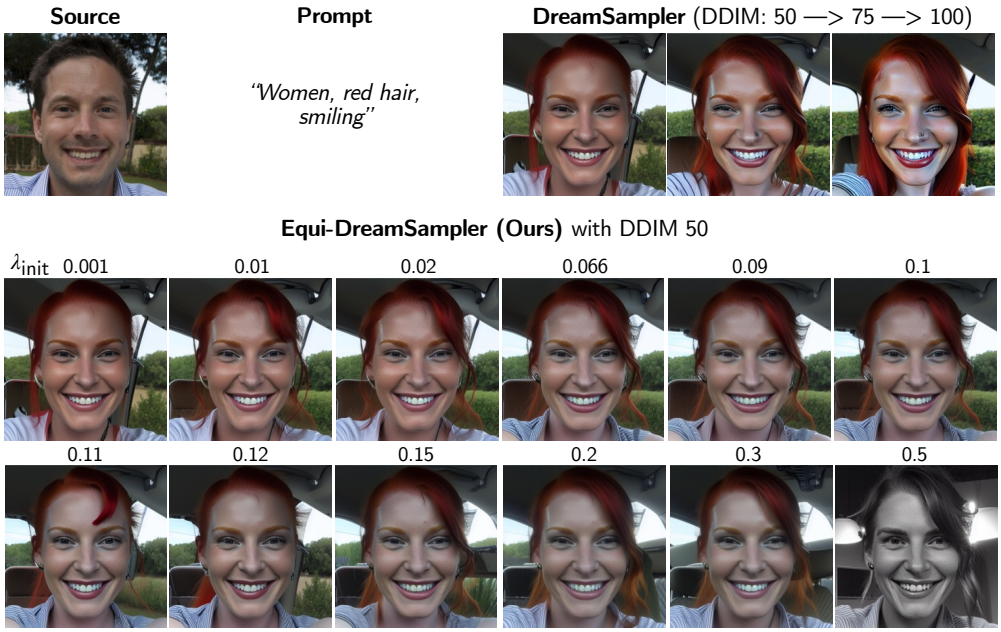


Figure 7: **Impact of EquiReg parameter λ_t, implicit acceleration, and introduction of more image details on FFHQ 512×512.** Women, red hair, smiling.

A EquiReg for Text-to-Image Guidance

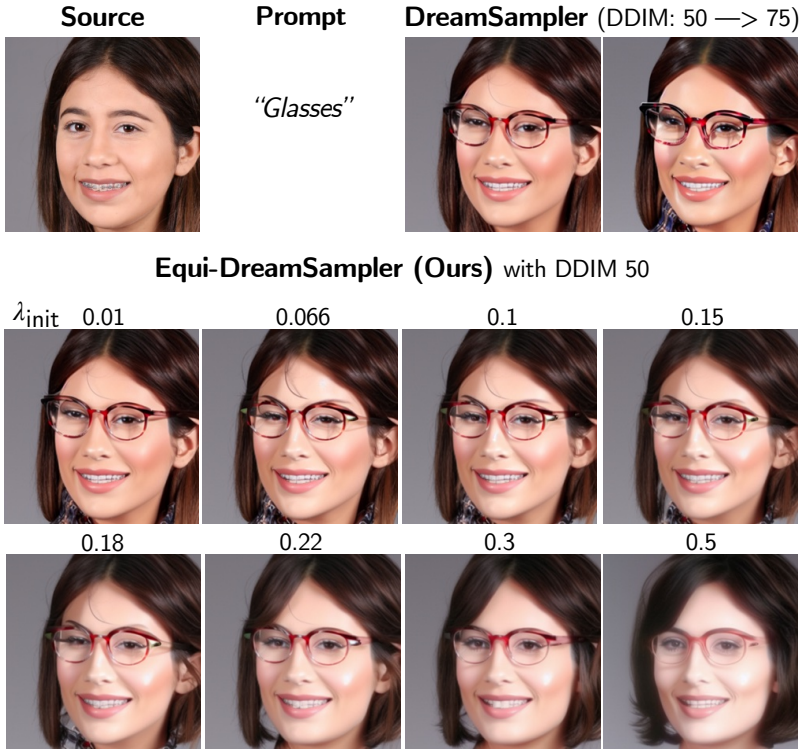


Figure 9: **Adding EquiReg into the text-to-image guidance method DreamSampler for improved performance.** Glasses.

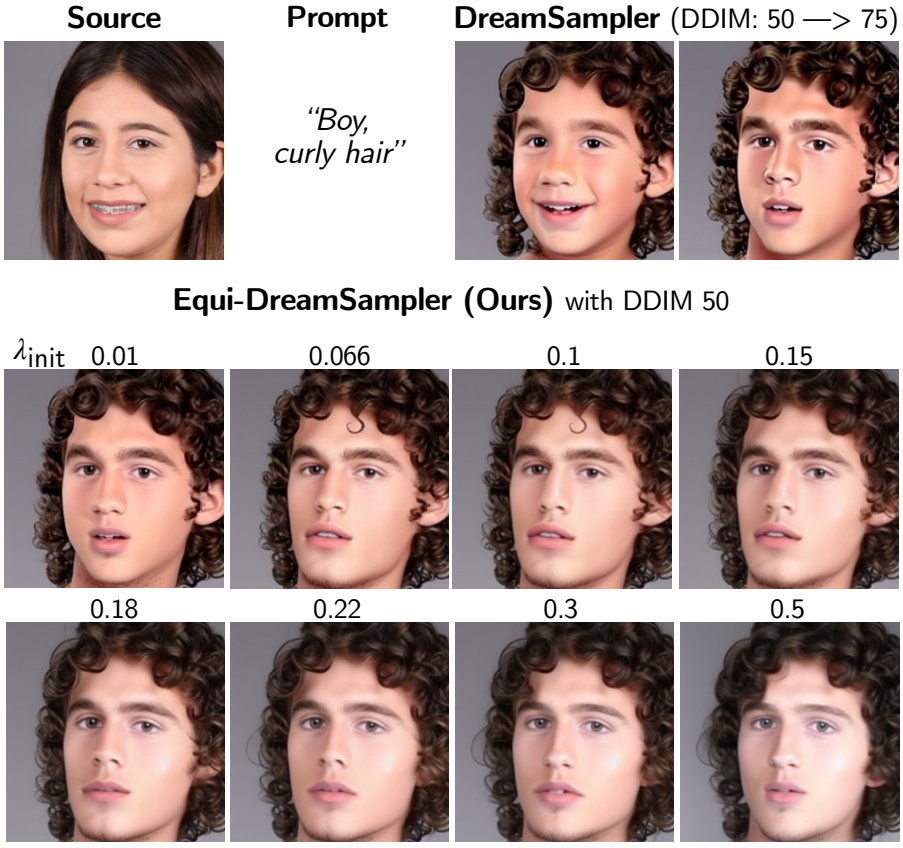


Figure 8: **Adding EquiReg into the text-to-image guidance method DreamSampler for improved performance on FFHQ 512×512.** Boy, curly hair.

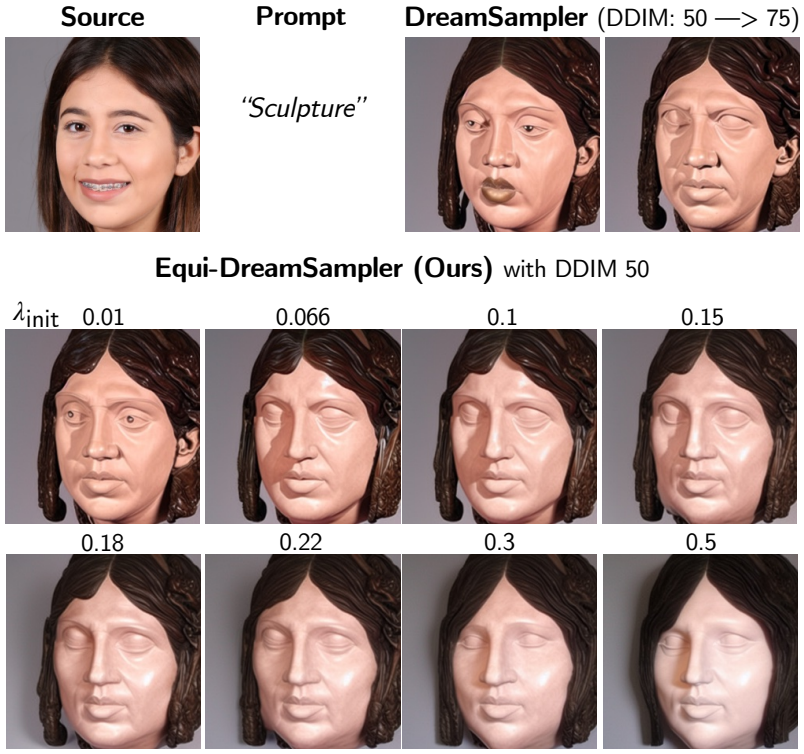


Figure 11: **Adding EquiReg into the text-to-image guidance method DreamSampler for improved performance.** Sculpture.

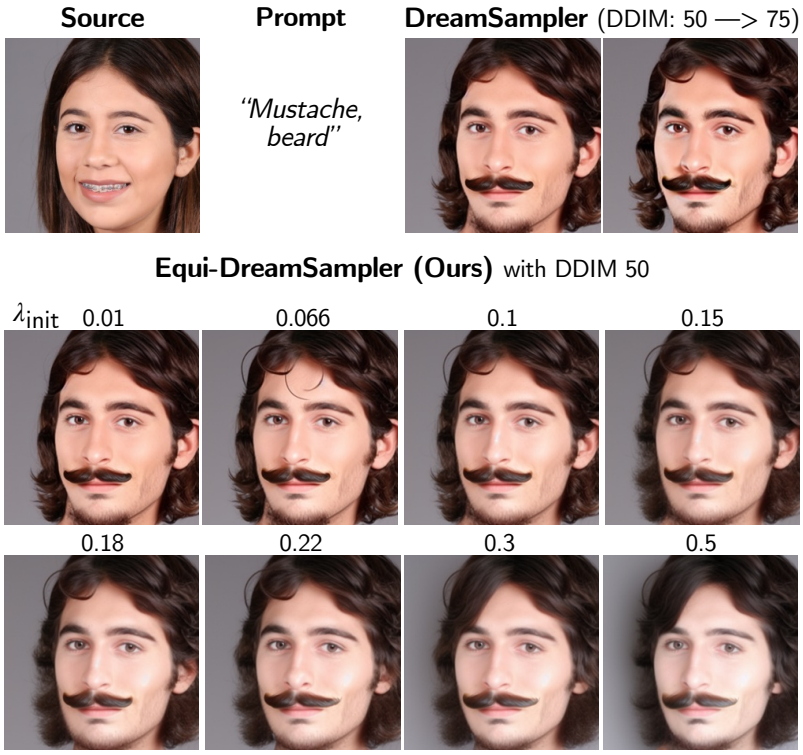


Figure 10: **Adding EquiReg into the text-to-image guidance method DreamSampler for improved performance.** Mustache, beard.

B Additional Experiments on Robustness

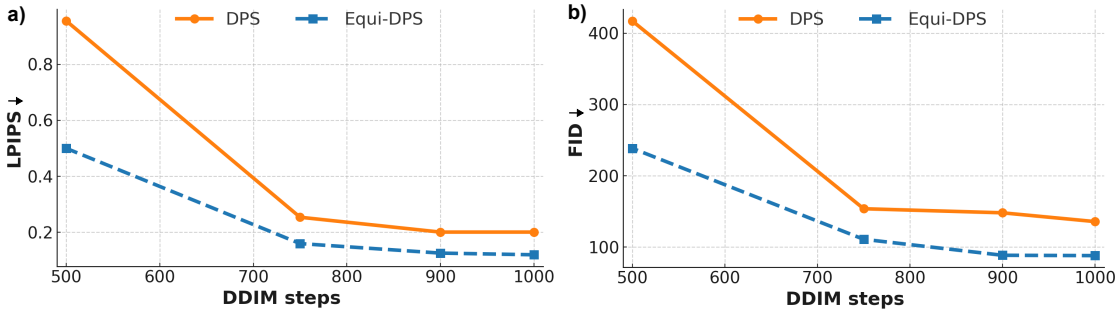


Figure 12: **Advantages of EquiReg under reduced DDIM steps.** Super-resolution on FFHQ.

Table 9: **EquiReg improves SITCOM under reduced measurement consistency steps (K_{meas}).** Motion deblur on FFHQ sampled with 50 DDIM steps.

$K_{\text{meas.}}$	K_{EquiReg}	PSNR↑	SSIM↑	Runtime (s)
10	N/A	28.06	0.81	21.57
10	1	28.71	0.82	21.07
5	5	**29.26**	**0.83**	**11.09**
20	N/A	27.04	0.79	38.85
20	1	28.54	**0.82**	37.74
10	10	**28.93**	**0.82**	**20.92**
30	N/A	27.79	0.80	58.84
30	1	28.35	0.81	55.51
15	15	**29.63**	**0.84**	**30.19**
40	N/A	30.40	**0.85**	78.08
40	1	**30.58**	**0.85**	69.83
20	20	29.50	0.83	**41.02**
60	N/A	28.35	0.81	108.57
60	1	27.02	0.78	95.62
30	30	**31.36**	**0.87**	**59.38**

Table 10: **EquiReg Effectiveness with Subset of Group Actions.**

PSLD		**Equi-PSLD** (90, 270 deg)	
PSNR↑	SSIM↑	PSNR↑	SSIM↑
15.86 (1.19)	0.77 (0.03)	17.60 (1.60)	0.79 (0.03)

C Visualizations for Image Restoration Experiments

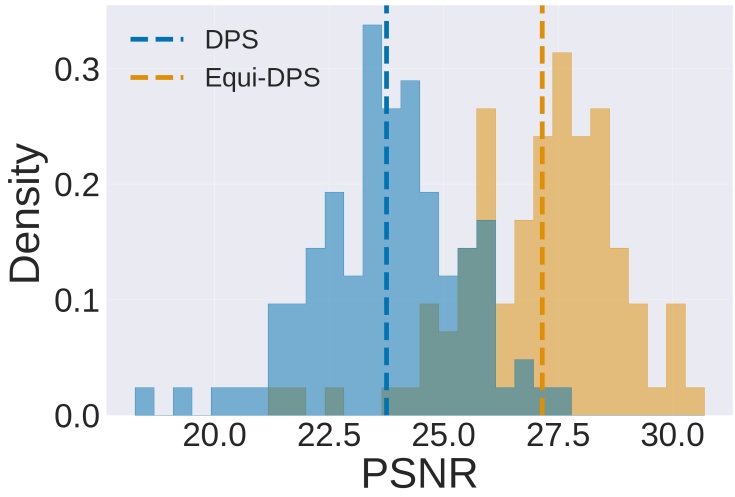
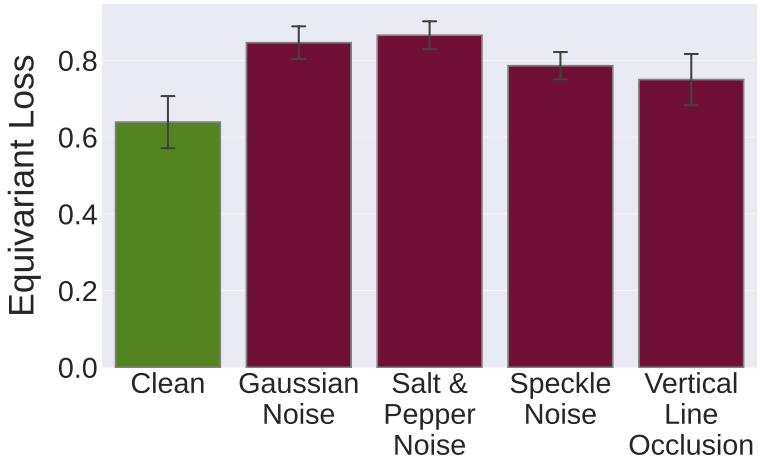
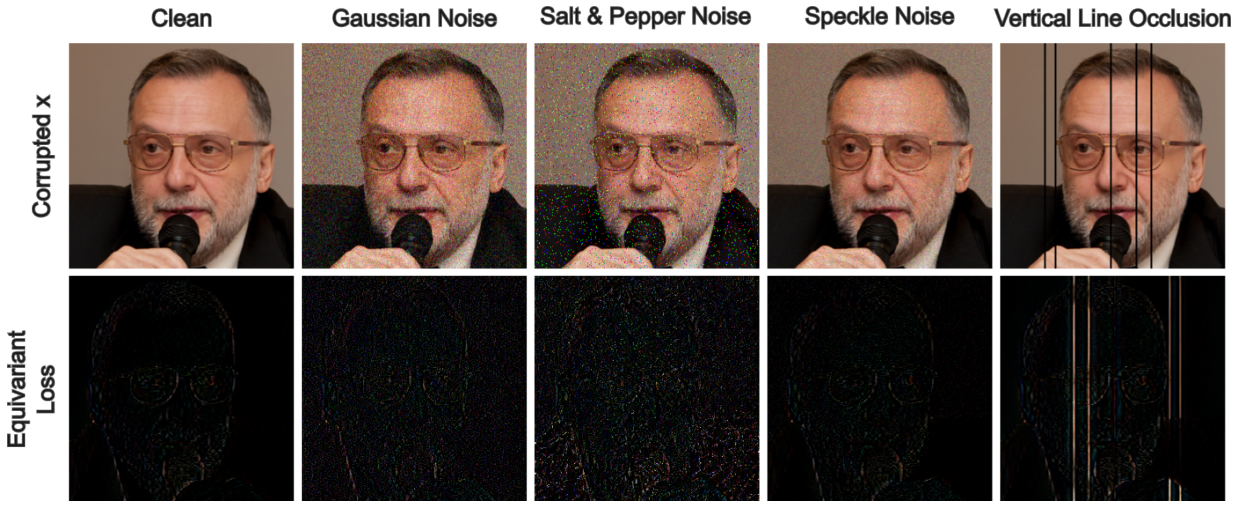


Figure 13: **Histogram of EquiReg improvement for DPS.** Super-resolution using FFHQ 256 × 256.

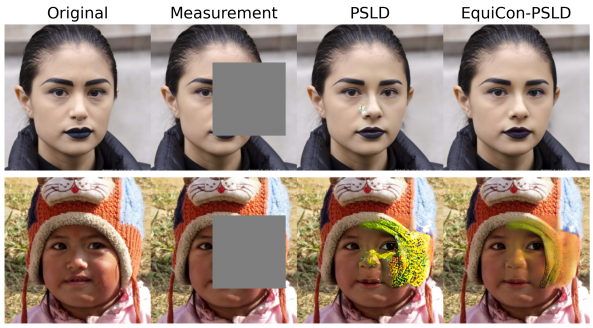


(a) The equivariance error of the encoder is lower on clean, natural images than corrupted ones.

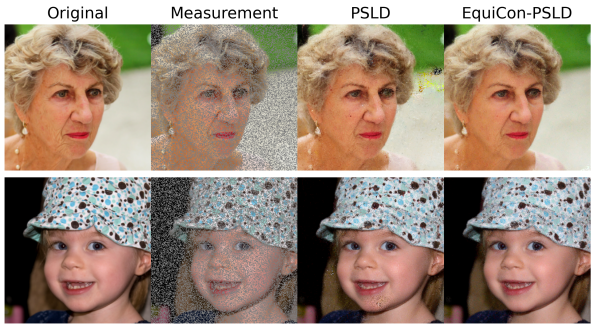


(b) Example visualizations of used images and corresponding equivariance error computed using the decoder (see Figure 4a).

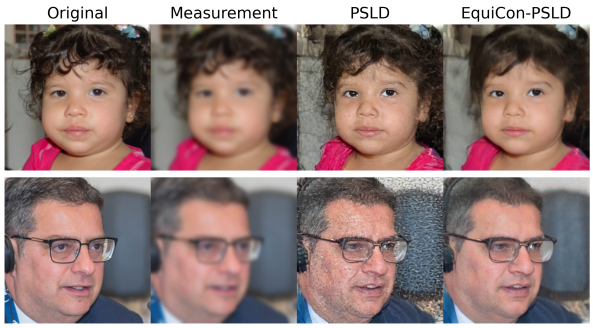
Figure 14: **Training induced equivariance for a pre-trained function.**



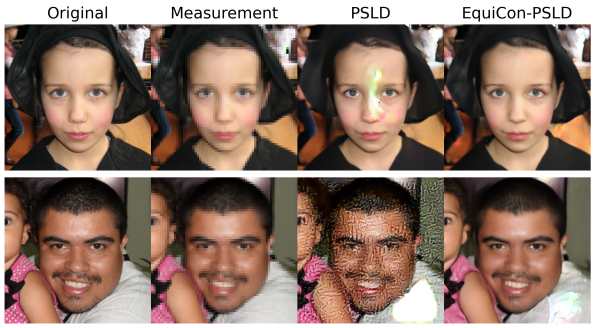
(a) Box inpainting.



(b) Random inpainting.

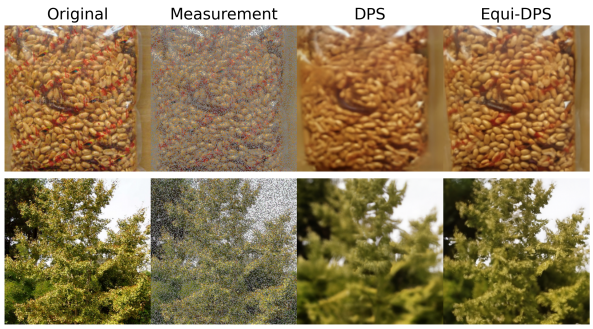


(c) Gaussian deblur.

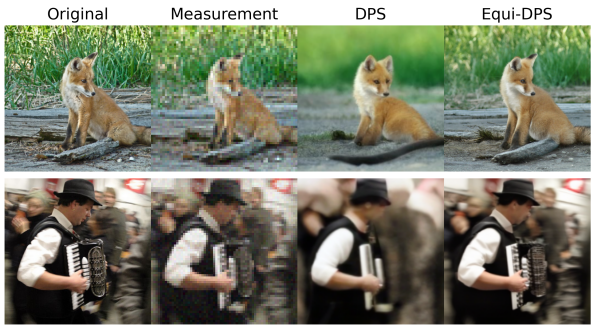


(d) Super-resolution (×4).

Figure 15: **Qualitative comparison of EquiCon-PSLD and PSLD on FFHQ 256 × 256.**



(a) Random inpainting.



(b) Super-resolution (×4).

Figure 16: **Qualitative comparison of Equi-DPS and DPS on ImageNet 256 × 256.**

D Diversity Analysis

In the Bayesian setting, the objective of solving inverse problems with diffusion models is to sample from high-probability regions of the posterior distribution. While the goal is not to maximize "diversity", the true diversity emerges when the posterior admits meaningful variability. In practice, diversity-related concerns in inverse problems arise when a method suffers from mode collapse, i.e., the sampler becomes biased and fails to explore multiple plausible modes of the posterior. Thus, the relevant question is whether a method properly explores the posterior rather than whether it maximizes diversity in an unconstrained sense.

Because closed-form posteriors are unavailable for real image restoration tasks, the standard practice in the diffusion inverse-problem literature is to evaluate diversity through variation among plausible reconstructions consistent with the measurement, without collapsing to a single solution. This is the notion of "diversity" our work adopts.

Given the goal of posterior sampling, EquiReg is not designed to maximize diversity for its own sake. Its objective is to incorporate data-inherent geometric structure (equivariance) to guide sampling toward high-probability regions of the posterior. Hence, diversity arises naturally from the ill-posedness of the inverse problem; it is a consequence of posterior uncertainty, not the goal of the regularizer.

To quantify this effect, in addition to reconstruction quality, we analyzed the diversity of posterior samples produced by EquiReg. We evaluate diversity metrics across multiple tasks and difficulty levels to characterize the sampling behavior of our method.

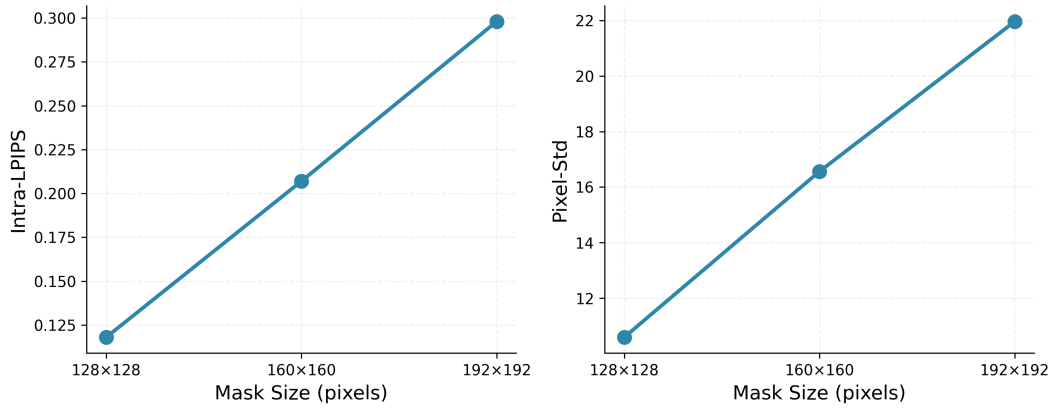
D.1 Experimental Setup

To evaluate diversity, we generate multiple posterior samples and measure variation across these samples. For each of 20 test images, we generate K=10 reconstructions using different random seeds. We evaluate diversity using two complementary metrics: Intra-LPIPS, which measures perceptual diversity by computing the average LPIPS distance between all pairs of samples, and Pixel-Std, which measures spatial diversity through pixel-wise standard deviation across samples. Higher values for both metrics indicate greater diversity. For Intra-LPIPS, we compute distances for all $\binom{K}{2} = 45$ pairs per image and average across all test images. For Pixel-Std, we compute the standard deviation at each pixel location across the K samples, then average across all pixels and test images. We evaluate diversity across three inverse problems (box inpainting, Gaussian deblurring, and 4× super-resolution) comparing EquiReg against DPS (Chung et al., 2023) without equivariance regularization. To investigate how diversity scales with task difficulty, we additionally vary the inpainting mask size from 128×128 (standard) to 160×160 to 192×192 pixels.

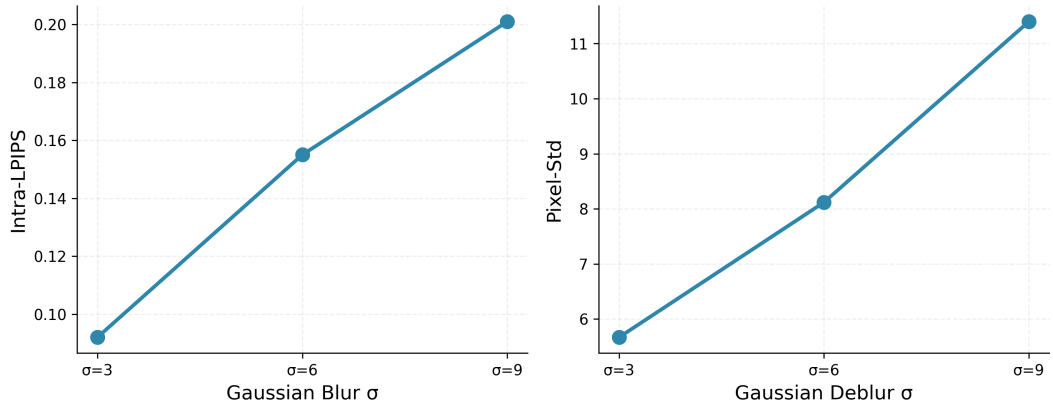
D.2 Results and Discussion

Table 7 shows that Equi-DPS achieves favorable fidelity-diversity trade-offs across three inverse problems. For box inpainting and super-resolution, equivariance regularization improves both fidelity and diversity simultaneously. For Gaussian deblurring, Equi-DPS achieves 15-20% better fidelity while retaining 80-85% of baseline diversity, representing a modest but justified trade-off. These results demonstrate that equivariance constraints do not inherently suppress diversity; rather, they can guide sampling toward regions of higher data fidelity while maintaining posterior exploration.

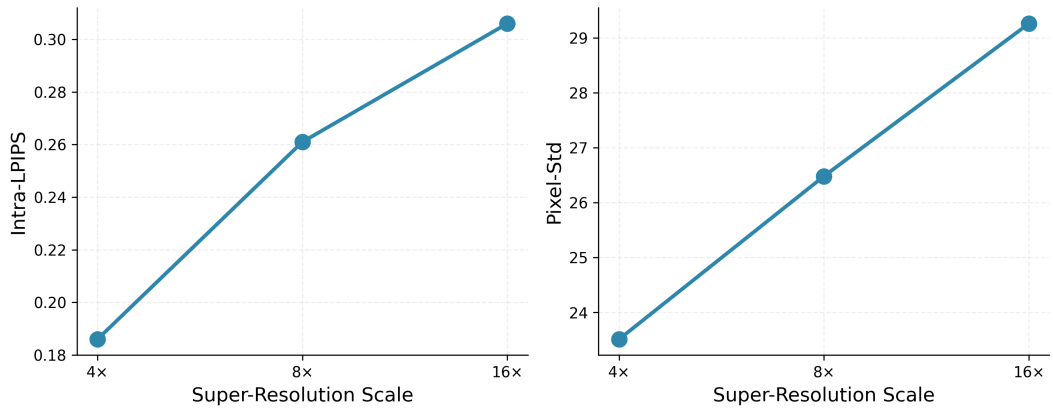
Figure 17 reveals linear diversity scaling with task difficulty. Diversity metrics grow proportionally with task difficulty, indicating Equi-DPS naturally expands sampling as problems become more ill-posed. This linear relationship demonstrates stable, predictable behavior across difficulty levels without artificial diversity suppression. Figures 6 and 18 provide qualitative results.



(a) Box inpainting



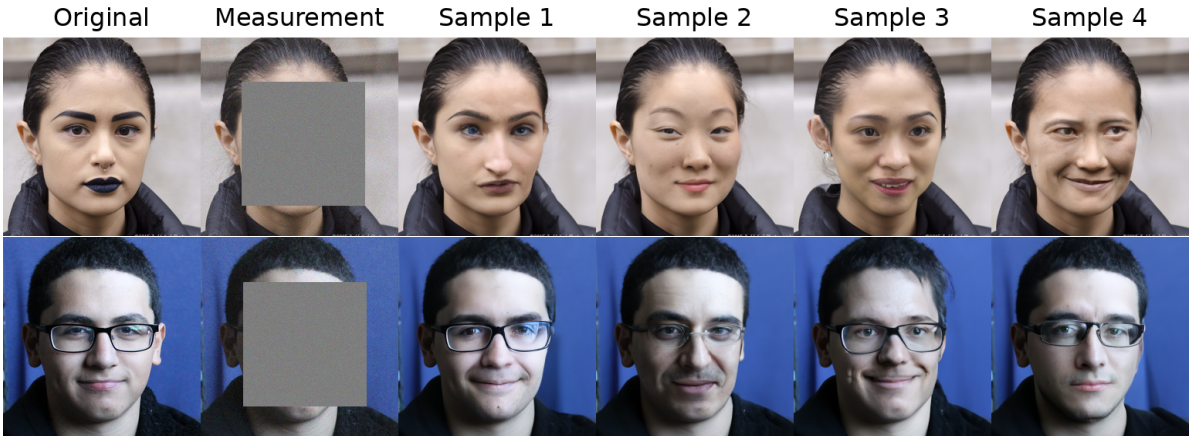
(b) Gaussian deblur



(c) Super-resolution

Figure 17: **Diversity vs task difficulty across three inverse problems.** As task difficulty increases (larger inpainting mask, stronger blur, higher SR scale), both diversity metrics increase proportionally, demonstrating that Equi-DPS maintains healthy posterior sampling behavior across a wide difficulty spectrum.

(a) Gaussian deblur



(b) Box inpainting

Figure 18: **Qualitative diversity examples across Gaussian deblur and box inpainting.** Each subfigure shows $K = 4$ posterior samples for two different test images. (a) Gaussian deblur: samples differ in facial expressions and accessories (i.e., earrings in first test image). (b) Box inpainting: Inputs were obstructed with 160×160 masks. Samples exhibit perceptually distinct facial features (i.e., expressions, eye gaze, facial structure). Across both tasks, EquiReg produces diverse plausible reconstructions rather than collapsing to a single mode.

D.3 Conclusion

Finally, we highlight that EquiReg improves both fidelity and diversity on 2 of the 3 considered tasks, an encouraging outcome that is uncommon given the general behavior of classical regularizers. Hand-crafted regularizers such as TV and ℓ_1 may suppress diversity by shrinking solutions toward simple structures. By contrast, EquiReg leverages data-dependent regularization that captures the richness and structural complexity of the underlying data manifold, enabling it to preserve manifold-consistent variability while suppressing implausible samples.

High diversity without fidelity is not meaningful for posterior sampling. A method that samples the entire solution space, including low-probability and artifacted regions, may score well on diversity but fail to provide useful reconstructions. Equi-DPS avoids this failure mode: it maintains meaningful diversity while reducing artifacts and improving perceptual quality. In the experiments conducted during the rebuttal, our goal was to demonstrate clearly that EquiReg preserves meaningful diversity, reflecting the posterior uncertainty, rather than unstructured or unconstrained variability.

E Implementation Details for Image Restoration Tasks

Experimental Setup. We evaluate EquiReg on a variety of linear and nonlinear restoration tasks for natural images. We fix sets of 100 images from FFHQ and ImageNet as our validation sets. All images are normalized from $[0, 1]$. For the majority of experiments, we use noise level $\sigma_{\boldsymbol{y}} = 0.05$ (we indicate $\sigma_{\boldsymbol{y}}$ in our tables). For linear inverse problems, we consider (1) box inpainting, (2) random inpainting, (3) Gaussian deblur, (4) motion deblur, and (5) super-resolution. We apply a random 128×128 pixel box for box inpainting, and a 70% random mask for random inpainting. For Gaussian and motion deblur, we use kernels of size 61×61, with standard deviations of 3.0 and 0.5, respectively. For super-resolution, we downscale images by a factor of 4 using a bicubic resizer. For nonlinear inverse problems, we consider (1) phase retrieval, (2) nonlinear deblur, and (3) high dynamic range (HDR). We use an oversampling rate of 2.0 for phase retrieval, and due to instability of the task, we generate four independent reconstructions and take the best result (as also done in DPS (Chung et al., 2023), DAPS (Zhang et al., 2025a), and DiffStateGrad (Zirvi et al., 2025)). We use the default setting from (Tran et al., 2021) for nonlinear deblur, and a scale factor of 2 for HDR.

Hyperparameters. Our method introduces a single hyperparameter λ_t that controls the amount of regularization applied. Below we include a table detailing the use of this hyperparameter in the main experiments (Table 11). For majority of experiments, we keep λ_t constant throughout iterations. For all unscaled experiments, we employ early stopping, setting $\lambda_t = 0$ for the last 10% of sampling.

Table 11: Equivariance regularization weight λ_t used in main experiments.

Method	Box Inpainting	Random Inpainting	Gaussian Deblur	Motion Deblur	Super-resolution ($\times$4)
FFHQ 256×256					
Equi-PSLD	0.05	0.05	0.03	0.03	0.02
EquiCon-PSLD	0.01	0.01	0.01	0.01	0.01
Equi-ReSample	0.03	0.05	0.02	0.02	0.05
EquiCon-ReSample	0.001	0.001	0.001	0.001	0.001
Equi-DPS	0.0001	0.001	0.001	0.001	0.1
ImageNet 256×256					
EquiCon-PSLD	0.0015	0.05	0.06	0.07	0.001

Algorithm 2 Equi-PSLD for Image Restoration Tasks

Require: $T, \boldsymbol{y}, \{\eta_t\}_{t=1}^T, \{\gamma_t\}_{t=1}^T, \{\tilde{\sigma}_t\}_{t=1}^T$
Require: $\mathcal{E}, \mathcal{D}, \mathcal{A}\boldsymbol{x}_0^*, \mathcal{A}, \boldsymbol{s}_\theta$, T_g and S_g, $\{\lambda_t\}_{t=1}^T$
$\boldsymbol{z}_T \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I})$
for $t = T - 1$ **to** 0 **do**
  $\hat{\boldsymbol{s}} \leftarrow \boldsymbol{s}_\theta(\boldsymbol{z}_t, t)$
  $\boldsymbol{z}_{0|t} \leftarrow \frac{1}{\sqrt{\bar{\alpha}_t}}(\boldsymbol{z}_t + (1 - \bar{\alpha}_t)\hat{\boldsymbol{s}})$
  $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I})$
  $\boldsymbol{z}'_{t-1} \leftarrow \frac{\sqrt{\alpha_t}(1-\bar{\alpha}_{t-1})}{1-\bar{\alpha}_t}\boldsymbol{z}_t + \frac{\sqrt{\bar{\alpha}_{t-1}}\beta_t}{1-\bar{\alpha}_t}\boldsymbol{z}_{0|t} + \tilde{\sigma}_t\boldsymbol{\epsilon}$
  $\boldsymbol{z}''_{t-1} \leftarrow \boldsymbol{z}'_{t-1} - \eta_t \nabla_{\boldsymbol{z}_t} \|\boldsymbol{y} - \mathcal{A}(\mathcal{D}(\boldsymbol{z}_{0|t}))\|_2^2$
  $\boldsymbol{z}_{t-1} \leftarrow \boldsymbol{z}''_{t-1} - \gamma_t \nabla_{\boldsymbol{z}_t} \|\boldsymbol{z}_{0|t} - \mathcal{E}(\mathcal{A}^T\mathcal{A}\boldsymbol{x}_0^* + (\boldsymbol{I} - \mathcal{A}^T\mathcal{A})\mathcal{D}(\boldsymbol{z}_{0|t}))\|_2^2$
  $\boldsymbol{z}_{t-1} \leftarrow \boldsymbol{z}_{t-1} - \lambda_t \nabla_{\boldsymbol{z}_t} \|S_g(\mathcal{D}(\boldsymbol{z}_{0|t})) - \mathcal{D}(T_g(\boldsymbol{z}_{0|t}))\|_2^2$
end for
return $\mathcal{D}(\boldsymbol{z}_{0|t})$

PSLD. We integrate EquiReg into PSLD by simply adding an additional gradient update step using our regularization term (Algorithms 2 and 3).

In our experiments, we use the official PSLD implementation from Rout et al. (2023), running with its default settings to reproduce the baseline results. We note that in our code, we do not square the norm when computing the gradient, aligning with PSLD's implementation.

Algorithm 3 EquiCon-PSLD for Image Restoration Tasks

Require: $T, \boldsymbol{y}, \{\eta_t\}_{t=1}^T, \{\gamma_t\}_{t=1}^T, \{\tilde{\sigma}_t\}_{t=1}^T$
Require: $\mathcal{E}, \mathcal{D}, \mathcal{A}\boldsymbol{x}_0^*, \mathcal{A}, \boldsymbol{s}_\theta$, T_g and S_g, $\{\lambda_t\}_{t=1}^T$
$\boldsymbol{z}_T \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I})$
for $t = T-1$ **to** 0 **do**
  $\hat{\boldsymbol{s}} \leftarrow \boldsymbol{s}_\theta(\boldsymbol{z}_t, t)$
  $\boldsymbol{z}_{0|t} \leftarrow \frac{1}{\sqrt{\bar{\alpha}_t}}(\boldsymbol{z}_t + (1-\bar{\alpha}_t)\hat{\boldsymbol{s}})$
  $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I})$
  $\boldsymbol{z}'_{t-1} \leftarrow \frac{\sqrt{\alpha_t}(1-\bar{\alpha}_{t-1})}{1-\bar{\alpha}_t}\boldsymbol{z}_t + \frac{\sqrt{\bar{\alpha}_{t-1}}\beta_t}{1-\bar{\alpha}_t}\boldsymbol{z}_{0|t} + \tilde{\sigma}_t\boldsymbol{\epsilon}$
  $\boldsymbol{z}''_{t-1} \leftarrow \boldsymbol{z}'_{t-1} - \eta_t \nabla_{\boldsymbol{z}_t} \|\boldsymbol{y} - \mathcal{A}(\mathcal{D}(\boldsymbol{z}_{0|t}))\|_2^2$
  $\boldsymbol{z}_{t-1} \leftarrow \boldsymbol{z}''_{t-1} - \gamma_t \nabla_{\boldsymbol{z}_t} \|\boldsymbol{z}_{0|t} - \mathcal{E}(\mathcal{A}^T\mathcal{A}\boldsymbol{x}_0^* + (\boldsymbol{I} - \mathcal{A}^T\mathcal{A})\mathcal{D}(\boldsymbol{z}_{0|t}))\|_2^2$
  $\boldsymbol{z}_{t-1} \leftarrow \boldsymbol{z}_{t-1} - \lambda_t \nabla_{\boldsymbol{z}_t} \|\boldsymbol{z}_{0|t} - \mathcal{E}(S_g^{-1}(\mathcal{D}(T_g(\boldsymbol{z}_{0|t}))))\|_2^2$
end for
return $\mathcal{D}(\boldsymbol{z}_{0|t})$

ReSample. We integrate EquiReg into ReSample by adding our regularization term into the hard data consistency step (Algorithms 4 and 5). We note that the ReSample algorithm employs a two-stage approach; initially, it performs pixel-space optimization, and later it performs latent-space optimization. We apply EquiReg in the latent-space optimization stage.

In our experiments, we use the official ReSample implementation from Song et al. (2023a), running with its default settings to reproduce the baseline results.

Algorithm 4 Equi-ReSample for Image Restoration Tasks

Require: Measurements $\boldsymbol{y}$, $\mathcal{A}(\cdot)$, Encoder $\mathcal{E}(\cdot)$, Decoder $\mathcal{D}(\cdot)$, Score function $\boldsymbol{s}_\theta(\cdot, t)$, Pretrained LDM Parameters β_t, $\bar{\alpha}_t$, η, δ, Hyperparameter γ to control σ_t^2, Time steps to perform resample C, T_g and S_g, $\{\lambda_t\}_{t=1}^T$
$\boldsymbol{z}_T \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I})$ ▷ Initial noise vector
for $t = T-1, \ldots, 0$ **do**
  $\boldsymbol{\epsilon}_1 \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I})$
  $\hat{\boldsymbol{\epsilon}}_{t+1} = \boldsymbol{s}_\theta(\boldsymbol{z}_{t+1}, t+1)$ ▷ Compute the score
  $\hat{\boldsymbol{z}}_0(\boldsymbol{z}_{t+1}) = \frac{1}{\sqrt{\bar{\alpha}_{t+1}}}(\boldsymbol{z}_{t+1} - \sqrt{1-\bar{\alpha}_{t+1}}\hat{\boldsymbol{\epsilon}}_{t+1})$ ▷ Predict $\hat{\boldsymbol{z}}_0$ using Tweedie's formula
  $\boldsymbol{z}'_t = \sqrt{\bar{\alpha}_t}\hat{\boldsymbol{z}}_0(\boldsymbol{z}_{t+1}) + \sqrt{1-\bar{\alpha}_t - \eta\delta^2}\hat{\boldsymbol{\epsilon}}_{t+1} + \eta\delta\boldsymbol{\epsilon}_1$ ▷ Unconditional DDIM step
  if $t \in C$ **then** ▷ ReSample time step
    Initialize $\hat{\boldsymbol{z}}_0(\boldsymbol{y})$ with $\hat{\boldsymbol{z}}_0(\boldsymbol{z}_{t+1})$
    for each step in gradient descent **do**
      $\boldsymbol{g} \leftarrow \nabla_{\hat{\boldsymbol{z}}_0(\boldsymbol{y})} \frac{1}{2}\|\boldsymbol{y} - \mathcal{A}(\mathcal{D}(\hat{\boldsymbol{z}}_0(\boldsymbol{y})))\|_2^2 + \lambda_t \nabla_{\hat{\boldsymbol{z}}_0(\boldsymbol{y})} \|S_g(\mathcal{D}(\hat{\boldsymbol{z}}_0(\boldsymbol{y}))) - \mathcal{D}(T_g(\hat{\boldsymbol{z}}_0(\boldsymbol{y})))\|_2^2$
      Update $\hat{\boldsymbol{z}}_0(\boldsymbol{y})$ using gradient $\boldsymbol{g}$
    end for
    $\boldsymbol{z}_t = \text{StochasticResample}(\hat{\boldsymbol{z}}_0(\boldsymbol{y}), \boldsymbol{z}'_t, \gamma)$ ▷ Map back to t
  else
    $\boldsymbol{z}_t = \boldsymbol{z}'_t$ ▷ Unconditional sampling if not resampling
  end if
end for
$\boldsymbol{x}_0 = \mathcal{D}(\boldsymbol{z}_0)$ ▷ Output reconstructed image
return $\boldsymbol{x}_0$

Algorithm 5 EquiCon-ReSample for Image Restoration Tasks

Require: Measurements $\boldsymbol{y}$, $\mathcal{A}(\cdot)$, Encoder $\mathcal{E}(\cdot)$, Decoder $\mathcal{D}(\cdot)$, Score function $\boldsymbol{s}_\theta(\cdot, t)$, Pretrained LDM Parameters β_t, $\bar{\alpha}_t$, η, δ, Hyperparameter γ to control σ_t^2, Time steps to perform resample C, T_g and S_g, $\{\lambda_t\}_{t=1}^T$

$\boldsymbol{z}_T \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I})$ ▷ Initial noise vector
for $t = T-1, \ldots, 0$ **do**
  $\boldsymbol{\epsilon}_1 \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I})$
  $\hat{\boldsymbol{\epsilon}}_{t+1} = \boldsymbol{s}_\theta(\boldsymbol{z}_{t+1}, t+1)$ ▷ Compute the score
  $\hat{\boldsymbol{z}}_0(\boldsymbol{z}_{t+1}) = \frac{1}{\sqrt{\bar{\alpha}_{t+1}}}(\boldsymbol{z}_{t+1} - \sqrt{1-\bar{\alpha}_{t+1}}\hat{\boldsymbol{\epsilon}}_{t+1})$ ▷ Predict $\hat{\boldsymbol{z}}_0$ using Tweedie's formula
  $\boldsymbol{z}'_t = \sqrt{\bar{\alpha}_t}\hat{\boldsymbol{z}}_0(\boldsymbol{z}_{t+1}) + \sqrt{1-\bar{\alpha}_t-\eta\delta^2}\hat{\boldsymbol{\epsilon}}_{t+1} + \eta\delta\boldsymbol{\epsilon}_1$ ▷ Unconditional DDIM step
  if $t \in C$ **then** ▷ ReSample time step
    Initialize $\hat{\boldsymbol{z}}_0(\boldsymbol{y})$ with $\hat{\boldsymbol{z}}_0(\boldsymbol{z}_{t+1})$
    for each step in gradient descent **do**
      $\boldsymbol{g} \leftarrow \nabla_{\hat{\boldsymbol{z}}_0(\boldsymbol{y})}\frac{1}{2}\|\boldsymbol{y} - \mathcal{A}(\mathcal{D}(\hat{\boldsymbol{z}}_0(\boldsymbol{y})))\|_2^2 + \lambda_t\nabla_{\hat{\boldsymbol{z}}_0(\boldsymbol{y})}\|\hat{\boldsymbol{z}}_0(\boldsymbol{y}) - \mathcal{E}(S_g^{-1}(\mathcal{D}(T_g(\hat{\boldsymbol{z}}_0(\boldsymbol{y})))))\|_2^2$
      Update $\hat{\boldsymbol{z}}_0(\boldsymbol{y})$ using gradient $\boldsymbol{g}$
    end for
    $\boldsymbol{z}_t = \text{StochasticResample}(\hat{\boldsymbol{z}}_0(\boldsymbol{y}), \boldsymbol{z}'_t, \gamma)$ ▷ Map back to t
  else
    $\boldsymbol{z}_t = \boldsymbol{z}'_t$ ▷ Unconditional sampling if not resampling
  end if
end for
$\boldsymbol{x}_0 = \mathcal{D}(\boldsymbol{z}_0)$ ▷ Output reconstructed image
return $\boldsymbol{x}_0$

DPS. Similar to PSLD, we integrate EquiReg into DPS by simply adding an additional gradient update step using our regularization term (Algorithm 6).

In our experiments, we use the official DPS implementation from Chung et al. (2023), running with its default settings to reproduce the baseline results.

Algorithm 6 Equi-DPS for Image Restoration Tasks

Require: $T, \boldsymbol{y}, \{\zeta_t\}_{t=1}^T, \{\tilde{\sigma}_t\}_{t=1}^T, \boldsymbol{s}_\theta$, $\mathcal{E}$, T_g and S_g, $\{\lambda_t\}_{t=1}^T$

$\boldsymbol{x}_T \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I})$
for $t = T-1$ **to** 0 **do**
  $\hat{\boldsymbol{s}} \leftarrow \boldsymbol{s}_\theta(\boldsymbol{x}_t, t)$
  $\boldsymbol{x}_{0|t} \leftarrow \frac{1}{\sqrt{\bar{\alpha}_t}}(\boldsymbol{x}_t + (1-\bar{\alpha}_t)\hat{\boldsymbol{s}})$
  $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I})$
  $\boldsymbol{x}'_{t-1} \leftarrow \frac{\sqrt{\alpha_t}(1-\bar{\alpha}_{t-1})}{1-\bar{\alpha}_t}\boldsymbol{x}_t + \frac{\sqrt{\bar{\alpha}_{t-1}}\beta_t}{1-\bar{\alpha}_t}\boldsymbol{x}_{0|t} + \tilde{\sigma}_t\boldsymbol{\epsilon}$
  $\boldsymbol{x}_{t-1} \leftarrow \boldsymbol{x}'_{t-1} - \zeta_t\nabla_{\boldsymbol{x}_t}\|\boldsymbol{y} - \mathcal{A}(\boldsymbol{x}_{0|t})\|_2^2$
  $\boldsymbol{x}_{t-1} \leftarrow \boldsymbol{x}_{t-1} - \lambda_t\nabla_{\boldsymbol{x}_t}\|S_g(\mathcal{E}(\boldsymbol{x}_{0|t})) - \mathcal{E}(T_g(\boldsymbol{x}_{0|t}))\|_2^2$
end for
return $\boldsymbol{x}_0$

SITCOM. We augment the original SITCOM algorithm by introducing an additional equivariant refinement stage at each reverse diffusion step. After completing the standard measurement and backward-consistency gradient updates, we perform a second optimization over the equivariance loss, enforcing consistency between $\mathcal{E}(T_g(v))$ and $T_g(\mathcal{E}(v))$ (Algorithm 7).

In our experiments, we use the official SITCOM implementation from Alkhouri et al. (2025), running with its default settings to reproduce the baseline results.

Algorithm 7 Equi-SITCOM for Image Restoration Tasks

Require: Measurements $\mathbf{y}$, forward operator $\mathcal{A}(\cdot)$, pre-trained DM $\epsilon_\theta(\cdot,\cdot)$, diffusion steps N, schedule $\bar{\alpha}_i$, measurement gradient steps K, equivariant gradient steps K_{equi}, stop δ, lr γ, reg. λ.
Ensure: Restored image $\hat{\mathbf{x}}$.

Initialize $\mathbf{x}_N \sim \mathcal{N}(\mathbf{0},\mathbf{I})$, $\Delta t = \lfloor \frac{T}{N} \rfloor$.
for $i = N, N-1, \ldots, 1$ **do** ▷ Reducing diffusion sampling steps
  $\mathbf{v}_i^{(0)} \leftarrow \mathbf{x}_i$ ▷ Init for closeness (C3)
  for $k = 1, \ldots, K$ **do** ▷ Adam on measurement/backward consistency (C1, C2)
    $\mathbf{v}_i^{(k)} \leftarrow \mathbf{v}_i^{(k-1)} - \gamma \nabla_{\mathbf{v}_i} \left[\left\| \mathcal{A}\left(\frac{1}{\sqrt{\bar{\alpha}_i}} \left(\mathbf{v}_i - \sqrt{1-\bar{\alpha}_i}\, \epsilon_\theta(\mathbf{v}_i, i\Delta t) \right) \right) - \mathbf{y} \right\|_2^2 + \lambda \|\mathbf{x}_i - \mathbf{v}_i\|_2^2 \right] \Big|_{\mathbf{v}_i = \mathbf{v}_i^{(k-1)}}$
    if $\left\| \mathcal{A}\left(\frac{1}{\sqrt{\bar{\alpha}_i}} \left(\mathbf{v}_i^{(k)} - \sqrt{1-\bar{\alpha}_i}\, \epsilon_\theta(\mathbf{v}_i^{(k)}, i\Delta t) \right) \right) - \mathbf{y} \right\|_2^2 < \delta^2$ **then**
      break ▷ Prevent noise overfitting
    end if
  end for
  $\mathbf{v}_i^{(0)} \leftarrow \mathbf{v}_i^{(k)}$ ▷ Initialize to optimized $\mathbf{v}_i$
  for $k = 1, \ldots, K_{\text{equi}}$ **do**
    $\mathbf{v}_i^{(k)} \leftarrow \mathbf{v}_i^{(k-1)} - \gamma \nabla_{\mathbf{v}_i} \left[\left\| \mathcal{E}\big(T_g(\mathbf{v}_i^{(k)})\big) - T_g\big(\mathcal{E}(\mathbf{v}_i^{(k)})\big) \right\|_2^2 \right] \Big|_{\mathbf{v}_i = \mathbf{v}_i^{(k-1)}}$
    if $\left\| \mathcal{E}\big(T_g(\mathbf{v}_i^{(k)})\big) - T_g\big(\mathcal{E}(\mathbf{v}_i^{(k)})\big) \right\|_2^2 < \delta^2$ **then**
      break
    end if
  end for
  $\hat{\mathbf{v}}_i \leftarrow \mathbf{v}_i^{(k)}$ ▷ Backward diffusion consistency (C2)
  $\hat{\mathbf{x}}_0' \leftarrow \frac{1}{\sqrt{\bar{\alpha}_i}} \left[\hat{\mathbf{v}}_i - \sqrt{1-\bar{\alpha}_i}\, \epsilon_\theta(\hat{\mathbf{v}}_i, i\Delta t) \right]$ ▷ Backward consistency (C2)
  $\mathbf{x}_{i-1} \leftarrow \sqrt{\bar{\alpha}_{i-1}}\, \hat{\mathbf{x}}_0' + \sqrt{1-\bar{\alpha}_{i-1}}\, \boldsymbol{\eta}_i, \quad \boldsymbol{\eta}_i \sim \mathcal{N}(\mathbf{0},\mathbf{I})$ ▷ Forward consistency (C3)
end for
return $\hat{\mathbf{x}} = \mathbf{x}_0$

F Experiment Setup for PDE Reconstructions

Helmholtz equation. The Helmholtz equation represents wave propagation in heterogeneous media:

$$\nabla^2 u(x) + k^2 u(x) = a(x), \quad x \in (0,1)^2, \tag{6}$$

with $k = 1$ and $u|_{\partial\Omega} = 0$. Coefficient fields $a(x)$ are generated according to $a \sim \mathcal{N}(0, (-\Delta + 9\mathbf{I})^2)$. We note that this system has reflection equivariance along $x_1 = \frac{1}{2}, x_2 = \frac{1}{2}, x_1 = x_2$ and rotation equivariance by $\frac{\pi}{2}, \pi, \frac{3\pi}{2}$.

Navier-Stokes equations. Following the methodology of (Li et al., 2020), we model the time evolution of a vorticity field, $u(x,t)$, governed by:

$$\partial_t u(x,t) + \boldsymbol{w}(x,t) \cdot \nabla u(x,t) = \nu \Delta u(x,t) + f(x), \quad x \in (0,1)^2,\ t \in (0,T], \tag{7}$$

$$\nabla \cdot \boldsymbol{w}(x,t) = 0, \quad x \in (0,1)^2,\ t \in [0,T], \tag{8}$$

$$u(x,0) = a(x), \quad x \in (0,1)^2, \tag{9}$$

where $\boldsymbol{w}$ is the velocity field; $\nu = \frac{1}{1000}$, viscosity; and f, a fixed forcing term. The initial condition $a(x)$ is drawn from $\mathcal{N}(0, 7^{3/2}(-\Delta + 49\mathbf{I})^{-5/2})$ under periodic boundary conditions. The forcing term is $f(x) = 0.1\,(\sin(2\pi(x_1 + x_2)) + \cos(2\pi(x_1 + x_2)))$. We borrow the dataset from (Huang et al., 2024). We note that this system has a reflection symmetry along the $x_1 = x_2$ axis.

Implementation details. EquiReg, as a regularizer for diffusion posterior sampling, can be adapted to many inverse solvers in a plug-and-play manner. For PDE experiments, we use the same model weights and configurations as FunDPS (Yao et al., 2025). Error rates are calculated using the L^2 relative error between the predicted and true solutions, averaged on 100 randomly selected test samples. We provide the information on the EquiReg scaling weights in Table 12.

Table 12: **EquiReg loss used in PDE experiments.**

	Helmholtz		Navier-Stokes	
	Forward	Inverse	Forward	Inverse
EquiReg Norm Type	MSE	L2	MSE	L2
EquiReg Weight λ	100	100	100	1000

G Theoretical Analysis

G.1 Summary of the Analysis

The theoretical framework presented in this paper is intended to motivate and guide the design of effective regularizers. This perspective, grounded in optimal transport theory (Ferreira & Valencia-Guevara, 2018), serves as an intuitive interpretation of the dynamics and motivates the design of regularization strategies such as EquiReg. We note that whether diffusion models follow exact Wasserstein dynamics still remains an open problem (Zheng et al., 2025).

Proposition G.1. *Let $\rho(\boldsymbol{x}, t)$ be the distribution of $\boldsymbol{x}_{T-t}$ driven by the ideal reverse dynamics (eq. (3)). Then, the evolution of ρ follows the Wasserstein-2 gradient flow associated with minimizing functional $\Phi(\rho, t)$ defined as $\beta_{T-t}\int[\rho\phi(\boldsymbol{x}, t) + \frac{1}{2}\rho\log\rho]d\boldsymbol{x}$, where $\phi(\boldsymbol{x}, t) = -(\log p_{T-t}(\boldsymbol{x}|\boldsymbol{y}) + \frac{1}{4}\|\boldsymbol{x}\|^2)$.*

The dynamics of ρ remain the same if we replace $\phi(\boldsymbol{x}, t)$ with $\phi_C(\boldsymbol{x}, t) := \phi(\boldsymbol{x}, t) - C(t)$ for arbitrary temporal function $C(t)$. Without loss of generality, we assume $\phi_C(\boldsymbol{x}, t) < 0$ for all $\boldsymbol{x}$ and t. In practice, the density function p_{T-t} is not available and thus $\phi_C(\boldsymbol{x}, t)$ is approximated as $\hat{\phi}$ with $p_{T-t}(\boldsymbol{x}_{T-t}|\boldsymbol{y}) \approx \tilde{C}p_{T-t}(\boldsymbol{x}_{T-t})p(\boldsymbol{y}|\mathbb{E}[\boldsymbol{x}_0|\boldsymbol{x}_{T-t}])$ where $\tilde{C}$ only depends on $\boldsymbol{y}$.

Because the conditional expectation $\mathbb{E}[\boldsymbol{x}_0|\boldsymbol{x}_{T-t}]$ is a linear combination of all candidate $\boldsymbol{x}_0$, the approximation remains relatively accurate when $T-t$ is small (i.e., $\boldsymbol{x}_{T-t}$ stays close to the data manifold under low noise) but may incur high error for larger $T-t$, as shown in Figure 1b. To mitigate this, we reweight the contributions to the first term of Φ, down-weighting unreliable estimates, and amplifying the reliable ones. The resulting reweighted functional is

$$\tilde{\Phi}(\rho, t) = \beta_{T-t}\big[{Z_t}^{-1}\int\rho(\boldsymbol{x})\hat{\phi}_c(\boldsymbol{x}, t)e^{\frac{\mathcal{R}(\boldsymbol{x})}{\hat{\phi}_c(\boldsymbol{x},t)}}\mathrm{d}\boldsymbol{x} + \tfrac{1}{2}\int\rho(\boldsymbol{x})\log\rho(\boldsymbol{x})d\boldsymbol{x}\big], \tag{10}$$

where $Z_t = \int e^{\frac{\mathcal{R}(\boldsymbol{x})}{\hat{\phi}_c(\boldsymbol{x},t)}}\mathrm{d}\boldsymbol{x}$ is the normalizing factor, and $\mathcal{R}(\boldsymbol{x})$ is a positive regularization that is nearly zero near the data manifold and much larger elsewhere. Intuitively, since $\hat{\phi}_C < 0$, the weight is nearly one for $\boldsymbol{x}$ near the data manifold and much smaller elsewhere.

Proposition G.2. *(Informal) The evolution of ρ, the probability distribution of $\boldsymbol{x}_{T-t}$ driven by the practical and regularized reverse dynamics (eq. (11)), is an approximation of the Wasserstein-2 gradient flow associated with minimizing $\tilde{\Phi}$.*

$$\mathrm{d}\boldsymbol{x} = [-\tfrac{\beta_t}{2}\boldsymbol{x}\mathrm{d}t - \beta_t\nabla_{\boldsymbol{x}_t}(\log p_t(\boldsymbol{x}_t) + \log\textstyle\int p(\boldsymbol{y}|\boldsymbol{x}_0)\tilde{p}_t(\boldsymbol{x}_0|\boldsymbol{x}_t)\mathrm{d}\boldsymbol{x}_0 - \mathcal{R}(\boldsymbol{x}_t))]\mathrm{d}t + \sqrt{\beta_t}\mathrm{d}\bar{\boldsymbol{w}} \tag{11}$$

G.2 Preliminary and Notations

We first remind the readers of gradient flow under the Wasserstein-2 metric and introduce the notations related to the diffusion model.

Wasserstein Gradient Flow Let $\mathcal{F} : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}\cup\{+\infty\}$ be a functional of probability distributions. The Wasserstein gradient flow of $\mathcal{F}$ is characterized by the minimizing movement scheme (also known as JKO scheme) introduced by (Jordan et al., 1998). For a fixed time step $\tau > 0$, the sequence $(\rho_k)_{k\in\mathbb{N}}$ of probability densities is defined recursively by:

$$\rho_{k+1} \in \arg\min_{\rho\in\mathcal{P}_2(\mathbb{R}^d)}\left\{\frac{1}{2\tau}W_2^2(\rho, \rho_k) + \mathcal{F}(\rho)\right\},$$

where W_2 denotes the 2-Wasserstein distance, and each ρ_k is a probability density representing the distribution at time $t = k\tau$. In the limit $\tau \to 0$, this discrete-time scheme recovers the continuous-time gradient flow of $\mathcal{F}$ under the W_2 metric.

Diffusion Model A diffusion model defines a forward stochastic process $(\boldsymbol{x}_t)_{t\in[0,T]}$ governed by the Itô SDE:

$$\mathrm{d}\boldsymbol{x}_t = f(\boldsymbol{x}_t, t)\,\mathrm{d}t + \sqrt{\beta_t}\,\mathrm{d}\boldsymbol{w}_t, \tag{12}$$

where $\boldsymbol{w}_t$ is standard Brownian motion, $\beta_t > 0$ is a time-dependent variance schedule, and $f(\boldsymbol{x}, t)$ is a drift term. For instance, $f \equiv 0$ for a variance-exploding SDE and $f(\boldsymbol{x}, t) = -\frac{\beta_t}{2}\boldsymbol{x}$ for a variance-preserving SDE defined in (Song et al., 2021). In this work, we carry out our analysis under a more general setting.

Assumption G.1. *The drift term is a gradient field, $f(\boldsymbol{x}, t) = \nabla h(\boldsymbol{x}, t)$ for a scalar function h.*

This process progressively transforms an initial data distribution $\boldsymbol{x}_0 \sim p_0$ into a tractable reference distribution (e.g., approximately a Gaussian $\mathcal{N}(0, I)$) at time T.

Sampling is performed by simulating the *reverse-time SDE*:

$$\mathrm{d}\boldsymbol{x}_t = [f(\boldsymbol{x}_t, t) - \beta_t \nabla_{\boldsymbol{x}} \log p_t(\boldsymbol{x}_t)]\,\mathrm{d}t + \sqrt{\beta_t}\,\mathrm{d}\bar{\boldsymbol{w}}_t, \tag{13}$$

where p_t is the marginal density of $\boldsymbol{x}_t$, and $\bar{\boldsymbol{w}}_t$ is a standard Brownian motion in reverse time.

In practice, the score function $\nabla_{\boldsymbol{x}} \log p_t(\boldsymbol{x})$ is approximated by a neural network $s_\theta(\boldsymbol{x}, t)$ trained to estimate the score of the forward process. For *conditional sampling*, where we sample $\boldsymbol{x}_0$ given some observed variable y, the score is replaced by $\nabla_{\boldsymbol{x}} \log p_t(\boldsymbol{x}|\boldsymbol{y})$ and decomposed as

$$\nabla_{\boldsymbol{x}} \log p_t(\boldsymbol{x}|\boldsymbol{y}) = \nabla_{\boldsymbol{x}} \log p_t(\boldsymbol{x}) + \nabla_{\boldsymbol{x}} \log p_t(\boldsymbol{y}|\boldsymbol{x}), \tag{14}$$

based on Bayes' rule.

To simplify notation in the sequel, we perform a time reparameterization $t = T - t'$, so that the reverse process is written as a forward SDE over $t \in [0, T]$:

$$\mathrm{d}\boldsymbol{x}_t = -\left[f(\boldsymbol{x}_t, T-t) - \beta_{T-t}[\nabla_{\boldsymbol{x}} \log p_{T-t}(\boldsymbol{x}_t) + \nabla_{\boldsymbol{x}} \log p_t(y|\boldsymbol{x}_t)]\right]\mathrm{d}t + \sqrt{\beta_{T-t}}\,\mathrm{d}\boldsymbol{w}_t, \tag{15}$$

This form describes the generative process as evolving forward from $t = 0$ to $t = T$, matching the usual direction of analysis in gradient flow frameworks.

G.3 Proof of Proposition G.1

In this work, we consider Wasserstein gradient flow under the setting where the functional $\mathcal{F}$ depends on time.

Lemma G.1. *Consider a time-dependent functional $\mathcal{F}(\rho, t) = \int \rho(\boldsymbol{x})V(\boldsymbol{x}, t)\mathrm{d}x + \int \alpha(t)\rho\log\rho\mathrm{d}\boldsymbol{x}$. Then the particle description of Wasserstein-2 gradient flow associated with this functional derived by JKO scheme is*

$$\mathrm{d}\boldsymbol{x}_t = -\nabla V(\boldsymbol{x}_t, t)\mathrm{d}t + \sqrt{2\alpha(t)}\mathrm{d}\boldsymbol{w}_t. \tag{16}$$

Proof. Consider the following optimization

$$\min_{\rho'} \mathcal{F}(\rho', t+\Delta t) - \mathcal{F}(\rho, t) + \frac{1}{2\Delta t} W_2^2(\rho, \rho'), \tag{17}$$

where the change of density is restricted to the Liouville equation

$$\partial_t \rho = -\nabla \cdot (\rho v(\boldsymbol{x}, t)), \text{ and } \rho'(x) = \rho(x) - \Delta t \nabla \cdot (\rho(\boldsymbol{x})v(\boldsymbol{x})) + o(\Delta t). \tag{18}$$

Using the static formulation of W_2 distance, we have

$$W_2^2(\rho, \rho') = \int \rho(\boldsymbol{x})\|\boldsymbol{x} - T^*(\boldsymbol{x})\|^2\,\mathrm{d}\boldsymbol{x} = \Delta t^2 \int \rho(\boldsymbol{x})\|v^*(\boldsymbol{x})\|^2\,\mathrm{d}\boldsymbol{x}, \tag{19}$$

where $T^*(\boldsymbol{x})$ is the optimal transport map, and $v^*(\boldsymbol{x})$ is the associated optimal velocity field.

Thus, we can rewrite the eq. (17) as

$$\inf_v \ \mathcal{F}(\rho,t) - \Delta t \int \nabla\cdot(\rho(\boldsymbol{x})v(\boldsymbol{x}))\,\frac{\delta\mathcal{F}(\rho,t)}{\delta\rho}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} + \Delta t\int\left[\rho(\boldsymbol{x})\partial_t V(\boldsymbol{x},t) + \dot{\alpha}(t)\rho\log\rho\right]\mathrm{d}\boldsymbol{x} \tag{20}$$

$$-\mathcal{F}(\rho,t) + \frac{\Delta t}{2}\int\rho(\boldsymbol{x})\|v(\boldsymbol{x})\|^2\,\mathrm{d}\boldsymbol{x}, \tag{21}$$

which simplifies to

$$\min_v \int\rho(\boldsymbol{x})\left\langle v(\boldsymbol{x}), \nabla\frac{\delta\mathcal{F}(\rho,t)}{\delta\rho}(\boldsymbol{x})\right\rangle\mathrm{d}\boldsymbol{x} + \frac{1}{2}\int\rho(\boldsymbol{x})\|v(\boldsymbol{x})\|^2\,\mathrm{d}\boldsymbol{x}, \tag{22}$$

since the last term in the first line of (20) does not depend on v. and further to

$$\min_v \int\rho(\boldsymbol{x})\left\|v(\boldsymbol{x}) + \nabla\frac{\delta\mathcal{F}(\rho,t)}{\delta\rho}(\boldsymbol{x})\right\|^2\mathrm{d}\boldsymbol{x}. \tag{23}$$

From the optimality condition of the above problem, we obtain

$$v(\boldsymbol{x},t) = -\nabla\frac{\delta\mathcal{F}(\rho,t)}{\delta\rho}(\boldsymbol{x}) = -(\nabla V(\boldsymbol{x},t) + \alpha(t)\nabla\log\rho(\boldsymbol{x},t)). \tag{24}$$

We note that By Hörmander's theorem, a smooth density $\rho(\boldsymbol{x},t)$ exists for $t>0$, ensuring that the above v is well-defined. The corresponding evolution of probability density is

$$\partial_t\rho(\boldsymbol{x},t) = -\nabla\cdot(\rho(\boldsymbol{x},t)v(\boldsymbol{x},t)) \tag{25}$$

$$= \nabla\cdot(\rho(\boldsymbol{x},t)(\nabla V(\boldsymbol{x},t) + \alpha(t)\frac{\nabla\rho(\boldsymbol{x},t)}{\rho})) \tag{26}$$

$$= -\nabla\cdot(\rho(\boldsymbol{x},t)(-\nabla V(\boldsymbol{x},t)) + \alpha(t)\Delta\rho(\boldsymbol{x},t)), \tag{27}$$

which is exactly the Fokker-Planck equation describing the evolution of the probability density describing the particles following

$$\mathrm{d}\boldsymbol{x}_t = -\nabla V(\boldsymbol{x}_t,t)\mathrm{d}t + \sqrt{2\alpha(t)}\mathrm{d}\boldsymbol{w}_t. \tag{28}$$

□

Now we come back to Proposition G.1. From eq. (15) we know that choosing

$$V(\boldsymbol{x},t) = h(\boldsymbol{x},T-t) - \beta_{T-t}[\log p_{T-t}(\boldsymbol{x}) + \log p_{T-t}(\boldsymbol{y}|\boldsymbol{x})] \ \text{ and } \alpha(t) = \frac{\beta_{T-t}}{2} \tag{29}$$

in Lemma G.1 completes the proof, where h is defined in Assumption G.1.

G.4 Detailed version of Proposition G.2

In practice, one does not have access to $\log p_t(\boldsymbol{y}|\boldsymbol{x}_t)$ which appears in the reverse SDE. The most popular approach is do the following approximation,

$$p_t(\boldsymbol{y}|\boldsymbol{x}_t) = \int p(\boldsymbol{y}|\boldsymbol{x}_0)p(\boldsymbol{x}_0|\boldsymbol{x}_t)\mathrm{d}\boldsymbol{x}_0 = \mathbb{E}_{\boldsymbol{x}_0\sim p(\boldsymbol{x}_0|\boldsymbol{x}_t)}[p(\boldsymbol{y}|\boldsymbol{x}_0)] \approx p(\boldsymbol{y}|\ \mathbb{E}[\boldsymbol{x}_0|\boldsymbol{x}_t]), \tag{30}$$

which can be interpreted as exchanging two operations, the conditional expectation and the measurement $p(\boldsymbol{y}|\cdot)$.

As discussed in the main text, since the conditional expectation is a linear combination over all possible values of $\boldsymbol{x}_0$, it may fall outside the data manifold, resulting in physically invalid samples. One of the central

challenges in diffusion-based inverse sampling is guiding the sampling trajectory, generated by the reverse SDE dynamics, toward the data manifold. A common strategy is to incorporate regularization into the reverse SDE to encourage manifold adherence. In this work, building on the perspective of Wasserstein gradient flow as outlined above, we provide a novel interpretation of the role played by such regularization terms.

We show that the regularizer serves to reweight the contribution of different regions in the calculation of the underlying functional being minimized, $\Phi(\rho, t)$ defined in Proposition G.1. Specifically, it amplifies the influence of regions where the density estimate is reliable (typically near the data manifold), while down-weighting regions with poor approximation quality of based on eq. (30), often corresponding to off-manifold samples.

Following from what we have shown in the main text, $\Phi(\rho, t)$ has the form of $\beta_{T-t}\int[\rho\phi(\boldsymbol{x}, t) + \frac{1}{2}\rho\log\rho]d\boldsymbol{x}$ for a function $\phi(\boldsymbol{x}, t)$, which can be derived by (29). The $\log p_t(\boldsymbol{y}|\boldsymbol{x})$ term in (29) or $\nabla\log p_t(\boldsymbol{y}|\boldsymbol{x})$ term in (28), equivalently, is computed based on approximation (30). We denote the corresponding approximation of $\phi(\boldsymbol{x}, t)$ as $\hat{\phi}(\boldsymbol{x}, t)$. As discussed in the main text, we can assume without loss of generality that $\phi(\boldsymbol{x}, t) < 0$ and $\hat{\phi}(\boldsymbol{x}, t) < 0$. We have

$$\hat{\Phi}(\rho, t) = \beta_{T-t}\big[\int_{\boldsymbol{x}\in N(\mathcal{M})}\rho(\boldsymbol{x})\hat{\phi}(\boldsymbol{x}, t)\mathrm{d}\boldsymbol{x} + \int_{\boldsymbol{x}\notin N(\mathcal{M})}\rho(\boldsymbol{x})\hat{\phi}(\boldsymbol{x}, t)\mathrm{d}\boldsymbol{x} + \frac{1}{2}\int\rho\log\rho\mathrm{d}\boldsymbol{x}\big], \tag{31}$$

where $N(\mathcal{M})$ denotes a neighborhood of the data manifold $\mathcal{M}$. Intuitively, we aim to focus on the contribution from regions near $\mathcal{M}$, which corresponds to the first term, while down-weighting the influence of points farther away, where the approximation tends to be unreliable. For instance, we can introduce two positive weights $A \gg B$ and adopt the modified functional

$$\tilde{\Phi}(\rho, t) = \beta_{T-t}\big[A\int_{\boldsymbol{x}\in N(\mathcal{M})}\rho(\boldsymbol{x})\hat{\phi}(\boldsymbol{x}, t)\mathrm{d}\boldsymbol{x} + B\int_{\boldsymbol{x}\notin N(\mathcal{M})}\rho(\boldsymbol{x})\hat{\phi}(\boldsymbol{x}, t)\mathrm{d}\boldsymbol{x} + \frac{1}{2}\int\rho\log\rho\mathrm{d}\boldsymbol{x}\big]. \tag{32}$$

In this work, we further generalize this idea and consider a continuous weight function,

$$\tilde{\Phi}(\rho, t) = \beta_{T-t}\big[\int\rho(\boldsymbol{x})\hat{\phi}(\boldsymbol{x}, t)\lambda(\boldsymbol{x})\mathrm{d}\boldsymbol{x} + \frac{1}{2}\int\rho\log\rho\mathrm{d}\boldsymbol{x}\big], \tag{33}$$

where the non-negative weight $\lambda(\boldsymbol{x})$ is large for $\boldsymbol{x} \in N(\mathcal{M})$ and small elsewhere.

In practice, a nonnegative regularization function $\mathcal{R}(\boldsymbol{x})$ is introduced, ideally being nearly zero for $\boldsymbol{x}$ near the data manifold and much larger elsewhere. We consider the following modified functional with weight function $\lambda(\boldsymbol{x}, t) := e^{\frac{\mathcal{R}(\boldsymbol{x})}{\hat{\phi}(\boldsymbol{x}, t)}}$,

$$\tilde{\Phi}(\rho, t) = \beta_{T-t}\big[\int\rho(\boldsymbol{x})\hat{\phi}(\boldsymbol{x}, t)e^{\frac{\mathcal{R}(\boldsymbol{x})}{\hat{\phi}(\boldsymbol{x}, t)}}\mathrm{d}\boldsymbol{x} + \tfrac{1}{2}\int\rho(\boldsymbol{x})\log\rho(\boldsymbol{x})d\boldsymbol{x}\big]. \tag{34}$$

Note that $\hat{\phi} < 0$, we have that

$$\mathcal{R}(\boldsymbol{x}) \approx \begin{cases} 0, & \boldsymbol{x} \in N(\mathcal{M}) \\ \gg 1, & \boldsymbol{x} \text{ far away from } N(\mathcal{M}) \end{cases} \quad \Rightarrow \quad \lambda(\boldsymbol{x}, t) \approx \begin{cases} 1, & \boldsymbol{x} \in N(\mathcal{M}) \\ 0, & \boldsymbol{x} \text{ far away from } N(\mathcal{M}) \end{cases}.$$

Next, we consider practical algorithms based on this reweighted functional. In practice, we only have the score function instead of the function value of $\log p_{T-t}(\boldsymbol{x})$. Thus, the Wasserstein gradient flow associated with (34) is intractable since we cannot evaluate the weight function. We consider the following approximation based on $e^{\delta} \approx 1 + \delta$ when δ is sufficiently small,

$$\tilde{\Phi}(\rho, t) \approx \beta_{T-t}\big[\int\rho(\boldsymbol{x})\hat{\phi}(\boldsymbol{x}, t)\big(1 + \frac{\mathcal{R}(\boldsymbol{x})}{\hat{\phi}(\boldsymbol{x}, t)}\big)\mathrm{d}\boldsymbol{x} + \tfrac{1}{2}\int\rho(\boldsymbol{x})\log\rho(\boldsymbol{x})d\boldsymbol{x}\big] \tag{35}$$

$$= \beta_{T-t}\big[\int\rho(\boldsymbol{x})\big(\hat{\phi}(\boldsymbol{x}, t) + \mathcal{R}(\boldsymbol{x})\big)\mathrm{d}\boldsymbol{x} + \tfrac{1}{2}\int\rho(\boldsymbol{x})\log\rho(\boldsymbol{x})d\boldsymbol{x}\big]. \tag{36}$$

By Lemma G.1, the dynamics of $\boldsymbol{x}$ driven by the Wasserstein gradient flow associated with the approximated functional above is

$$\mathrm{d}\boldsymbol{x} = [-f(\boldsymbol{x}, T-t) - \beta_{T-t}\nabla_{\boldsymbol{x}}\big(\log p_{T-t}(\boldsymbol{x}) + \log \hat{p}_{T-t}(\boldsymbol{y}|\boldsymbol{x}) + \mathcal{R}(\boldsymbol{x})\big)]\mathrm{d}t + \sqrt{\beta_{T-t}}\mathrm{d}\bar{\boldsymbol{w}}. \tag{37}$$

This completes the proof.

Remark 1. *Since $\hat{\phi} < 0$, and $e^A \geq 1 + A$ for any $A \in \mathbb{R}$, the dynamics derived by the approximated functional in (36) is evolving to minimize an upper bound of the reweighted functional $\tilde{\Phi}$.*

H Additional Background Information

Solving inverse problems with deep learning prior to diffusion models. Earlier works (Metzler et al., 2016; Romano et al., 2017; Zhang et al., 2017; Metzler et al., 2017) used deep neural networks as denoisers to solve inverse problems. Furthermore, deep generative models such as variational autoencoders (VAEs) (Kingma, 2013), and generative adversarial networks (GANs) (Goodfellow et al., 2014) were employed. Notable applications include compressed sensing (Bora et al., 2017) and MRI (Jalal et al., 2021).

Applications on diffusion models to solve inverse problems. Most popular applications include image restoration (Chung et al., 2023; 2022b; Kawar et al., 2022; Lugmayr et al., 2022; Saharia et al., 2022; Song et al., 2023a; Rout et al., 2023; Zhu et al., 2023; Zhang et al., 2025a; Zirvi et al., 2025), medical imaging (Song et al., 2022; Chung & Ye, 2022; Chung et al., 2022a; Hung et al., 2023; Dorjsembe et al., 2024; Li et al., 2024; Kazerouni et al., 2023; Bian et al., 2024), and solving partial differential equations (PDEs) (Isakov, 2006; Huang et al., 2024; Shysheya et al., 2024; Liu et al., 2023; Li et al., 2025; Baldassari et al., 2023; Mammadov et al., 2024a; Yao et al., 2025). On the methodology side, there has been numerous advancements (Chung et al., 2023; 2022b; Kawar et al., 2022; Lugmayr et al., 2022; Saharia et al., 2022; Song et al., 2023a; Rout et al., 2023; Zhu et al., 2023; Zhang et al., 2025a; Zirvi et al., 2025; Song et al., 2022; Chung & Ye, 2022; Chung et al., 2022a; Hung et al., 2023; Dorjsembe et al., 2024; Li et al., 2024; Kazerouni et al., 2023; Bian et al., 2024; Huang et al., 2024; Shysheya et al., 2024; Mammadov et al., 2024b; Cardoso et al., 2024).

Resources for Definition 3.2 on vanishing-error autoencoders. Manifold constrained distribution-dependent equivariance error uses the notion of *vanishing-error autoencoders* (Shao et al., 2018; Anders et al., 2020; He et al., 2024) (Definition H.1), also known as an asymptotically-trained autoencoder (Anders et al., 2020) or a perfect autoencoder (He et al., 2024). Vanishing-error autoencoders have previously been employed by diffusion-based inverse solvers to preserve the diffusion process on the manifold (He et al., 2024).

Definition H.1 (Vanishing-Error Autoencoder)**.** *A vanishing-error autoencoder under the manifold $\mathcal{M}$ with encoder $\mathcal{E} : \mathcal{X} \to \mathcal{Z}$ and decoder $\mathcal{D} : \mathcal{Z} \to \mathcal{X}$ with $\mathcal{Z} = \mathbb{R}^k$ where $k < d$, has zero reconstruction error under the support of the data distribution $\mathcal{X}$, i.e., $\forall \boldsymbol{x} \in \mathcal{X} \subset \mathcal{M}$, $\boldsymbol{x} = \mathcal{D}(\mathcal{E}(\boldsymbol{x}))$. It follows that the decoder is surjective on the data manifold, $\mathcal{D} : \mathcal{Z} \to \mathcal{M}$ (He et al., 2024), and the encoder-decoder composition forms an identity map, i.e., $\forall \boldsymbol{z} \in \mathcal{M}, \boldsymbol{z} = \mathcal{E}(\mathcal{D}(\boldsymbol{z}))$.*

Equivariance. Let $\boldsymbol{z} \in \mathbb{R}^d$ and $\boldsymbol{x} = f(\boldsymbol{z}) \in \mathbb{R}^d$. For rotation and reflection equivariance, the transformations T_g and S_g can be defined by a rotation matrix $\boldsymbol{R} \in \mathbb{R}^{d\times d}$; then, a function f with the rotation equivariant property would satisfy $\boldsymbol{R}\boldsymbol{x} = f(\boldsymbol{R}\boldsymbol{z})$. For translation equivariance, the transformations would be $T_g(\boldsymbol{z}) = \boldsymbol{z} + g$ and $S_g(\boldsymbol{x}) = \boldsymbol{x} + g$, where $g \in \mathbb{R}^d$. Hence, for a translation equivariance function f, we would have $\boldsymbol{x} + g = f(\boldsymbol{z} + g)$. For the case where the output dimension is larger than the input, $f : \mathbb{R}^k \to \mathbb{R}^d$ with $d > k$, translation equivariance can be defined up to a discrete scale, i.e., $T_g(\boldsymbol{z}) = \boldsymbol{z} + g$ and $S_g(\boldsymbol{x}) = T_{sg}(\boldsymbol{z})$ where $s = {}^d\!/_k$. The equivariance properties of translation, rotation, and reflections, combined, are referred to as E(3) symmetries. Without reflections, the symmetries form a Euclidean group SE(3) (Thomas et al., 2018; Fuchs et al., 2020).

E(3), SE(3), and SO(3) are important symmetry groups in 3D Euclidean space, with well-established applications in physics and chemistry, computer vision, and reinforcement learning (Cohen & Welling, 2016; Thomas et al., 2018; Hoogeboom et al., 2022; Xu et al., 2024; Park et al., 2025). Finally, our contributions are complementary to, and can be combined with, the growing literature on meta-learning and automatic symmetry discovery to learn symmetry groups and their actions directly from data (Zhou et al., 2021; Quessard et al., 2020; Dehmamy et al., 2021; Mohapatra et al., 2025).

Table 13: DPS superresolution with $\lambda = 0.01$ using different MPE functions.

(a) FFHQ 256.

MPE function	PSNR	SSIM	LPIPS	FID
None	23.160 (1.923)	0.657 (0.072)	0.193 (0.057)	129.528
LDM Encoder (FFHQ)	26.581 (2.457)	0.773 (0.044)	0.120 (0.030)	87.437
CNN Autoencoder (FFHQ)	26.866 (1.943)	0.771 (0.044)	0.116 (0.029)	85.352
Pretrained ResNet50	26.873 (1.941)	0.771 (0.044)	0.116 (0.029)	85.138
Pretrained CLIP	26.860 (1.942)	0.771 (0.044)	0.116 (0.029)	85.495

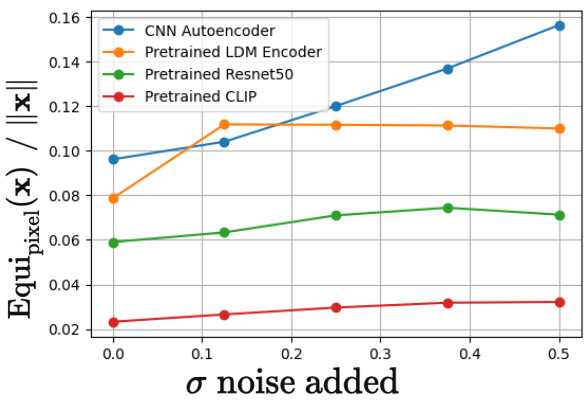
(b) ImageNet.

MPE function	PSNR	SSIM	LPIPS	FID
None	19.727 (4.292)	0.407 (0.180)	0.541 (0.182)	446.829
LDM Encoder (ImageNet)	22.200 (4.295)	0.568 (0.146)	0.384 (0.130)	311.636
CNN Autoencoder (ImageNet)	22.178 (4.294)	0.568 (0.148)	0.375 (0.125)	312.530
Pretrained ResNet50	22.176 (4.290)	0.568 (0.148)	0.375 (0.125)	314.590
Pretrained CLIP	22.177 (4.293)	0.568 (0.148)	0.376 (0.125)	313.468

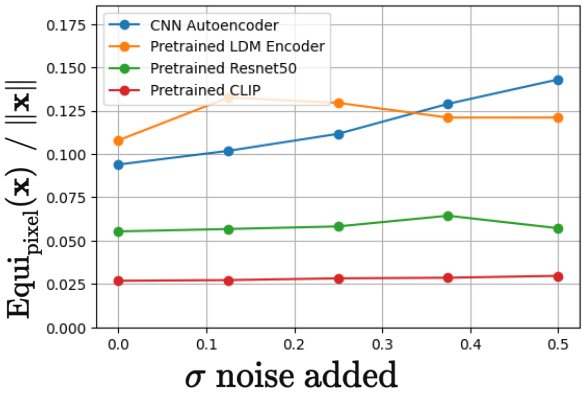
Data manifold hypothesis. Let data $\boldsymbol{x} \in \mathcal{X} \subset \mathbb{R}^d$ be in an ambient space of dimension d with support $\mathcal{X}$ distribution. We assume that data are sampled from a low-dimensional manifold $\mathcal{M}$ (Cayton et al., 2005; Ma & Fu, 2012) embedded in a high-dimensional space (Assumption H.1). This hypothesis is popular in machine learning (Bordt et al., 2023), and has been studied mathematically in the literature (Narayanan & Mitter, 2010; Bortoli, 2022). Moreover, empirical evidence in image processing supports the manifold hypothesis (Weinberger & Saul, 2006; Fefferman et al., 2016), and diffusion-based solvers assume this property (He et al., 2024; Chung et al., 2022b; 2023).

Assumption H.1 (Manifold Hypothesis)**.** *Let $\boldsymbol{x} \in \mathcal{X} \subset \mathbb{R}^d$ be a data sample. The support $\mathcal{X}$ of the data distribution lies on a k dimensional manifold $\mathcal{M}$ within an ambient space $\mathbb{R}^d$ where $k \ll d$.*

I Additional figures on MPE functions



(a) FFHQ 256.



(b) ImageNet.

Figure 19: **Equivariance error vs. σ noise added.** As more noise is added, equivariance error, computed with all MPE functions, increases.

J Computing Resources

We conduct experiments on two NVIDIA GeForce RTX 4090 GPUs with 24 GB of VRAM. We note that we use pre-trained models and perform inference, so not much compute is required.

K Assets

We use the publicly available code from PSLD (https://github.com/LituRout/PSLD), ReSample (https://github.com/soominkwon/resample), DPS (https://github.com/DPS2022/diffusion-posterior-sampling), and SITCOM (https://github.com/sjames40/SITCOM).

L Responsible Release

Our approach uses only publicly available datasets and standard pre-trained diffusion models, introducing no novel dual-use or privacy risks. Consequently, no additional safeguards are required.