Local Epistemic Uncertainty Guided Active Sampling for Plug-and-play Diffusive Image Restoration

Anonymous Author(s)
Submission Id: 7960

Abstract

Diffusion models have demonstrated remarkable effectiveness in image restoration tasks. However, when guiding image reconstruction, existing Diffusion Model-based Image Restoration (DMIR) methods typically rely on fixed data constraints and uniform step sizes, thereby overlooking the dynamic nature of the generative process. Such rigid designs render the models vulnerable to spatially non-uniform degradations, thus resulting in structural distortions and loss of fine details. Meanwhile, uniform step sizes introduce computational redundancy, whereas naïve step reduction strategies tend to accumulate approximation errors. To address these limitations, we propose a Local Epistemic Uncertainty Guided Active Sampling framework (LEADer). In the spatial domain, LEADer leverages pixel-wise uncertainty to dynamically modulate the prior strength within the null space, which effectively balances detail preservation and artifact suppression. In the temporal domain, it quantifies sampling stability via the uncertainty trace to enable adaptive trajectory pruning, thereby accelerating convergence. Theoretical proofs demonstrate that our framework achieves strict data consistency, while the trajectory pruning strategy admits a deterministic error bound, thereby guaranteeing stable convergence under skip sampling. Notably, our plug-and-play method can be seamlessly integrated into various DMIR baselines. Extensive experiments show that LEADer improves the performance of multiple state-of-the-art DMIR methods, while significantly reducing sampling time with negligible memory overhead.

CCS Concepts

• **Computing methodologies** → **Computer vision**; *Machine learning*.

Keywords

Diffusion model, Plug-and-play, Image restoration, Local epistemic uncertainty

1 Introduction

Image restoration [9, 11, 15, 20, 22] aims to reconstruct clear, natural images from degraded observations corrupted by noise, blur, or low resolution, serving as a fundamental problem in computer vision and medical imaging [27, 40, 42]. Recently, deep learning-based methods have achieved remarkable performance on various specific image restoration tasks [4, 11, 22, 48, 52]. However, these methods, optimized for specific degradation patterns, typically rely on massive paired training data. Moreover, this task-specific paradigm often suffers from limited generalization capabilities [51].

Recently, diffusion models [8, 17, 31, 34, 36, 46] have demonstrated powerful capabilities in modeling complex data distributions and have been successfully applied to diverse image generation tasks [10, 29]. Consequently, researchers have begun exploring pre-trained unconditional diffusion models as universal priors to solve image restoration inverse problems [13]. Diffusion Model-based Image Restoration (DMIR) methods typically adopt a posterior sampling paradigm: they directly utilize pre-trained diffusion models to fit the prior distribution of clear natural images, incorporating likelihood guidance [2, 5, 7, 26, 35, 39, 44, 45] or data consistency projection [6, 14, 19, 25, 38, 41, 43, 47, 50, 54] during the reverse sampling process. This enables a single pre-trained model to adapt to diverse degradation scenarios, which effectively enhances the generalization capability of restoration methods.

However, during reverse sampling, existing DMIR methods typically employ fixed data constraint weights and uniform time steps. This setting overlooks the dynamic nature of image generation and introduce two main challenges. Spatially, since local degradations in real images are often uneven, globally fixed weights struggle to accommodate varying regional restoration needs, thus leading to compromised details or inaccurate local structures. Temporally, fixed time steps cause additional computational cost during the stationary phase of sampling, whereas directly reducing steps introduces additional discretization errors that compromise overall restoration quality [18, 24].

In response to these challenges, we propose an active diffusion sampling framework guided by local epistemic uncertainty, termed LEADer. Specifically, to ensure structural consistency, we propose Uncertainty-Calibrated Prior Modulation (UCPM), which leverages pixel-level uncertainty to dynamically regulate the strength of prior within the null space, thereby adaptively balancing detail preservation and artifact suppression. Furthermore, to improve inference efficiency, we propose State-Aware Trajectory Pruning (SATP). This technique evaluates the stability of the sampling process via the trace of uncertainty and executes step-skipping within an controlled error tolerance, which effectively reduces computational redundancy and accelerates convergence, as illustrated in Figure 1. In addition, theoretical analysis demonstrates that our method guarantees data fidelity while providing a controlled error bound for the skip sampling process. The main contributions of this paper are summarized as follows:

- We propose a novel perspective on the diffusion reverse process based on local epistemic uncertainty, which reveals the spatiotemporal limitations of the sampling strategies in existing DMIR methods.
- We propose an active diffusion sampling framework that proactively resolves local distortion and sampling redundancy via uncertainty-calibrated prior modulation and state-aware trajectory pruning.
- Extensive experimental results demonstrate that our proposed plug-and-play framework effectively enhances image restoration quality and performance while improving inference efficiency.

2 Background

DMIR methods have emerged as the mainstream for zero-shot solvers in image inverse problems. The core challenge lies in injecting specific measurement constraints into the reverse denoising process. Depending on the intervention mechanisms, existing methods are primarily categorized into the following two classes:

Gradient-based Optimization Guidance: Rooted in Bayesian posterior sampling theory [5, 26, 35], these methods approximate the measurement likelihood gradient using the clean image estimate at the current time step to guide reverse sampling. Notably, DPS [5] directly enforces data consistency using the gradient of the l_2-norm of the measurement error. Subsequent works have focused on improving guidance accuracy and stability. For example, MCG [7] incorporates manifold projection constraints, while ΠGDM [33] enhances robustness via pseudo-inverse matrices and Jacobian computations. ZAPS [1] further proposes an alternative guidance scheme to improve performance over DPS [5]. More recent studies investigate the numerical instability and bias arising during the guidance process. For instance, SITCOM [2] and SPGD [45] introduce explicit gradient control mechanisms to stabilize optimization. In parallel, DCDP [21] and DAPS [49] refine the sampling trajectory by decoupling the purification process to mitigate prior bias, while DPPS [44] employs proximal operators to handle non-smooth constraints, thereby improving flexibility under complex degradations.

Null-Space Orthogonal Projection: In contrast to gradient-based guidance, these methods decompose the image space into range and null spaces via orthogonal projection, and iteratively refine the null-space components during reverse diffusion. Representative approaches such as DDNM [41] and DDRM [19] achieve strict data consistency through pseudo-inverse projection and singular value decomposition (SVD), respectively. Building upon this foundation, subsequent works extend the framework to improve flexibility and guidance quality. DiffPIR [54] integrates Half-Quadratic Splitting (HQS) to reformulate measurement constraints as proximal mappings, while DDPG [14] introduces a preconditioned guidance trajectory that evolves with sampling steps, which effectively bridges back-projection and least-squares gradient formulations. More recent studies focus on enhancing global consistency and structural regularity. For example, ProjDiff [50] constructs a bivariate constrained optimization framework to better exploit denoising priors, EquS [43] utilizes equivariant sampling to constrain manifold regularity, and PIRP [47] incorporates parameterized gradient priors to elevate overall perception and fidelity.

3 Preliminary

Problem Formulation: The image restoration task aims to recover a target clean image $\mathbf{x}_0 \in \mathbb{R}^N$ from a degraded observation $\mathbf{y} \in \mathbb{R}^M$. This physical degradation process is typically modeled as a linear inverse problem:

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{n}, \tag{1}$$

where $\mathbf{y} \in \mathbb{R}^M$ is the degraded image, and $\mathbf{x}_0 \in \mathbb{R}^N$ represents the high-quality image to be recovered. $\mathbf{A} \in \mathbb{R}^{M\times N}$ is a known linear operator, such as a bicubic downsampler in image super-resolution. Finally, $\mathbf{n} \sim \mathcal{N}(\mathbf{0}, \sigma_y^2\mathbf{I})$ is additive Gaussian noise [19], which represents uncertainty or disturbance in the observation $\mathbf{y}$.

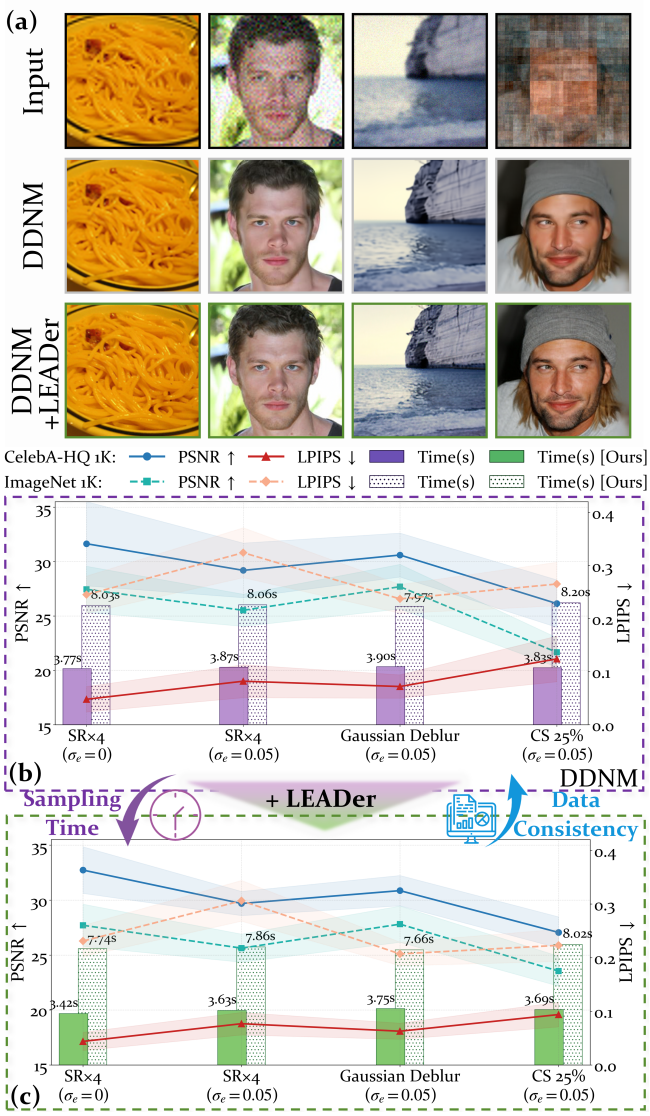

Figure 1: (a) Qualitative comparisons of the proposed method with plug-and-play adaptation across different tasks. (b) and (c) Quantitative comparisons under the same settings. The results demonstrate that our method facilitates restoration accuracy and efficiency.

Within the zero-shot diffusion framework, restoration is achieved without retraining on paired degraded data. Instead, a pre-trained unconditional diffusion model serves as a generative prior. At any time step $t \in (0, T]$ during reverse sampling, given the current noisy variable $\mathbf{x}_t$, the diffusion model predicts the noise component via a denoising neural network $\boldsymbol{\epsilon}_\theta(\mathbf{x}_t, t)$. Based on Tweedie's formula [12, 37], the corresponding estimate of the clean image is:

$$\mathbf{x}_{0|t} = \frac{\mathbf{x}_t - \sqrt{1-\bar{\alpha}_t}\boldsymbol{\epsilon}_\theta(\mathbf{x}_t, t)}{\sqrt{\bar{\alpha}_t}}, \tag{2}$$

where $\bar{\alpha}_t$ is a predefined noise schedule parameter.

To enforce data consistency with the physical observation, the Denoising Diffusion Null-Space Model [41] introduces an orthogonal decomposition of the image space into range and null spaces:

$$\hat{\mathbf{x}}_{0|t} = \mathbf{A}^{\dagger}\mathbf{y} + (\mathbf{I} - \mathbf{A}^{\dagger}\mathbf{A})\mathbf{x}_{0|t}, \tag{3}$$

where $\mathbf{A}^{\dagger}$ is the pseudo-inverse of the degradation matrix. $\mathbf{A}^{\dagger}\mathbf{y}$ enforces data fidelity in the range space, while $(\mathbf{I}-\mathbf{A}^{\dagger}\mathbf{A})\mathbf{x}_{0|t}$ projects $\mathbf{x}_{0|t}$ onto the null space for prior-guided refinement.

Diffusion Models: Denoising Diffusion Probabilistic Models (DDPM) [17] learn data generation by reversing a diffusion process that progressively corrupts data through noise addition. This framework consists of a forward process and a reverse process.

The forward process is a fixed Markov chain that gradually perturbs the data over T time steps. Let β_t denote the noise schedule. At each step t, the noisy state $\mathbf{x}_t$ is given by:

$$\mathbf{x}_t = \sqrt{1-\beta_t}\mathbf{x}_{t-1} + \sqrt{\beta_t}\boldsymbol{\epsilon}_{t-1}, \tag{4}$$

where $\boldsymbol{\epsilon}_{t-1} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$. Let the cumulative noise up to time step t be $\bar{\alpha}_t = \prod_{i=1}^{t} \alpha_i$, then we have:

$$\mathbf{x}_t = \sqrt{\bar{\alpha}_t}\mathbf{x}_0 + \sqrt{1-\bar{\alpha}_t}\boldsymbol{\epsilon}. \tag{5}$$

The reverse process aims to reconstruct the original clean image $\mathbf{x}_0$ from the noisy state $\mathbf{x}_t$. Mathematically, this is described by:

$$\mathbf{x}_{t-1} = \frac{1}{\sqrt{\alpha_t}}\left(\mathbf{x}_t - \frac{1-\alpha_t}{\sqrt{1-\bar{\alpha}_t}}\boldsymbol{\epsilon}_\theta(\mathbf{x}_t, t)\right) + \sigma_t\boldsymbol{\epsilon}_t, \tag{6}$$

where $\boldsymbol{\epsilon}_\theta(\mathbf{x}_t, t)$ denotes the noise predicted by a neural network parameterized by θ, and $\boldsymbol{\epsilon}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ is the Gaussian noise.

Despite their strong generative capability, DDPM suffers from high computational cost due to the large number of sampling steps. To address this limitation, [32] introduces Denoising Diffusion Implicit Models (DDIM), which relax the Markovian assumption, and enable more efficient sampling with fewer steps. Based on the estimate $\mathbf{x}_{0|t}$ in Eq. (2), the DDIM update rule is:

$$\mathbf{x}_{t-1} = \sqrt{\bar{\alpha}_{t-1}}\mathbf{x}_{0|t} + \sqrt{1-\bar{\alpha}_{t-1}-\sigma_t^2}\boldsymbol{\epsilon}_\theta(\mathbf{x}_t, t) + \sigma_t\boldsymbol{\epsilon}_t. \tag{7}$$

4 Local Epistemic Uncertainty Guided Active Sampling

4.1 Motivation

Existing zero-shot image restoration solvers typically employ globally fixed prior weights and uniform time steps when intervening in the reverse diffusion process. Such a rigid design implicitly treats sampling as a static procedure and overlooks the inherently dynamic characteristics of diffusion models across different spatial regions and temporal evolution stages. This limitation gives rise to two key issues:

Perception-Distortion Trade-off: The level of degradation and the difficulty of restoration varies significantly across different image regions [3]. Applying globally uniform constraint weights fails to account for this heterogeneity. Specifically, in regions where the model can already recover fine details, overly strong priors tend to over-smooth structures. On the other hand, in severely degraded areas where prior-driven inference is less reliable, insufficient constraints can lead to artifacts that deviate from the true underlying structures [16, 28].

Sampling Redundancy: From the perspective of ordinary differential equation (ODE), the reverse diffusion trajectory evolves with non-uniform dynamics. However, existing methods typically adopt fixed, uniform step sizes. This static temporal scheduling ignores the varying demands of different sampling stages. Once the global structure is established, continuing dense step-by-step updates introduces substantial computational redundancy; whereas naively increasing step sizes can accumulate errors during detail reconstruction, which ultimately degrades the quality [18].

To address these issues, we argue that the diffusion sampling steps should be state-aware. Specifically, we quantify local epistemic uncertainty at each spatial location and sampling step, then we use it as a unified modulation signal. In the spatial dimension, this uncertainty facilitates pixel-level on-demand regularization modulation to balance detail preservation and artifact suppression (see Section 4.3). In the temporal dimension, it serves as an error control metric to guide adaptive step sizing (see Section 4.4). The overall architecture of our method is illustrated in Figure 2.

4.2 Local Epistemic Uncertainty Quantification

At any time step t in the reverse dynamics, the noisy observation $\mathbf{x}_t$ embodies two fundamentally different types of uncertainty: the inherent aleatoric uncertainty introduced by the forward Gaussian diffusion process, and the epistemic uncertainty arising from missing information, particularly within the null space. In this paper, we focus on quantifying the latter from an information geometry perspective.

Based on the marginal distribution $q(\mathbf{x}_t|\mathbf{x}_0)$ of the forward noise-adding process, the Observed Fisher Information Matrix (FIM) $\mathcal{F}_t(\mathbf{x}_t) \in \mathbb{R}^{N\times N}$ of the pre-trained diffusion model at the current state $\mathbf{x}_t$ is defined as:

$$\mathcal{F}_t(\mathbf{x}_t) = -\nabla^2_{\mathbf{x}_t} \log p_t(\mathbf{x}_t), \tag{8}$$

where $\mathcal{F}_t(\mathbf{x}_t)$ represents the model's information sensitivity and structural certainty regarding the local features at this state. A higher value indicates that the model has stronger prior confidence in high-frequency details, whereas a lower value implies higher generative ambiguity in that region. As detailed in Section 3, the posterior expectation estimate $\mathbf{x}_{0|t}$ of the target clean image $\mathbf{x}_0$ is:

$$\mathbf{x}_{0|t} = \mathbb{E}[\mathbf{x}_0|\mathbf{x}_t] = \frac{1}{\sqrt{\bar{\alpha}_t}}\left(\mathbf{x}_t + (1-\bar{\alpha}_t)\nabla_{\mathbf{x}_t}\log p_t(\mathbf{x}_t)\right). \tag{9}$$

To quantify the prediction variance of the model given $\mathbf{x}_t$, we perform a second-order Taylor expansion on the posterior distribution to derive the analytical relationship between the posterior covariance matrix $\boldsymbol{\Sigma}_t$ and FIM:

$$\boldsymbol{\Sigma}_t = \mathrm{Cov}(\mathbf{x}_0|\mathbf{x}_t) = \frac{1-\bar{\alpha}_t}{\bar{\alpha}_t}\left[\mathbf{I} - (1-\bar{\alpha}_t)\mathcal{F}_t(\mathbf{x}_t)\right]. \tag{10}$$

This decomposition provides a clear interpretation. The leading term $\frac{1-\bar{\alpha}_t}{\bar{\alpha}_t}\mathbf{I}$ represents the inherent upper bound of the aleatoric uncertainty at time step t, while the second term $(1-\bar{\alpha}_t)\mathcal{F}_t(\mathbf{x}_t)$ reflects the Fisher information gain provided by the pre-trained model via the score function. Consequently, the diagonal matrix $\boldsymbol{\Sigma}_t$ serve as an effective measure of local epistemic uncertainty. Specifically, when $\mathcal{F}_t(\mathbf{x}_t)$ is large, the model exhibits high confidence in the reconstructed structures, thus leading to an epistemic uncertainty

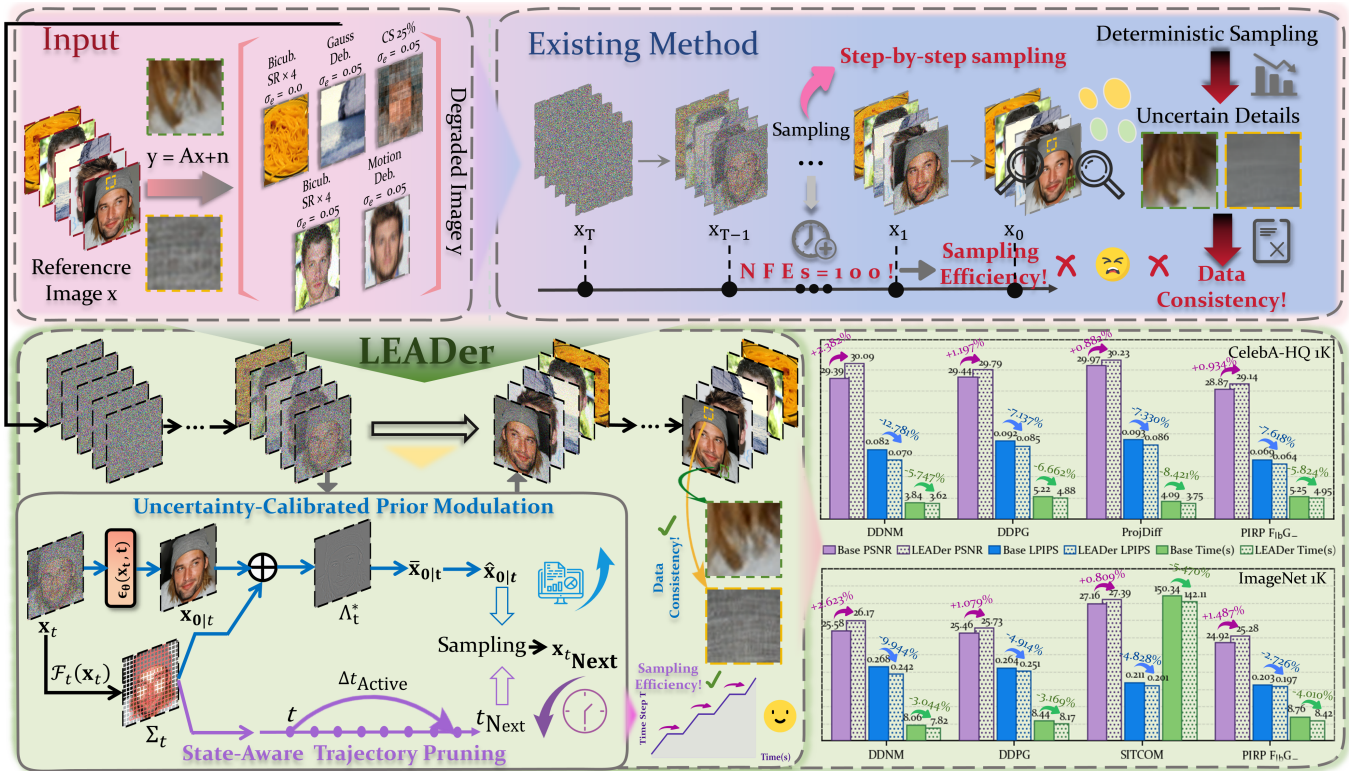

Figure 2: Architecture of LEADer. Existing DMIR methods rely on fixed step-by-step sampling with uniform constraints, leading to computational redundancy and detail distortion in uncertain regions. To address this, LEADer quantifies local epistemic uncertainty Σ_t to guide reverse sampling. Spatially, UCPM utilizes Σ_t to adaptively modulate prior constraints for strict data consistency. Temporally, SATP computes an adaptive step size Δt_{Active} based on the uncertainty trace to skip redundant iterations. As a result, LEADer improves the restoration quality of various baselines while reducing sampling time.

of $\Sigma_t \to \mathbf{0}$. Conversely, in regions with limited observational constraints, particularly within the null-space, the Fisher information diminishes, thus resulting in a higher Σ_t.

4.3 Uncertainty-Calibrated Prior Modulation

To address the limitations of globally uniform weighting, we propose Uncertainty-Calibrated Prior Modulation (UCPM). By leveraging the extracted local epistemic uncertainty Σ_t as a spatial guidance signal, UCPM dynamically modulates the strength of physical priors $\Phi(\mathbf{x})$ (e.g., gradient and structural priors) at the pixel level. Specifically, we cast the problem as a Maximum A Posteriori (MAP) optimization objective within the local state space:

$$\mathcal{J}(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \mathbf{x}_{0|t})^T \Sigma_t^{-1}(\mathbf{x} - \mathbf{x}_{0|t}) + \Phi(\mathbf{x}). \tag{11}$$

Setting the first-order optimality condition $\nabla \mathcal{J}(\mathbf{x}) = \mathbf{0}$ yields:

$$\Sigma_t^{-1}(\mathbf{x} - \mathbf{x}_{0|t}) + \nabla\Phi(\mathbf{x}) = \mathbf{0}. \tag{12}$$

To obtain a tractable solution, we linearize the non-linear prior term via a first-order Taylor expansion around $\mathbf{x}_{0|t}$:

$$\nabla\Phi(\mathbf{x}) \approx \nabla\Phi(\mathbf{x}_{0|t}) + \mathbf{H}_\Phi(\mathbf{x}_{0|t})(\mathbf{x} - \mathbf{x}_{0|t}), \tag{13}$$

where $\mathbf{H}_\Phi(\mathbf{x}_{0|t}) \in \mathbb{R}^{N\times N}$ is the Hessian matrix of the prior functional. Substituting this into the optimality condition, we derive the closed-form solution for the optimal image estimate $\bar{\mathbf{x}}_{0|t}$:

$$\bar{\mathbf{x}}_{0|t} = \mathbf{x}_{0|t} - \Lambda_t^* \nabla\Phi(\mathbf{x}_{0|t}), \tag{14}$$

where the spatially heterogeneous modulation matrix Λ_t^* is defined as:

$$\Lambda_t^* = \left[\Sigma_t^{-1} + \mathbf{H}_\Phi(\mathbf{x}_{0|t})\right]^{-1} \equiv \Sigma_t \left(\mathbf{I} + \mathbf{H}_\Phi(\mathbf{x}_{0|t})\Sigma_t\right)^{-1}. \tag{15}$$

The matrix Λ_t^* acts as an adaptive preconditioner, with its modulation strength governed by Σ_t. When $\Sigma_t \to \mathbf{0}$, we have $\Lambda_t^* \to \mathbf{0}$, which indicates that prior intervention is attenuated to preserve high-frequency details generated by the diffusion model. In contrast, a larger value of Σ_t leads to stronger modulation, which enforces more aggressive regularization to suppress artifacts and structural deviations.

4.4 State-Aware Trajectory Pruning

To alleviate the computational redundancy caused by fixed time steps during the stable evolution phase, LEADer introduces State-Aware Trajectory Pruning (SATP), which incorporates epistemic uncertainty into the temporal dimension to enable adaptive sampling acceleration. In the reverse sampling of diffusion models, a single-step state transition can be approximated via Taylor expansion. Specifically, the state transition equation from time step t to

$t - \Delta t$ can be evaluated as:

$$\mathbf{x}_{t-\Delta t} = \mathbf{x}_t - \Delta t \frac{d\mathbf{x}_t}{dt} + O\left((\Delta t)^2 \left\| \frac{d^2\mathbf{x}_t}{dt^2} \right\|\right). \tag{16}$$

The higher-order residual term characterizes the local evolution error, whose magnitude reflects the rate of change in the image state. We relate this quantity to the epistemic uncertainty Σ_t derived in Section 4.2. Specifically, the norm of the second-order derivative is proportional to the trace of Σ_t:

$$\left\| \frac{d^2\mathbf{x}_t}{dt^2} \right\| \propto \mathrm{Tr}\left(\mathbf{I} + (1-\bar{\alpha}_t)\nabla^2_{\mathbf{x}_t} \log p_t(\mathbf{x}_t)\right) \propto \mathrm{Tr}(\Sigma_t) \triangleq U_t, \tag{17}$$

where $U_t = \frac{1}{N}\mathrm{Tr}\left(\Sigma_t\right)$ denotes the global epistemic uncertainty at time step t. To balance reconstruction fidelity and computational efficiency, we introduce a constant information loss budget B that constrains the permissible local truncation error. for each step. Specifically, the error induced by a step of size Δt is bounded as:

$$\delta(\Delta t) \approx (\Delta t)^{\rho} \cdot U_t \leq B, \tag{18}$$

where $\rho \geq 1$ is the local error order of the sampler. In our implementation, we use the DDIM solver and set $\rho = 1$. By solving this error constraint, the maximum admissible step size Δt_{Active} can be derived, and the next evolution state node t_{Next} can be updated accordingly:

$$\begin{aligned} \Delta t_{\text{Active}} &= \max\left\{1, \lfloor \eta \left(\frac{B}{U_t + \boldsymbol{\epsilon}}\right)^{\frac{1}{\rho}} \rfloor\right\}, \\ t_{\text{Next}} &= \max\{0, t - \Delta t_{\text{Active}}\}, \end{aligned} \tag{19}$$

where η is a scaling factor, and $\boldsymbol{\epsilon}$ is a constant that ensures numerical stability. This formulation enables a principled trade-off between efficiency and accuracy. When the model is confident in the current estimate, it indicates that the image is in a stable restoration phase. At this point, the induced local truncation error remains small, thereby permitting larger step sizes and the skipping of redundant iterations. Conversely, in high-uncertainty regions associated with complex structural reconstruction, the step size is reduced to maintain reconstruction accuracy.

4.5 More Analysis

Algorithm Summary: In summary, the spatially optimal modulated prior estimate is computed via Eq.(14). To strictly ensure physical measurement consistency, we project it onto the null-space and fuse it with the range-space observation $\mathbf{y}$, which yields the clean image estimate:

$$\hat{\mathbf{x}}_{0|t} = \mathbf{A}^{\dagger}\mathbf{y} + (\mathbf{I} - \mathbf{A}^{\dagger}\mathbf{A})\bar{\mathbf{x}}_{0|t}. \tag{20}$$

Finally, following the non-Markovian formulation of DDIM, the next state $\mathbf{x}_{t_{\text{Next}}}$ can be derived using $\hat{\mathbf{x}}_{0|t}$ and the adaptively selected time step t_{Next}:

$$\mathbf{x}_{t_{\text{Next}}} = \sqrt{\bar{\alpha}_{t_{\text{Next}}}}\hat{\mathbf{x}}_{0|t} + \sqrt{1 - \bar{\alpha}_{t_{\text{Next}}} - \sigma^2_{t_{\text{Next}}}}\boldsymbol{\epsilon}_{\theta}(\mathbf{x}_t, t) + \sigma_{t_{\text{Next}}}\boldsymbol{\epsilon}_t, \tag{21}$$

where $\boldsymbol{\epsilon}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ is the sampling noise. The overall LEADer algorithm is summarized in Algorithm 1.

Plug-and-Play Analysis: Notably, the LEADer framework can be seamlessly integrated into existing baseline sampling methods as a plug-and-play module. The computations of UCPM and SATP rely entirely on the posterior statistics ($\mathbf{x}_{0|t}$ and Σ_t) of the diffusion model during the pure inference phase. This eliminates the need for any degradation-specific fine-tuning or retraining of the pre-trained denoising network $\boldsymbol{\epsilon}_{\theta}$. This design ensures broad compatibility with existing zero-shot image restoration frameworks, facilitating its application across diverse inverse imaging tasks without modifying the underlying model parameters.

Algorithm 1 Local Epistemic Uncertainty Guided Active Sampling

Require: Noise estimator $\boldsymbol{\epsilon}_{\theta}(\cdot, t)$, T, $\mathbf{y}$, $\mathbf{A}$, Φ, B, η, ρ, $\boldsymbol{\epsilon}$

Initialize $\mathbf{x}_T \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$
while $t > 0$ **do**
  // Pre-computation
  $\mathbf{x}_{0|t} = \frac{1}{\sqrt{\bar{\alpha}_t}}(\mathbf{x}_t - \sqrt{1-\bar{\alpha}_t}\boldsymbol{\epsilon}_{\theta}(\mathbf{x}_t, t))$
  $\mathcal{F}_t(\mathbf{x}_t) = -\nabla^2_{\mathbf{x}_t} \log p_t(\mathbf{x}_t)$
  $\Sigma_t = \frac{1-\bar{\alpha}_t}{\bar{\alpha}_t}[\mathbf{I} - (1-\bar{\alpha}_t)\mathcal{F}_t(\mathbf{x}_t)]$
  // Uncertainty-Calibrated Prior Modulation (UCPM)
  $\boldsymbol{\Lambda}^*_t = \Sigma_t(\mathbf{I} + \mathbf{H}_{\Phi}(\mathbf{x}_{0|t})\Sigma_t)^{-1}$
  $\bar{\mathbf{x}}_{0|t} = \mathbf{x}_{0|t} - \boldsymbol{\Lambda}^*_t \nabla\Phi(\mathbf{x}_{0|t})$
  $\hat{\mathbf{x}}_{0|t} = \mathbf{A}^{\dagger}\mathbf{y} + (\mathbf{I} - \mathbf{A}^{\dagger}\mathbf{A})\bar{\mathbf{x}}_{0|t}$
  // State-Aware Trajectory Pruning (SATP)
  $U_t = \frac{1}{N}\mathrm{Tr}(\Sigma_t)$
  $\Delta t_{\text{Active}} = \max\left\{1, \lfloor \eta \left(\frac{B}{U_t + \boldsymbol{\epsilon}}\right)^{\frac{1}{\rho}} \rfloor\right\}$
  $t_{\text{Next}} = \max\{0, t - \Delta t_{\text{Active}}\}$
  // Generative Sampling Step
  $\boldsymbol{\epsilon}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$
  $\mathbf{x}_{t_{\text{Next}}} = \sqrt{\bar{\alpha}_{t_{\text{Next}}}}\hat{\mathbf{x}}_{0|t} + \sqrt{1 - \bar{\alpha}_{t_{\text{Next}}} - \sigma^2_{t_{\text{Next}}}}\boldsymbol{\epsilon}_{\theta}(\mathbf{x}_t, t) + \sigma_{t_{\text{Next}}}\boldsymbol{\epsilon}_t$
  $t \leftarrow t_{\text{Next}}$
end while
return $\hat{\mathbf{x}}$

Theoretical Analysis: We now provide theoretical insights into the behavior of the UCPM and SATP mechanisms within the LEADer framework. In particular, we analyze how UCPM maintains data fidelity while introducing spatial prior modulation (Proposition 4.1), and how SATP guarantees a bounded global evolution error during adaptive temporal acceleration (Proposition 4.2). The proofs of the two Propositions are provided in the Supplementary Material.

Proposition 4.1. *For a linear degradation operator* $\mathbf{A}$ *and an observation* $\mathbf{y}$*, let the null-space projection be* $P_N = \mathbf{I} - \mathbf{A}^{\dagger}\mathbf{A}$*. Since the modulated estimation of UCPM is defined as* $\hat{\mathbf{x}} = \mathbf{A}^{\dagger}\mathbf{y} + P_N\mathbf{x}_{Prior}$*, under the noise-free assumption, the final generated image strictly satisfies the data fidelity condition:*

$$\|\mathbf{y} - \mathbf{A}\hat{\mathbf{x}}\|_2^2 = 0. \tag{22}$$

Since $\mathbf{A}P_N \equiv \mathbf{0}$, all prior modulation induced by the epistemic uncertainty Σ_t is strictly constrained within the null space. This implies that UCPM does not interfere with the known range-space components when suppressing artifacts or synthesizing details, thereby circumventing the common issue of data fidelity corruption caused by prior-driven updates.

Proposition 4.2. *Let* $\mathbf{x}^{Dense}$ *denote the continuous exact solution of the ODE and* $\mathbf{x}^{LEADer}$ *the discrete solution accelerated by SATP. Assuming the drift field is L-Lipschitz continuous and SATP bounds the single-step truncation error within the budget B. Based on Grönwall's inequality, the global accumulated error possesses a deterministic*

Table 1: Quantitative results for five image restoration tasks on CelebA-HQ 1K (top) and ImageNet 1K (bottom). The Avg. Δ column reports the average performance improvements and time reductions achieved by our method over the baselines.

Method	Bicub. SR × 4 ($\sigma_e = 0$)			Bicub. SR × 4 ($\sigma_e = 0.05$)			Gaussian Deb. ($\sigma_e = 0.05$)			Motion Deb. ($\sigma_e = 0.05$)			CS 25% ($\sigma_e = 0.05$)			Avg. Δ		
	PSNR ↑	LPIPS ↓	Time ↓	PSNR ↑	LPIPS ↓	Time ↓	PSNR ↑	LPIPS ↓	Time ↓	PSNR ↑	LPIPS ↓	Time ↓	PSNR ↑	LPIPS ↓	Time ↓	PSNR ↑	LPIPS ↓	Time ↓
CelebA-HQ 1K																		
DDRM [19] [NeurIPS2023]	$31.64_{\pm 2.43}$	$0.054_{\pm 0.021}$	3.87 s	$29.26_{\pm 1.86}$	$0.090_{\pm 0.028}$	4.05 s	$30.53_{\pm 1.74}$	$0.074_{\pm 0.024}$	4.06 s	N/A	N/A	N/A	$26.07_{\pm 2.08}$	$0.128_{\pm 0.037}$	3.86 s	N/A	N/A	N/A
DPS [5] [ICLR2023]	$29.39_{\pm 3.16}$	$0.065_{\pm 0.026}$	134.20 s	$27.49_{\pm 2.84}$	$0.086_{\pm 0.032}$	138.25 s	$27.75_{\pm 2.57}$	$0.084_{\pm 0.029}$	143.69 s	$19.63_{\pm 3.41}$	$0.227_{\pm 0.043}$	148.69 s	N/A	N/A	N/A	N/A	N/A	N/A
DiffPIR [54] [CVPR2023]	$30.36_{\pm 2.27}$	$0.051_{\pm 0.019}$	7.45 s	$27.44_{\pm 1.93}$	$0.085_{\pm 0.027}$	7.68 s	$28.89_{\pm 1.78}$	$0.074_{\pm 0.022}$	7.66 s	$27.96_{\pm 1.61}$	$0.102_{\pm 0.031}$	7.58 s	N/A	N/A	N/A	N/A	N/A	N/A
IDPG [14] [CVPR2024]	$32.66_{\pm 2.18}$	$0.111_{\pm 0.023}$	4.75 s	$29.89_{\pm 1.64}$	$0.155_{\pm 0.031}$	4.88 s	$31.08_{\pm 1.52}$	$0.150_{\pm 0.027}$	4.81 s	$29.73_{\pm 1.73}$	$0.134_{\pm 0.025}$	5.59 s	$26.26_{\pm 1.95}$	$0.143_{\pm 0.033}$	5.00 s	N/A	N/A	N/A
DDNM [41] [ICLR2023]	$31.64_{\pm 3.87}$	$0.048_{\pm 0.024}$	3.77 s	$29.19_{\pm 2.51}$	$0.082_{\pm 0.031}$	3.87 s	$30.60_{\pm 2.02}$	$0.072_{\pm 0.022}$	3.90 s	N/A	N/A	N/A	$26.14_{\pm 2.18}$	$0.124_{\pm 0.043}$	3.83 s	+2.382%	−12.781%	−5.747%
+ EquS* [43] [WACV2026]	$32.44_{\pm 2.58}$	$0.053_{\pm 0.019}$	3.83 s	$29.20_{\pm 1.38}$	$0.093_{\pm 0.027}$	3.87 s	$30.83_{\pm 1.56}$	$0.069_{\pm 0.021}$	3.96 s	N/A	N/A	N/A	$26.31_{\pm 2.45}$	$0.119_{\pm 0.035}$	4.02 s			
+ LEADer [Ours]	$32.73_{\pm 2.12}$	$0.044_{\pm 0.017}$	3.42 s	$29.69_{\pm 1.10}$	$0.077_{\pm 0.021}$	3.63 s	$30.86_{\pm 1.38}$	$0.063_{\pm 0.015}$	3.75 s	N/A	N/A	N/A	$27.06_{\pm 1.42}$	$0.094_{\pm 0.023}$	3.69 s			
DDPG [14] [CVPR2024]	$31.60_{\pm 2.36}$	$0.052_{\pm 0.021}$	4.96 s	$29.39_{\pm 1.74}$	$0.105_{\pm 0.029}$	5.09 s	$30.41_{\pm 1.58}$	$0.068_{\pm 0.023}$	5.11 s	$29.02_{\pm 1.69}$	$0.082_{\pm 0.026}$	5.72 s	$26.76_{\pm 1.95}$	$0.152_{\pm 0.034}$	5.24 s	+1.197%	−7.137%	−6.662%
+ EquS [43] [WACV2026]	$31.73_{\pm 2.08}$	$0.054_{\pm 0.019}$	4.88 s	$29.52_{\pm 1.53}$	$0.109_{\pm 0.027}$	5.01 s	$30.47_{\pm 1.42}$	$0.071_{\pm 0.021}$	5.03 s	$29.17_{\pm 1.55}$	$0.088_{\pm 0.024}$	5.62 s	$26.82_{\pm 1.76}$	$0.155_{\pm 0.032}$	5.15 s			
+ LEADer [Ours]	$31.94_{\pm 1.79}$	$0.047_{\pm 0.017}$	4.63 s	$29.65_{\pm 1.31}$	$0.098_{\pm 0.023}$	4.75 s	$30.77_{\pm 1.24}$	$0.062_{\pm 0.018}$	4.77 s	$29.40_{\pm 1.37}$	$0.076_{\pm 0.021}$	5.34 s	$27.17_{\pm 1.58}$	$0.144_{\pm 0.028}$	4.89 s			
ProjDiff [50] [NeurIPS2024]	$32.57_{\pm 1.88}$	$0.091_{\pm 0.022}$	3.94 s	$29.49_{\pm 1.42}$	$0.080_{\pm 0.024}$	4.12 s	$31.41_{\pm 1.35}$	$0.068_{\pm 0.020}$	4.27 s	N/A	N/A	N/A	$26.41_{\pm 1.57}$	$0.132_{\pm 0.031}$	4.05 s	+0.882%	−7.330%	−8.421%
+ EquS* [43] [WACV2026]	$32.64_{\pm 1.63}$	$0.094_{\pm 0.020}$	3.87 s	$29.53_{\pm 1.21}$	$0.080_{\pm 0.021}$	4.04 s	$31.45_{\pm 1.18}$	$0.071_{\pm 0.018}$	4.18 s	N/A	N/A	N/A	$26.45_{\pm 1.39}$	$0.135_{\pm 0.028}$	3.97 s			
+ LEADer [Ours]	$32.86_{\pm 1.34}$	$0.084_{\pm 0.017}$	3.61 s	$29.75_{\pm 1.05}$	$0.074_{\pm 0.018}$	3.77 s	$31.64_{\pm 0.98}$	$0.062_{\pm 0.015}$	3.90 s	N/A	N/A	N/A	$26.68_{\pm 1.16}$	$0.125_{\pm 0.023}$	3.72 s			
PIRP $F_{lb}G_-$ [47] [KBS2026]	$30.80_{\pm 2.11}$	$0.042_{\pm 0.018}$	5.01 s	$28.55_{\pm 1.57}$	$0.079_{\pm 0.024}$	5.14 s	$29.90_{\pm 1.46}$	$0.061_{\pm 0.020}$	5.12 s	$28.50_{\pm 1.63}$	$0.074_{\pm 0.022}$	5.77 s	$26.59_{\pm 1.74}$	$0.091_{\pm 0.027}$	5.22 s	+0.934%	−7.618%	−5.824%
+ EquS [43] [WACV2026]	$30.85_{\pm 1.86}$	$0.043_{\pm 0.016}$	4.94 s	$28.60_{\pm 1.39}$	$0.078_{\pm 0.021}$	5.06 s	$29.94_{\pm 1.28}$	$0.062_{\pm 0.018}$	5.04 s	$28.55_{\pm 1.47}$	$0.075_{\pm 0.020}$	5.67 s	$26.63_{\pm 1.58}$	$0.092_{\pm 0.024}$	5.14 s			
+ LEADer [Ours]	$31.02_{\pm 1.61}$	$0.039_{\pm 0.014}$	4.72 s	$28.86_{\pm 1.18}$	$0.072_{\pm 0.018}$	4.84 s	$30.16_{\pm 1.09}$	$0.055_{\pm 0.015}$	4.83 s	$28.77_{\pm 1.26}$	$0.069_{\pm 0.017}$	5.43 s	$26.87_{\pm 1.37}$	$0.086_{\pm 0.021}$	4.91 s			
ImageNet 1K																		
DDRM [19] [NeurIPS2023]	$27.38_{\pm 2.12}$	$0.270_{\pm 0.045}$	8.16 s	$25.54_{\pm 1.52}$	$0.333_{\pm 0.047}$	8.27 s	$27.71_{\pm 1.83}$	$0.243_{\pm 0.032}$	8.02 s	N/A	N/A	N/A	$21.58_{\pm 1.75}$	$0.301_{\pm 0.038}$	8.07 s	N/A	N/A	N/A
DPS [5] [ICLR2023]	$25.56_{\pm 1.85}$	$0.236_{\pm 0.029}$	246.14 s	$24.05_{\pm 1.43}$	$0.271_{\pm 0.045}$	263.95 s	$26.64_{\pm 1.57}$	$0.240_{\pm 0.032}$	248.26 s	$17.52_{\pm 1.78}$	$0.468_{\pm 0.066}$	275.41 s	N/A	N/A	N/A	N/A	N/A	N/A
DiffPIR [54] [CVPR2023]	$26.99_{\pm 1.92}$	$0.255_{\pm 0.033}$	15.40 s	$24.65_{\pm 1.32}$	$0.318_{\pm 0.044}$	15.62 s	$26.64_{\pm 1.74}$	$0.240_{\pm 0.029}$	15.49 s	$25.34_{\pm 1.52}$	$0.284_{\pm 0.038}$	18.19 s	N/A	N/A	N/A	N/A	N/A	N/A
IDPG [14] [CVPR2024]	$27.20_{\pm 1.68}$	$0.326_{\pm 0.047}$	8.42 s	$25.51_{\pm 1.22}$	$0.411_{\pm 0.054}$	8.53 s	$27.47_{\pm 1.83}$	$0.313_{\pm 0.045}$	8.46 s	$26.02_{\pm 1.56}$	$0.354_{\pm 0.048}$	9.77 s	$22.37_{\pm 1.29}$	$0.261_{\pm 0.033}$	8.48 s	N/A	N/A	N/A
ProjDiff [50] [NeurIPS2024]	$27.09_{\pm 1.53}$	$0.242_{\pm 0.031}$	8.13 s	$25.73_{\pm 1.11}$	$0.336_{\pm 0.046}$	8.15 s	$27.91_{\pm 1.78}$	$0.238_{\pm 0.029}$	8.11 s	N/A	N/A	N/A	$22.06_{\pm 1.26}$	$0.248_{\pm 0.033}$	8.32 s	N/A	N/A	N/A
DDNM [41] [ICLR2023]	$27.45_{\pm 2.12}$	$0.245_{\pm 0.035}$	8.03 s	$25.52_{\pm 1.48}$	$0.324_{\pm 0.047}$	8.06 s	$27.69_{\pm 2.05}$	$0.237_{\pm 0.027}$	7.97 s	N/A	N/A	N/A	$21.66_{\pm 1.35}$	$0.265_{\pm 0.041}$	8.20 s	+2.623%	−9.944%	−3.044%
+ EquS* [43] [WACV2026]	$27.63_{\pm 2.03}$	$0.242_{\pm 0.033}$	8.00 s	$25.56_{\pm 1.37}$	$0.327_{\pm 0.045}$	8.12 s	$27.72_{\pm 1.87}$	$0.234_{\pm 0.025}$	8.01 s	N/A	N/A	N/A	$22.14_{\pm 1.25}$	$0.246_{\pm 0.038}$	8.21 s			
+ LEADer [Ours]	$27.71_{\pm 1.92}$	$0.231_{\pm 0.029}$	7.74 s	$25.63_{\pm 1.28}$	$0.306_{\pm 0.038}$	7.86 s	$27.81_{\pm 1.65}$	$0.207_{\pm 0.022}$	7.66 s	N/A	N/A	N/A	$23.54_{\pm 1.36}$	$0.223_{\pm 0.030}$	8.02 s			
DDPG [14] [CVPR2024]	$27.41_{\pm 1.86}$	$0.255_{\pm 0.031}$	8.43 s	$25.55_{\pm 1.42}$	$0.354_{\pm 0.044}$	8.17 s	$27.73_{\pm 1.74}$	$0.205_{\pm 0.026}$	8.24 s	$25.94_{\pm 1.58}$	$0.249_{\pm 0.033}$	8.79 s	$20.69_{\pm 1.27}$	$0.258_{\pm 0.037}$	8.56 s	+1.079%	−4.914%	−3.169%
+ EquS [43] [WACV2026]	$27.44_{\pm 1.73}$	$0.252_{\pm 0.029}$	8.49 s	$25.63_{\pm 1.35}$	$0.351_{\pm 0.041}$	8.23 s	$27.73_{\pm 1.61}$	$0.203_{\pm 0.024}$	8.31 s	$25.99_{\pm 1.46}$	$0.247_{\pm 0.031}$	8.86 s	$20.71_{\pm 1.18}$	$0.255_{\pm 0.034}$	8.63 s			
+ LEADer [Ours]	$27.62_{\pm 1.52}$	$0.243_{\pm 0.024}$	8.18 s	$25.75_{\pm 1.21}$	$0.339_{\pm 0.036}$	7.95 s	$28.02_{\pm 1.43}$	$0.194_{\pm 0.021}$	8.00 s	$26.19_{\pm 1.33}$	$0.236_{\pm 0.028}$	8.51 s	$21.07_{\pm 1.09}$	$0.245_{\pm 0.030}$	8.21 s			
SITCOM [2] [ICML2025]	$27.26_{\pm 1.32}$	$0.186_{\pm 0.026}$	154.64 s	$25.35_{\pm 1.21}$	$0.232_{\pm 0.038}$	163.50 s	$27.40_{\pm 0.45}$	$0.236_{\pm 0.039}$	142.36 s	$28.65_{\pm 0.34}$	$0.189_{\pm 0.036}$	140.87 s	N/A	N/A	N/A	+0.809%	−4.828%	−5.470%
+ EquS [43] [WACV2026]	$27.30_{\pm 1.30}$	$0.181_{\pm 0.024}$	151.88 s	$25.36_{\pm 1.19}$	$0.228_{\pm 0.036}$	160.42 s	$27.38_{\pm 0.44}$	$0.232_{\pm 0.037}$	139.93 s	$28.66_{\pm 0.33}$	$0.185_{\pm 0.034}$	138.51 s	N/A	N/A	N/A			
+ LEADer [Ours]	$27.56_{\pm 1.28}$	$0.176_{\pm 0.022}$	145.97 s	$25.53_{\pm 1.16}$	$0.223_{\pm 0.033}$	154.36 s	$27.59_{\pm 0.43}$	$0.226_{\pm 0.035}$	134.68 s	$28.86_{\pm 0.32}$	$0.178_{\pm 0.031}$	133.42 s	N/A	N/A	N/A			
PIRP $F_{lb}G_-$ [47] [KBS2026]	$24.86_{\pm 1.34}$	$0.151_{\pm 0.024}$	8.55 s	$25.07_{\pm 1.52}$	$0.290_{\pm 0.041}$	8.56 s	$26.17_{\pm 1.28}$	$0.150_{\pm 0.022}$	8.69 s	$24.81_{\pm 1.67}$	$0.209_{\pm 0.035}$	9.83 s	$23.68_{\pm 1.45}$	$0.217_{\pm 0.031}$	8.16 s	+1.487%	−2.726%	−4.010%
+ EquS [43] [WACV2026]	$24.91_{\pm 1.21}$	$0.153_{\pm 0.021}$	8.58 s	$25.02_{\pm 1.39}$	$0.294_{\pm 0.038}$	8.61 s	$26.11_{\pm 1.16}$	$0.155_{\pm 0.020}$	8.73 s	$24.76_{\pm 1.48}$	$0.211_{\pm 0.032}$	9.88 s	$23.71_{\pm 1.33}$	$0.220_{\pm 0.029}$	8.19 s			
+ LEADer [Ours]	$25.37_{\pm 1.05}$	$0.146_{\pm 0.018}$	8.29 s	$25.34_{\pm 1.12}$	$0.278_{\pm 0.030}$	8.32 s	$26.28_{\pm 0.98}$	$0.152_{\pm 0.019}$	8.41 s	$25.24_{\pm 1.21}$	$0.201_{\pm 0.028}$	9.63 s	$24.19_{\pm 1.10}$	$0.209_{\pm 0.026}$	7.43 s			

analytical upper bound:

$$\|\mathbf{x}^{Dense} - \mathbf{x}^{LEADer}\|_2 \leq C \cdot B \cdot \frac{e^{LT} - 1}{L}, \tag{23}$$

where $C > 0$ is a constant, T is the total samplings steps of the reverse process, and L is the Lipschitz constant.

In the discrete ODE solution process, due to Lipschitz continuity, the local truncation error accumulates exponentially over time steps, yielding the e^{LT} term. By introducing the uncertainty trace U_t to adjust the sampling step size, SATP constrains the local error at each step within the predefined tolerance B. This state-aware local error control ensures that the global accumulated error remains bounded under non-uniform sampling, thereby guaranteeing the convergence of the accelerated generation trajectory.

5 Experiments

5.1 Experimental Setup

We conduct experiments on five typical IR tasks: 4× Super-Resolution (SR) with a bicubic downsampler ($\sigma_e = 0$ and $\sigma_e = 0.05$), Gaussian deblurring ($\sigma_e = 0.05$), motion deblurring ($\sigma_e = 0.05$), and Compressed Sensing (CS) using a Walsh-Hadamard sampling matrix with a 0.25 compression ratio. For validation, we use two standard image restoration benchmarks, i.e., CelebA-HQ 1K [23] and ImageNet 1K [30], both at 256×256 resolution. For a fair comparison, all methods utilize the same pre-trained DDM denoisers, i.e., one trained on CelebA-HQ [25] and one trained on ImageNet [10]. Sampling steps are uniformly set to $T = 100$ for all methods, except for DPS [5] ($T = 1000$) and SITCOM [2]. Note that due to the step-skipping capability of SATP, our total Number of Function Evaluations (NFEs) is less than 100. All experiments were performed on an NVIDIA GeForce RTX 3090 GPU.

Evaluation Metrics. We employ three prevalent metrics to evaluate the proposed method on image restoration tasks, which include Peak Signal-to-Noise Ratio (PSNR) for distortion evaluation, Learned Perceptual Image Patch Similarity (LPIPS) [53] for perceptual quality, and average per-image Run-time (Time) for computational efficiency. We additionally report memory consumption to assess resource usage. Standard deviations of PSNR and LPIPS are provided to reflect performance stability.

Comparison with Other Methods. To evaluate the effectiveness of LEADer, we integrate it as a plug-and-play module into several advanced zero-shot DMIR approaches, including DDNM [41], IDPG [14], DDPG [14], ProjDiff [50], SITCOM [2], and PIRP [47]. We apply the same integration for EquS [43] and compare their respective effects on quality and efficiency across these baselines. Additionally, we compare LEADer against other SOTA zero-shot DMIR methods, including DDRM [19], DPS [5] and DiffPIR [54].

5.2 Improvements to DMIR Methods

Quantitative Results. To validate the effectiveness of LEADer as a plug-and-play method for DMIR, we conduct detailed quantitative evaluations across five image restoration tasks on CelebA-HQ 1K and ImageNet 1K. As shown in Table 1, experimental results confirm the superior performance of LEADer. When integrated into various zero-shot DMIR methods, such as DDNM [41], DDPG [14], ProjDiff [50], SITCOM [2], and PIRP [47], LEADer consistently enhances restoration quality. Specifically, on CelebA-HQ 1K, LEADer achieves average PSNR gains ranging from 0.88% to 2.38%, while

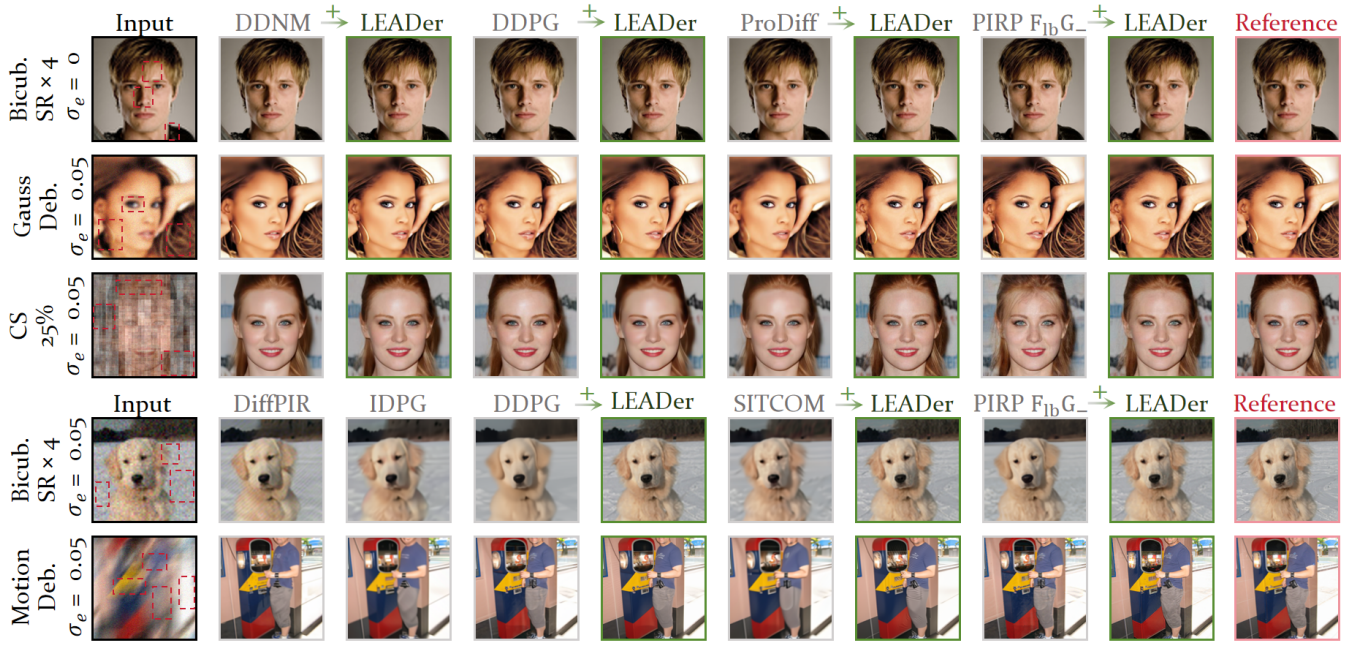

Figure 3: Qualitative comparisons on five image restoration tasks.

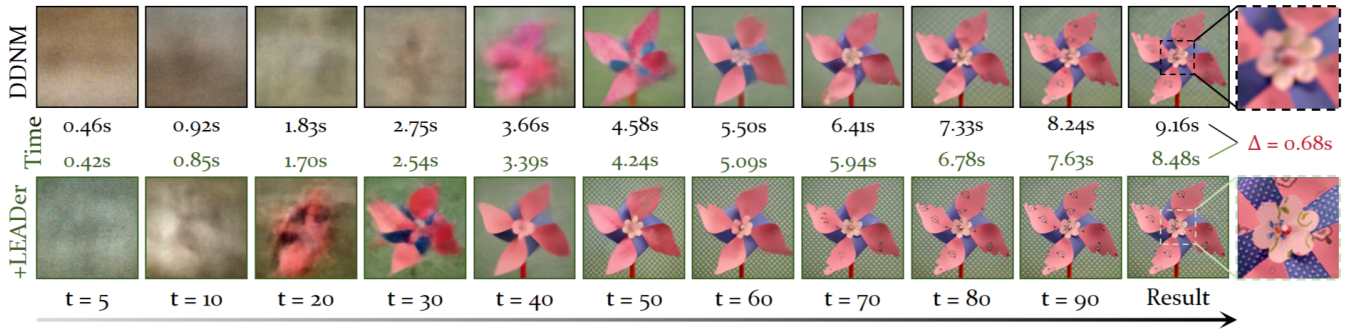

Figure 4: Comparison of $\hat{x}_{0|t}$ between DDNM and LEADer at different NFEs.

significantly reducing LPIPS by 7.13% to 12.78%. Notably, unlike EquS [43], LEADer leverages a state-aware trajectory pruning strategy to improve generation quality while achieving substantial sampling speedups, saving 3.04% to 8.42% of sampling time on average. These results demonstrate that our local epistemic uncertainty quantification directly strengthens the adaptive adjustment for complex local structures while significantly boosting inference efficiency.

Qualitative Results. To assess the visual restoration capability of LEADer under complex degradations, we compare the visual results of various baseline methods before and after its integration, as shown in Figure 3. When confronted with complex localized degradations, baseline methods often suffer from detail loss or over-smoothing due to their fixed prior weights. By integrating LEADer, the model can adaptively modulate the prior strength, which effectively suppresses artifacts while recovering sharper edges and more realistic textures, thus leading to a substantial improvement in overall visual fidelity.

To further illustrate the effect of the uncertainty-based active sampling strategy, we visualize the restoration trajectories of $\hat{x}_{0|t}$ across different steps on the 4× super-resolution task, as shown in Figure 4. Due to its fixed step size, the baseline DDNM recovers details through a slow and uniform process, where both intermediate states and final outputs often exhibit high uncertainty. In contrast, LEADer perceives the current state and dynamically adjusts the sampling pace, which enables earlier convergence to clear local textures at lower steps. This enhanced sensitivity to detail not only improves data consistency but also allows the SATP strategy to safely skip steps, thereby reducing inference time.

5.3 Ablation Study

Impact of Core Strategies. To validate the effectiveness of the key strategies in the LEADer framework, we conduct ablation studies on three image restoration tasks using the DDNM [41] baseline. As shown in Table 2, introducing UCPM alone significantly improves

Table 2: Ablation study results of each component in LEADer.

DDNM			Bicub. SR × 4 ($\sigma_e = 0.05$)			Gaussian Deb. ($\sigma_e = 0.05$)			CS 25% ($\sigma_e = 0.05$)		
UCPM	SATP	T	PSNR ↑	LPIPS ↓	Time ↓	PSNR ↑	LPIPS ↓	Time ↓	PSNR ↑	LPIPS ↓	Time ↓
×	×	100	29.19	0.082	3.87 s	30.60	0.072	3.90 s	26.14	0.124	3.83 s
✓	×	100	29.69	0.075	3.96 s	30.90	0.061	4.01 s	27.14.	0.090	3.97 s
×	×	93	28.85	0.098	3.67 s	30.31	0.089	3.83 s	25.79	0.134	3.75 s
×	✓	100	29.17	0.084	3.66 s	30.58	0.077	3.80 s	26.10	0.125	3.73 s
✓	×	93	29.35	0.090	3.65 s	30.58	0.074	3.75 s	26.74	0.106	3.70 s
✓	✓	100	29.69	0.077	3.63 s	30.86	0.063	3.75 s	27.06	0.094	3.69 s

Table 3: Hyperparameter sensitivity analysis on the information loss budget B.

ImageNet 1K	Bicub. SR × 4 ($\sigma_e = 0$)			Motion Deb. ($\sigma_e = 0.05$)		
Setting	PSNR↑	LPIPS↓	Time↓	PSNR↑	LPIPS↓	Time↓
$B = 0$	25.41	0.143	8.56	25.31	0.196	9.83
$B = 0.001$	25.41	0.143	8.56	25.30	0.196	9.83
$B = 0.005$	25.38	0.144	8.45	25.26	0.199	9.74
$B = 0.01$	25.37	0.146	8.29	25.24	0.201	9.63
$B = 0.05$	24.83	0.161	4.54	24.71	0.226	5.21
$B = 0.1$	23.96	0.187	2.76	23.85	0.258	3.08

the restoration quality, albeit with a slight increase in sampling time. Conversely, employing SATP alone effectively reduces inference time with virtually no quality degradation. When combined, they exhibit excellent complementarity, achieving superior restoration quality while simultaneously reducing the overall sampling time. This indicates that the effective local guidance provided by UCPM enhances the stability of reverse sampling, thereby enabling SATP to perform step-skipping more safely. Furthermore, to demonstrate the advantage of the SATP strategy over directly reducing sampling steps, we conduct a step ablation study. The results show that, under the condition of roughly equivalent sampling time, our active sampling strategy demonstrates significant robustness. These experimental results demonstrate that both UCPM and SATP are indispensable, yielding clear synergy and mutual enhancement across spatiotemporal dimensions.

Impact of Sampling Steps. To evaluate the robustness of our method across varying sampling steps, we compare the performance of DDNM [41] and DDPG [14] with and without LEADer at $T \in \{10, 20, 50, 100\}$. As shown in Figure 5, introducing LEADer consistently improves image quality across all settings, thus demonstrating the effectiveness of our UCPM. In terms of efficiency, SATP consistently reduces inference time. However, under extremely low step counts, the runtime reduction becomes marginal. This is because the overall epistemic uncertainty remains excessively high under such sparse sampling; consequently, SATP adaptively performs fewer skip operations to preserve restoration quality. Overall, these results verify that LEADer exhibits strong robustness across different sampling steps.

Impact of B. The information loss budget B in Equation (18) governs the trade-off between restoration quality and computational efficiency. A larger B tolerates higher local truncation errors, thereby encouraging SATP to execute more aggressive trajectory pruning. To investigate the impact of B, we conduct a sensitivity analysis using PIRP $F_{lb}G_{-}$ [47] as the baseline on the ImageNet 1K dataset for the Bicubic SR ×4 ($\sigma_e = 0$) and Motion Deblurring ($\sigma_e = 0.05$) tasks. As shown in Table 3, a conservative budget ($B = 0.001$) yields quality improvements while restricting acceleration. Conversely, an overly aggressive budget ($B = 0.1$), despite significantly reducing inference time, leads to notable performance degradation. Based on these observations, we set $B = 0.01$ as the default configuration for our SATP strategy to optimally balance restoration fidelity and sampling efficiency.

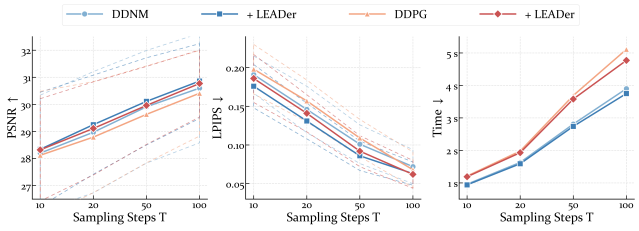

Figure 5: Ablation study on different sampling steps. (Dataset: CelebA-HQ 1K, Task: Gaussion Deb. & $\sigma = 0.05$)

Table 4: Memory consumption comparison across five tasks.

Methods	Bicub. SR × 4 ($\sigma_e = 0$)	Bicub. SR × 4 ($\sigma_e = 0.05$)	Gaussian Deb. ($\sigma_e = 0.05$)	Motion Deb. ($\sigma_e = 0.05$)	CS 25% ($\sigma_e = 0.05$)
DDNM [41]	2507 MB	2513 MB	2509 MB	N/A	2541 MB
+ LEADer	2509 MB	2516 MB	2515 MB	N/A	2543 MB
DDPG [14]	2521 MB	2524 MB	2523 MB	2959 MB	2564 MB
+ LEADer	2526 MB	2526 MB	2530 MB	2965 MB	2567 MB
SITCOM [2]	15831 MB	15863 MB	15912 MB	18478 MB	N/A
+ LEADer	15842 MB	15872 MB	15923 MB	18485 MB	N/A
PIRP $F_{lb}G_{-}$ [47]	2513 MB	2520 MB	2515 MB	2948 MB	2557 MB
+ LEADer	2516 MB	2521 MB	2521 MB	2952 MB	2561 MB

5.4 Memory Consumption

To evaluate the practical overhead of LEADer as a plug-and-play module, we compare the GPU memory consumption of various baseline methods before and after its integration, as shown in Table 4. Evaluations across five tasks on the ImageNet 1K dataset reveal that despite introducing additional computations, LEADer incurs a marginal overall GPU memory increase of merely 0.07% to 0.28%. This indicates that LEADer can robustly enhance both the restoration quality and sampling efficiency of existing methods at an almost negligible memory cost.

6 Conclusion

In this paper, we identify a fundamental limitation of existing zero-shot DMIR methods, i.e., their reliance on fixed prior constraints and uniform sampling schedules, which often leads to local structural distortions and redundant computations. To address this, we propose LEADer, an active diffusion sampling framework guided by local epistemic uncertainty. Spatially, we leverage pixel-level uncertainty to dynamically modulate prior strength, which effectively balances detail preservation and artifact suppression. Temporally, we quantify the stability of the generation process via uncertainty traces and adaptively prune trajectories within deterministic error bounds. Both theoretical analysis and extensive experiments demonstrate that LEADer can be seamlessly integrated as a plug-and-play module to consistently improve restoration quality and sampling efficiency, while preserving strict data consistency. In the future, it would be also interesting to investigate the application of this uncertainty-based active sampling mechanism to diffusion-model-based generative tasks.

References

[1] Yasar Utku Alçalar and Mehmet Akçakaya. 2024. Zero-Shot Adaptation for Approximate Posterior Sampling of Diffusion Models in Inverse Problems. In *European Conference on Computer Vision (Lecture Notes in Computer Science, Vol. 15141)*. Springer, 444–460.

[2] Ismail Alkhouri, Shijun Liang, Cheng-Han Huang, Jimmy Dai, Qing Qu, Saiprasad Ravishankar, and Rongrong Wang. 2025. SITCOM: Step-wise Triple-Consistent Diffusion Sampling For Inverse Problems. In *International Conference on Machine Learning*, Vol. 267. PMLR / OpenReview.net.

[3] Yochai Blau and Tomer Michaeli. 2018. The Perception-Distortion Tradeoff. In *IEEE Conference on Computer Vision and Pattern Recognition*. Computer Vision Foundation / IEEE Computer Society, 6228–6237.

[4] Liangyu Chen, Xiaojie Chu, Xiangyu Zhang, and Jian Sun. 2022. Simple Baselines for Image Restoration. In *European Conference on Computer Vision (Lecture Notes in Computer Science, Vol. 13667)*. Springer, 17–33.

[5] Hyungjin Chung, Jeongsol Kim, Michael Thompson McCann, Marc Louis Klasky, and Jong Chul Ye. 2023. Diffusion Posterior Sampling for General Noisy Inverse Problems. In *International Conference on Learning Representations*. OpenReview.net.

[6] Hyungjin Chung, Jeongsol Kim, and Jong Chul Ye. 2023. Direct Diffusion Bridge using Data Consistency for Inverse Problems. In *Annual Conference on Neural Information Processing Systems*.

[7] Hyungjin Chung, Byeongsu Sim, Dohoon Ryu, and Jong Chul Ye. 2022. Improving Diffusion Models for Inverse Problems using Manifold Constraints. In *Annual Conference on Neural Information Processing Systems*.

[8] Florinel-Alin Croitoru, Vlad Hondru, Radu Tudor Ionescu, and Mubarak Shah. 2023. Diffusion Models in Vision: A Survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 45, 9 (2023), 10850–10869.

[9] Kostadin Dabov, Alessandro Foi, Vladimir Katkovnik, and Karen O. Egiazarian. 2007. Image Denoising by Sparse 3-D Transform-Domain Collaborative Filtering. *IEEE Transactions on Image Processing* 16, 8 (2007), 2080–2095.

[10] Prafulla Dhariwal and Alexander Quinn Nichol. 2021. Diffusion Models Beat GANs on Image Synthesis. In *Annual Conference on Neural Information Processing Systems*. 8780–8794.

[11] Chao Dong, Chen Change Loy, Kaiming He, and Xiaoou Tang. 2016. Image Super-Resolution Using Deep Convolutional Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 38, 2 (2016), 295–307.

[12] Bradley Efron. 2011. Tweedie's formula and selection bias. *J. Amer. Statist. Assoc.* 106, 496 (2011), 1602–1614.

[13] Ben Fei, Zhaoyang Lyu, Liang Pan, Junzhe Zhang, Weidong Yang, Tianyue Luo, Bo Zhang, and Bo Dai. 2023. Generative Diffusion Prior for Unified Image Restoration and Enhancement. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition*. IEEE, 9935–9946.

[14] Tomer Garber and Tom Tirer. 2024. Image Restoration by Denoising Diffusion Models with Iteratively Preconditioned Guidance. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition*. IEEE, 25245–25254.

[15] Shuhang Gu, Lei Zhang, Wangmeng Zuo, and Xiangchu Feng. 2014. Weighted Nuclear Norm Minimization with Application to Image Denoising. In *IEEE Conference on Computer Vision and Pattern Recognition*. IEEE Computer Society, 2862–2869.

[16] Linchao He, Hongyu Yan, Mengting Luo, Hongjie Wu, Kunming Luo, Wang Wang, Wenchao Du, Hu Chen, Hongyu Yang, Yi Zhang, et al. 2023. Fast and stable diffusion inverse solver with history gradient update. *CoRR* abs/2307.12070 (2023).

[17] Jonathan Ho, Ajay Jain, and Pieter Abbeel. 2020. Denoising Diffusion Probabilistic Models. In *Annual Conference on Neural Information Processing Systems*.

[18] Tero Karras, Miika Aittala, Timo Aila, and Samuli Laine. 2022. Elucidating the Design Space of Diffusion-Based Generative Models. In *Annual Conference on Neural Information Processing Systems*.

[19] Bahjat Kawar, Michael Elad, Stefano Ermon, and Jiaming Song. 2022. Denoising Diffusion Restoration Models. In *Annual Conference on Neural Information Processing Systems*.

[20] Christian Ledig, Lucas Theis, Ferenc Huszar, Jose Caballero, Andrew Cunningham, Alejandro Acosta, Andrew P. Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, and Wenzhe Shi. 2017. Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network. In *IEEE Conference on Computer Vision and Pattern Recognition*. IEEE Computer Society, 105–114.

[21] Xiang Li, Soo Min Kwon, Ismail R. Alkhouri, Saiprasad Ravishankar, and Qing Qu. 2024. Decoupled Data Consistency with Diffusion Purification for Image Restoration. *CoRR* abs/2403.06054 (2024).

[22] Jingyun Liang, Jiezhang Cao, Guolei Sun, Kai Zhang, Luc Van Gool, and Radu Timofte. 2021. SwinIR: Image Restoration Using Swin Transformer. In *IEEE/CVF International Conference on Computer Vision Workshops*. IEEE, 1833–1844.

[23] Ziwei Liu, Ping Luo, Xiaogang Wang, and Xiaoou Tang. 2015. Deep Learning Face Attributes in the Wild. In *International Conference on Computer Vision*. IEEE Computer Society, 3730–3738.

[24] Cheng Lu, Yuhao Zhou, Fan Bao, Jianfei Chen, Chongxuan Li, and Jun Zhu. 2022. DPM-Solver: A Fast ODE Solver for Diffusion Probabilistic Model Sampling in Around 10 Steps. In *Annual Conference on Neural Information Processing Systems*.

[25] Andreas Lugmayr, Martin Danelljan, Andrés Romero, Fisher Yu, Radu Timofte, and Luc Van Gool. 2022. RePaint: Inpainting using Denoising Diffusion Probabilistic Models. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition*. IEEE, 11451–11461.

[26] Morteza Mardani, Jiaming Song, Jan Kautz, and Arash Vahdat. 2024. A Variational Perspective on Solving Inverse Problems with Diffusion Models. In *International Conference on Learning Representations*. OpenReview.net.

[27] Michael T. McCann, Kyong Hwan Jin, and Michael Unser. 2017. Convolutional Neural Networks for Inverse Problems in Imaging: A Review. *IEEE Signal Processing Magazine* 34, 6 (2017), 85–95.

[28] Mang Ning, Enver Sangineto, Angelo Porrello, Simone Calderara, and Rita Cucchiara. 2023. Input Perturbation Reduces Exposure Bias in Diffusion Models. In *International Conference on Machine Learning (Proceedings of Machine Learning Research)*. PMLR, 26245–26265.

[29] Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. 2022. High-Resolution Image Synthesis with Latent Diffusion Models. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition*. IEEE, 10674–10685.

[30] Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael S. Bernstein, Alexander C. Berg, and Li Fei-Fei. 2015. ImageNet Large Scale Visual Recognition Challenge. *Int. J. Comput. Vis.* 115 (2015), 211–252.

[31] Jascha Sohl-Dickstein, Eric A. Weiss, Niru Maheswaranathan, and Surya Ganguli. 2015. Deep Unsupervised Learning using Nonequilibrium Thermodynamics. In *International Conference on Machine Learning (JMLR Workshop and Conference Proceedings, Vol. 37)*, Francis R. Bach and David M. Blei (Eds.). JMLR.org, 2256–2265.

[32] Jiaming Song, Chenlin Meng, and Stefano Ermon. 2021. Denoising Diffusion Implicit Models. In *International Conference on Learning Representations*. OpenReview.net.

[33] Jiaming Song, Arash Vahdat, Morteza Mardani, and Jan Kautz. 2023. Pseudoinverse-Guided Diffusion Models for Inverse Problems. In *International Conference on Learning Representations*. OpenReview.net.

[34] Yang Song and Stefano Ermon. 2019. Generative Modeling by Estimating Gradients of the Data Distribution. In *Annual Conference on Neural Information Processing Systems*. 11895–11907.

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

[36] Yang Song, Jascha Sohl-Dickstein, Diederik P. Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. 2021. Score-Based Generative Modeling through Stochastic Differential Equations. In *International Conference on Learning Representations*. OpenReview.net.

[37] Charles M Stein. 1981. Estimation of the mean of a multivariate normal distribution. *The Annals of Statistics* (1981), 1135–1151.

[38] Darshan Thaker, Abhishek Goyal, and René Vidal. 2024. Frequency-Guided Posterior Sampling for Diffusion-Based Image Restoration. *CoRR* abs/2411.15295 (2024).

[39] Chong Wang, Lanqing Guo, Zixuan Fu, Siyuan Yang, Hao Cheng, Alex C. Kot, and Bihan Wen. 2025. Reconciling Stochastic and Deterministic Strategies for Zero-shot Image Restoration using Diffusion Model in Dual. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Computer Vision Foundation / IEEE, 23207–23216.

[40] Ge Wang, Jong Chul Ye, Klaus Mueller, and Jeffrey A. Fessler. 2018. Image Reconstruction is a New Frontier of Machine Learning. *IEEE Transactions on Medical Imaging* 37, 6 (2018), 1289–1296.

[41] Yinhuai Wang, Jiwen Yu, and Jian Zhang. 2023. Zero-Shot Image Restoration Using Denoising Diffusion Null-Space Model. In *International Conference on Learning Representations*. OpenReview.net.

[42] Zhihao Wang, Jian Chen, and Steven C. H. Hoi. 2021. Deep Learning for Image Super-Resolution: A Survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 43, 10 (2021), 3365–3387.

[43] Chenxu Wu, Qingpeng Kong, Peiang Zhao, Wendi Yang, Wenxin Ma, Fenghe Tang, Zihang Jiang, and S. Kevin Zhou. 2025. Equivariant Sampling for Improving Diffusion Model-based Image Restoration. *CoRR* abs/2511.09965 (2025).

[44] Hongjie Wu, Linchao He, Mingqin Zhang, Dongdong Chen, Kunming Luo, Mengting Luo, Jizhe Zhou, Hu Chen, and Jiancheng Lv. 2024. Diffusion Posterior Proximal Sampling for Image Restoration. In *ACM International Conference on Multimedia*. ACM, 214–223.

[45] Hongjie Wu, Mingqin Zhang, Linchao He, Ji-Zhe Zhou, and Jiancheng Lv. 2025. Enhancing Diffusion Model Stability for Image Restoration via Gradient Management. In *ACM International Conference on Multimedia*. ACM, 10768–10777.

[46] Ling Yang, Zhilong Zhang, Yang Song, Shenda Hong, Runsheng Xu, Yue Zhao, Wentao Zhang, Bin Cui, and Ming-Hsuan Yang. 2024. Diffusion Models: A Comprehensive Survey of Methods and Applications. *Comput. Surveys* 56, 4

(2024), 105:1–105:39.

[47] Yang Yang, Xi Zhang, Jiaqi Zhang, and Lanling Zeng. 2026. Parameterized image restoration with diffusion and gradient priors. *Knowledge Based Systems* 338 (2026), 115488.

[48] Syed Waqas Zamir, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, and Ming-Hsuan Yang. 2022. Restormer: Efficient Transformer for High-Resolution Image Restoration. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition*. IEEE, 5718–5729.

[49] Bingliang Zhang, Wenda Chu, Julius Berner, Chenlin Meng, Anima Anandkumar, and Yang Song. 2025. Improving Diffusion Inverse Problem Solving with Decoupled Noise Annealing. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition*. IEEE, 20895–20905.

[50] Jiawei Zhang, Jiaxin Zhuang, Cheng Jin, Gen Li, and Yuantao Gu. 2024. Unleashing the Denoising Capability of Diffusion Prior for Solving Inverse Problems. In *Annual Conference on Neural Information Processing Systems*.

[51] Kai Zhang, Yawei Li, Wangmeng Zuo, Lei Zhang, Luc Van Gool, and Radu Timofte. 2022. Plug-and-Play Image Restoration With Deep Denoiser Prior. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 44, 10 (2022), 6360–6376.

[52] Kai Zhang, Wangmeng Zuo, Yunjin Chen, Deyu Meng, and Lei Zhang. 2017. Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising. *IEEE Transactions on Image Processing* 26, 7 (2017), 3142–3155.

[53] Richard Zhang, Phillip Isola, Alexei A. Efros, Eli Shechtman, and Oliver Wang. 2018. The Unreasonable Effectiveness of Deep Features as a Perceptual Metric. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 586–595.

[54] Yuanzhi Zhu, Kai Zhang, Jingyun Liang, Jiezhang Cao, Bihan Wen, Radu Timofte, and Luc Van Gool. 2023. Denoising Diffusion Models for Plug-and-Play Image Restoration. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition*. IEEE, 1219–1229.