Analyzing and Guiding Zero-Shot Posterior Sampling in Diffusion Models

Roi Benita
Department of Electrical and Computer Engineering
Technion, Haifa, Israel
roibenita@campus.technion.ac.il

Michael Elad
Department of Computer Science
Technion, Haifa, Israel
elad@cs.technion.ac.il

Joseph Keshet
Department of Electrical and Computer Engineering
Technion, Haifa, Israel
jkeshet@technion.ac.il

Abstract

Recovering a signal from its degraded measurements is a long standing challenge in science and engineering. Recently, zero-shot diffusion based methods have been proposed for such inverse problems, offering a posterior sampling based solution that leverages prior knowledge. Such algorithms incorporate the observations through inference, often leaning on manual tuning and heuristics. In this work we propose a rigorous analysis of these approximate posterior samplers, relying on a Gaussianity assumption of the prior. Under this regime, we show that both sampling with the ideal posterior denoiser and diffusion-based reconstruction algorithms can be expressed in closed-form, enabling their thorough analysis and comparisons in the spectral domain. Building on these representations, we introduce a principled framework for parameter design, replacing heuristic selection strategies used to date. The proposed approach is method-agnostic and yields tailored parameter choices that jointly account for the characteristics of the prior, the degraded signal, and the diffusion dynamics. We show that our spectral recommendations differ structurally from standard heuristics and vary with the diffusion step size, resulting in a consistent balance between perceptual quality and signal fidelity.

1 Introduction

Linear inverse problems, a common task across many scientific fields, refer to the recovery of an unknown signal from its noisy, degraded measurements. Traditionally, for a given degradation model, one can learn a direct mapping from observations to clean signals using large paired datasets [14, 56, 33]. However, real-world applications, such as medical imaging, computational photography, and audio processing, often require more adaptable approaches for handling a potentially wide range of degradations [20, 45, 50, 55, 7]. Recently, diffusion-based zero-shot posterior samplers have been proposed to bridge this gap, avoiding task-specific training [11].

In this unsupervised setting, sampling is based on a pretrained diffusion model learned on the clean signal manifold, with different strategies for incorporating observations during inference. One perspective steers the synthesis process by explicitly projecting intermediate samples onto geometric constraints induced by the measurements [8, 39, 58]. Another common approach adopts a Bayesian formulation, where posterior estimation involves a likelihood component [2, 48, 9, 26, 31]. This

40th Conference on Neural Information Processing Systems (NeurIPS 2026).

guidance term is method-dependent and reflects the underlying degradation model, directing the denoising process toward consistency with the measurements.[1]

Alongside their remarkable success, a fundamental question within these methods is how to determine the relative weight between both forces, the likelihood and the prior information. One strategy is using small guidance step sizes, allowing samples to be gradually influenced by the observations [9, 52, 2]. However, this typically requires a large number of diffusion steps, slows inference and may result in poor measurement consistency [51, 54]. Alternatively, large guidance weights can induce unstable or unpredictable behavior, potentially driving samples away from the posterior [19].

In the absence of a principled solution, existing methods choose these weighting parameters manually or heuristically [11]. One approach relies on empirically chosen, dataset-dependent scaling factors adjusted based on the discrepancy with the observations [9]. Another approach instead uses globally defined parameters determined by the diffusion noise levels [35, 2, 58]. Notably, current design choices do not explicitly account for the interaction between prior characteristics, sampling dynamics, and the observed information. This can affect performance, reflected in the balance between sample naturalness and measurement fidelity [5], and may lead to longer sampling trajectories.

In this work, we propose a structured analysis of zero-shot posterior samplers, inspired by the spectral perspective introduced in [4]. Under a Gaussian prior assumption, we derive closed-form expressions for the ideal posterior denoiser and diffusion-based training-free reconstruction methods. These analytical representations provide a sound basis for posterior-level comparison across existing samplers. In addition, we introduce a principled way of designing the guidance weighting parameters while jointly accounting for the prior characteristics, the degraded observations, and the diffusion dynamics. Finally, we validate our framework on real-world datasets, including FFHQ [24], ImageNet [12], and fastMRI [53, 27], comparing our spectral recommendations with existing heuristics while evaluating image naturalness and measurement fidelity across varied diffusion step sizes.

In summary, our contributions are the following: (i) We introduce a structured spectral analysis of zero-shot posterior sampling methods, providing a unified view under a common analytical framework. (ii) Under a Gaussian prior assumption, we derive closed-form expressions for the output distributions induced by the optimal posterior denoiser and diffusion-based training-free reconstruction methods, enabling their direct posterior-level comparison and analysis. (iii) We formulate a principled optimization problem for selecting guidance weighting parameters which jointly accounts for the prior characteristics, the degraded observations, and the diffusion dynamics. We demonstrate how it can be instantiated across different posterior sampling algorithms and solved either for individual realizations or efficiently for a general degradation system. (iv) We evaluate our spectral recommendations relative to existing heuristics, considering both averaged and fixed-observation settings, and assessing sample naturalness, measurement fidelity, and dependence on the number of diffusion steps.

2 Background

Linear inverse problems. A general linear inverse problem is defined as:

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

where the goal is to recover an unknown signal $\mathbf{x} \in \mathbb{R}^d$ from a degraded observation $\mathbf{y} \in \mathbb{R}^m$. Here, $\mathbf{H} \in \mathbb{R}^{m \times d}$ is a known linear degradation matrix, and $\mathbf{n} \sim \mathcal{N}(\mathbf{0}, \sigma_y^2 \mathbf{I})$ denotes i.i.d. additive Gaussian noise. A common property of these problems is the loss of information induced by the degradation operator $\mathbf{H}$, making perfect recovery generally impossible and rendering the problem ill-posed.

Denoising diffusion probabilistic models. Diffusion models form a generative framework that transforms a simple normal distribution into a target data distribution $p(\mathbf{x}_0)$, thereby enabling the synthesis of samples $\mathbf{x}_0 \in \mathbb{R}^d$. As originally formulated in DDPM [22], the diffusion process consists of two Markovian paths, the *forward* and *reverse* processes, each comprising a sequence of T latent variables. The *forward* process gradually corrupts a clean signal $\mathbf{x}_0$ through additive Gaussian noise:

$$\mathbf{x}_t = \sqrt{\alpha_t}\mathbf{x}_{t-1} + \sqrt{1-\alpha_t}\boldsymbol{\epsilon}_t, \tag{2}$$

[1] Throughout this work, we adopt a unified prior-likelihood terminology [11]

where $\{\alpha_t\}_{t=1}^T$ denotes a monotonically decreasing noise schedule and $\boldsymbol{\epsilon}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$. A key property of the forward process is the marginal distribution of $\mathbf{x}_t$ given $\mathbf{x}_0$:

$$\mathbf{x}_t = \sqrt{\bar{\alpha}_t}\mathbf{x}_0 + \sqrt{1-\bar{\alpha}_t}\boldsymbol{\epsilon} \quad \boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}) \,, \tag{3}$$

with $\bar{\alpha}_t = \prod_{i=1}^t \alpha_i$, and for sufficiently large T, the terminal state approaches $\mathbf{x}_T \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$.

The *reverse* process aims to reconstruct $\mathbf{x}_0$ from $\mathbf{x}_T$ by iteratively denoising it. Extending the stochastic DDPM formulation [22], DDIM [34] relaxes the Markovian assumption to enable deterministic and accelerated sampling. As a result, the reverse trajectory is defined by:

$$\mathbf{x}_{s-1} = \sqrt{\bar{\alpha}_{s-1}}\left(\frac{\mathbf{x}_s - \sqrt{1-\bar{\alpha}_s}\cdot\boldsymbol{\epsilon}_\theta(\mathbf{x}_s, s)}{\sqrt{\bar{\alpha}_s}}\right) + \sqrt{1-\bar{\alpha}_{s-1}-\sigma_s^2}\cdot\boldsymbol{\epsilon}_\theta(\mathbf{x}_s, s) + \sigma_s\epsilon_s, \tag{4}$$

where $s \in \{0, 1, \ldots, S\}$ indexes a subsequence of the original T diffusion steps, $\sigma_s = \eta\sqrt{(1-\alpha_{s-1})/(1-\alpha_s)}\sqrt{1-\alpha_s/\alpha_{s-1}}$, and $\boldsymbol{\epsilon}_\theta(\mathbf{x}_s, s)$ denotes a neural network parameterized by θ, trained to estimate $\boldsymbol{\epsilon}$ from a given $\mathbf{x}_s$ at time s. When $\eta = 1$, the generative process reduces to DDPM, whereas $\eta = 0$ yields the deterministic DDIM sampler. In this work, we denote the cumulative noise schedule by $\bar{\boldsymbol{\alpha}} = \{\bar{\alpha}_t\}_{t=1}^T$ for DDPM and by $\bar{\boldsymbol{\alpha}} = \{\bar{\alpha}_s\}_{s=1}^S$ for DDIM sampling.

Zero-shot diffusion-based methods for inverse problems Training-free methods address inverse problems by adapting a pretrained diffusion prior at inference time, thereby avoiding task-specific retraining [11]. One class of methods relies on projection-based updates, in which the sampling process is modified to enforce data consistency at each time step [39, 8, 58]. In contrast, Bayesian methods [26, 2, 48, 9, 35] formulate inverse problems through posterior inference by combining a pretrained diffusion prior with a measurement likelihood, leading to the following score decomposition:

$$\nabla_{\mathbf{x}_t}\log p_t(\mathbf{x}_t|\mathbf{y}) = \nabla_{\mathbf{x}_t}\log p_t(\mathbf{x}_t) + \nabla_{\mathbf{x}_t}\log p_t(\mathbf{y}|\mathbf{x}_t). \tag{5}$$

We detail the Bayesian formulations here, with projection-based methods described in Appendix I. Based on the Tweedie's formula [15], the prior score is obtained from the learned noise estimator as:

$$\nabla_{\mathbf{x}_t}\log p_t(\mathbf{x}_t) \approx -(1/\sqrt{1-\bar{\alpha}_t})\boldsymbol{\epsilon}_\theta(\mathbf{x}_t, t), \tag{6}$$

and the likelihood term is approximated in a method-dependent manner to guide sampling toward consistency with the measurements. A common Bayesian approach is DPS [9], which estimates the likelihood using $p(\mathbf{y}|\mathbf{x}_t) = \int p(\mathbf{y}|\mathbf{x}_0)p(\mathbf{x}_0|\mathbf{x}_t)d\mathbf{x}_0$ and then approximates $p(\mathbf{x}_0|\mathbf{x}_t)$ by a point mass at $\hat{\mathbf{x}}_0$, where $\hat{\mathbf{x}}_0 := \mathbb{E}[\mathbf{x}_0|\mathbf{x}_t]$. Incorporating this approximation within the deterministic DDIM formulation (4) leads to the following update rule:

$$\mathbf{x}_{s-1} = \sqrt{\bar{\alpha}_{s-1}}\hat{\mathbf{x}}_0 + \sqrt{1-\bar{\alpha}_{s-1}}\boldsymbol{\epsilon}_\theta(\mathbf{x}_s, s) - \zeta_s\nabla_{\mathbf{x}_s}\|\mathbf{y} - \mathbf{H}\hat{\mathbf{x}}_0\|_2^2, \tag{7}$$

where $\zeta_s = \zeta'/|\mathbf{y} - \mathbf{H}\hat{\mathbf{x}}_0(\mathbf{x}_s)|$ is a heuristic guidance step size which scales inversely with the measurement error norm, and ζ' is a user-defined constant. Alternatively, ΠGDM [35] models $p(\mathbf{x}_0|\mathbf{x}_t)$ as a Gaussian distribution with mean $\hat{\mathbf{x}}_0$ and variance r_t^2, governing the uncertainty at diffusion step t. This yields the likelihood approximation $p_t(\mathbf{y}|\mathbf{x}_t) \approx \mathcal{N}(\mathbf{H}\hat{\mathbf{x}}_0, r_t^2\mathbf{H}\mathbf{H}^T + \sigma_{\mathbf{y}}^2\mathbf{I})$, where the variance parameter r_s serves as a weighting term in the deterministic DDIM update:

$$\mathbf{x}_{s-1} = \sqrt{\bar{\alpha}_{s-1}}\hat{\mathbf{x}}_0 + \sqrt{1-\bar{\alpha}_{s-1}}\,\boldsymbol{\epsilon}_\theta(\mathbf{x}_s, s) + r_s^2\nabla_{\mathbf{x}_s}(\hat{\mathbf{x}}_0)^T\mathbf{H}^T((r_s^2\mathbf{H}\mathbf{H}^T + \sigma_{\mathbf{y}}^2\mathbf{I})^{-1})^T(\mathbf{y} - \mathbf{H}\hat{\mathbf{x}}_0). \tag{8}$$

Specifically, r_s is heuristically defined according to the diffusion noise schedule as $r_s = \sqrt{1-\bar{\alpha}_s}$. Throughout the paper, we denote by $\boldsymbol{\zeta}$ and $\boldsymbol{r}$ the heuristic parameter sets $\{\zeta_s\}_{s=1}^S$ and $\{r_s\}_{s=1}^S$, respectively.

3 Guided zero-shot posterior samplers

In this section, we bring diffusion-based zero-shot posterior sampling methods together under a unified spectral perspective. Our framework applies to a broad class of such approaches, enabling simplified analysis and principled design of their underlying heuristics. We use DPS [9] as a representative example and further apply our analysis to ΠGDM [35] and the projection-based DiffPIR [58] in Appendices H and I, respectively.

3.1 The optimal denoiser for a gaussian prior

As reviewed in Section 2, training-free posterior samplers leverage a pretrained denoiser to incorporate prior information during the inference phase. We assume a Gaussian prior on the clean signal, $\mathbf{x}_0 \sim \mathcal{N}(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$, where $\boldsymbol{\mu}_0 \in \mathbb{R}^d$ and $\boldsymbol{\Sigma}_0 \in \mathbb{R}^{d\times d}$. Under this modeling choice, the corresponding optimal denoiser for a noisy signal $\mathbf{x}_t$ at diffusion step t and noise schedule $\bar{\boldsymbol{\alpha}}$ admits the following closed-form expression [4]:

$$\mathbf{x}_0^* = (\bar{\alpha}_t \boldsymbol{\Sigma}_0 + (1-\bar{\alpha}_t)\mathbf{I})^{-1}(\sqrt{\bar{\alpha}_t}\boldsymbol{\Sigma}_0 \mathbf{x}_t + (1-\bar{\alpha}_t)\boldsymbol{\mu}_0) \tag{9}$$

3.2 Reverse process formulation

We next describe the DPS [9] sampling procedure under the Gaussian setting. For brevity, we highlight the key steps using the DDIM [34] formulation, which enables an accelerated sampling framework. Based on the marginal relation in (3), the deterministic update in (7) can be expressed in terms of the optimal denoiser as

$$\mathbf{x}_{s-1} = a_s \mathbf{x}_s + b_s \mathbf{x}_0^* - \zeta_s \nabla_{\mathbf{x}_s} \|\mathbf{y} - \mathbf{H}\mathbf{x}_0^*\|_2^2, \tag{10}$$

where $\mathbf{y} = \mathbf{H}\mathbf{x}_0 + \mathbf{n}$ denotes the linear degradation model, a_s and b_s are determined by the noise schedule $\bar{\boldsymbol{\alpha}}$, and ζ_s is the corresponding heuristic guidance weight. The full derivation is provided in Appendix E.

3.3 Spectral form of the update rule

We adopt a spectral perspective which allows a simplified analysis of the diffusion dynamics [4]. In particular, we consider the case where the covariance matrix $\boldsymbol{\Sigma}_0$ is shift-invariant, a reasonable assumption for natural images [38, 16], and the degradation operator $\mathbf{H}$ shares this property. This structure induces diagonalization in the Fourier basis and naturally aligns with common linear inverse problems such as deblurring, denoising, and accelerated acquisition (see Appendix K.4).

Under this setting, the prior distribution remains Gaussian in the spectral domain, $\mathbf{x}_0^{\mathcal{F}} \sim \mathcal{N}(\boldsymbol{\mu}_0^{\mathcal{F}}, \boldsymbol{\Lambda}_0)$, where $\boldsymbol{\mu}_0^{\mathcal{F}}$ is the transformed mean vector, and $\boldsymbol{\Lambda}_0 \in \mathbb{R}^{d\times d}$ is a positive semidefinite diagonal matrix containing the eigenvalues of $\boldsymbol{\Sigma}_0$, denoted by $\{\lambda_i\}_{i=1}^d$. Furthermore, $\mathbf{H} \in \mathbb{R}^{d\times d}$ admits a diagonal spectral representation $\boldsymbol{\Lambda}_h$, whose diagonal entries are the eigenvalues of $\mathbf{H}$, with $\boldsymbol{\Lambda}_{h^*}$ denoting its complex conjugate. Accordingly, the posterior $p(\mathbf{x}_0^{\mathcal{F}} \mid \mathbf{y}^{\mathcal{F}})$ remains Gaussian and admits a closed-form expression in the spectral domain, where $\mathbf{y}^{\mathcal{F}}$ denotes the spectral domain measurements. A detailed derivation is provided in Appendix D.

By projecting Equation (10) onto the unitary Fourier basis $\mathbf{F} \in \mathbb{C}^{d\times d}$, we obtain the following relation between consecutive diffusion steps in the spectral domain. where the superscript $\mathcal{F}$ denotes the Fourier-domain representation.

$$\mathbf{x}_{s-1}^{\mathcal{F}} = [a_s + b_s c_s - 2\zeta_s c_s \boldsymbol{\Lambda}_{h^*}\boldsymbol{\Lambda}_h c_s]\mathbf{x}_s^{\mathcal{F}} + [2\zeta_s c_s \boldsymbol{\Lambda}_{h^*}]\mathbf{y}^{\mathcal{F}} + [b_s d_s - 2\zeta_s c_s \boldsymbol{\Lambda}_{h^*}\boldsymbol{\Lambda}_h d_s]\boldsymbol{\mu}_0^{\mathcal{F}} \tag{11}$$

Here, a_s and b_s are diagonal matrices determined by the noise schedule $\bar{\boldsymbol{\alpha}}$, and c_s, d_s further reflect the spectral properties encoded in $\boldsymbol{\Lambda}_0$. The full definitions are provided in Appendix E.2.

Using the diagonalization property of Equation (11), the update rule decouples into d independent scalar equations, which can be recursively substituted. This yields a closed-form expression for the estimated output signal over S diffusion steps:

$$\hat{\mathbf{x}}_{0,\text{DPS}}^{\mathcal{F}} = \mathbf{D}_1 \mathbf{x}_S^{\mathcal{F}} + \mathbf{D}_2 \mathbf{y}^{\mathcal{F}} + \mathbf{D}_3 \boldsymbol{\mu}_0^{\mathcal{F}} \tag{12}$$

where $\mathbf{D}_1$, $\mathbf{D}_2$, and $\mathbf{D}_3$ are diagonal matrices determined by the noise schedule $\bar{\boldsymbol{\alpha}}$, the number of diffusion steps S, the spectral characteristics of the prior $\boldsymbol{\Lambda}_0$, the heuristic guidance weights $\boldsymbol{\zeta}$, and the degradation operator $\boldsymbol{\Lambda}_h$. Interestingly, Equation (12) can be interpreted as a set of d scalar *transfer functions* mapping the Gaussian input noise $\mathbf{x}_S \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ and the degraded measurement $\mathbf{y}^{\mathcal{F}}$ to the clean output estimate.

Finally, from a probabilistic perspective, the resulting posterior distribution is

$$\hat{\mathbf{x}}_{0,\text{DPS}}^{\mathcal{F}} \,\big|\, \mathbf{y}^{\mathcal{F}} \sim \mathcal{N}\big(\mathbf{D}_2 \mathbf{y}^{\mathcal{F}} + \mathbf{D}_3 \boldsymbol{\mu}_0^{\mathcal{F}} \;,\; \mathbf{D}_1^2\big)\,. \tag{13}$$

This reconstruction captures information from the prior structure, the diffusion dynamics and the degraded measurements $\mathbf{y}^{\mathcal{F}}$, with the latter incorporated through the DPS update, contributing to the posterior via its first-order statistics.

More generally, the scope of our framework covers a range of posterior sampling methods whose updates admit an affine spectral recursion under the Gaussian-optimal denoiser. Since the analysis operates at the distribution level rather than individual sampling trajectories, it can also be extended to stochastic sampling, such as DDPM. Additional details are provided in Appendices E, I and H.

3.4 Optimal spectral guidance weights

By leveraging the closed-form expressions for the reconstructed distributions such as (13), we further analyze the dynamics of diffusion-based posterior sampling methods from a spectral perspective. A key aspect relates to the design choices of the prior-likelihood weighting heuristics. We consider a setting with a prior covariance matrix $\mathbf{\Sigma}_0$, characterized by eigenvalues $\{\lambda_i\}_{i=1}^d$, a fixed linear degradation matrix $\mathbf{H}$, an observation $\mathbf{y}$, and a diffusion process with S diffusion steps and noise schedule $\bar{\boldsymbol{\alpha}}$. Our objective is to bring the estimated probability density induced by such posterior sampling methods as close as possible to the true posterior distribution. Accordingly, we seek a set of weights $\boldsymbol{\zeta}$ that minimizes a discrepancy measure $\mathcal{D}$ between the estimated reconstructed distribution, $p(\hat{\mathbf{x}}^{\mathcal{F}}_{0,\text{DPS}} \mid \mathbf{y}^{\mathcal{F}}; \boldsymbol{\zeta}, \bar{\boldsymbol{\alpha}})$, and the true posterior distribution $p(\mathbf{x}_0^{\mathcal{F}} \mid \mathbf{y}^{\mathcal{F}})$, leading to the following optimization problem.[2]

$$\boldsymbol{\zeta}^* = \arg\min_{\boldsymbol{\zeta}} \mathcal{D}\left(p(\hat{\mathbf{x}}^{\mathcal{F}}_{0,\text{DPS}} \,|\, \mathbf{y}^{\mathcal{F}}; \boldsymbol{\zeta}, \bar{\boldsymbol{\alpha}}), p(\mathbf{x}_0^{\mathcal{F}} \,|\, \mathbf{y}^{\mathcal{F}})\right) \tag{14}$$

Choosing the discrepancy measure $\mathcal{D}$ is a key design choice, as it governs how measurement fidelity is balanced with prior information and ultimately shapes the trade-off between perceptual quality and distortion [5]. In principle, for a given observation $\mathbf{y}$, one could define various discrepancy measures $\mathcal{D}$ between the corresponding posterior distributions, such as the *Wasserstein-2* [1] or *KL divergence* [29] (see Appendix G for further details). In this work, we propose a more general formulation based on the *Wasserstein-2* distance, referred to as the *averaged Wasserstein distance* and denoted by:

$$\left[\mathcal{D}^2_{W_2}\right]_{\text{Avg}} = \frac{1}{K}\sum_{k=1}^{K} \mathcal{D}^2_{W_2}\left(p(\hat{\mathbf{x}}_0^{\mathcal{F}} \,|\, \mathbf{y}_k^{\mathcal{F}}; \boldsymbol{\zeta}, \bar{\boldsymbol{\alpha}}), p(\mathbf{x}_0^{\mathcal{F}} \,|\, \mathbf{y}_k^{\mathcal{F}})\right). \tag{15}$$

This formulation accommodates two regimes. For $K = 1$, the optimization is performed for a single realization $\mathbf{y}$, reducing to the standard *Wasserstein-2* distance:

$$\left[\mathcal{D}^2_{W_2}\right]_{\text{Avg}} = \sum_{i=1}^{d} \left(\sqrt{\lambda_i} - [\mathbf{D}_1]_i\right)^2 + \sum_{i=1}^{d} \left([\mathbf{D}_2 - \mathbf{A}]_i [\mathbf{y}^{\mathcal{F}}]_i + [\mathbf{D}_3 - \mathbf{I} + \mathbf{A}\mathbf{\Lambda}_h]_i [\boldsymbol{\mu}_0^{\mathcal{F}}]_i\right)^2 \tag{16}$$

where $\mathbf{A} = \mathbf{\Lambda}_0 \mathbf{\Lambda}_{h^*} (\mathbf{\Lambda}_h \mathbf{\Lambda}_0 \mathbf{\Lambda}_{h^*} + \sigma_n^2 \mathbf{I})^{-1}$.

Alternatively, when multiple realizations are considered, the weight parameters can be adjusted to capture the expected behavior of the posterior $p(\mathbf{x}_0^{\mathcal{F}} \,|\, \mathbf{y}^{\mathcal{F}})$ across $\mathbf{y}^{\mathcal{F}}$, thereby eliminating the need to solve the optimization for each observation. In this setting, the optimal solution admits a closed-form expression in terms of the prior characteristics:

$$[\mathcal{D}^2_{W_2}]_{\text{Avg}} = \sum_{i=1}^{d} \left(\sqrt{\lambda_i} - [\mathbf{D}_1]_i\right)^2 + \text{Tr}\left(\mathbf{M}^T \mathbf{M} \left(\mathbf{\Lambda}_h \, \mathbf{\Lambda}_0 \, \mathbf{\Lambda}_{h^*} + \sigma_n^2 \, \mathbf{I}\right)\right) + \|\mathbf{M}\mathbf{\Lambda}_h \, \boldsymbol{\mu}_0^{\mathcal{F}} + \mathbf{b}\|_2^2 \tag{17}$$

where $\mathbf{M} = \mathbf{D}_2 - \mathbf{A}$ and $\mathbf{b} = (\mathbf{D}_3 - \mathbf{I} + \mathbf{A}\mathbf{\Lambda}_h) \, \boldsymbol{\mu}_0^{\mathcal{F}}$, with $\mathbf{D}_2$ and $\mathbf{D}_3$ defined in (12). The full derivations are provided in Appendix F.

4 Posterior optimal denoising

In this section, we formulate a closed-form expression for the posterior optimal denoiser under the Gaussian setting and analyze the output induced by the diffusion process. The derived sampler provides a principled reference for evaluating training-free posterior sampling methods and for assessing the effectiveness of our proposed recommendations under identical diffusion dynamics. Such an ideal denoiser is tailored to a specific measurement operator $\mathbf{H}$ and observation $\mathbf{y}$, and generally does not admit an analytic form beyond the Gaussian setting, hence necessitating approximations in practice.

[2] While $\bar{\boldsymbol{\alpha}}$ could also be treated as an optimized argument, we focus on the current formulation.

4.1 The posterior optimal denoiser

We adopt the notation introduced in Section 3.1. The clean signal $\mathbf{x}_0 \in \mathbb{R}^d$ is modeled by a Gaussian prior, $\mathbf{x}_0 \sim \mathcal{N}(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$. Observations $\mathbf{y} \in \mathbb{R}^m$ follow $\mathbf{y} = \mathbf{H}x_0 + \mathbf{n}$ with a general linear operator $\mathbf{H} \in \mathbb{R}^{m\times d}$ and $\mathbf{n} \sim \mathcal{N}(\mathbf{0}, \sigma_y^2\mathbf{I})$. Under this modeling framework, the true posterior distribution is also Gaussian, $p(\mathbf{x}_0|\mathbf{y}) \sim \mathcal{N}(\boldsymbol{\mu}_{\mathbf{x}_0|\mathbf{y}}, \boldsymbol{\Sigma}_{\mathbf{x}_0|\mathbf{y}})$, with the resulting expressions for the posterior mean and covariance provided in Appendix D.

We begin by deriving the exact optimal denoiser for the posterior. Under the Gaussian model described above, this denoiser coincides with both the posterior mean and the MAP estimator, and reduces to the classical Wiener filter in closed form. A detailed proof is provided in Appendix A.

Theorem 4.1. *Let $\mathbf{x}_0 \sim \mathcal{N}(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$ and let $\mathbf{x}_t$ denote the noisy signal obtained from the forward diffusion process, as defined in (3). Given the linear measurement model $\mathbf{y} = \mathbf{H}\mathbf{x}_0 + \mathbf{n}$, and letting $\bar{\boldsymbol{\alpha}}$ denote the noise schedule parameters, the denoised signal obtained from the MAP estimator is given by:*

$$\tilde{\mathbf{x}}_0 = \arg\max_{\mathbf{x}_0} \log p(\mathbf{x}_0|\mathbf{x}_t, \mathbf{y}) \tag{18}$$

and admits the following closed-form:

$$\begin{aligned}\tilde{\mathbf{x}}_0 = &\left((1-\bar{\alpha}_t)\boldsymbol{\Sigma}_0\mathbf{H}^T\mathbf{H} + \sigma_y^2\bar{\alpha}_t\boldsymbol{\Sigma}_0 + \sigma_y^2(1-\bar{\alpha}_t)\mathbf{I}\right)^{-1} \\ &\left((1-\bar{\alpha}_t)\boldsymbol{\Sigma}_0\mathbf{H}^T\mathbf{y} + \sigma_y^2\sqrt{\bar{\alpha}_t}\boldsymbol{\Sigma}_0\mathbf{x}_t + \sigma_y^2(1-\bar{\alpha}_t)\boldsymbol{\mu}_0\right)\end{aligned} \tag{19}$$

4.2 Spectral form of the update rule

We now describe the discrete sampling procedure, as introduced in [22, 34] instantiated using optimal denoiser derived in (19). In particular, we focus on the DDIM formulation [34], as it facilitates an accelerated sampling framework. For brevity, we only present the final reconstructed distribution of the output $\hat{\mathbf{x}}^{\mathcal{F}}_{0,\text{opt}}$, where the superscript $\mathcal{F}$ denotes the Fourier-domain representation. A full derivation is provided in Appendices B and C.

Lemma 4.2. *Assume $\mathbf{x}_0 \sim \mathcal{N}(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$, where the covariance matrix $\boldsymbol{\Sigma}_0$ is shift-invariant, and let $\mathbf{y} = \mathbf{H}\mathbf{x}_0 + \mathbf{n}$ denote a linear measurement model with a shift-invariant operator. The frequency domain reconstruction obtained using DDIM with the optimal posterior denoiser in (19) follows a Gaussian distribution and is given by:*

$$p(\hat{\mathbf{x}}^{\mathcal{F}}_{0,opt}|\mathbf{y}^{\mathcal{F}}) \sim \mathcal{N}(\mathbf{V}_2\mathbf{y}^{\mathcal{F}} + \mathbf{V}_3\boldsymbol{\mu}_0^{\mathcal{F}}, \mathbf{V}_1^2) \tag{20}$$

The matrices $\mathbf{V}_1$, $\mathbf{V}_2$, and $\mathbf{V}_3$ are diagonal and depend on the noise schedule parameters $\bar{\boldsymbol{\alpha}}$, the degradation operator $\boldsymbol{\Lambda}_h$, and the Fourier-domain prior, $\mathbf{x}_0^{\mathcal{F}} \sim \mathcal{N}(\boldsymbol{\mu}_0^{\mathcal{F}}, \boldsymbol{\Lambda}_0)$.

Consequently, under the Gaussianity assumption, the resulting sampling distribution admits a closed-form expression that decouples into d independent scalar *transfer functions*. This formulation provides a principled reference for analyzing and interpreting existing posterior sampling methods.

5 Related work

Diffusion models have been widely applied to inverse problems in a zero-shot setting [11]. One line of work incorporates measurements by enforcing data consistency through projection-based geometric operations, such as null-space decompositions and pseudoinverse operators [8, 39]. Among these methods, DiffPIR [58] builds on the half-quadratic-splitting (HQS) algorithm [17] to decouple the data and the prior terms during inference. Another line adopts a Bayesian perspective, integrating measurements through likelihood-based guidance [26, 2, 48, 31]. Within this framework, DPS [9] employs a point-estimate approximation of the clean signal motivated by Jensen's inequality, while ΠGDM [35] further introduces time-dependent variances.

Alongside their effectiveness, these approaches face practical challenges. A key aspect is their reliance on manual hyperparameter tuning, which affect both stability and performance. While stability can be improved by using small guidance step sizes, this comes at the cost of increased sampling steps and slower inference [19]. Recently, DSG [51] introduced the notion of manifold deviation during sampling and proposed constraints that enable larger guidance step sizes. Similarly,

DAPS [54] employs an annealing strategy that decouples consecutive steps in the diffusion trajectory, improving stability at larger step sizes. Related works further explore how inference-time dynamics, such as manifold constraints or noise modulation, affect the tradeoff between perceptual quality and data fidelity in diffusion models [10, 19, 40]. Nevertheless, most training-free posterior sampling methods still rely on manual tuning to control these trade-offs.

Recently, the design of heuristics for diffusion models has received significant attention [44, 47, 41, 37, 6, 49, 46]. In particular, the authors of [4] proposed a spectral approach for determining the noise schedule in prior sampling, expressing the diffusion process as a transfer function and formulating an associated optimization problem. In the context of posterior sampling, the authors of [3] addressed the purely denoising task using the prior score, and considered their solution transferable to related inverse problems. While their approach provided heuristic recommendations for DPS, it involves per-instance optimization during inference, increasing synthesis time. In a related direction, C-DPS [30] seeks to avoid heuristic likelihood guidance by coupling data- and measurement-space dynamics and deriving a Gaussian posterior transition under sufficiently small diffusion steps. Our work instead provides a unified spectral perspective on existing posterior sampling algorithms, enabling principled guidance weight design without modifying the underlying sampling dynamics.

6 Experiments

We turn to validate our approach through a series of experiments under multiple degradation settings, examining our spectral recommendations compared to existing heuristics, and analyzing current sampling methods relative to the derived optimal posterior denoiser.

6.1 Synthetic Gaussian distribution

In this section, we consider a Gaussian prior distribution $\mathbf{x}_0 \sim \mathcal{N}(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$, where $\mathbf{x}_0 \in \mathbb{R}^d$ and $\boldsymbol{\Sigma}_0 \in \mathbb{R}^{d\times d}$ is a circulant covariance matrix. The measurements $\mathbf{y} \in \mathbb{R}^d$ are generated by the linear degradation model $\mathbf{y} = \mathbf{H}\mathbf{x}_0 + \mathbf{n}$, where the operator $\mathbf{H}$ acts as a low-pass filter (*LPF*), preserving a fraction $\mathcal{V} \in (0, 1]$ of the frequency components, and $\mathbf{n} \sim \mathcal{N}(\mathbf{0}, \sigma_y^2\mathbf{I})$ is additive Gaussian noise. Further details regarding the experimental configuration are provided in Appendix J.

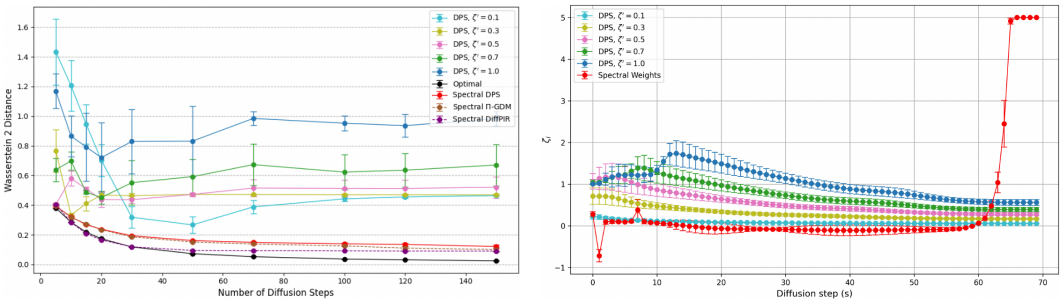
We begin by solving the optimization problem in (16) for DPS [9] using varying numbers of diffusion steps. Specifically, we repeat the procedure independently for $N=5$ observations $\{\mathbf{y}^{(i)}\}_{i=1}^N$, where each realization corresponds to an independently sampled signal $\mathbf{x}_0^{(i)}$ and noise realization $\mathbf{n}^{(i)}$. We set $\mathcal{V}=0.5$, $\sigma_y=0.1$, and consider a prior with mean $\boldsymbol{\mu}_0=\mathbf{0}$ and a covariance dimension $d=50$.

Figure (4) in Appendix J.2, shows the resulting spectral recommendations for the weighting coefficients $\boldsymbol{\zeta}$, along with the mean and standard deviation across realizations. While the recommendations vary with the number of diffusion steps, their overall structure is preserved. Notably, larger weights are assigned at the early stages of the diffusion process and gradually decrease toward later steps, with greater variability across realizations observed at the beginning of the reverse process.

We compare our spectral recommendations with the guidance weights used in DPS under the same experimental setting. As these weights are selected heuristically, we evaluate several values of ζ' to assess the effect of different choices. Figure 1b illustrates the comparison for 70 diffusion steps. Interestingly, due to their inverse dependence on the measurement error norm, the DPS weights exhibit an opposing trend, increasing toward the end of the diffusion process.

While the true posterior is often intractable for general priors, as only a single sample $\mathbf{x}_0^{(i)}$ is available for a given observation $\mathbf{y}^{(i)}$, the Gaussian setting admits a closed-form posterior, enabling direct distributional evaluation. Leveraging this property, Figure 1a compares the *Wasserstein-2* distance between the reconstructed and the true posterior distributions for several heuristic choices and the proposed spectral recommendation across diffusion steps and realizations. As performance varies across heuristic choices, the spectral recommendation consistently surpasses all hand-crafted alternatives and exhibits greater stability across realizations.

Figure 1a also includes the distance obtained using the analytically derived posterior-optimal denoiser from Equation (20). This solution serves as a reference for evaluating the different posterior samplers and is generally unavailable beyond the Gaussian setting. The figure further compares the spectral recommendations for DPS [9], ΠGDM [35], and DiffPIR [58], highlighting the applicability of our



(a) (b)

Figure 1: (a) Comparison of the Wasserstein-2 distance for DPS heuristic values $\zeta' \in \{0.1, 0.3, 0.5, 0.7, 1.0\}$, the spectral recommendations applied to DPS (red) and to ΠGDM (brown, dotted), and sampling with the posterior optimal denoiser (black), across different numbers of diffusion steps $S \in \{5, 10, 15, 20, 30, 50, 70, 100, 120, 150\}$. (b) Comparison of spectral recommendations (red) and DPS guidance coefficients for 70 diffusion steps, evaluated across heuristic values.

approach. While ΠGDM achieves marginally lower Wasserstein-2 values than DPS, likely due to incorporating an additional uncertainty term, the DiffPIR recommendations outperform all other methods and, for a small number of diffusion steps, slightly surpass the exact posterior denoiser. Notably, the optimality established in Theorem 4.1 applies to the single-step denoising operation rather than the resulting finite-step posterior sampler. For brevity, Appendices J.3 and J.4 provide the spectral recommendations for ΠGDM and DiffPIR, and include further comparison with their heuristic counterparts and the optimal denoiser.

6.2 Empirical distribution

We turn to evaluate our method on real-world datasets, including FFHQ 256×256 [24], ImageNet 256×256 [12] and fastMRI [53, 27], a single-coil knee magnetic resonance imaging (MRI) dataset centered and cropped to a spatial size of 320×320. In order to derive the spectral recommendations, we fit a Gaussian distribution to each dataset by estimating the corresponding mean and shift-invariant covariance matrix (see Appendix M).

We begin by solving the optimization problem in (17) within the DPS framework. Figure 2 shows the spectral weights obtained for the FFHQ dataset under LPF degradation with $\mathcal{V} = 0.1$ and $\sigma_y = 0.1$, across different diffusion step counts. For comparison, we also include the DPS heuristic weights, averaged over 100 syntheses, with variability indicated. Interestingly, the optimization reveals a markedly different weighting structure, with guidance values increasing over the reverse diffusion process. While the heuristic weights remain consistently small, the spectral recommendations place greater emphasis on the observations, suggesting a different balance between prior information and the likelihood term. Spectral recommendations and analogous analyses for the projection-based DiffPIR sampler and ΠGDM, along with additional ImageNet results, are provided in Appendix K.

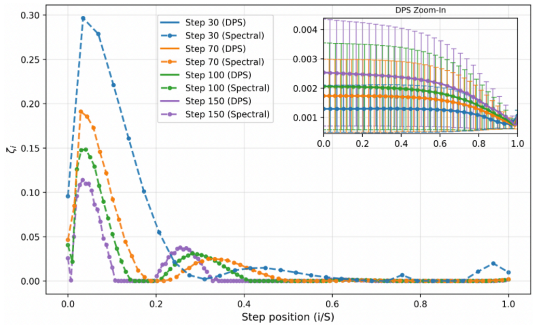


Figure 2: Comparison of spectral recommendations on the FFHQ dataset, with a magnified y-axis view of the DPS heuristics. Results are shown for diffusion steps $S \in \{30, 70, 100, 150\}$.

As an additional degradation setting, we consider accelerated MRI acquisition for the fastMRI dataset, using the horizontal frequency subsampling protocol introduced in [23]. Since obtaining full Fourier measurements in MRI is time-consuming and expensive, this setting constitutes a highly relevant and challenging inverse problem that has attracted considerable research attention [45, 43, 18]. Additional details are provided in Appendix K.4.

Table 1: Quantitative evaluation on FFHQ, comparing spectral recommendations with DPS heuristics ($\mathcal{V}=0.1$ and $\sigma_y=0.1$).

Method	Steps	PSNR↑	SSIM↑	LPIPS↓	FID↓
DPS	50	16.00	0.35	0.57	69.13
	100	18.31	0.43	0.51	62.68
	150	19.82	0.49	0.45	58.12
	200	20.76	0.53	0.43	55.88
Spectral	50	**25.55**	**0.69**	**0.32**	**50.86**
	100	**27.82**	**0.80**	**0.23**	**41.31**
	150	**28.10**	**0.79**	**0.23**	**37.51**
	200	**28.31**	**0.80**	**0.22**	**35.27**

Table 2: PSNR↑ results on the fastMRI dataset under LPF degradation and accelerated MRI acquisition.

Method	Steps	LPF	R=4	R=8	R=12
DPS	50	17.98	18.32	17.62	17.10
	100	19.56	19.78	19.06	18.26
	150	20.39	20.29	19.79	18.96
	200	20.82	20.57	19.99	19.27
Spectral	50	**23.20**	**21.66**	**19.21**	**17.57**
	100	**24.62**	**23.44**	**20.61**	**18.82**
	150	**24.90**	**24.08**	**21.26**	**19.31**
	200	**24.95**	**24.44**	**21.76**	**19.61**

We further examine the spectral recommendations during inference, evaluating their generalization to real-world datasets beyond the Gaussian setting. Since the optimal prior denoiser (9) is unavailable in these settings, neural denoisers are employed instead, introducing approximation error. For the FFHQ and ImageNet datasets, we use the pretrained models from [9], while for fastMRI we train a similar U-Net [32] architecture with adaptations to the MRI setting (see Appendix K.4).

In our first experiment, we assess prior characteristics using $1,000$ randomly selected validation images, each independently degraded and reconstructed once with DPS using both the spectral and the heuristic weights. To maintain consistency with the denoiser training, we employ a linear noise schedule. For evaluation, we report PSNR, SSIM [42], LPIPS [57], and FID [21], all computed using the piq library [25], with images normalized to the $[0,1]$ range.

Table 1 reports the results on the FFHQ dataset. For this setting, we evaluate a broad range of DPS heuristics in Appendix K.1 and report here the best-performing value, $\zeta' = 0.1$. Overall, the spectral recommendations achieve a more balanced tradeoff between measurement fidelity and perceptual quality compared to the DPS heuristic. At relatively smaller numbers of diffusion steps, DPS heuristic attains closer FID values; however, qualitative results in Appendix K, suggest that this behavior is associated with its small likelihood weights, leading the samples to be dominated by the pretrained prior. While this can enhance perceptual appearance, the resulting images may deviate substantially from the observed measurements, which is undesirable in inverse problems.

Table 2 presents PSNR results on the fastMRI dataset under both LPF degradation with $\mathcal{V} = 0.1$ and accelerated MRI acquisition using horizontal frequency subsampling [23]. Here, R denotes the acceleration factor, retaining the central $120/R$ frequencies together with an additional uniformly sampled set of $200/R$ frequencies. In both settings, additive Gaussian noise with $\sigma_y = 0.1$ is applied, and $\zeta' = 0.1$ is used for the DPS heuristic, as selected in Appendix K. Notably, the spectral recommendations consistently outperform the heuristic weighting strategies, indicating improved adherence to the underlying signal, a valuable property in MRI reconstruction. Appendix K provides more details on the degradation operator, additional perceptual metrics, broader comparisons across $\mathcal{V}$ and σ_y values, a range of heuristic configurations, and ImageNet results.

While our recommendation are analytically derived for shift-invariance degradation models, we evaluate their effectiveness on settings where this property is relaxed, including random inpainting with 70% pixel removal and box inpainting using a 128×128 mask, both with $\sigma_y = 0.1$. We compare the results with DPS and recent heuristic-based methods, including DAPS [54] and DiffPIR [58]. Figures 13 and 14 show consistent improvements over DPS heuristics across both degradations along with frequent gains over DAPS and DiffPIR. For further discussion and images refer to appendix K.3.

We proceed with an analysis of the posterior behavior under fixed observations by selecting three representative images and applying the same LPF degradation setting. We consider both optimization formulations introduced in (16) and (17) . The first is solved independently for each observation and yields realization-specific recommendations, whereas the second captures the expected posterior behavior across degraded measurements and is solved once for a fixed degradation operator. For each image, we generate $1{,}000$ reconstructions across different numbers of diffusion steps and compute the corresponding performance metrics.

Unlike the prior analysis, this experiment fixes the observation and probes the posterior through multiple reconstructions per image. The results, summarized in Table 6, indicate that both spectral optimization strategies generally offer a balance between measurement fidelity and sample naturalness, with the realization-specific approach providing a modest but consistent improvement by adapting to individual observations. This benefit comes at the cost of solving a separate optimization problem for each realization. However, in both settings, the optimization overhead remains modest relative to the repeated synthesis and evaluation required for heuristic tuning. Further qualitative results and details on optimization acceleration strategies are provided in Appendices L and N.

Limitations Our analytical derivation assumes a Gaussian prior and relies on shift-invariance of both the covariance matrix and the degradation operator, enabling a common diagonalizing basis. While one could alternatively work in the eigenbasis of a single matrix and represent the other accordingly, this may lead to poor approximation quality, particularly under spatially varying degradations. Nevertheless, we demonstrate that the proposed recommendations remain effective beyond this setting, including on real-world datasets and under non-shift-invariant operators, as well as on practical scenarios such as accelerated MRI, which naturally align with the required conditions. Extending the framework to richer priors, such as Gaussian mixture models, or to representations that preserve tractability without joint diagonalization remains an interesting direction for future work.

7 Conclusion

In this work, we propose a spectral approach for analyzing zero-shot diffusion posterior sampling methods. Under a Gaussian assumption, we derive the output distributions induced by these samplers as well as by sampling with the ideal posterior denoiser, which provides a theoretical reference for analysis. We demonstrate how our approach enables a principled design of the heuristic components used by these schemes, while jointly accounting for properties of the prior distribution, the degraded observations, and the dynamics of the diffusion process. Finally, we examine our recommendation on existing methods across various real-world scenarios and characterize their behavior in terms of sample naturalness and measurement fidelity. We hope this work offers a useful perspective to inform future research on developing and analyzing zero-shot posterior sampling methods.

8 Acknowledgment

This research was partially supported by the Israel Science Foundation (ISF) under Grants 951/24 and 409/24, and by the Council for Higher Education– Planning and Budgeting Committee.

References

[1] Martin Arjovsky, Soumith Chintala, and Léon Bottou. Wasserstein generative adversarial networks. In *International conference on machine learning*, pages 214–223. PMLR, 2017.

[2] Arpit Bansal, Hong-Min Chu, Avi Schwarzschild, Soumyadip Sengupta, Micah Goldblum, Jonas Geiping, and Tom Goldstein. Universal guidance for diffusion models. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 843–852, 2023.

[3] Gregory Bellchambers. Exploiting the exact denoising posterior score in training-free guidance of diffusion models. *arXiv preprint arXiv:2506.13614*, 2025.

[4] Roi Benita, Michael Elad, and Joseph Keshet. Spectral analysis of diffusion models with application to schedule design. *arXiv preprint arXiv:2502.00180*, 2025.

[5] Yochai Blau and Tomer Michaeli. The perception-distortion tradeoff. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 6228–6237, 2018.

[6] Defang Chen, Zhenyu Zhou, Can Wang, Chunhua Shen, and Siwei Lyu. On the trajectory regularity of ode-based diffusion sampling. *arXiv preprint arXiv:2405.11326*, 2024.

[7] Ke Chen, Xingjian Du, Bilei Zhu, Zejun Ma, Taylor Berg-Kirkpatrick, and Shlomo Dubnov. Zero-shot audio source separation through query-based learning from weakly-labeled data. In

Proceedings of the AAAI Conference on Artificial Intelligence, volume 36, pages 4441–4449, 2022.

[8] Jooyoung Choi, Sungwon Kim, Yonghyun Jeong, Youngjune Gwon, and Sungroh Yoon. Ilvr: Conditioning method for denoising diffusion probabilistic models. *arXiv preprint arXiv:2108.02938*, 2021.

[9] Hyungjin Chung, Jeongsol Kim, Michael T Mccann, Marc L Klasky, and Jong Chul Ye. Diffusion posterior sampling for general noisy inverse problems. *arXiv preprint arXiv:2209.14687*, 2022.

[10] 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, 2022.

[11] Giannis Daras, Hyungjin Chung, Chieh-Hsin Lai, Yuki Mitsufuji, Jong Chul Ye, Peyman Milanfar, Alexandros G Dimakis, and Mauricio Delbracio. A survey on diffusion models for inverse problems. *arXiv preprint arXiv:2410.00083*, 2024.

[12] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. Imagenet: A large-scale hierarchical image database. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2009.

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

[14] Chao Dong, Chen Change Loy, Kaiming He, and Xiaoou Tang. Image super-resolution using deep convolutional networks. *IEEE transactions on pattern analysis and machine intelligence*, 38(2):295–307, 2015.

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

[16] Dror Freirich, Tomer Michaeli, and Ron Meir. A theory of the distortion-perception tradeoff in wasserstein space. *Advances in Neural Information Processing Systems*, 34:25661–25672, 2021.

[17] Donald Geman and Chengda Yang. Nonlinear image recovery with half-quadratic regularization. *IEEE transactions on Image Processing*, 4(7):932–946, 1995.

[18] Yoseo Han, Leonard Sunwoo, and Jong Chul Ye. k-space deep learning for accelerated mri. *IEEE transactions on medical imaging*, 39(2):377–386, 2019.

[19] Yutong He, Naoki Murata, Chieh-Hsin Lai, Yuhta Takida, Toshimitsu Uesaka, Dongjun Kim, Wei-Hsiang Liao, Yuki Mitsufuji, J Zico Kolter, Ruslan Salakhutdinov, et al. Manifold preserving guided diffusion. *arXiv preprint arXiv:2311.16424*, 2023.

[20] John R Hershey, Zhuo Chen, Jonathan Le Roux, and Shinji Watanabe. Deep clustering: Discriminative embeddings for segmentation and separation. In *2016 IEEE international conference on acoustics, speech and signal processing (ICASSP)*, pages 31–35. IEEE, 2016.

[21] Martin Heusel, Hubert Ramsauer, Thomas Unterthiner, Bernhard Nessler, and Sepp Hochreiter. Gans trained by a two time-scale update rule converge to a local nash equilibrium. *Advances in neural information processing systems*, 30, 2017.

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

[23] Ajil Jalal, Marius Arvinte, Giannis Daras, Eric Price, Alexandros G Dimakis, and Jon Tamir. Robust compressed sensing mri with deep generative priors. *Advances in neural information processing systems*, 34:14938–14954, 2021.

[24] Tero Karras, Samuli Laine, and Timo Aila. A style-based generator architecture for generative adversarial networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019.

[25] Sergey Kastryulin, Jamil Zakirov, Denis Prokopenko, and Dmitry V Dylov. Pytorch image quality: Metrics for image quality assessment. *arXiv preprint arXiv:2208.14818*, 2022.

[26] Bahjat Kawar, Michael Elad, Stefano Ermon, and Jiaming Song. Denoising diffusion restoration models. *Advances in neural information processing systems*, 35:23593–23606, 2022.

[27] Florian Knoll, Jure Zbontar, Anuroop Sriram, Matthew J Muckley, Mary Bruno, Aaron Defazio, Marc Parente, Krzysztof J Geras, Joe Katsnelson, Hersh Chandarana, et al. fastmri: A publicly available raw k-space and dicom dataset of knee images for accelerated mr image reconstruction using machine learning. *Radiology: Artificial Intelligence*, 2(1):e190007, 2020.

[28] Dieter Kraft. A software package for sequential quadratic programming. *Forschungsbericht-Deutsche Forschungs- und Versuchsanstalt fur Luft- und Raumfahrt*, 1988.

[29] Solomon Kullback and Richard A Leibler. On information and sufficiency. *The annals of mathematical statistics*, 22(1):79–86, 1951.

[30] Shayan Mohajer Hamidi, Ben Liang, and En-Hui Yang. Coupled data and measurement space dynamics for enhanced diffusion posterior sampling. *Advances in Neural Information Processing Systems*, 38:55554–55592, 2026.

[31] Badr Moufad, Yazid Janati, Lisa Bedin, Alain Durmus, Randal Douc, Eric Moulines, and Jimmy Olsson. Variational diffusion posterior sampling with midpoint guidance. *arXiv preprint arXiv:2410.09945*, 2024.

[32] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.

[33] Chitwan Saharia, William Chan, Huiwen Chang, Chris Lee, Jonathan Ho, Tim Salimans, David Fleet, and Mohammad Norouzi. Palette: Image-to-image diffusion models. In *ACM SIGGRAPH 2022 conference proceedings*, pages 1–10, 2022.

[34] Jiaming Song, Chenlin Meng, and Stefano Ermon. Denoising diffusion implicit models. In *International Conference on Learning Representations (ICLR)*, 2021.

[35] Jiaming Song, Arash Vahdat, Morteza Mardani, and Jan Kautz. Pseudoinverse-guided diffusion models for inverse problems. In *International Conference on Learning Representations*, 2023.

[36] Yang Song, Jascha Sohl-Dickstein, Diederik P Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-based generative modeling through stochastic differential equations. *arXiv preprint arXiv:2011.13456*, 2020.

[37] Vinh Tong, Hoang Trung-Dung, Anji Liu, Guy Van den Broeck, and Mathias Niepert. Learning to discretize denoising diffusion odes. *arXiv preprint arXiv:2405.15506*, 2024.

[38] Michael Unser. On the approximation of the discrete karhunen-loeve transform for stationary processes. *Signal Processing*, 7(3):231–249, 1984.

[39] Yinhuai Wang, Jiwen Yu, and Jian Zhang. Zero-shot image restoration using denoising diffusion null-space model. *arXiv preprint arXiv:2212.00490*, 2022.

[40] Yuhan Wang, Suzhi Bi, Ying-Jun Angela Zhang, and Xiaojun Yuan. Traversing distortion-perception tradeoff using a single score-based generative model. In *Proceedings of the Computer Vision and Pattern Recognition Conference*, pages 2377–2386, 2025.

[41] Yunke Wang, Xiyu Wang, Anh-Dung Dinh, Bo Du, and Charles Xu. Learning to schedule in diffusion probabilistic models. In *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pages 2478–2488, 2023.

[42] Zhou Wang, Alan C Bovik, Hamid R Sheikh, and Eero P Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing*, 13(4):600–612, 2004.

[43] Zi Wang, Chen Qian, Di Guo, Hongwei Sun, Rushuai Li, Bo Zhao, and Xiaobo Qu. One-dimensional deep low-rank and sparse network for accelerated mri. *IEEE Transactions on Medical Imaging*, 42(1):79–90, 2022.

[44] Daniel Watson, William Chan, Jonathan Ho, and Mohammad Norouzi. Learning fast samplers for diffusion models by differentiating through sample quality. In *International Conference on Learning Representations*, 2021.

[45] Tomer Weiss, Ortal Senouf, Sanketh Vedula, Oleg Michailovich, Michael Zibulevsky, and Alex Bronstein. Pilot: Physics-informed learned optimized trajectories for accelerated mri. *arXiv preprint arXiv:1909.05773*, 2019.

[46] Christopher Williams, Andrew Campbell, Arnaud Doucet, and Saifuddin Syed. Score-optimal diffusion schedules. *Advances in Neural Information Processing Systems*, 37:107960–107983, 2024.

[47] Mengfei Xia, Yujun Shen, Changsong Lei, Yu Zhou, Ran Yi, Deli Zhao, Wenping Wang, and Yong-jin Liu. Towards more accurate diffusion model acceleration with a timestep aligner. *arXiv preprint arXiv:2310.09469*, 2023.

[48] Xingyu Xu and Yuejie Chi. Provably robust score-based diffusion posterior sampling for plug-and-play image reconstruction. *Advances in Neural Information Processing Systems*, 37:36148–36184, 2024.

[49] Shuchen Xue, Zhaoqiang Liu, Fei Chen, Shifeng Zhang, Tianyang Hu, Enze Xie, and Zhenguo Li. Accelerating diffusion sampling with optimized time steps. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 8292–8301, 2024.

[50] Burhaneddin Yaman, Seyed Amir Hossein Hosseini, and Mehmet Akçakaya. Zero-shot self-supervised learning for mri reconstruction. *arXiv preprint arXiv:2102.07737*, 2021.

[51] Lingxiao Yang, Shutong Ding, Yifan Cai, Jingyi Yu, Jingya Wang, and Ye Shi. Guidance with spherical gaussian constraint for conditional diffusion. *arXiv preprint arXiv:2402.03201*, 2024.

[52] Jiwen Yu, Yinhuai Wang, Chen Zhao, Bernard Ghanem, and Jian Zhang. Freedom: Training-free energy-guided conditional diffusion model. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 23174–23184, 2023.

[53] Jure Zbontar, Florian Knoll, Anuroop Sriram, Tullie Murrell, Zhengnan Huang, Matthew J Muckley, Aaron Defazio, Ruben Stern, Patricia Johnson, Mary Bruno, et al. fastmri: An open dataset and benchmarks for accelerated mri. *arXiv preprint arXiv:1811.08839*, 2018.

[54] 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*, pages 20895–20905, 2025.

[55] Kai Zhang, Yawei Li, Wangmeng Zuo, Lei Zhang, Luc Van Gool, and Radu Timofte. Plug-and-play image restoration with deep denoiser prior. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(10):6360–6376, 2021.

[56] Richard Zhang, Phillip Isola, and Alexei A Efros. Colorful image colorization. In *European conference on computer vision*, pages 649–666. Springer, 2016.

[57] Richard Zhang, Phillip Isola, Alexei A Efros, Eli Shechtman, and Oliver Wang. The unreasonable effectiveness of deep features as a perceptual metric. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 586–595, 2018.

[58] 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*, pages 1219–1229, 2023.

Appendix Contents

A The optimal denoiser for a Gaussian posterior **15**

B The reverse process in the time domain **16**

C Spectral form of the update rule **17**

D Gaussian posterior distribution **19**

E DPS algorithm **21**

E.1 The reverse process in the time domain 21

E.2 Spectral form of the update rule 23

E.3 Stochastic DDPM formulation 24

F Averaged Wasserstein distance **27**

G Kullback-Leibler divergence **29**

H ΠGDM algorithm **30**

H.1 The reverse process in the time domain 30

H.2 Spectral form of the update rule 32

I DiffPIR algorithm **33**

I.1 The reverse process in the time domain 33

I.2 Spectral form of the update rule 34

J Additional results: synthetic Gaussian distribution **36**

J.1 Optimization procedure 36

J.2 DPS algorithm 36

J.3 ΠGDM algorithm 36

J.4 DiffPIR algorithm 37

K Additional results: empirical distribution **39**

K.1 FFHQ dataset 39

K.2 ImageNet dataset 39

K.3 General degradation settings 44

K.4 fastMRI dataset 48

K.5 DiffPIR algorithm 54

L Additional results: realization-specific optimization **57**

M Covariance matrix estimation **60**

N Optimization time analysis **60**

A The optimal denoiser for a Gaussian posterior

Let $\mathbf{x}_0 \sim \mathcal{N}(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$ represent the distribution of the original dataset, where $\mathbf{x}_0 \in \mathbb{R}^d$. A general linear inverse problem is posed as: $\mathbf{y} = \mathbf{H}\mathbf{x}_0 + \mathbf{n}$ with $\mathbf{y} \in \mathbb{R}^m$, $\mathbf{H} \in \mathbb{R}^{m \times d}$ a known linear degradation matrix, and $\mathbf{n} \sim \mathcal{N}(\mathbf{0}, \sigma_y^2 \mathbf{I})$ i.i.d. additive Gaussian noise with known variance.

Under the Gaussian prior and linear measurement model, the posterior distribution of $\mathbf{x}_0$ given $\mathbf{y}$ is Gaussian; consequently, the MAP estimator coincides with the posterior mean and takes the form of a Wiener filter, which is optimal under mean-squared error.

For the Maximum A Posteriori (MAP) estimation, we seek to maximize the posterior distribution:

$$\max_{x_0} \log p(\mathbf{x}_0|\mathbf{x}_t, \mathbf{y})$$

Using Bayes' rule, this can be written as:

$$\min_{x_0} -\log\left[\frac{p(\mathbf{y}|\mathbf{x}_0, \mathbf{x}_t)p(\mathbf{x}_0|\mathbf{x}_t)}{p(\mathbf{y}|\mathbf{x}_t)}\right]$$

$$\min_{x_0} -\log p(\mathbf{y}|\mathbf{x}_0, \mathbf{x}_t) - \log p(\mathbf{x}_0|\mathbf{x}_t) \tag{21}$$

Since $\mathbf{y}$ is conditionally independent of $\mathbf{x}_t$ given $\mathbf{x}_0$, The conditional log-likelihood $\log p(\mathbf{y}|\mathbf{x}_0, \mathbf{x}_t)$ is given by:

$$p(\mathbf{y}|\mathbf{x}_0, \mathbf{x}_t) = \mathcal{N}(\mathbf{y}; \mathbf{H}\mathbf{x}_0, \sigma_y^2\mathbf{I}) = \frac{1}{\sqrt{\left(2\pi\sigma_y^2\right)^d}} \exp\left\{-\frac{1}{2\sigma_y^2}\left(\mathbf{y} - \mathbf{H}\mathbf{x}_0\right)^T\left(\mathbf{y} - \mathbf{H}\mathbf{x}_0\right)\right\}$$

$$\log p(\mathbf{y}|\mathbf{x}_0, \mathbf{x}_t) = -\frac{1}{2}\log\left(2\pi\sigma_y^2\right)^d - \frac{1}{2\sigma_y^2}\left(\mathbf{y} - \mathbf{H}\mathbf{x}_0\right)^T\left(\mathbf{y} - \mathbf{H}\mathbf{x}_0\right) \tag{22}$$

We next consider the second term in Eq. (21), $p(\mathbf{x}_0|\mathbf{x}_t)$. Applying Bayes' rule, we obtain

$$\log p(\mathbf{x}_0|\mathbf{x}_t) \propto \log p(\mathbf{x}_t|\mathbf{x}_0) + \log p(\mathbf{x}_0).$$

Through the diffusion process, the signal undergoes noise contamination, leading to the following marginal expression for $\mathbf{x}_t$:

$$\mathbf{x}_t = \sqrt{\bar{\alpha}_t}\mathbf{x}_0 + \sqrt{1-\bar{\alpha}_t}\boldsymbol{\epsilon} \quad \boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I}) \tag{23}$$

where α_t for $t \in [1, T]$ is referred to as the incremental noise schedule and $\bar{\alpha}_t = \prod_{i=1}^{t} \alpha_i$.

Hence, the conditional log-likelihood $\log p(\mathbf{x}_t|\mathbf{x}_0)$ is given by:

$$\begin{aligned} p(\mathbf{x}_t|\mathbf{x}_0) =& \mathcal{N}(\mathbf{x}_t; \sqrt{\bar{\alpha}_t}\,\mathbf{x}_0, (1-\bar{\alpha}_t)\mathbf{I}) \\ =& \frac{1}{\sqrt{\left(2\pi(1-\bar{\alpha}_t)\right)^d)}} \exp\left\{-\frac{1}{2(1-\bar{\alpha}_t)}\left(\mathbf{x}_t - \sqrt{\bar{\alpha}_t}\mathbf{x}_0\right)^T\left(\mathbf{x}_t - \sqrt{\bar{\alpha}_t}\mathbf{x}_0\right)\right\} \end{aligned}$$

$$\log\, p(\mathbf{x}_t|\mathbf{x}_0) = -\frac{1}{2}\log\left(2\pi(1-\bar{\alpha}_t)\right)^d - \frac{1}{2(1-\bar{\alpha}_t)}\left(\mathbf{x}_t - \sqrt{\bar{\alpha}_t}\mathbf{x}_0\right)^T\left(\mathbf{x}_t - \sqrt{\bar{\alpha}_t}\mathbf{x}_0\right) \tag{24}$$

The log-likelihood $\log p(\mathbf{x}_0)$ is given by:

$$p(\mathbf{x}_0) = \mathcal{N}(\mathbf{x}_0; \boldsymbol{\mu}, \boldsymbol{\Sigma}_0) = \frac{1}{\sqrt{(2\pi)^d\left|\boldsymbol{\Sigma}_0\right|}} \exp\left\{-\frac{1}{2}\left(\mathbf{x}_0 - \boldsymbol{\mu}_0\right)^T\boldsymbol{\Sigma}_0^{-1}\left(\mathbf{x}_0 - \boldsymbol{\mu}_0\right)\right\} \tag{25}$$

$$\log p(\mathbf{x}_0) = -\frac{1}{2}\log\left(2\pi\right)^d\left|\boldsymbol{\Sigma}_0\right| - \frac{1}{2}\left(\mathbf{x}_0 - \boldsymbol{\mu}_0\right)^T\boldsymbol{\Sigma}_0^{-1}\left(\mathbf{x}_0 - \boldsymbol{\mu}_0\right) \tag{26}$$

By substituting the expressions from Equations 22, 24, and 26 into Equation 21 and differentiating with respect to $\mathbf{x}_0$, we obtain:

$$-\frac{2\mathbf{H}^T\left(\mathbf{y}-\mathbf{H}\mathbf{x}_0\right)}{2{\sigma_y}^2}-\frac{2\sqrt{\bar{\alpha}_t}\left(\mathbf{x}_t-\sqrt{\bar{\alpha}_t}\mathbf{x}_0\right)}{2(1-\bar{\alpha}_t)}+\frac{2\boldsymbol{\Sigma}_0^{-1}\left(\mathbf{x}_0-\boldsymbol{\mu}_0\right)}{2}=0$$

This simplifies to:

$$-\frac{\mathbf{H}^T\left(\mathbf{y}-\mathbf{H}\mathbf{x}_0\right)}{{\sigma_y}^2}-\frac{\sqrt{\bar{\alpha}_t}\left(\mathbf{x}_t-\sqrt{\bar{\alpha}_t}\mathbf{x}_0\right)}{(1-\bar{\alpha}_t)}+\boldsymbol{\Sigma}_0^{-1}\left(\mathbf{x}_0-\boldsymbol{\mu}_0\right)=0$$

Resulting in:

$$-(1-\bar{\alpha}_t)\boldsymbol{\Sigma}_0\mathbf{H}^T\left(\mathbf{y}-\mathbf{H}\mathbf{x}_0\right)-\sigma_y^2\sqrt{\bar{\alpha}_t}\boldsymbol{\Sigma}_0\left(\mathbf{x}_t-\sqrt{\bar{\alpha}_t}\mathbf{x}_0\right)+\sigma_y^2(1-\bar{\alpha}_t)\left(\mathbf{x}_0-\boldsymbol{\mu}_0\right)=0$$

Thus:

$$\left((1-\bar{\alpha}_t)\boldsymbol{\Sigma}_0\mathbf{H}^T\mathbf{H}+\sigma_y^2\bar{\alpha}_t\boldsymbol{\Sigma}_0+I\sigma_y^2(1-\bar{\alpha}_t)\right)\mathbf{x}_0=(1-\bar{\alpha}_t)\boldsymbol{\Sigma}_0\mathbf{H}^T\mathbf{y}+\sigma_y^2\sqrt{\bar{\alpha}_t}\boldsymbol{\Sigma}_0\mathbf{x}_t+\sigma_y^2(1-\bar{\alpha}_t)\boldsymbol{\mu}_0$$

Finally:

$$\begin{aligned}\mathbf{x}_0^*=&\left((1-\bar{\alpha}_t)\boldsymbol{\Sigma}_0\mathbf{H}^T\mathbf{H}+\sigma_y^2\bar{\alpha}_t\boldsymbol{\Sigma}_0+I\sigma_y^2(1-\bar{\alpha}_t)\right)^{-1}\\&\left((1-\bar{\alpha}_t)\boldsymbol{\Sigma}_0\mathbf{H}^T\mathbf{y}+\sigma_y^2\sqrt{\bar{\alpha}_t}\boldsymbol{\Sigma}_0\mathbf{x}_t+\sigma_y^2(1-\bar{\alpha}_t)\boldsymbol{\mu}_0\right)\end{aligned}\tag{27}$$

B The reverse process in the time domain

Here, we present the reverse process in the time domain for DDIM [34]. Let $\mathbf{x}_0$ follow the distribution:

$$\mathbf{x}_0\sim\mathcal{N}(\boldsymbol{\mu}_0,\boldsymbol{\Sigma}_0),\quad \mathbf{x}_0\in\mathbb{R}^d$$

Using the procedure outlined in [34], the inference process is initialized from $\mathbf{x}_S\sim\mathcal{N}(\mathbf{0},\mathbf{I})$, where $\mathbf{x}_S\in\mathbb{R}^d$, and proceeds through the following iterative denoising process:[3]

$$\mathbf{x}_{s-1}(\eta)=\sqrt{\bar{\alpha}_{s-1}}\left(\frac{\mathbf{x}_s-\sqrt{1-\bar{\alpha}_s}\cdot\boldsymbol{\epsilon}_\theta(\mathbf{x}_s,s)}{\sqrt{\bar{\alpha}_s}}\right)+\sqrt{1-\bar{\alpha}_{s-1}-\sigma_s^2(\eta)}\cdot\boldsymbol{\epsilon}_\theta(\mathbf{x}_s,s)+\sigma_s(\eta)\mathbf{z}_s\tag{28}$$

The stochasticity of the sampling process is controlled by η, with

$$\sigma_s(\eta)=\eta\sqrt{\frac{1-\bar{\alpha}_{s-1}}{1-\bar{\alpha}_s}}\sqrt{1-\frac{\bar{\alpha}_s}{\bar{\alpha}_{s-1}}}.\tag{29}$$

Using the marginal relation in (3), we obtain

$$\boldsymbol{\epsilon}_\theta(\mathbf{x}_s,s)=\frac{\mathbf{x}_s-\sqrt{\bar{\alpha}_s}\hat{\mathbf{x}}_0}{\sqrt{1-\bar{\alpha}_s}}\qquad\hat{\mathbf{x}}_0=\frac{\mathbf{x}_s-\sqrt{1-\bar{\alpha}_s}\cdot\boldsymbol{\epsilon}_\theta(\mathbf{x}_s,s)}{\sqrt{\bar{\alpha}_s}}.\tag{30}$$

Substituting these expressions into the DDIM update gives

$$\mathbf{x}_{s-1}(\eta)=\sqrt{\bar{\alpha}_{s-1}}\hat{\mathbf{x}}_0+\sqrt{1-\bar{\alpha}_{s-1}-\sigma_s^2(\eta)}\left(\frac{\mathbf{x}_s-\sqrt{\bar{\alpha}_s}\hat{\mathbf{x}}_0}{\sqrt{1-\bar{\alpha}_s}}\right)+\sigma_s(\eta)\mathbf{z}_s\tag{31}$$

[3]We follow the DDIM notation in which the sampling steps $s\in\{1,\ldots,S\}$ correspond to a subsequence of the DDPM diffusion steps $t\in\{1,\ldots,T\}$. The DDIM subsequence is chosen such that its terminal timestep coincides with that of the DDPM process, i.e., $\tau_S=\tau_T$.

For deterministic DDIM sampling, we set $\eta = 0$ in (29), yielding $\sigma_s(0) = 0$. Therefore,

$$\begin{aligned}\mathbf{x}_{s-1}(\eta = 0) &= \sqrt{\bar{\alpha}_{s-1}}\hat{\mathbf{x}}_0 + \sqrt{1-\bar{\alpha}_{s-1}}\left(\frac{\mathbf{x}_s - \sqrt{\bar{\alpha}_s}\hat{\mathbf{x}}_0}{\sqrt{1-\bar{\alpha}_s}}\right) \\ &= \frac{\sqrt{1-\bar{\alpha}_{s-1}}}{\sqrt{1-\bar{\alpha}_s}}\mathbf{x}_s + \left[\sqrt{\bar{\alpha}_{s-1}} - \frac{\sqrt{\bar{\alpha}_s}\sqrt{1-\bar{\alpha}_{s-1}}}{\sqrt{1-\bar{\alpha}_s}}\right]\hat{\mathbf{x}}_0. \end{aligned} \tag{32}$$

We define the coefficients

$$a_s = \left[\frac{\sqrt{1-\bar{\alpha}_{s-1}}}{\sqrt{1-\bar{\alpha}_s}}\right]\mathbf{I} \qquad b_s = \left[\sqrt{\bar{\alpha}_{s-1}} - \frac{\sqrt{\bar{\alpha}_s}\sqrt{1-\bar{\alpha}_{s-1}}}{\sqrt{1-\bar{\alpha}_s}}\right]\mathbf{I}.$$

Accordingly, substituting the MAP estimator $\mathbf{x}_0^*$ from Equation 27 into the deterministic DDIM update yields

$$\mathbf{x}_{s-1} = a_s\mathbf{x}_s + b_s\mathbf{x}_0^*\,.$$

Expanding the MAP estimator gives

$$\begin{aligned}\mathbf{x}_{s-1} &= a_s\mathbf{x}_s + \\ &\quad b_s\left[\left((1-\bar{\alpha}_s)\boldsymbol{\Sigma}_0\mathbf{H}^T\mathbf{H} + \sigma_y^2\bar{\alpha}_s\boldsymbol{\Sigma}_0 + \mathbf{I}\sigma_y^2(1-\bar{\alpha}_s)\right)^{-1}\left((1-\bar{\alpha}_s)\boldsymbol{\Sigma}_0\mathbf{H}^T\mathbf{y} + \sigma_y^2\sqrt{\bar{\alpha}_s}\boldsymbol{\Sigma}_0\mathbf{x}_s + \sigma_y^2(1-\bar{\alpha}_s)\boldsymbol{\mu}_0\right)\right] \\ \mathbf{x}_{s-1} &= \left[a_s + b_s\left[\left((1-\bar{\alpha}_s)\boldsymbol{\Sigma}_0\mathbf{H}^T\mathbf{H} + \sigma_y^2\bar{\alpha}_s\boldsymbol{\Sigma}_0 + \mathbf{I}\sigma_y^2(1-\bar{\alpha}_s)\right)^{-1}\right]\sigma_y^2\sqrt{\bar{\alpha}_s}\boldsymbol{\Sigma}_0\right]\mathbf{x}_s \\ &\quad + b_s\left((1-\bar{\alpha}_s)\boldsymbol{\Sigma}_0\mathbf{H}^T\mathbf{H} + \sigma_y^2\bar{\alpha}_s\boldsymbol{\Sigma}_0 + \mathbf{I}\sigma_y^2(1-\bar{\alpha}_s)\right)^{-1}\left((1-\bar{\alpha}_s)\boldsymbol{\Sigma}_0\mathbf{H}^T\mathbf{y}\right) \\ &\quad + b_s\left((1-\bar{\alpha}_s)\boldsymbol{\Sigma}_0\mathbf{H}^T\mathbf{H} + \sigma_y^2\bar{\alpha}_s\boldsymbol{\Sigma}_0 + \mathbf{I}\sigma_y^2(1-\bar{\alpha}_s)\right)^{-1}\left(\sigma_y^2(1-\bar{\alpha}_s)\boldsymbol{\mu}_0\right).\end{aligned}$$

We define

$$\bar{\boldsymbol{\Sigma}}_{0,s} = (1-\bar{\alpha}_s)\boldsymbol{\Sigma}_0\mathbf{H}^T\mathbf{H} + \sigma_y^2\bar{\alpha}_s\boldsymbol{\Sigma}_0 + \mathbf{I}\sigma_y^2(1-\bar{\alpha}_s).$$

Using this notation, the update becomes

$$\begin{aligned}\mathbf{x}_{s-1} &= \left(a_s + b_s\left(\bar{\boldsymbol{\Sigma}}_{0,s}\right)^{-1}\sigma_y^2\sqrt{\bar{\alpha}_s}\boldsymbol{\Sigma}_0\right)\mathbf{x}_s + b_s\left(\bar{\boldsymbol{\Sigma}}_{0,s}\right)^{-1}\left((1-\bar{\alpha}_s)\boldsymbol{\Sigma}_0\mathbf{H}^T\mathbf{y}\right) \\ &\quad + b_s\left(\bar{\boldsymbol{\Sigma}}_{0,s}\right)^{-1}\left(\sigma_y^2(1-\bar{\alpha}_s)\boldsymbol{\mu}_0\right).\end{aligned} \tag{33}$$

C Spectral form of the update rule

To migrate to the spectral domain, we apply the Discrete Fourier Transform (DFT), represented by the unitary matrix $\mathbf{F}$, to both sides of the DDIM update:

$$\begin{aligned}\mathbf{x}_{s-1}^{\mathcal{F}} &= \mathbf{F}\left(a_s + b_s\bar{\boldsymbol{\Sigma}}_{0,s}^{-1}\sigma_y^2\sqrt{\bar{\alpha}_s}\boldsymbol{\Sigma}_0\right)\mathbf{x}_s\mathbf{F}^T \\ &\quad + \mathbf{F}b_s\bar{\boldsymbol{\Sigma}}_{0,s}^{-1}\left((1-\bar{\alpha}_s)\boldsymbol{\Sigma}_0\mathbf{H}^T\mathbf{y} + \sigma_y^2(1-\bar{\alpha}_s)\boldsymbol{\mu}_0\right)\mathbf{F}^T,\end{aligned} \tag{34}$$

where the superscript $\mathcal{F}$ indicates the Fourier-domain representation of a vector, e.g., $\mathbf{x}_s^{\mathcal{F}} = \mathbf{F}\mathbf{x}_s$.

$$\begin{aligned}\mathbf{x}_{s-1}^{\mathcal{F}} &= a_s\mathbf{x}_s^{\mathcal{F}} + \mathbf{F}b_s\left(\bar{\boldsymbol{\Sigma}}_{0,s}\right)^{-1}\mathbf{F}^T\mathbf{F}\sigma_y^2\sqrt{\bar{\alpha}_s}\boldsymbol{\Sigma}_0\mathbf{F}^T\mathbf{F}\mathbf{x}_s \\ &\quad + b_s\mathbf{F}\left(\bar{\boldsymbol{\Sigma}}_{0,s}\right)^{-1}\mathbf{F}^T\mathbf{F}(1-\bar{\alpha}_s)\boldsymbol{\Sigma}_0\mathbf{F}^T\mathbf{F}\mathbf{H}^T\mathbf{F}^T\mathbf{F}\mathbf{y} + b_s\mathbf{F}\left(\bar{\boldsymbol{\Sigma}}_{0,s}\right)^{-1}\mathbf{F}^T\mathbf{F}\sigma_y^2(1-\bar{\alpha}_s)\boldsymbol{\mu}_0\end{aligned}$$

$$\begin{aligned}\mathbf{x}_{s-1}^{\mathcal{F}} &= a_s\mathbf{x}_s^{\mathcal{F}} + b_s\mathbf{F}\left(\bar{\boldsymbol{\Sigma}}_{0,s}\right)^{-1}\mathbf{F}^T\sqrt{\bar{\alpha}_s}\sigma_y^2\mathbf{F}\boldsymbol{\Sigma}_0\mathbf{F}^T\mathbf{x}_s^{\mathcal{F}} \\ &\quad + b_s\mathbf{F}\left(\bar{\boldsymbol{\Sigma}}_{0,s}\right)^{-1}\mathbf{F}^T(1-\bar{\alpha}_s)\left(\mathbf{F}\boldsymbol{\Sigma}_0\mathbf{F}^T\right)\left(\mathbf{F}\mathbf{H}^T\mathbf{F}^T\right)\mathbf{y}^{\mathcal{F}} + b_s\mathbf{F}\left(\bar{\boldsymbol{\Sigma}}_{0,s}\right)^{-1}\mathbf{F}^T\sigma_y^2(1-\bar{\alpha}_s)\boldsymbol{\mu}_0^{\mathcal{F}}\end{aligned} \tag{35}$$

Assuming that $\mathbf{H}$ is circulant with eigenvalues $\{h_i\}_{i=1}^d$, and that $\boldsymbol{\Sigma}_0$ is also circulant, both matrices commute and can be diagonalized via the Discrete Fourier Transform (DFT) matrix $\mathbf{F}$, which greatly simplifies the matrix inversion. Specifically:

- $\mathbf{\Lambda}_0 = \mathbf{F}\mathbf{\Sigma}_0\mathbf{F}^T, \quad \mathbf{\Sigma}_0 = \mathbf{F}^T\mathbf{\Lambda}_0\mathbf{F}$
- $\mathbf{H} = \mathbf{F}^T\mathbf{\Lambda}_\mathbf{H}\mathbf{F}, \quad \mathbf{H^T} = \mathbf{F}^T\mathbf{\Lambda}_{\mathbf{H}^\mathbf{T}}\mathbf{F}, \quad \mathbf{H}^T\mathbf{H} = \mathbf{F}^T\mathbf{\Lambda}_{\mathbf{H}^T\mathbf{H}}\mathbf{F}$
- $a\mathbf{\Sigma}_0 + b\mathbf{I} = a\mathbf{F}\mathbf{\Lambda}_0\mathbf{F}^T + b\mathbf{F}\mathbf{I}\mathbf{F}^T = \mathbf{F}\left(a\mathbf{\Lambda}_0 + b\mathbf{I}\right)\mathbf{F}^T$
- $\mathbf{\Sigma}_0^{-1} = \mathbf{F}\mathbf{\Lambda}_0^{-1}\mathbf{F}^T, \quad \mathbf{\Lambda}_0^{-1} = \mathbf{F}^T\mathbf{\Sigma}_0^{-1}\mathbf{F}$

Therefore, we obtain:

$$\mathbf{\Sigma}_0\mathbf{H}^T\mathbf{H} = \mathbf{F}^T\mathbf{\Lambda}_0\mathbf{F}\mathbf{F}^T\mathbf{\Lambda}_{\mathbf{H}^T\mathbf{H}}\mathbf{F} = \mathbf{F}^T\mathbf{\Lambda}_0\mathbf{\Lambda}_{\mathbf{H}^T\mathbf{H}}\mathbf{F}$$

and:

$$\mathbf{F}\left(\bar{\mathbf{\Sigma}}_{0,s}\right)^{-1}\mathbf{F}^T = \mathbf{F}\left[(1-\bar{\alpha}_s)\mathbf{\Sigma}_0\mathbf{H}^T\mathbf{H} + \sigma_y^2\bar{\alpha}_s\mathbf{\Sigma}_0 + \sigma_y^2(1-\bar{\alpha}_s)\mathbf{I}\right]^{-1}\mathbf{F}^T$$

$$= \left((1-\bar{\alpha}_s)\mathbf{\Lambda}_0\mathbf{\Lambda}_{\mathbf{H}^T\mathbf{H}} + \sigma_y^2\bar{\alpha}_s\mathbf{\Lambda}_0 + \sigma_y^2(1-\bar{\alpha}_s)\mathbf{I}\right)^{-1}$$

$$= diag\left((1-\bar{\alpha}_s)\lambda_i h_i^2 + \sigma_y^2\bar{\alpha}_s\lambda_i + \sigma_y^2(1-\bar{\alpha}_s)\mathbf{I}\right)^{-1}$$

$$\mathbf{F}\left(\bar{\mathbf{\Sigma}}_{0,s}\right)^{-1}\mathbf{F}^T = diag\left(\frac{1}{(1-\bar{\alpha}_s)\lambda_i h_i^2 + \sigma_y^2\bar{\alpha}_s\lambda_i + \sigma_y^2(1-\bar{\alpha}_s)}\right)$$

We denote the expression as:

$$\mathbf{F}\left(\bar{\mathbf{\Sigma}}_{0,s}\right)^{-1}\mathbf{F}^T = diag\left(\frac{1}{\lambda_{sum}}\right) = \mathbf{\Lambda}_{sum}^{-1} \tag{36}$$

Substitute 36 into 35

$$\mathbf{x}_{s-1}^{\mathcal{F}} = a_s\mathbf{x}_s^{\mathcal{F}} + b_s\mathbf{\Lambda}_{sum}^{-1}\sqrt{\bar{\alpha}_s}\sigma_y^2\mathbf{\Lambda}_0\mathbf{x}_s^{\mathcal{F}} + b_s\mathbf{\Lambda}_{sum}^{-1}(1-\bar{\alpha}_s)\mathbf{\Lambda}_0\mathbf{\Lambda}_{\mathbf{H}^T}\mathbf{y}^{\mathcal{F}} + b_s\mathbf{\Lambda}_{sum}^{-1}\sigma_y^2(1-\bar{\alpha}_s)\boldsymbol{\mu}_0^{\mathcal{F}}$$

$$\mathbf{x}_{s-1}^{\mathcal{F}} = \left[a_s + b_s\mathbf{\Lambda}_{sum}^{-1}\sqrt{\bar{\alpha}_s}\sigma_y^2\mathbf{\Lambda}_0\right]\mathbf{x}_s^{\mathcal{F}} + b_s\mathbf{\Lambda}_{sum}^{-1}(1-\bar{\alpha}_s)\mathbf{\Lambda}_0\mathbf{\Lambda}_{\mathbf{H}^T}\mathbf{y}^{\mathcal{F}} + b_s\sigma_y^2(1-\bar{\alpha}_s)\mathbf{\Lambda}_{sum}^{-1}\boldsymbol{\mu}_0^{\mathcal{F}}$$

We can then recursively obtain $\mathbf{x}_m^{\mathcal{F}}$ for a general m:

$$\mathbf{x}_m^{\mathcal{F}} = \prod_{s=m+1}^{S}\left[a_s + b_s\mathbf{\Lambda}_{sum}^{-1}\sqrt{\bar{\alpha}_s}\sigma_y^2\mathbf{\Lambda}_0\right]\mathbf{x}_S^{\mathcal{F}} + \left[\sum_{i=m+1}^{S}\left[b_i\mathbf{\Lambda}_{sum}^{-1}(1-\bar{\alpha}_s)\mathbf{\Lambda}_0\mathbf{\Lambda}_{\mathbf{H}^T}\right]\prod_{j=m+1}^{i-1}\left[a_j + b_j\mathbf{\Lambda}_{sum}^{-1}\sqrt{\bar{\alpha}_s}\sigma_y^2\mathbf{\Lambda}_0\right]\right]\mathbf{y}^{\mathcal{F}}$$
$$+ \left[\sum_{i=m+1}^{S}\left[b_i\sigma_y^2(1-\bar{\alpha}_s)\mathbf{\Lambda}_{sum}^{-1}\right]\prod_{j=m+1}^{i-1}\left[a_j + b_j\mathbf{\Lambda}_{sum}^{-1}\sqrt{\bar{\alpha}_s}\sigma_y^2\mathbf{\Lambda}_0\right]\right]\boldsymbol{\mu}_0^{\mathcal{F}}$$

specifically for m=0:

$$\mathbf{x}_0^{\mathcal{F}} = \prod_{s=1}^{S}\left[a_s + b_s\mathbf{\Lambda}_{sum}^{-1}\sqrt{\bar{\alpha}_s}\sigma_y^2\mathbf{\Lambda}_0\right]\mathbf{x}_S^{\mathcal{F}} + \left[\sum_{i=1}^{S}\left[b_i\mathbf{\Lambda}_{sum}^{-1}(1-\bar{\alpha}_s)\mathbf{\Lambda}_0\mathbf{\Lambda}_{\mathbf{H}^T}\right]\prod_{j=1}^{i-1}\left[a_j + b_j\mathbf{\Lambda}_{sum}^{-1}\sqrt{\bar{\alpha}_s}\sigma_y^2\mathbf{\Lambda}_0\right]\right]\mathbf{y}^{\mathcal{F}}$$
$$+ \left[\sum_{i=1}^{S}\left[b_i\sigma_y^2(1-\bar{\alpha}_s)\mathbf{\Lambda}_{sum}^{-1}\right]\prod_{j=1}^{i-1}\left[a_j + b_j\mathbf{\Lambda}_{sum}^{-1}\sqrt{\bar{\alpha}_s}\sigma_y^2\mathbf{\Lambda}_0\right]\right]\boldsymbol{\mu}_0^{\mathcal{F}}$$

we will denote the following:

$$\mathbf{V}_1 = \prod_{s=1}^{S} \left[a_s + b_s \mathbf{\Lambda}_{sum}^{-1} \sqrt{\bar{\alpha}_s} \sigma_y^2 \mathbf{\Lambda}_0 \right]$$

$$\mathbf{V}_2 = \left[\sum_{i=1}^{S} \left[b_i \mathbf{\Lambda}_{sum}^{-1} (1 - \bar{\alpha}_s) \mathbf{\Lambda}_0 \mathbf{\Lambda}_{\mathbf{H}^T} \right] \prod_{j=1}^{i-1} \left[a_j + b_j \mathbf{\Lambda}_{sum}^{-1} \sqrt{\bar{\alpha}_s} \sigma_y^2 \mathbf{\Lambda}_0 \right] \right]$$

$$\mathbf{V}_3 = \left[\sum_{i=1}^{S} \left[b_i \sigma_y^2 (1 - \bar{\alpha}_s) \mathbf{\Lambda}_{sum}^{-1} \right] \prod_{j=1}^{i-1} \left[a_j + b_j \mathbf{\Lambda}_{sum}^{-1} \sqrt{\bar{\alpha}_s} \sigma_y^2 \mathbf{\Lambda}_0 \right] \right]$$

$$\mathbf{x}_0^{\mathcal{F}} = \mathbf{V}_1 \mathbf{x}_S^{\mathcal{F}} + \mathbf{V}_2 \mathbf{y}^{\mathcal{F}} + \mathbf{V}_3 \boldsymbol{\mu}_0^{\mathcal{F}} \tag{37}$$

Given the closed-form expression for $\mathbf{x}_0^{\mathcal{F}}$ in Equation 37, we derive its mean and covariance as functions of the problem settings.

Mean:

$$\begin{aligned} \boldsymbol{\mu}_{opt} = \mathbb{E}\left[\mathbf{V}_1 \mathbf{x}_S^{\mathcal{F}} + \mathbf{V}_2 \mathbf{y}^{\mathcal{F}} + \mathbf{V}_3 \boldsymbol{\mu}_0^{\mathcal{F}} \mid \mathbf{y} \right] &= \mathbf{V}_1 \mathbf{F} \mathbb{E}\left[\mathbf{x}_S \mid \mathbf{y} \right] + \mathbf{V}_2 \mathbf{F} \mathbb{E}\left[\mathbf{y} \mid \mathbf{y} \right] + \mathbf{V}_3 \mathbf{F} \mathbb{E}\left[\boldsymbol{\mu}_0 \mid \mathbf{y} \right] \\ &= \mathbf{V}_1 \mathbf{F} \mathbb{E}\left[\mathbf{x}_S \right] + \mathbf{V}_2 \mathbf{F} \mathbf{y} + \mathbf{V}_3 \mathbf{F} \boldsymbol{\mu}_0 \\ &= \mathbf{V}_2 \mathbf{y}^{\mathcal{F}} + \mathbf{V}_3 \boldsymbol{\mu}_0^{\mathcal{F}} \end{aligned}$$

$$\boxed{\boldsymbol{\mu}_{opt} = \mathbb{E}\left[\mathbf{V}_1 \mathbf{x}_S^{\mathcal{F}} + \mathbf{V}_2 \mathbf{y}^{\mathcal{F}} + \mathbf{V}_3 \boldsymbol{\mu}_0^{\mathcal{F}} \right] = \mathbf{V}_2 \mathbf{y}^{\mathcal{F}} + \mathbf{V}_3 \boldsymbol{\mu}_0^{\mathcal{F}}} \tag{38}$$

Covariance:

$$\begin{aligned} \mathbf{\Lambda}_{opt} =& \mathbb{E}[\left(\mathbf{V}_1 \mathbf{x}_S^{\mathcal{F}} + \mathbf{V}_2 \mathbf{y}^{\mathcal{F}} + \mathbf{V}_3 \boldsymbol{\mu}_0^{\mathcal{F}} - \mathbb{E}\left[\mathbf{V}_1 \mathbf{x}_S^{\mathcal{F}} + \mathbf{V}_2 \mathbf{y}^{\mathcal{F}} + \mathbf{V}_3 \boldsymbol{\mu}_0^{\mathcal{F}} | \mathbf{y} \right] \right) \\ & \left(\mathbf{V}_1 \mathbf{x}_S^{\mathcal{F}} + \mathbf{V}_2 \mathbf{y}^{\mathcal{F}} + \mathbf{V}_3 \boldsymbol{\mu}_0^{\mathcal{F}} - \mathbb{E}\left[\mathbf{V}_1 \mathbf{x}_S^{\mathcal{F}} + \mathbf{V}_2 \mathbf{y}^{\mathcal{F}} + \mathbf{V}_3 \boldsymbol{\mu}_0^{\mathcal{F}} | \mathbf{y} \right] \right)^T \mid \mathbf{y}] \\ &= \mathbb{E}\left[\left(\mathbf{V}_1 \mathbf{x}_S^{\mathcal{F}} + \mathbf{V}_2 \mathbf{y}^{\mathcal{F}} + \mathbf{V}_3 \boldsymbol{\mu}_0^{\mathcal{F}} - \mathbf{V}_2 \mathbf{y}^{\mathcal{F}} - \mathbf{V}_3 \boldsymbol{\mu}_0^{\mathcal{F}} \right) \left(\mathbf{V}_1 \mathbf{x}_S^{\mathcal{F}} + \mathbf{V}_2 \mathbf{y}^{\mathcal{F}} + \mathbf{V}_3 \boldsymbol{\mu}_0^{\mathcal{F}} - \mathbf{V}_2 \mathbf{y}^{\mathcal{F}} - \mathbf{V}_3 \boldsymbol{\mu}_0^{\mathcal{F}} \right)^T \mid \mathbf{y} \right] \\ &= \mathbb{E}\left[\left(\mathbf{V}_1 \mathbf{x}_S^{\mathcal{F}} \right) \left(\mathbf{V}_1 \mathbf{x}_S^{\mathcal{F}} \right)^T \mid \mathbf{y} \right] \\ &= \mathbf{V}_1 \mathbb{E}\left[\mathbf{x}_S^{\mathcal{F}} \left(\mathbf{x}_S^{\mathcal{F}} \right)^T \right] {\mathbf{V}_1}^T \\ &= \mathbf{V}_1 {\mathbf{V}_1}^T \end{aligned}$$

$$\boxed{\mathbf{\Lambda}_{opt} = \mathbf{V}_1^2} \tag{39}$$

D Gaussian posterior distribution

Here, we derive a closed-form expression for $p(\mathbf{x} \mid \mathbf{y})$ under a linear degradation model. Specifically, we assume that the clean signal $\mathbf{x} \sim \mathcal{N}(\boldsymbol{\mu}_0, \mathbf{\Sigma}_0)$ and the noise $\mathbf{n} \sim \mathcal{N}(\mathbf{0}, \sigma_n^2 \mathbf{I})$ are independent Gaussian vectors, and that the observation is given by a linear transformation $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}$, where $\mathbf{H}$ is a deterministic matrix. Under these assumptions, the joint distribution $p(\mathbf{x}, \mathbf{y})$ is Gaussian.

Define the concatenated vector $\mathbf{z} = \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} = \begin{pmatrix} \mathbf{x} \\ \mathbf{H}\mathbf{x} + \mathbf{n} \end{pmatrix}$. The mean and covariance of the jointly Gaussian vector $\mathbf{z}$ are given by

$$\mathbb{E}[\mathbf{z}] = \boldsymbol{\mu}_{(x,y)} = \begin{pmatrix} \boldsymbol{\mu}_x \\ \boldsymbol{\mu}_y \end{pmatrix}, \quad \mathbf{\Sigma}_{\mathbf{z}} = \mathbf{\Sigma}_{(x,y)} = \begin{pmatrix} \mathbf{\Sigma}_x & \mathbf{\Sigma}_{xy} \\ \mathbf{\Sigma}_{yx} & \mathbf{\Sigma}_y \end{pmatrix}$$

where:

$$\boldsymbol{\mu}_x = \boldsymbol{\mu}_0$$
$$\boldsymbol{\mu}_y = \mathbb{E}[\mathbf{y}] = \mathbb{E}[\mathbf{H}\mathbf{x} + \mathbf{n}] = \mathbf{H}\mathbb{E}[\mathbf{x}] + \mathbb{E}[\mathbf{n}] = \mathbf{H}\boldsymbol{\mu}_0.$$

Recall that, by the definition of the prior,

$$\boldsymbol{\Sigma}_x = \boldsymbol{\Sigma}_0.$$

The covariance of $\mathbf{y}$ is therefore

$$\boldsymbol{\Sigma}_y = \mathrm{Cov}[\mathbf{y}] = \mathbf{H}\boldsymbol{\Sigma}_x\mathbf{H}^T + \sigma_n^2\mathbf{I},$$

Hence,

$$\boxed{p(\mathbf{y}) = \mathcal{N}(\mathbf{y}; \mathbf{H}\boldsymbol{\mu}_0, \mathbf{H}\boldsymbol{\Sigma}_0\mathbf{H}^T + \sigma_n^2\mathbf{I})} \tag{40}$$

The cross-covariance between $\mathbf{x}$ and $\mathbf{y}$ is

$$\boldsymbol{\Sigma}_{xy} = \mathbb{E}[(\mathbf{x} - \boldsymbol{\mu}_x)(\mathbf{y} - \boldsymbol{\mu}_y)^T]$$

which gives

$$\boldsymbol{\Sigma}_{xy} = \mathrm{Cov}[\mathbf{x}, \mathbf{y}] = \boldsymbol{\Sigma}_x\mathbf{H}^T, \quad \boldsymbol{\Sigma}_{yx} = \mathrm{Cov}[\mathbf{y}, \mathbf{x}] = \mathbf{H}\boldsymbol{\Sigma}_x.$$

Therefore, the joint mean and covariance of $p(\mathbf{x}, \mathbf{y})$ are

$$\boldsymbol{\mu}_{(x,y)} = \begin{pmatrix} \boldsymbol{\mu}_0 \\ \mathbf{H}\boldsymbol{\mu}_0 \end{pmatrix}, \quad \text{and} \qquad \boldsymbol{\Sigma}_{(x,y)} = \begin{pmatrix} \boldsymbol{\Sigma}_0 & \boldsymbol{\Sigma}_0\mathbf{H}^T \\ \mathbf{H}^T\boldsymbol{\Sigma}_0 & \mathbf{H}\boldsymbol{\Sigma}_0\mathbf{H}^T + \sigma_n^2\mathbf{I} \end{pmatrix}.$$

To obtain $p(\mathbf{x} \mid \mathbf{y})$, we use the standard conditioning formula for jointly Gaussian random vectors. The resulting conditional mean and covariance are:

Mean of $p(\mathbf{x} \mid \mathbf{y})$:

$$\begin{aligned} \boldsymbol{\mu}_{x|y} &= \boldsymbol{\mu}_x + \boldsymbol{\Sigma}_{xy}\boldsymbol{\Sigma}_y^{-1}(\mathbf{y} - \boldsymbol{\mu}_y) \\ &= \boldsymbol{\mu}_0 + \boldsymbol{\Sigma}_0\mathbf{H}^T(\mathbf{H}\boldsymbol{\Sigma}_0\mathbf{H}^T + \sigma_n^2\mathbf{I})^{-1}(\mathbf{y} - \mathbf{H}\boldsymbol{\mu}_0) \end{aligned} \tag{41}$$

Covariance of $p(\mathbf{x} \mid \mathbf{y})$:

$$\begin{aligned} \boldsymbol{\Sigma}_{x|y} &= \boldsymbol{\Sigma}_x - \boldsymbol{\Sigma}_{xy}\boldsymbol{\Sigma}_y^{-1}\boldsymbol{\Sigma}_{yx} \\ &= \boldsymbol{\Sigma}_0 - \boldsymbol{\Sigma}_0\mathbf{H}^T(\mathbf{H}\boldsymbol{\Sigma}_0\mathbf{H}^T + \sigma_n^2\mathbf{I})^{-1}\mathbf{H}\boldsymbol{\Sigma}_0 \end{aligned} \tag{42}$$

Therefore, the conditional probability distribution $p(\mathbf{x} \mid \mathbf{y})$ is:

$$\boxed{\begin{aligned} p(\mathbf{x} \mid \mathbf{y}) = \mathcal{N}\big(\mathbf{x};\, &\boldsymbol{\mu}_0 + \boldsymbol{\Sigma}_0\mathbf{H}^T(\mathbf{H}\boldsymbol{\Sigma}_0\mathbf{H}^T + \sigma_n^2\mathbf{I})^{-1}(\mathbf{y} - \mathbf{H}\boldsymbol{\mu}_0), \\ &\boldsymbol{\Sigma}_0 - \boldsymbol{\Sigma}_0\mathbf{H}^T(\mathbf{H}\boldsymbol{\Sigma}_0\mathbf{H}^T + \sigma_n^2\mathbf{I})^{-1}\mathbf{H}\boldsymbol{\Sigma}_0\big) \end{aligned}}$$

Migrating to the spectral domain:

We next express the posterior mean and covariance in the spectral domain. Let $\mathbf{F}$ denote the unitary Discrete Fourier Transform (DFT) matrix, with the superscript $\mathcal{F}$ denoting the Fourier-domain representation. The posterior mean in the spectral domain is given by

$$\begin{aligned} \boldsymbol{\mu}^{\mathcal{F}}_{\mathbf{x}_0|\mathbf{y}} =& \mathbf{F}\mathbb{E}\left[\mathbf{x}_0 \mid \mathbf{y}\right] \\ =& \mathbf{F}\left[\boldsymbol{\mu}_0 + \boldsymbol{\Sigma}_0\mathbf{H}^T(\mathbf{H}\Sigma_0\mathbf{H}^T + \sigma_n^2\mathbf{I})^{-1}(\mathbf{y} - \mathbf{H}\boldsymbol{\mu}_0)\right] \\ =& \mathbf{F}\boldsymbol{\mu}_0 + \mathbf{F}\boldsymbol{\Sigma}_0\mathbf{F}^T\mathbf{F}\mathbf{H}^T\mathbf{F}^T\mathbf{F}(\mathbf{H}\boldsymbol{\Sigma}_0\mathbf{H}^T + \sigma_n^2\mathbf{I})^{-1}\mathbf{F}^T\mathbf{F}(\mathbf{y} - \mathbf{H}\boldsymbol{\mu}_0) \\ =& \boldsymbol{\mu}_0^{\mathcal{F}} + \boldsymbol{\Lambda}_0\boldsymbol{\Lambda}_{H^T}(\boldsymbol{\Lambda}_H\boldsymbol{\Lambda}_0\boldsymbol{\Lambda}_{H^T} + \sigma_n^2\mathbf{I})^{-1}(\mathbf{y}^{\mathcal{F}} - \boldsymbol{\Lambda}_H\boldsymbol{\mu}_0^{\mathcal{F}}) \end{aligned}$$

Using

$$\mathbf{H} = \mathbf{F}^T\boldsymbol{\Lambda}_{\mathbf{H}}\mathbf{F}, \quad \mathbf{H}^T = \mathbf{F}^T\boldsymbol{\Lambda}_{\mathbf{H}^T}\mathbf{F}, \quad \mathbf{H}^T\mathbf{H} = \mathbf{F}^T\boldsymbol{\Lambda}_{\mathbf{H}^T\mathbf{H}}\mathbf{F}$$

we obtain

$$\mathbf{F}\left[\mathbf{H}\boldsymbol{\Sigma}_0\mathbf{H}^T\right]\mathbf{F}^T = \mathbf{F}\mathbf{H}\mathbf{F}^T\mathbf{F}\boldsymbol{\Sigma}_0\mathbf{F}^T\mathbf{F}\mathbf{H}^T\mathbf{F}^T = \boldsymbol{\Lambda}_H\boldsymbol{\Lambda}_0\boldsymbol{\Lambda}_{H^T}.$$

Thus,

$$\boxed{\boldsymbol{\mu}^{\mathcal{F}}_{\mathbf{x}_0|\mathbf{y}} = \boldsymbol{\mu}^{\mathcal{F}}_0 + \boldsymbol{\Lambda}_0\boldsymbol{\Lambda}_{H^T}(\boldsymbol{\Lambda}_H\boldsymbol{\Lambda}_0\boldsymbol{\Lambda}_{H^T} + \sigma_n^2\mathbf{I})^{-1}(\mathbf{y}^{\mathcal{F}} - \boldsymbol{\Lambda}_H\boldsymbol{\mu}^{\mathcal{F}}_0)}} \tag{43}$$

The posterior covariance in the spectral domain is given by

$$\begin{aligned}
\boldsymbol{\Lambda}_{x_0|y} =& \mathbf{F}\Sigma_{\mathbf{x}_0|\mathbf{y}}\mathbf{F}^T \\
=& \mathbf{F}\left[\boldsymbol{\Sigma}_0 - \boldsymbol{\Sigma}_0\mathbf{H}^T(\mathbf{H}\Sigma_0\mathbf{H}^T + \sigma_n^2\mathbf{I})^{-1}\mathbf{H}\boldsymbol{\Sigma}_0\right]\mathbf{F}^T \\
=& \mathbf{F}\boldsymbol{\Sigma}_0\mathbf{F}^T - \mathbf{F}\boldsymbol{\Sigma}_0\mathbf{F}^T\mathbf{F}\mathbf{H}^T\mathbf{F}^T\mathbf{F}(\mathbf{H}\boldsymbol{\Sigma}_0\mathbf{H}^T + \sigma_n^2\mathbf{I})^{-1}\mathbf{F}^T\mathbf{F}\mathbf{H}\mathbf{F}^T\mathbf{F}\boldsymbol{\Sigma}_0\mathbf{F}^T \\
=& \boldsymbol{\Lambda}_0 - \boldsymbol{\Lambda}_0\boldsymbol{\Lambda}_{H^T}(\boldsymbol{\Lambda}_H\boldsymbol{\Lambda}_0\boldsymbol{\Lambda}_{H^T} + \sigma_n^2\mathbf{I})^{-1}\boldsymbol{\Lambda}_H\boldsymbol{\Lambda}_0
\end{aligned}$$

Therefore,

$$\boxed{\boldsymbol{\Lambda}_{x_0|y} = \boldsymbol{\Lambda}_0 - \boldsymbol{\Lambda}_0\boldsymbol{\Lambda}_{H^T}(\boldsymbol{\Lambda}_H\boldsymbol{\Lambda}_0\boldsymbol{\Lambda}_{H^T} + \sigma_n^2\mathbf{I})^{-1}\boldsymbol{\Lambda}_H\boldsymbol{\Lambda}_0} \tag{44}$$

E DPS algorithm

In this section, we present the DPS algorithm [9] from a spectral perspective describe its integration into our framework.

As discussed in Section 2, to avoid retraining a model for a specific degradation matrix $\mathbf{H}$, DPS [9] uses a Bayesian formulation, in which the posterior score gradient is expressed as

$$\nabla_{\mathbf{x}_t}\log p_t(\mathbf{x}_t|\mathbf{y}) = \nabla_{\mathbf{x}_t}\log p_t(\mathbf{x}_t) + \nabla_{\mathbf{x}_t}\log p_t(\mathbf{y}|\mathbf{x}_t). \tag{45}$$

While the first term is obtained directly from the pretrained diffusion model, the second, likelihood term is expressed by marginalizing over $\mathbf{x}_0$:

$$p(\mathbf{y}|\mathbf{x}_t) = \int_{\mathbf{x}_0} p(\mathbf{x}_0|\mathbf{x}_t)p(\mathbf{y}|\mathbf{x}_0)\,d\mathbf{x}_0 \tag{46}$$

Under the linear observation model $\mathbf{y} = \mathbf{H}\mathbf{x}_0 + \mathbf{n}$, with Gaussian measurement noise, the likelihood is

$$p(\mathbf{y}|\mathbf{x}_0) = \mathcal{N}(\mathbf{H}\mathbf{x}_0, \sigma_y^2\mathbf{I}),$$

or equivalently,

$$p(\mathbf{y}|\mathbf{x}_0) = \frac{1}{\sqrt{(2\pi)^d\sigma_y^2}}\exp\left[-\frac{||\mathbf{y} - \mathbf{H}\mathbf{x}_0||_2^2}{2\sigma^2}\right].$$

A key approximation in DPS is modeling the distribution $p(\mathbf{x}_0|\mathbf{x}_t)$ as a deterministic Gaussian centered at the estimated mean:

$$p(\mathbf{x}_0|\mathbf{x}_t) \approx \mathcal{N}(\hat{\mathbf{x}}_0, \mathbf{0})$$

Substituting this approximation into the marginalization in (46) yields

$$p_t(\mathbf{y}|\mathbf{x}_t) \approx \mathcal{N}(\mathbf{H}\hat{\mathbf{x}}_0, \sigma_{\mathbf{y}}^2\mathbf{I}),$$

which leads to the likelihood gradient

$$\nabla_{\mathbf{x}_t}\log p_t(\mathbf{y}|\mathbf{x}_t) \simeq -\frac{1}{\sigma_y^2}\nabla_{\mathbf{x}_t}\|\mathbf{y} - \mathbf{H}\hat{\mathbf{x}}_0\|_2^2.$$

The authors of [9] reweight this term using the heuristic $\zeta_i = \zeta'/|\mathbf{y} - \mathbf{H}\hat{\mathbf{x}}_0(\mathbf{x}_s)|$, which scales the step size inversely with the norm of the measurement error, where ζ' is a user-defined constant.

$$\nabla_{\mathbf{x}_t}\log p_t(\mathbf{y}|\mathbf{x}_t) \simeq -\zeta_s\nabla_{\mathbf{x}_t}\|\mathbf{y} - \mathbf{H}\hat{\mathbf{x}}_0\|_2^2.$$

E.1 The reverse process in the time domain

The original formulation of the DPS [9] algorithm is based on DDPM [22]. Here, we present it within the DDIM [34] framework under the Variance-Preserving (VP) parameterization [36].

$$\mathbf{x}_{s-1} = \sqrt{\bar{\alpha}_{s-1}}\hat{\mathbf{x}}_0 + \sqrt{1 - \bar{\alpha}_{s-1} - \sigma_s^2(\eta)}\,\boldsymbol{\epsilon}_\theta(\mathbf{x}_s, s) - \zeta_s\nabla_{\mathbf{x}_s}\|\mathbf{y} - \mathbf{H}\hat{\mathbf{x}}_0\|_2^2 + \sigma_s(\eta)\mathbf{z}_s. \tag{47}$$

where:

$$\sigma_s(\eta) = \eta\sqrt{\frac{1-\bar{\alpha}_{s-1}}{1-\bar{\alpha}_s}}\sqrt{1-\frac{\bar{\alpha}_s}{\bar{\alpha}_{s-1}}}.$$

We consider the deterministic setting, corresponding to $\eta = 0$.

$$\mathbf{x}_{s-1} = \sqrt{\bar{\alpha}_{s-1}}\hat{\mathbf{x}}_0 + \sqrt{1-\bar{\alpha}_{s-1}}\boldsymbol{\epsilon}_\theta(\mathbf{x}_s, s) - \zeta_s\nabla_{\mathbf{x}_s}\|\mathbf{y} - \mathbf{H}\hat{\mathbf{x}}_0\|_2^2$$

Using the marginal property,

$$\boldsymbol{\epsilon}_\theta(\mathbf{x}_s, s) = \frac{\mathbf{x}_s - \sqrt{\bar{\alpha}_s}\hat{\mathbf{x}}0}{\sqrt{1-\bar{\alpha}_s}}. \tag{48}$$

which, substituted into Equation 47, gives:

$$\mathbf{x}_{s-1} = \sqrt{\bar{\alpha}_{s-1}}\hat{\mathbf{x}}_0 + \sqrt{1-\bar{\alpha}_{s-1}}\left[\frac{\mathbf{x}_s - \sqrt{\bar{\alpha}_s}\hat{\mathbf{x}}_0}{\sqrt{1-\bar{\alpha}_s}}\right] - \zeta_s\nabla_{\mathbf{x}_s}\|\mathbf{y} - \mathbf{H}\hat{\mathbf{x}}_0\|_2^2$$

$$\mathbf{x}_{s-1} = \frac{\sqrt{1-\bar{\alpha}_{s-1}}}{\sqrt{1-\bar{\alpha}_s}}\mathbf{x}_s + \left[\sqrt{\bar{\alpha}_{s-1}} - \frac{\sqrt{\bar{\alpha}_s}\sqrt{1-\bar{\alpha}_{s-1}}}{\sqrt{1-\bar{\alpha}_s}}\right]\hat{\mathbf{x}}_0 - \zeta_s\nabla_{\mathbf{x}_t}\|\mathbf{y} - \mathbf{H}\hat{\mathbf{x}}_0\|_2^2$$

Introducing:

$$a_s = \left[\frac{\sqrt{1-\bar{\alpha}_{s-1}}}{\sqrt{1-\bar{\alpha}_s}}\right]\mathbf{I}$$

$$b_s = \left[\sqrt{\bar{\alpha}_{s-1}} - \frac{\sqrt{\bar{\alpha}_s}\sqrt{1-\bar{\alpha}_{s-1}}}{\sqrt{1-\bar{\alpha}_s}}\right]\mathbf{I}$$

we obtain:

$$\mathbf{x}_{s-1} = a_s\mathbf{x}_s + b_s\hat{\mathbf{x}}_0 - \zeta_s\nabla_{\mathbf{x}_s}\|\mathbf{y} - \mathbf{H}\hat{\mathbf{x}}_0\|_2^2$$

Where $\hat{\mathbf{x}}_0$ denotes the MMSE denoiser, corresponds to the denoised signal produced by a pretrained network. Assuming a Gaussian prior, the optimal solution $\mathbf{x}_0^*$ is given by [4]:

$$\mathbf{x}_0^* = \left(\bar{\alpha}_s\boldsymbol{\Sigma}_0 + (1-\bar{\alpha}_s)\mathbf{I}\right)^{-1}\left(\sqrt{\bar{\alpha}_s}\boldsymbol{\Sigma}_0\mathbf{x}_s + (1-\bar{\alpha}_s)\boldsymbol{\mu}_0\right)$$

Therefore we get:

$$\boxed{\mathbf{x}_{s-1} = a_s\mathbf{x}_s + b_s\mathbf{x}_0^* - \zeta_s\nabla_{\mathbf{x}_s}\|\mathbf{y} - \mathbf{H}\mathbf{x}_0^*\|_2^2} \tag{49}$$

Since $\mathbf{x}_0^*$ depends on $\mathbf{x}_s$ we denote it as $\mathbf{x}_0^*(\mathbf{x}_s)$:

$$\zeta_s\nabla_{x_s}||\mathbf{y} - \mathbf{H}\mathbf{x}_0^*||_2^2 = \zeta_s\nabla_{x_s}||\mathbf{y} - \mathbf{H}\mathbf{x}_0^*(\mathbf{x}_s)||_2^2$$

Hence, the gradient can be written as:

$$\zeta_s\nabla_{x_t}||\mathbf{y} - \mathbf{H}\mathbf{x}_0^*||_2^2 = \zeta_s\left[-2\left[\nabla_{x_s}\mathbf{x}_0^*(\mathbf{x}_s)\right]^T\mathbf{H}^T\left(\mathbf{y} - \mathbf{H}\mathbf{x}_0^*(\mathbf{x}_s)\right)\right]$$

Where:

$$\nabla_{x_s}\mathbf{x}_0^*(\mathbf{x}_s) = \nabla_{x_s}\left[\left(\bar{\alpha}_s\boldsymbol{\Sigma}_0 + (1-\bar{\alpha}_s)\mathbf{I}\right)^{-1}\left(\sqrt{\bar{\alpha}_s}\boldsymbol{\Sigma}_0\mathbf{x}_s + (1-\bar{\alpha}_s)\boldsymbol{\mu}_0\right)\right]$$

$$\nabla_{x_s}\mathbf{x}_0^*(\mathbf{x}_s) = \left(\bar{\alpha}_s\boldsymbol{\Sigma}_0 + (1-\bar{\alpha}_s)\mathbf{I}\right)^{-1}\sqrt{\bar{\alpha}_s}\boldsymbol{\Sigma}_0 \tag{50}$$

This yields:

$$\zeta_s\nabla_{x_s}||\mathbf{y} - \mathbf{H}\mathbf{x}_0^*||_2^2 = \zeta_s\left[-2\sqrt{\bar{\alpha}_s}\boldsymbol{\Sigma}_0\left(\bar{\alpha}_s\boldsymbol{\Sigma}_0 + (1-\bar{\alpha}_s)\mathbf{I}\right)^{-1}\mathbf{H}^T\left(\mathbf{y} - \mathbf{H}\mathbf{x}_0^*(\mathbf{x}_s)\right)\right] \tag{51}$$

Substituting into (49) gives:

$$\boxed{\mathbf{x}_{s-1} = a_s\mathbf{x}_s + b_s\mathbf{x}_0^* + 2\zeta_s\left[\sqrt{\bar{\alpha}_s}\boldsymbol{\Sigma}_0\left(\bar{\alpha}_s\boldsymbol{\Sigma}_0 + (1-\bar{\alpha}_s)\mathbf{I}\right)^{-1}\mathbf{H}^T\left(\mathbf{y} - \mathbf{H}\mathbf{x}_0^*\right)\right]} \tag{52}$$

E.2 Spectral form of the update rule

Next, we apply the Discrete Fourier Transform (DFT), denoted by $\mathcal{F}$, to Equation (52). Assuming that $\mathbf{\Sigma}_0$ and $\mathbf{H}$ are shift-invariant, both matrices are diagonalized by the DFT. Specifically we denote:

$$\begin{aligned}\mathbf{F}\mathbf{H}\mathbf{F}^T &= \mathbf{\Lambda}_h = \text{diag}(h_1, h_2, \ldots, h_d),\\ \mathbf{F}\mathbf{H}^T\mathbf{F}^T &= \mathbf{\Lambda}_h^* = \text{diag}(\bar{h}_1, \bar{h}_2, \ldots, \bar{h}_d),\\ \mathbf{F}\mathbf{H}^T\mathbf{H}\mathbf{F}^T &= \mathbf{\Lambda}_{|h|^2} = \text{diag}(|h_1|^2, |h_2|^2, \ldots, |h_d|^2).\end{aligned}$$

Applying the unitary Fourier operator $\mathbf{F}$ transforms Equation (52) into the frequency domain.

$$\mathbf{F}\mathbf{x}_{s-1} = \mathbf{F}\left[a_s\mathbf{x}_s + b_s\mathbf{x}_0^* + 2\zeta_s\left[\sqrt{\bar{\alpha}_s}\mathbf{\Sigma}_0\left(\bar{\alpha}_s\mathbf{\Sigma}_0 + (1-\bar{\alpha}_s)\mathbf{I}\right)^{-1}\mathbf{H}^T\left(\mathbf{y} - \mathbf{H}\mathbf{x}_0^*\right)\right]\right]$$

$$\mathbf{x}_{s-1}^{\mathcal{F}} = a_s\mathbf{x}_s^{\mathcal{F}} + b_s\mathbf{F}\mathbf{x}_0^* + 2\zeta_s\mathbf{F}\left[\sqrt{\bar{\alpha}_s}\mathbf{\Sigma}_0\left(\bar{\alpha}_s\mathbf{\Sigma}_0 + (1-\bar{\alpha}_s)\mathbf{I}\right)^{-1}\mathbf{H}^T\left(\mathbf{y} - \mathbf{H}\mathbf{x}_0^*\right)\right] \tag{53}$$

Where:

$$\begin{aligned}\mathbf{F}\mathbf{x}_0^* &= \mathbf{F}\left[\left(\bar{\alpha}_s\mathbf{\Sigma}_0 + (1-\bar{\alpha}_s)\mathbf{I}\right)^{-1}\left(\sqrt{\bar{\alpha}_s}\mathbf{\Sigma}_0\mathbf{x}_s + (1-\bar{\alpha}_s)\boldsymbol{\mu}_0\right)\right]\\ &= \mathbf{F}\left(\bar{\alpha}_s\mathbf{\Sigma}_0 + (1-\bar{\alpha}_s)\mathbf{I}\right)^{-1}\mathbf{F}^T\mathbf{F}\left(\sqrt{\bar{\alpha}_s}\mathbf{\Sigma}_0\mathbf{x}_s + (1-\bar{\alpha}_s)\boldsymbol{\mu}_0\right)\\ &= \left[\bar{\alpha}_s\mathbf{\Lambda}_0 + (1-\bar{\alpha}_s)\mathbf{I}\right]^{-1}\left(\sqrt{\bar{\alpha}_s}\mathbf{F}\mathbf{\Sigma}_0\mathbf{F}^T\mathbf{F}\mathbf{x}_s + (1-\bar{\alpha}_s)\mathbf{F}\boldsymbol{\mu}_0\right)\end{aligned}$$

$$\mathbf{F}\mathbf{x}_0^* = \left[\bar{\alpha}_s\mathbf{\Lambda}_0 + (1-\bar{\alpha}_s)\mathbf{I}\right]^{-1}\left(\sqrt{\bar{\alpha}_s}\mathbf{\Lambda}_0\mathbf{x}_s^{\mathcal{F}} + (1-\bar{\alpha}_s)\boldsymbol{\mu}_0^{\mathcal{F}}\right) \tag{54}$$

and:

$$\begin{aligned}&\mathbf{F}\left[\sqrt{\bar{\alpha}_s}\mathbf{\Sigma}_0\left(\bar{\alpha}_s\mathbf{\Sigma}_0 + (1-\bar{\alpha}_s)\mathbf{I}\right)^{-1}\mathbf{H}^T\left(\mathbf{y} - \mathbf{H}\mathbf{x}_0^*\right)\right]\\ &= \left[\sqrt{\bar{\alpha}_s}\mathbf{F}\mathbf{\Sigma}_0\mathbf{F}^T\mathbf{F}\left(\bar{\alpha}_s\mathbf{\Sigma}_0 + (1-\bar{\alpha}_s)\mathbf{I}\right)^{-1}\mathbf{F}^T\mathbf{F}\mathbf{H}^T\mathbf{F}^T\mathbf{F}\left(\mathbf{y} - \mathbf{H}\mathbf{x}_0^*\right)\right]\\ &= \left[\sqrt{\bar{\alpha}_s}\mathbf{\Lambda}_0\left[\bar{\alpha}_s\mathbf{\Lambda}_0 + (1-\bar{\alpha}_s)\mathbf{I}\right]^{-1}\mathbf{\Lambda}_{h^*}\left(\mathbf{F}\mathbf{y} - \mathbf{F}\mathbf{H}\mathbf{F}^T\mathbf{F}\mathbf{x}_0^*\right)\right]\\ &= \left[\sqrt{\bar{\alpha}_s}\mathbf{\Lambda}_0\left[\bar{\alpha}_s\mathbf{\Lambda}_0 + (1-\bar{\alpha}_s)\mathbf{I}\right]^{-1}\mathbf{\Lambda}_{h^*}\left(\mathbf{F}\mathbf{y} - \mathbf{F}\mathbf{H}\mathbf{F}^T\mathbf{F}\mathbf{x}_0^*\right)\right]\\ &= \left[\sqrt{\bar{\alpha}_s}\mathbf{\Lambda}_0\left[\bar{\alpha}_s\mathbf{\Lambda}_0 + (1-\bar{\alpha}_s)\mathbf{I}\right]^{-1}\mathbf{\Lambda}_{h^*}\left(\mathbf{y}^{\mathcal{F}} - \mathbf{\Lambda}_h\mathbf{F}\mathbf{x}_0^*\right)\right].\end{aligned} \tag{55}$$

By defining:

$$\begin{aligned}c_s &= \sqrt{\bar{\alpha}_s}\mathbf{\Lambda}_0\left[\bar{\alpha}_s\mathbf{\Lambda}_0 + (1-\bar{\alpha}_s)\mathbf{I}\right]^{-1}\\ d_s &= (1-\bar{\alpha}_s)\left[\bar{\alpha}_s\mathbf{\Lambda}_0 + (1-\bar{\alpha}_s)\mathbf{I}\right]^{-1}\end{aligned}$$

Equation 54 can be rewritten as:

$$\mathbf{F}\mathbf{x}_0^* = c_s\mathbf{x}_s^{\mathcal{F}} + d_s\boldsymbol{\mu}_0^{\mathcal{F}} \tag{56}$$

Substitute it into Equation 55, we get:

$$\mathbf{F}\left[\sqrt{\bar{\alpha}_s}\mathbf{\Sigma}_0\left(\bar{\alpha}_s\mathbf{\Sigma}_0 + (1-\bar{\alpha}_s)\mathbf{I}\right)^{-1}\mathbf{H}^T\left(\mathbf{y} - \mathbf{H}\mathbf{x}_0^*\right)\right] = \left[c_s\mathbf{\Lambda}_{h^*}\left(\mathbf{y}^{\mathcal{F}} - \mathbf{\Lambda}_h(c_s\mathbf{x}_s^{\mathcal{F}} + d_s\boldsymbol{\mu}_0^{\mathcal{F}})\right)\right] \tag{57}$$

Finally, Equation 53 can be rewritten as:

$$\mathbf{x}_{t-1}^{\mathcal{F}} = a_s\mathbf{x}_s^{\mathcal{F}} + b_s[c_s\mathbf{x}_s^{\mathcal{F}} + d_s\boldsymbol{\mu}_0^{\mathcal{F}}] + 2\zeta_s\left[c_s\mathbf{\Lambda}_{h^*}\left(\mathbf{y}^{\mathcal{F}} - \mathbf{\Lambda}_h(c_s\mathbf{x}_s^{\mathcal{F}} + d_s\boldsymbol{\mu}_0^{\mathcal{F}})\right)\right]$$

$$\mathbf{x}_{t-1}^{\mathcal{F}} = [a_s + b_sc_s - 2\zeta_sc_s\mathbf{\Lambda}_{h^*}\mathbf{\Lambda}_hc_s]\mathbf{x}_s^{\mathcal{F}} + [2\zeta_sc_s\mathbf{\Lambda}_{h^*}]\mathbf{y}^{\mathcal{F}} + [b_sd_s - 2\zeta_sc_s\mathbf{\Lambda}_{h^*}\mathbf{\Lambda}_hd_s]\boldsymbol{\mu}_0^{\mathcal{F}}$$

Denoting:

$$\begin{aligned}\mathbf{G}(s) &= [a_s + b_sc_s - 2\zeta_sc_s\mathbf{\Lambda}_{h^*}\mathbf{\Lambda}_hc_s]\\ \mathbf{Q}(s) &= [2\zeta_sc_s\mathbf{\Lambda}_{h^*}]\\ \mathbf{M}(s) &= [b_sd_s - 2\zeta_sc_s\mathbf{\Lambda}_{h^*}\mathbf{\Lambda}_hd_s]\end{aligned}$$

$$\mathbf{x}_{s-1}^{\mathcal{F}} = \mathbf{G}(s)\mathbf{x}_s^{\mathcal{F}} + \mathbf{Q}(s)\mathbf{y}^{\mathcal{F}} + \mathbf{M}(s)\boldsymbol{\mu}_0^{\mathcal{F}} \tag{58}$$

By recursively applying Equation (58), we obtain a frequency-domain representation for each diffusion step $\mathbf{x}_l^{\mathcal{F}}$:

$$\mathbf{x}_l^{\mathcal{F}} = \left[\prod_{j=l+1}^{S} \mathbf{G}(j)\right] \mathbf{x}_S^{\mathcal{F}} + \left[\sum_{i=l+1}^{S} \left(\prod_{j=l+1}^{i-1} \mathbf{G}(j)\right) \mathbf{Q}(i)\right] \mathbf{y}^{\mathcal{F}} + \left[\sum_{i=l+1}^{S} \left(\prod_{j=l+1}^{i-1} \mathbf{G}(j)\right) \mathbf{M}(i)\right] {\boldsymbol{\mu}_\mathbf{0}}^{\mathcal{F}}$$

Specifically for $l = 0$:

$$\mathbf{x}_0^{\mathcal{F}} = \left[\prod_{j=1}^{S} \mathbf{G}(j)\right] \mathbf{x}_S^{\mathcal{F}} + \left[\sum_{i=1}^{S} \left(\prod_{j=1}^{i-1} \mathbf{G}(j)\right) \mathbf{Q}(i)\right] \mathbf{y}^{\mathcal{F}} + \left[\sum_{i=1}^{S} \left(\prod_{j=1}^{i-1} \mathbf{G}(j)\right) \mathbf{M}(i)\right] {\boldsymbol{\mu}_\mathbf{0}}^{\mathcal{F}}$$

Denoting:

$$\mathbf{D}_1 = \prod_{j=l+1}^{S} \mathbf{G}(j)$$

$$\mathbf{D}_2 = \sum_{i=l+1}^{S} \left(\prod_{j=l+1}^{i-1} \mathbf{G}(j)\right) \mathbf{Q}(i)$$

$$\mathbf{D}_3 = \sum_{i=l+1}^{S} \left(\prod_{j=l+1}^{i-1} \mathbf{G}(j)\right) \mathbf{M}(i)$$

$$\mathbf{x}_0^{\mathcal{F}} = \mathbf{D}_1 \mathbf{x}_S^{\mathcal{F}} + \mathbf{D}_2 \mathbf{y}^{\mathcal{F}} + \mathbf{D}_3 {\boldsymbol{\mu}_\mathbf{0}}^{\mathcal{F}} \tag{59}$$

Equation 59 can be interpreted as a transfer function, describing the relationship between the input Gaussian noise $\mathbf{x}_S^{\mathcal{F}} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ and the estimated output signal $\mathbf{x}_{0,\text{DPS}}^{\mathcal{F}}$. The output signal is also Gaussian, with the following distributional properties:

Mean:

$$\mathbb{E}\left[\mathbf{x}_0^{\mathcal{F}}\right] = \mathbf{D}_2 \mathbf{y}^{\mathcal{F}} + \mathbf{D}_3 \boldsymbol{\mu}_0^{\mathcal{F}}$$

Covariance:

$$\boldsymbol{\Sigma}_{\mathbf{x}_0^{\mathcal{F}}} = \mathbb{E}\left[\left[\mathbf{x}_0^{\mathcal{F}} - \mathbb{E}\left[\mathbf{x}_0^{\mathcal{F}}\right]\right]\left[\mathbf{x}_0^{\mathcal{F}} - \mathbb{E}\left[\mathbf{x}_0^{\mathcal{F}}\right]\right]^T\right] = [\mathbf{D}_1]^2$$

$$\boxed{\hat{\mathbf{x}}_0^{\mathcal{F}} \sim \mathcal{N}\big(\mathbf{D}_2 \mathbf{y}^{\mathcal{F}} + \mathbf{D}_3 \boldsymbol{\mu}_0^{\mathcal{F}}\ ,\ \mathbf{D}_1^2\big)} \tag{60}$$

E.3 Stochastic DDPM formulation

We provide below the main steps in the analytical derivation of the stochastic DDPM [22] transition within the DPS framework [9]. A key property of our framework is that the optimization is formulated through a probabilistic characterization of the posterior distribution, rather than individual sampling trajectories. Consequently, the overall derivation remains unchanged, apart from modified coefficients, indicated by tildes, and the addition of the stochastic term. Since the derivation closely follows the deterministic case, we focus only on the main modifications.

E.3.1 The reverse process in the time domain

The original DPS [9] formulation is based on the stochastic DDPM [22] reverse process and can be written as

$$\mathbf{x}_{t-1} = \frac{1}{\sqrt{\alpha_t}} \left(\frac{\alpha_t - \bar{\alpha}_t}{1 - \bar{\alpha}_t}\right) \mathbf{x}_t + \sqrt{\bar{\alpha}_{t-1}} \left(\frac{1 - \alpha_t}{1 - \bar{\alpha}_t}\right) \hat{\mathbf{x}}_0 + \sigma_t \mathbf{z}_t - \zeta_t \nabla_{\mathbf{x}_t} \|\mathbf{y} - \mathbf{H}\hat{\mathbf{x}}_0\|_2^2 \tag{61}$$

Where

$$\sigma_t = \sqrt{\frac{1 - \bar{\alpha}_{t-1}}{1 - \bar{\alpha}_t} (1 - \alpha_t)},$$

and $\mathbf{z}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ denotes the stochastic noise term.

Defining

$$\tilde{a}_t = \frac{1}{\sqrt{\alpha_t}} \left(\frac{\alpha_t - \bar{\alpha}_t}{1 - \bar{\alpha}_t} \right), \qquad \tilde{b}_t = \sqrt{\bar{\alpha}_{t-1}} \left(\frac{1 - \alpha_t}{1 - \bar{\alpha}_t} \right),$$

the DPS update can be written as

$$\mathbf{x}_{t-1} = \tilde{a}_s \mathbf{x}_t + \tilde{b}_s \hat{\mathbf{x}}_0 + \sigma_t \mathbf{z}_t - \zeta_t \nabla_{\mathbf{x}_t} \|\mathbf{y} - \mathbf{H}\hat{\mathbf{x}}_0\|_2^2 \tag{62}$$

where $\hat{\mathbf{x}}_0$ denotes the denoised estimate of the clean signal obtained from the pretrained diffusion model. Under the Gaussian prior assumption, the MMSE estimator of $\mathbf{x}_0$ is given in closed form [4] by

$$\mathbf{x}_0^* = (\bar{\alpha}_t \boldsymbol{\Sigma}_0 + (1 - \bar{\alpha}_t)\mathbf{I})^{-1} \left(\sqrt{\bar{\alpha}_t} \boldsymbol{\Sigma}_0 \mathbf{x}_t + (1 - \bar{\alpha}_t) \boldsymbol{\mu}_0 \right).$$

Substituting this estimate into the DPS update yields

$$\boxed{\mathbf{x}_{t-1} = \tilde{a}_t \mathbf{x}_t + \tilde{b}_t \mathbf{x}_0^* + \sigma_t \mathbf{z}_t - \zeta_t \nabla_{\mathbf{x}_t} \|\mathbf{y} - \mathbf{H}\mathbf{x}_0^*\|_2^2} \tag{63}$$

Substituting the gradient expression from (51) into (63) yields

$$\boxed{\mathbf{x}_{t-1} = \tilde{a}_t \mathbf{x}_t + \tilde{b}_t \mathbf{x}_0^* + \sigma_t \mathbf{z}_t + 2\zeta_t \left[\sqrt{\bar{\alpha}_t} \boldsymbol{\Sigma}_0 \left(\bar{\alpha}_t \boldsymbol{\Sigma}_0 + (1 - \bar{\alpha}_t)\mathbf{I} \right)^{-1} \mathbf{H}^T \left(\mathbf{y} - \mathbf{H}\mathbf{x}_0^* \right) \right].} \tag{64}$$

E.3.2 Spectral form of the update rule

Next, we apply the Discrete Fourier Transform (DFT), denoted by $\mathcal{F}$, to Equation (64). Assuming that $\boldsymbol{\Sigma}_0$ and $\mathbf{H}$ are shift-invariant, both matrices are diagonalized by the DFT.

Applying the unitary Fourier operator $\mathbf{F}$ transforms Equation (64) into the frequency domain.

$$\mathbf{F}\mathbf{x}_{t-1} = \mathbf{F} \left[\tilde{a}_t \mathbf{x}_t + \tilde{b}_t \mathbf{x}_0^* + \sigma_t \mathbf{z}_t + 2\zeta_t \left[\sqrt{\bar{\alpha}_t} \boldsymbol{\Sigma}_0 \left(\bar{\alpha}_t \boldsymbol{\Sigma}_0 + (1 - \bar{\alpha}_t)\mathbf{I} \right)^{-1} \mathbf{H}^T \left(\mathbf{y} - \mathbf{H}\mathbf{x}_0^* \right) \right] \right].$$

$$\mathbf{x}_{t-1}^{\mathcal{F}} = \tilde{a}_t \mathbf{x}_t^{\mathcal{F}} + \tilde{b}_t \mathbf{F}\mathbf{x}_0^* + \sigma_t \mathbf{z}_t^{\mathcal{F}} + 2\zeta_t \mathbf{F} \left[\sqrt{\bar{\alpha}_t} \boldsymbol{\Sigma}_0 \left(\bar{\alpha}_t \boldsymbol{\Sigma}_0 + (1 - \bar{\alpha}_t)\mathbf{I} \right)^{-1} \mathbf{H}^T \left(\mathbf{y} - \mathbf{H}\mathbf{x}_0^* \right) \right] \tag{65}$$

By substiting the expression from Equesions (56) and (57) Equation (65) can be rewritten as:

$$\mathbf{x}_{t-1}^{\mathcal{F}} = \tilde{a}_t \mathbf{x}_t^{\mathcal{F}} + \tilde{b}_t [c_t \mathbf{x}_t^{\mathcal{F}} + d_t \boldsymbol{\mu}_0^{\mathcal{F}}] + 2\zeta_t \left[c_t \boldsymbol{\Lambda}_{h^*} \left(\mathbf{y}^{\mathcal{F}} - \boldsymbol{\Lambda}_h (c_t \mathbf{x}_t^{\mathcal{F}} + d_t \boldsymbol{\mu}_0^{\mathcal{F}}) \right) \right] + \sigma_t \mathbf{z}_t^{\mathcal{F}}$$

$$\mathbf{x}_{t-1}^{\mathcal{F}} = [\tilde{a}_t + \tilde{b}_t c_t - 2\zeta_t c_t \boldsymbol{\Lambda}_{h^*} \boldsymbol{\Lambda}_h c_t] \mathbf{x}_t^{\mathcal{F}} + [2\zeta_t c_t \boldsymbol{\Lambda}_{h^*}] \mathbf{y}^{\mathcal{F}} + [\tilde{b}_t d_t - 2\zeta_t c_t \boldsymbol{\Lambda}_{h^*} \boldsymbol{\Lambda}_h d_t] \boldsymbol{\mu}_0^{\mathcal{F}} + \sigma_t \mathbf{z}_t^{\mathcal{F}}$$

where

$$c_t = \sqrt{\bar{\alpha}_t} \boldsymbol{\Lambda}_0 \left[\bar{\alpha}_t \boldsymbol{\Lambda}_0 + (1 - \bar{\alpha}_t)\mathbf{I} \right]^{-1}$$

$$d_s = (1 - \bar{\alpha}_t) \left[\bar{\alpha}_t \boldsymbol{\Lambda}_0 + (1 - \bar{\alpha}_t)\mathbf{I} \right]^{-1}$$

Denoting:

$$\tilde{\mathbf{G}}(t) = [\tilde{a}_t + \tilde{b}_t c_t - 2\zeta_t c_t \boldsymbol{\Lambda}_{h^*} \boldsymbol{\Lambda}_h c_t]$$

$$\tilde{\mathbf{Q}}(t) = [2\zeta_t c_t \boldsymbol{\Lambda}_{h^*}]$$

$$\tilde{\mathbf{M}}(t) = [\tilde{b}_t d_t - 2\zeta_t c_t \boldsymbol{\Lambda}_{h^*} \boldsymbol{\Lambda}_h d_t]$$

$$\mathbf{x}_{t-1}^{\mathcal{F}} = \tilde{\mathbf{G}}(t) \mathbf{x}_t^{\mathcal{F}} + \tilde{\mathbf{Q}}(t) \mathbf{y}^{\mathcal{F}} + \tilde{\mathbf{M}}(t) \boldsymbol{\mu}_0^{\mathcal{F}} + \sigma_t \mathbf{z}_t^{\mathcal{F}} \tag{66}$$

By recursively applying Equation (66), we obtain a frequency-domain representation for each diffusion step $\mathbf{x}_l^{\mathcal{F}}$

$$\mathbf{x}_l^{\mathcal{F}} = \left[\prod_{j=l+1}^{T} \tilde{\mathbf{G}}(j)\right] \mathbf{x}_T^{\mathcal{F}} + \left[\sum_{i=l+1}^{T} \left(\prod_{j=l+1}^{i-1} \tilde{\mathbf{G}}(j)\right) \tilde{\mathbf{Q}}(i)\right] \mathbf{y}^{\mathcal{F}} + \left[\sum_{i=l+1}^{T} \left(\prod_{j=l+1}^{i-1} \tilde{\mathbf{G}}(j)\right) \tilde{\mathbf{M}}(i)\right] \boldsymbol{\mu_0}^{\mathcal{F}} + \left[\sum_{i=l+1}^{T} \left(\prod_{j=l+1}^{i-1} \tilde{\mathbf{G}}(j)\right) \sigma_i \mathbf{z}_i^{\mathcal{F}}\right]. \tag{67}$$

Specifically for $l = 0$:

$$\mathbf{x}_0^{\mathcal{F}} = \left[\prod_{j=1}^{T} \tilde{\mathbf{G}}(j)\right] \mathbf{x}_T^{\mathcal{F}} + \left[\sum_{i=1}^{T} \left(\prod_{j=1}^{i-1} \tilde{\mathbf{G}}(j)\right) \tilde{\mathbf{Q}}(i)\right] \mathbf{y}^{\mathcal{F}} + \left[\sum_{i=1}^{T} \left(\prod_{j=1}^{i-1} \tilde{\mathbf{G}}(j)\right) \tilde{\mathbf{M}}(i)\right] \boldsymbol{\mu_0}^{\mathcal{F}} + \left[\sum_{i=1}^{T} \left(\prod_{j=1}^{i-1} \tilde{\mathbf{G}}(j)\right) \sigma_i \mathbf{z}_i^{\mathcal{F}}\right]. \tag{68}$$

Denoting:

$$\tilde{\mathbf{D}}_1 = \prod_{j=l+1}^{T} \tilde{\mathbf{G}}(j)$$

$$\tilde{\mathbf{D}}_2 = \sum_{i=l+1}^{T} \left(\prod_{j=l+1}^{i-1} \tilde{\mathbf{G}}(j)\right) \tilde{\mathbf{Q}}(i)$$

$$\tilde{\mathbf{D}}_3 = \sum_{i=l+1}^{T} \left(\prod_{j=l+1}^{i-1} \tilde{\mathbf{G}}(j)\right) \tilde{\mathbf{M}}(i)$$

We get:

$$\mathbf{x}_0^{\mathcal{F}} = \tilde{\mathbf{D}}_1 \mathbf{x}_T^{\mathcal{F}} + \tilde{\mathbf{D}}_2 \mathbf{y}^{\mathcal{F}} + \tilde{\mathbf{D}}_3 \boldsymbol{\mu_0}^{\mathcal{F}} + \left[\sum_{i=1}^{T} \left(\prod_{j=1}^{i-1} \tilde{\mathbf{G}}(j)\right) \sigma_i \mathbf{z}_i^{\mathcal{F}}\right] \tag{69}$$

Equation (69) can be interpreted as a transfer function, describing the relationship between the initial Gaussian noise $\mathbf{x}_T^{\mathcal{F}} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ and the estimated output signal $\mathbf{x}_0^{\mathcal{F}}$. Compared with the corresponding deterministic expression in Equation (59), the DDPM formulation includes an additional stochastic term arising from the noise injected throughout the reverse process. Nevertheless, the output remains Gaussian, with the following distributional properties:

Mean:

$$\mathbb{E}\left[\mathbf{x}_0^{\mathcal{F}}\right] = \tilde{\mathbf{D}}_2 \mathbf{y}^{\mathcal{F}} + \tilde{\mathbf{D}}_3 \boldsymbol{\mu}_0^{\mathcal{F}}$$

Covariance:

$$\boldsymbol{\Sigma}_{\mathbf{x}_0^{\mathcal{F}}} = \mathbb{E}\left[\left(\mathbf{x}_0^{\mathcal{F}} - \mathbb{E}\left[\mathbf{x}_0^{\mathcal{F}}\right]\right)\left(\mathbf{x}_0^{\mathcal{F}} - \mathbb{E}\left[\mathbf{x}_0^{\mathcal{F}}\right]\right)^T\right]. \tag{70}$$

Substituting the expression for $\mathbf{x}_0^{\mathcal{F}}$ and using the corresponding mean, the deterministic terms involving $\mathbf{y}^{\mathcal{F}}$ and $\boldsymbol{\mu}_0^{\mathcal{F}}$ cancel, yielding

$$\boldsymbol{\Sigma}_{\mathbf{x}_0^{\mathcal{F}}} = \mathbb{E}\left[\left(\tilde{\mathbf{D}}_1 \mathbf{x}_T^{\mathcal{F}} + \left[\sum_{i=1}^{T} \left(\prod_{j=1}^{i-1} \tilde{\mathbf{G}}(j)\right) \sigma_i \mathbf{z}_i^{\mathcal{F}}\right]\right)\left(\tilde{\mathbf{D}}_1 \mathbf{x}_T^{\mathcal{F}} + \left[\sum_{i=1}^{T} \left(\prod_{j=1}^{i-1} \tilde{\mathbf{G}}(j)\right) \sigma_i \mathbf{z}_i^{\mathcal{F}}\right]\right)^T\right].$$

Since $\mathbf{x}_T \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ and $\mathbf{z}_i \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ are mutually independent for all i, their Fourier-domain representations also have identity covariance. In particular,

$$\mathbb{E}\left[\mathbf{x}_T^{\mathcal{F}}\mathbf{x}_T^{\mathcal{F}^T}\right] = \mathbb{E}\left[\mathbf{F}^T\mathbf{x}_T\mathbf{x}_T^T\mathbf{F}\right] = \mathbf{F}^T\mathbb{E}\left[\mathbf{x}_T\mathbf{x}_T^T\right]\mathbf{F} = \mathbf{I},$$

and, for all i:

$$\mathbb{E}\left[\mathbf{z}_i{\mathbf{z}_i}^T\right] = \mathbf{I}$$

Moreover, all cross-covariance terms vanish due to independence. Therefore, the covariance is given by

$$\mathbf{\Sigma}_{\mathbf{x}_0^{\mathcal{F}}} = \left[\tilde{\mathbf{D}}_1\right]^2 + \sum_{i=1}^{T}\left[\left(\prod_{j=1}^{i-1}\tilde{\mathbf{G}}(j)\right)\sigma_i\mathbf{I}\right]^2.$$

Therefore, the resulting distribution is

$$\boxed{\mathbf{x}_0^{\mathcal{F}} \sim \mathcal{N}\left(\tilde{\mathbf{D}}_2\mathbf{y}^{\mathcal{F}} + \tilde{\mathbf{D}}_3\boldsymbol{\mu}_0^{\mathcal{F}}\ ,\ \tilde{\mathbf{D}}_1^2 + \sum_{i=1}^{T}\left[\left(\prod_{j=1}^{i-1}\tilde{\mathbf{G}}(j)\right)\sigma_i\mathbf{I}\right]^2\right)} \tag{71}$$

This derivation illustrates that the distribution-level probabilistic perspective adopted by our framework naturally accommodates stochastic sampling processes such as DDPM, as the analysis is performed over the output distribution rather than individual sampling trajectories.

F Averaged Wasserstein distance

To bridge the analytical spectral recommendations with practical implementation, we formulate the optimization in terms of an average loss over the distribution of observations $\mathbf{y}$, modeled as $\mathbf{y} \sim \mathcal{N}(\boldsymbol{\mu}_y, \mathbf{\Sigma}_y)$ with $\mathbf{y} = \mathbf{H}\mathbf{x} + \boldsymbol{n}$. We propose the Average Wasserstein distance between the original and estimated posterior distributions as the loss function.

This formulation naturally accommodates two regimes. For a single observation ($K = 1$), the optimization reduces to the standard Wasserstein-2 loss associated with that realization. Alternatively, when multiple realizations are considered, the weight parameters are optimized to capture the expected behavior of the posterior $p(\mathbf{x} \mid \mathbf{y})$ across $\mathbf{y}$, eliminating the need to solve the optimization individually for each sample. In this average setting, the optimal solution can be expressed directly in terms of the statistical properties of $p(\mathbf{y})$ rather than for a specific observation. In what follows, we derive the resulting average loss formulation for both the single-sample case ($K=1$) and the regime of large K.

The Wasserstein-2 distance between two Gaussian distributions with means μ_1 and μ_2, and covariance matrices Σ_1 and Σ_2, and the corresponding eigenvalues $\{\lambda_1^{(i)}\}_{i=1}^d$ and $\{\lambda_2^{(i)}\}_{i=1}^d$ is given by:

$$W_2(\mathcal{N}_1, \mathcal{N}_2) = \sqrt{(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)^T(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2) + \sum_i\left(\sqrt{\lambda_1^{(i)}} - \sqrt{\lambda_2^{(i)}}\right)^2} \tag{72}$$

In addition, We define the average Wasserstein distance as:

$$\left[\mathcal{L}_{W_2}^2\right]_{\text{Avg}} = \frac{1}{K}\sum_{k=1}^{K} W_2^2\left(p(\mathbf{x}_0^{\mathcal{F}}|\mathbf{y}), p(\hat{\mathbf{x}}_0^{\mathcal{F}} \mid \mathbf{y}^{\mathcal{F}})\right) \tag{73}$$

Using the true posterior distribution derived in Section D and, as a representative example, the DPS posterior from Equation (60), we have:

$$\hat{\mathbf{x}}_0^{\mathcal{F}} \mid \mathbf{y}^{\mathcal{F}} \sim \mathcal{N}\big(\mathbf{D}_2\mathbf{y}^{\mathcal{F}} + \mathbf{D}_3\boldsymbol{\mu}_0^{\mathcal{F}}\ ,\ \mathbf{D}_1^2\big)\ , \quad \hat{\mathbf{x}}_0^{\mathcal{F}} \in \mathbb{R}^d \quad \text{and} \quad \mathbf{x}_0^{\mathcal{F}}|\mathbf{y} \sim \mathcal{N}(\boldsymbol{\mu}_{\mathbf{x}_0|\mathbf{y}}^{\mathcal{F}}, \mathbf{\Lambda}_{x_0|y}), \quad \mathbf{x}_0^{\mathcal{F}} \in \mathbb{R}^d$$

Substituting these expressions into the Wasserstein formula yields:

$$\left[\mathcal{L}_{W_2}^2\right]_{\text{Avg}} = \frac{1}{K}\sum_{k=1}^{K}\left(\sum_i\left(\sqrt{\lambda_0^{(i)}} - \mathbf{D}_1^{(i)}\right)^2 + \sum_i\left(\boldsymbol{\mu}_{\mathbf{x}^{\mathcal{F}}|\mathbf{y},k}^{(i)} - \boldsymbol{\mu}_{\hat{\mathbf{x}}^{\mathcal{F}}|\mathbf{y},k}^{(i)}\right)^2\right) \tag{74}$$

Where $\left\{\lambda_0^{(i)}\right\}_{i=1}^{d}$ are the eigenvalues of $p(\mathbf{x}_0^{\mathcal{F}}|\mathbf{y})$ located on the diagonal of the matrix $\mathbf{\Lambda}_{x_0|y}$ and $\mathbf{D}_1^{(i)}$ denotes the ith diagonal entry of the matrix $\mathbf{D}_1$.

We now derive the mean discrepancy between the two posteriors. The mean of the true posterior from Equation (43) is:

$$\boldsymbol{\mu}_{\mathbf{x}^{\mathcal{F}}|\mathbf{y}} = \boldsymbol{\mu}_0^{\mathcal{F}} + \mathbf{A}(\mathbf{y}^{\mathcal{F}} - \mathbf{\Lambda}_H \boldsymbol{\mu}_0^{\mathcal{F}}) = \mathbf{A}\mathbf{y}^{\mathcal{F}} + (\mathbf{I} - \mathbf{A}\mathbf{\Lambda}_H)\,\boldsymbol{\mu}_0^{\mathcal{F}} \tag{75}$$

where:

$$\mathbf{A} = \mathbf{\Lambda}_0 \mathbf{\Lambda}_{H^T} (\mathbf{\Lambda}_H \mathbf{\Lambda}_0 \mathbf{\Lambda}_{H^T} + \sigma_n^2 \mathbf{I})^{-1}$$

Similarly, the mean of the DPS posterior from Equation (60) is:

$$\boldsymbol{\mu}_{\hat{\mathbf{x}}^{\mathcal{F}}|\mathbf{y}} = \mathbf{D}_2 \mathbf{y}^{\mathcal{F}} + \mathbf{D}_3 \boldsymbol{\mu}_0^{\mathcal{F}}$$

The mean discrepancy between the two posteriors is therefore:

$$\Delta\boldsymbol{\mu}(\mathbf{y}) = \boldsymbol{\mu}_{\hat{\mathbf{x}}^{\mathcal{F}}|\mathbf{y}} - \boldsymbol{\mu}_{\mathbf{x}^{\mathcal{F}}|\mathbf{y}} = (\mathbf{D}_2 - \mathbf{A})\,\mathbf{y}^{\mathcal{F}} + (\mathbf{D}_3 - \mathbf{I} + \mathbf{A}\mathbf{\Lambda}_H)\,\boldsymbol{\mu}_0^{\mathcal{F}} \tag{76}$$

As for $K = 1$ the average Wasserstein distance trivially reduces to the standard Wasserstein distance, substituting $\Delta\boldsymbol{\mu}(\mathbf{y})$ into Equation (74) for $K = 1$ yields:

$$\left[\mathcal{D}_{W_2}^2\right]_{\text{Avg}} = \sum_{i=1}^{d} \left(\sqrt{\lambda_i} - [\mathbf{D}_1]_i\right)^2 + \sum_{i=1}^{d} \left([\mathbf{D}_2 - \mathbf{A}]_i [\mathbf{y}^{\mathcal{F}}]_i + [\mathbf{D}_3 - \mathbf{I} + \mathbf{A}\mathbf{\Lambda}_h]_i [\boldsymbol{\mu}_0^{\mathcal{F}}]_i\right)^2 \tag{77}$$

We now turn to the regime of large K Introducing the shorthand:

$$\mathbf{M} = (\mathbf{D}_2 - \mathbf{A}) \quad \text{and} \quad \mathbf{b} = (\mathbf{D}_3 - \mathbf{I} + \mathbf{A}\mathbf{\Lambda}_H)\,\boldsymbol{\mu}_0^{\mathcal{F}}$$

the mean discrepancy simplifies to:

$$\Delta\boldsymbol{\mu}(\mathbf{y}) = \mathbf{M}\mathbf{y}^{\mathcal{F}} + \mathbf{b}$$

Focusing on the mean term of the Wasserstein-2 loss:

$$\begin{aligned}
&\frac{1}{K}\sum_{k=1}^{K} \|\boldsymbol{\mu}_{\mathbf{x}^{\mathcal{F}}|\mathbf{y},k} - \boldsymbol{\mu}_{\hat{\mathbf{x}}^{\mathcal{F}}|\mathbf{y},k}\|_2^2 \\
&= \frac{1}{K}\sum_{k=1}^{K} \|\Delta\boldsymbol{\mu}(\mathbf{y}_k)\|_2^2 = \frac{1}{K}\sum_{k=1}^{K} \|\mathbf{M}\mathbf{y}_k^{\mathcal{F}} + \mathbf{b}\|_2^2 \\
&= \frac{1}{K}\sum_{k=1}^{K} (\mathbf{y}_k^{\mathcal{F}})^T \mathbf{M}^T \mathbf{M} \mathbf{y}_k^{\mathcal{F}} + 2\mathbf{b}^T \mathbf{M}\mathbf{y}_k^{\mathcal{F}} + \mathbf{b}^T\mathbf{b} \\
&= \mathrm{Tr}\left(\mathbf{M}^T\mathbf{M}\frac{1}{K}\sum_{k=1}^{K} \mathbf{y}_k^{\mathcal{F}} (\mathbf{y}_k^{\mathcal{F}})^T\right) + 2\mathbf{b}^T\mathbf{M}\boldsymbol{\mu}_y^{\mathcal{F}} + \mathbf{b}^T\mathbf{b} \\
&= \mathrm{Tr}\left(\mathbf{M}^T\mathbf{M}(\mathbf{\Lambda}_y + \boldsymbol{\mu}_y^{\mathcal{F}} (\boldsymbol{\mu}_y^{\mathcal{F}})^T)\right) + 2\mathbf{b}^T\mathbf{M}\boldsymbol{\mu}_y^{\mathcal{F}} + \mathbf{b}^T\mathbf{b} \\
&= \mathrm{Tr}(\mathbf{M}^T\mathbf{M}\mathbf{\Lambda}_y) + \|\mathbf{M}\boldsymbol{\mu}_y^{\mathcal{F}} + \mathbf{b}\|_2^2
\end{aligned} \tag{78}$$

where

$$\boldsymbol{\mu}_y^{\mathcal{F}} = \frac{1}{K}\sum_{k=1}^{K} \mathbf{y}_k^{\mathcal{F}}, \qquad \mathbf{\Lambda}_y = \frac{1}{K}\sum_{k=1}^{K} \left(\mathbf{y}_k^{\mathcal{F}} - \boldsymbol{\mu}_y^{\mathcal{F}}\right)\left(\mathbf{y}_k^{\mathcal{F}} - \boldsymbol{\mu}_y^{\mathcal{F}}\right)^T$$

denote the empirical mean and covariance of $\mathbf{y}$ in the frequency domain. Substituting 78 into the average Wasserstein distance 74 yields:

$$\begin{aligned}\left[\mathcal{L}^2_{W_2}\right]_{\text{Avg}} &= \frac{1}{K}\sum_{k=1}^{K}\left(\sum_i\left(\sqrt{\lambda_0^{(i)}}-\mathbf{D}_1^{(i)}\right)^2+\sum_i\left(\boldsymbol{\mu}^{(i)}_{\mathbf{x}^{\mathcal{F}}|\mathbf{y},k}-\boldsymbol{\mu}^{(i)}_{\hat{\mathbf{x}}^{\mathcal{F}}|\mathbf{y},k}\right)^2\right) \\ &= \sum_i\left(\sqrt{\lambda_0^{(i)}}-\mathbf{D}_1^{(i)}\right)^2+\operatorname{Tr}(\mathbf{M}^T\mathbf{M}\boldsymbol{\Lambda}_y)+\|\mathbf{M}\boldsymbol{\mu}_y^{\mathcal{F}}+\mathbf{b}\|_2^2\end{aligned} \tag{79}$$

Finally, using the true distribution of $\mathbf{y}$ in terms of the prior from Equation (40):

$$\boldsymbol{\mu}_y^{\mathcal{F}}=\boldsymbol{\Lambda}_h\,\boldsymbol{\mu}_x^{\mathcal{F}},\qquad \boldsymbol{\Lambda}_y=\boldsymbol{\Lambda}_h\,\boldsymbol{\Lambda}_0\,\boldsymbol{\Lambda}_{h^*}+\sigma_n^2\,\mathbf{I}$$

substituting into Equation (79) gives the final closed-form expression for the averaged Wasserstein-2 distance:

$$\boxed{\left[\mathcal{L}^2_{W_2}\right]_{\text{Avg}}=\sum_i\left(\sqrt{\lambda_0^{(i)}}-\mathbf{D}_1^{(i)}\right)^2+\operatorname{Tr}\left(\mathbf{M}^T\mathbf{M}\left(\boldsymbol{\Lambda}_h\,\boldsymbol{\Lambda}_0\,\boldsymbol{\Lambda}_{h^*}+\sigma_n^2\,\mathbf{I}\right)\right)+\|\mathbf{M}\boldsymbol{\Lambda}_h\,\boldsymbol{\mu}_x^{\mathcal{F}}+\mathbf{b}\|_2^2} \tag{80}$$

G Kullback-Leibler divergence

The Kullback-Leibler (KL) divergence between two Gaussian distributions with means $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$, covariance matrices $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_2$, and corresponding eigenvalues $\{\lambda_1^{(i)}\}_{i=1}^d$ and $\{\lambda_2^{(i)}\}_{i=1}^d$, is given by:

$$D_{\text{KL}}\left(\mathcal{N}(\boldsymbol{\mu}_1,\boldsymbol{\Sigma}_1)\parallel\mathcal{N}(\boldsymbol{\mu}_2,\boldsymbol{\Sigma}_2)\right)=\frac{1}{2}\left(\log\frac{|\boldsymbol{\Sigma}_2|}{|\boldsymbol{\Sigma}_1|}-d+\operatorname{tr}\left(\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\Sigma}_1\right)+(\boldsymbol{\mu}_2-\boldsymbol{\mu}_1)^T\boldsymbol{\Sigma}_2^{-1}(\boldsymbol{\mu}_2-\boldsymbol{\mu}_1)\right)$$

In our setting,

$$\hat{\mathbf{x}}_0^{\mathcal{F}}\,\big|\,\mathbf{y}^{\mathcal{F}}\sim\mathcal{N}\big(\mathbf{D}_2\mathbf{y}^{\mathcal{F}}+\mathbf{D}_3\boldsymbol{\mu}_0^{\mathcal{F}}\,,\,\mathbf{D}_1^2\big)\,,\quad \hat{\mathbf{x}}_0^{\mathcal{F}}\in\mathbb{R}^d\quad\text{and}\quad \mathbf{x}_0^{\mathcal{F}}|\mathbf{y}\sim\mathcal{N}(\boldsymbol{\mu}^{\mathcal{F}}_{\mathbf{x}_0|\mathbf{y}},\boldsymbol{\Lambda}_{x_0|y}),\quad \mathbf{x}_0^{\mathcal{F}}\in\mathbb{R}^d$$

Accordingly, the KL divergence between the true posterior distribution and the estimated posterior distribution is

$$D_{\text{KL}}\left(\mathbf{x}_0^{\mathcal{F}}|\mathbf{y}\parallel\hat{\mathbf{x}}_0^{\mathcal{F}}\,\big|\,\mathbf{y}^{\mathcal{F}}\right)=D_{\text{KL}}\left(\mathcal{N}(\boldsymbol{\mu}^{\mathcal{F}}_{\mathbf{x}_0|\mathbf{y}},\boldsymbol{\Lambda}_{x_0|y}),\,\mathcal{N}\big(\mathbf{D}_2\mathbf{y}^{\mathcal{F}}+\mathbf{D}_3\boldsymbol{\mu}_0^{\mathcal{F}}\,,\,\mathbf{D}_1^2\big)\right)$$

Denote by $\left\{\lambda_0^{(i)}\right\}_{i=1}^d$ the eigenvalues of $p(\mathbf{x}_0^{\mathcal{F}}|\mathbf{y})$, corresponding to the diagonal entries of $\boldsymbol{\Lambda}_{x_0|y}$.

We now derive each term in the KL divergence separately.

- The determinant of the estimated posterior covariance is

$$|\boldsymbol{\Sigma}_2|=|\mathbf{D}_1^T\mathbf{D}_1|=|\mathbf{D}_1^2|=\prod_{i=1}^d{\mathbf{D}_1^{(i)}}^2$$

- The determinant of the true posterior covariance is

$$|\boldsymbol{\Sigma}_1|=\prod_{i=1}^d\lambda_0^{(i)}$$

- The trace term becomes:

$$\operatorname{tr}\left(\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\Sigma}_1\right) = \sum_{i=1}^{d} \frac{\lambda_0^{(i)}}{{\mathbf{D}_1^{(i)}}^2}$$

- The quadratic mean term is given by:

$$(\boldsymbol{\mu}_2-\boldsymbol{\mu}_1)^T\boldsymbol{\Sigma}_2^{-1}(\boldsymbol{\mu}_2-\boldsymbol{\mu}_{\mathbf{x}_0}) = (\mathbf{D}_2\mathbf{y}^{\mathcal{F}}+\mathbf{D}_3\boldsymbol{\mu}_0^{\mathcal{F}}-\boldsymbol{\mu}_{\mathbf{x}_0|\mathbf{y}}^{\mathcal{F}})^T(\mathbf{D}_1^2)^{-1}(\mathbf{D}_2\mathbf{y}^{\mathcal{F}}+\mathbf{D}_3\boldsymbol{\mu}_0^{\mathcal{F}}-\boldsymbol{\mu}_{\mathbf{x}_0|\mathbf{y}}^{\mathcal{F}})$$

By substituting $\boldsymbol{\mu}_{\mathbf{x}^{\mathcal{F}}|\mathbf{y}} = \mathbf{A}\mathbf{y}^{\mathcal{F}} + (\mathbf{I} - \mathbf{A}\boldsymbol{\Lambda}_H)\,\boldsymbol{\mu}_0^{\mathcal{F}}$ from Equation (75) and assuming $\boldsymbol{\mu}_0 = 0$, we obtain:

$$\begin{aligned}(\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1)^T\boldsymbol{\Sigma}_2^{-1}(\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1) &= \left(\mathbf{D}_2\mathbf{y}^{\mathcal{F}} - \mathbf{A}\mathbf{y}^{\mathcal{F}}\right)^T(\mathbf{D}_1^2)^{-1}\left(\mathbf{D}_2\mathbf{y}^{\mathcal{F}} - \mathbf{A}\mathbf{y}^{\mathcal{F}}\right)\\ &= (\mathbf{y}^{\mathcal{F}})^T(\mathbf{D}_2^T - \mathbf{A}^T)(\mathbf{D}_1^2)^{-1}(\mathbf{D}_2 - \mathbf{A})(\mathbf{y}^{\mathcal{F}}).\end{aligned}$$

Rearranging terms gives:

$$(\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1)^T\boldsymbol{\Sigma}_2^{-1}(\boldsymbol{\mu}_2 - \boldsymbol{\mu}_{\mathbf{x}_0}) = \sum_{i=1}^{d} \frac{\left(\mathbf{D}_2^{(i)} - \mathbf{A}^{(i)}\right)^2}{{\mathbf{D}_1^{(i)}}^2}{(\mathbf{y}_i^{\mathcal{F}})}^2.$$

Substituting these terms into the KL divergence gives:

$$D_{\mathrm{KL}}\left(\mathbf{x}_0^{\mathcal{F}}|\mathbf{y} \parallel \hat{\mathbf{x}}_0^{\mathcal{F}} \,\middle|\, \mathbf{y}^{\mathcal{F}}\right) = D_{\mathrm{KL}}\left(\mathcal{N}(\boldsymbol{\mu}_{\mathbf{x}_0|\mathbf{y}}^{\mathcal{F}}, \boldsymbol{\Lambda}_{x_0|y}), \mathcal{N}\big(\mathbf{D}_2\mathbf{y}^{\mathcal{F}} + \mathbf{D}_3\boldsymbol{\mu}_0^{\mathcal{F}}\,,\, \mathbf{D}_1^2\big)\right)$$

$$= \frac{1}{2}\left[\sum_{i=1}^{d}\log {\mathbf{D}_1^{(i)}}^2 - \sum_{i=1}^{d}\log\lambda_0^{(i)} \; - d + \sum_{i=1}^{d}\frac{\lambda_0^{(i)}}{{\mathbf{D}_1^{(i)}}^2} + \sum_{i=1}^{d}\frac{\left(\mathbf{D}_2^{(i)} - \mathbf{A}^{(i)}\right)^2}{{\mathbf{D}_1^{(i)}}^2}{(\mathbf{y}_i^{\mathcal{F}})}^2\right]$$

$$\begin{aligned}&D_{\mathrm{KL}}\left(\mathbf{x}_0^{\mathcal{F}}|\mathbf{y} \parallel \hat{\mathbf{x}}_0^{\mathcal{F}} \,\middle|\, \mathbf{y}^{\mathcal{F}}\right) =\\ &\frac{1}{2}\left[\sum_{i=1}^{d} 2\log\mathbf{D}_1^{(i)} - \sum_{i=1}^{d}\log\lambda_0^{(i)} \; - d + \sum_{i=1}^{d}\frac{\lambda_0^{(i)} + \left(\mathbf{D}_2^{(i)} - \mathbf{A}^{(i)}\right)^2{(\mathbf{y}_i^{\mathcal{F}})}^2}{{\mathbf{D}_1^{(i)}}^2}\right]\end{aligned} \tag{81}$$

H ΠGDM algorithm

The ΠGDM method shares the same Bayesian principle as DPS, but differs in its approximation of the likelihood term. Specifically, $p(\mathbf{y}|\mathbf{x}_t)$ is modeled as a Gaussian distribution with covariance that varies across diffusion time steps t:

$$p_t(\mathbf{y}|\mathbf{x}_t) \approx \mathcal{N}(\mathbf{H}\hat{\mathbf{x}}_0, r_t^2\mathbf{H}\mathbf{H}^T + \sigma_{\mathbf{y}}^2\mathbf{I})$$

Thus, the gradient of the log-probability is given by:

$$\nabla_{\mathbf{x}_t}\log p_t(\mathbf{y}|\mathbf{x}_t) \approx \nabla_{\mathbf{x}_t}(\hat{\mathbf{x}}_0)^T\mathbf{H}^T\left((r_t^2\mathbf{H}\mathbf{H}^T + \sigma_{\mathbf{y}}^2\mathbf{I})^{-1}\right)^T(\mathbf{y} - \mathbf{H}\hat{\mathbf{x}}_0)$$

H.1 The reverse process in the time domain

Based on the DDIM [34] inference process and the Variance-Preserving (VP) framework, the inference procedure using ΠGDM is formulated as:

$$\begin{aligned}\mathbf{x}_{s-1} = &\sqrt{\bar{\alpha}_{s-1}}\hat{\mathbf{x}}_0 + \sqrt{1 - \bar{\alpha}_{s-1} - \sigma_s^2(\eta)}\,\boldsymbol{\epsilon}_\theta(\mathbf{x}_s, s)\\ &+ \sqrt{\bar{\alpha}_s}\nabla_{\mathbf{x}_s}(\hat{\mathbf{x}}_0)^T\mathbf{H}^T((r_s^2\mathbf{H}\mathbf{H}^T + \sigma_{\mathbf{y}}^2\mathbf{I})^{-1})^T(\mathbf{y} - \mathbf{H}\hat{\mathbf{x}}_0) + \sigma_s(\eta)\mathbf{z}_s.\end{aligned} \tag{82}$$

where:

$$\sigma_s(\eta) = \eta\sqrt{\frac{1-\bar{\alpha}_{s-1}}{1-\bar{\alpha}_s}}\sqrt{1-\frac{\bar{\alpha}_s}{\bar{\alpha}_{s-1}}}.$$

Since the VP formulation is already used for $\hat{\mathbf{x}}_0$, we omit the conversion factor $\sqrt{\bar{\alpha}_t}$ in the gradient term, leading to:

$$\begin{aligned}\mathbf{x}_{s-1} = &\sqrt{\bar{\alpha}_{s-1}}\hat{\mathbf{x}}_0 + \sqrt{1-\bar{\alpha}_{s-1}-\sigma_s^2(\eta)}\,\boldsymbol{\epsilon}_\theta(\mathbf{x}_s, s) \\ &+ \nabla_{\mathbf{x}_s}(\hat{\mathbf{x}}_0)^T\mathbf{H}^T((r_s^2\mathbf{H}\mathbf{H}^T + \sigma_{\mathbf{y}}^2\mathbf{I})^{-1})^T(\mathbf{y}-\mathbf{H}\hat{\mathbf{x}}_0) + \sigma_s(\eta)\mathbf{z}_s.\end{aligned} \tag{83}$$

By using the deterministic process ($\eta = 0$):

$$\begin{aligned}\mathbf{x}_{s-1} = &\sqrt{\bar{\alpha}_{s-1}}\hat{\mathbf{x}}_0 + \sqrt{1-\bar{\alpha}_{s-1}}\,\boldsymbol{\epsilon}_\theta(\mathbf{x}_s, s) \\ &+ \nabla_{\mathbf{x}_s}(\hat{\mathbf{x}}_0)^T\mathbf{H}^T((r_s^2\mathbf{H}\mathbf{H}^T + \sigma_{\mathbf{y}}^2\mathbf{I})^{-1})^T(\mathbf{y}-\mathbf{H}\hat{\mathbf{x}}_0)\end{aligned} \tag{84}$$

Using the marginal property,

$$\boldsymbol{\epsilon}_\theta(\mathbf{x}_s, s) = \frac{\mathbf{x}_s - \sqrt{\bar{\alpha}_s}\hat{\mathbf{x}}0}{\sqrt{1-\bar{\alpha}_s}}. \tag{85}$$

which, substituted into Equation 84, gives:

$$\begin{aligned}\mathbf{x}_{s-1} &= \sqrt{\bar{\alpha}_{s-1}}\hat{\mathbf{x}}_0 + \sqrt{1-\bar{\alpha}_{s-1}}\left[\frac{\mathbf{x}_s - \sqrt{\bar{\alpha}_s}\hat{\mathbf{x}}_0}{\sqrt{1-\bar{\alpha}_s}}\right] + \nabla_{\mathbf{x}_s}(\hat{\mathbf{x}}_0)^T\mathbf{H}^T((r_s^2\mathbf{H}\mathbf{H}^T + \sigma_{\mathbf{y}}^2\mathbf{I})^{-1})^T(\mathbf{y}-\mathbf{H}\hat{\mathbf{x}}_0) \\ &= \frac{\sqrt{1-\bar{\alpha}_{s-1}}}{\sqrt{1-\bar{\alpha}_s}}\mathbf{x}_s + \left[\sqrt{\bar{\alpha}_{s-1}} - \frac{\sqrt{\bar{\alpha}_s}\sqrt{1-\bar{\alpha}_{s-1}}}{\sqrt{1-\bar{\alpha}_s}}\right]\hat{\mathbf{x}}_0 + \nabla_{\mathbf{x}_s}(\hat{\mathbf{x}}_0)^T\mathbf{H}^T((r_s^2\mathbf{H}\mathbf{H}^T + \sigma_{\mathbf{y}}^2\mathbf{I})^{-1})^T(\mathbf{y}-\mathbf{H}\hat{\mathbf{x}}_0)\end{aligned}$$

Introducing:

$$a_s = \left[\frac{\sqrt{1-\bar{\alpha}_{s-1}}}{\sqrt{1-\bar{\alpha}_s}}\right]\mathbf{I}$$

$$b_s = \left[\sqrt{\bar{\alpha}_{s-1}} - \frac{\sqrt{\bar{\alpha}_s}\sqrt{1-\bar{\alpha}_{s-1}}}{\sqrt{1-\bar{\alpha}_s}}\right]\mathbf{I}$$

we obtain:

$$\boxed{\mathbf{x}_{s-1} = a_s\mathbf{x}_s + b_s\hat{\mathbf{x}}_0 + \nabla_{\mathbf{x}_s}(\hat{\mathbf{x}}_0)^T\mathbf{H}^T((r_s^2\mathbf{H}\mathbf{H}^T + \sigma_{\mathbf{y}}^2\mathbf{I})^{-1})^T(\mathbf{y}-\mathbf{H}\hat{\mathbf{x}}_0)} \tag{86}$$

Where $\hat{\mathbf{x}}_0$ denotes the MMSE denoiser, corresponds to the denoised signal produced by a pretrained network. Assuming a Gaussian prior, the optimal solution $\mathbf{x}_0^*$ is given by [4]:

$$\mathbf{x}_0^* = (\bar{\alpha}_s\boldsymbol{\Sigma}_0 + (1-\bar{\alpha}_s)\mathbf{I})^{-1}\left(\sqrt{\bar{\alpha}_s}\boldsymbol{\Sigma}_0\mathbf{x}_s + (1-\bar{\alpha}_s)\boldsymbol{\mu}_0\right)$$

Therefore we get:

$$\mathbf{x}_{s-1} = a_s\mathbf{x}_s + b_s\mathbf{x}_0^* + \nabla_{\mathbf{x}_s}(\mathbf{x}_0^*(\mathbf{x}_s))^T\mathbf{H}^T((r_s^2\mathbf{H}\mathbf{H}^T + \sigma_{\mathbf{y}}^2\mathbf{I})^{-1})^T(\mathbf{y}-\mathbf{H}\mathbf{x}_0^*)$$

from Equation 50:

$$\nabla_{x_s}\mathbf{x}_0^*(\mathbf{x}_s) = (\bar{\alpha}_s\boldsymbol{\Sigma}_0 + (1-\bar{\alpha}_s)\mathbf{I})^{-1}\sqrt{\bar{\alpha}_s}\boldsymbol{\Sigma}_0$$

Using the property $(\mathbf{A}^{-1})^T = (\mathbf{A}^T)^{-1}$ for any invertible matrix $\mathbf{A}$, with $\mathbf{A} = \alpha_s\boldsymbol{\Sigma}_0 + (1-\alpha_s)\mathbf{I}$, it can be rewritten as:

$$\nabla_{x_s}\mathbf{x}_0^*(\mathbf{x}_s)^T = \sqrt{\bar{\alpha}_s}\boldsymbol{\Sigma}_0\,(\bar{\alpha}_s\boldsymbol{\Sigma}_0 + (1-\bar{\alpha}_s)\mathbf{I})^{-1}$$

Substitute it into Equation 86, we get:

$$\boxed{\mathbf{x}_{s-1} = a_s\mathbf{x}_s + b_s\mathbf{x}_0^* + \sqrt{\bar{\alpha}_s}\boldsymbol{\Sigma}_0\,(\bar{\alpha}_t\boldsymbol{\Sigma}_0 + (1-\bar{\alpha}_s)\mathbf{I})^{-1}\,\mathbf{H}^T(r_s^2\mathbf{H}\mathbf{H}^T + \sigma_{\mathbf{y}}^2\mathbf{I})^{-1}(\mathbf{y}-\mathbf{H}\mathbf{x}_0^*)} \tag{87}$$

H.2 Spectral form of the update rule

Next, we apply the Discrete Fourier Transform (DFT), denoted by $\mathcal{F}\cdot$, to Equation (87). Assuming that $\mathbf{\Sigma}_0$ and $\mathbf{H}$ are shift-invariant, both matrices are diagonalized by the DFT.

$$\mathbf{F}\mathbf{x}_{s-1} = \mathbf{F}[a_s\mathbf{x}_s + b_s\mathbf{x}_0^* + \sqrt{\bar{\alpha}_s}\mathbf{\Sigma}_0\left(\bar{\alpha}_t\mathbf{\Sigma}_0 + (1-\bar{\alpha}_s)\mathbf{I}\right)^{-1}\mathbf{H}^T(r_s^2\mathbf{H}\mathbf{H}^T + \sigma_{\mathbf{y}}^2\mathbf{I})^{-1}(\mathbf{y} - \mathbf{H}\mathbf{x}_0^*)]$$

$$\mathbf{x}_{s-1}^{\mathcal{F}} = a_s\mathbf{x}_s^{\mathcal{F}} + b_s\mathbf{F}\mathbf{x}_0^* + \sqrt{\bar{\alpha}_s}\mathbf{F}[\mathbf{\Sigma}_0\left(\bar{\alpha}_t\mathbf{\Sigma}_0 + (1-\bar{\alpha}_s)\mathbf{I}\right)^{-1}\mathbf{H}^T(r_s^2\mathbf{H}\mathbf{H}^T + \sigma_{\mathbf{y}}^2\mathbf{I})^{-1})(\mathbf{y} - \mathbf{H}\mathbf{x}_0^*)] \tag{88}$$

From Equation 56, we obtain

$$\mathbf{F}\mathbf{x}_0^* = c_s\mathbf{x}_s^{\mathcal{F}} + d_s\boldsymbol{\mu}_0^{\mathcal{F}} \tag{89}$$

Where:

$$c_s = \sqrt{\bar{\alpha}_s}\mathbf{\Lambda}_0\left[\bar{\alpha}_s\mathbf{\Lambda}_0 + (1-\bar{\alpha}_s)\mathbf{I}\right]^{-1}$$
$$d_s = (1-\bar{\alpha}_s)\left[\bar{\alpha}_s\mathbf{\Lambda}_0 + (1-\bar{\alpha}_s)\mathbf{I}\right]^{-1}$$

$$= \sqrt{\bar{\alpha}_s}\mathbf{F}[\mathbf{\Sigma}_0\left(\bar{\alpha}_s\mathbf{\Sigma}_0 + (1-\bar{\alpha}_s)\mathbf{I}\right)^{-1}\mathbf{H}^T(r_s^2\mathbf{H}\mathbf{H}^T + \sigma_{\mathbf{y}}^2\mathbf{I})^{-1})(\mathbf{y} - \mathbf{H}\mathbf{x}_0^*)] \tag{90}$$
$$= \sqrt{\bar{\alpha}_s}[\mathbf{F}\mathbf{\Sigma}_0\mathbf{F}^T\mathbf{F}\left(\bar{\alpha}_s\mathbf{\Sigma}_0 + (1-\bar{\alpha}_s)\mathbf{I}\right)^{-1}\mathbf{F}^T\mathbf{F}\mathbf{H}^T\mathbf{F}^T\mathbf{F}(r_s^2\mathbf{H}\mathbf{H}^T + \sigma_{\mathbf{y}}^2\mathbf{I})^{-1})\mathbf{F}^T\mathbf{F}(\mathbf{y} - \mathbf{H}\mathbf{x}_0^*)]$$
$$= \sqrt{\bar{\alpha}_s}[\mathbf{\Lambda}_0\left[\bar{\alpha}_t\mathbf{\Lambda}_0 + (1-\bar{\alpha}_s)\mathbf{I}\right]^{-1}\mathbf{\Lambda}_{h^*}\mathbf{F}(r_s^2\mathbf{H}\mathbf{H}^T + \sigma_{\mathbf{y}}^2\mathbf{I})^{-1})\mathbf{F}^T(\mathbf{F}\mathbf{y} - \mathbf{F}\mathbf{H}\mathbf{F}^T\mathbf{F}\mathbf{x}_0^*)]$$

Since:

$$\mathbf{F}(r_s^2\mathbf{H}\mathbf{H}^T + \sigma_{\mathbf{y}}^2\mathbf{I})^{-1})\mathbf{F}^T = (r_s^2\mathbf{\Lambda}_{hh^T} + \sigma_{\mathbf{y}}^2\mathbf{I})^{-1}$$

and denoting:

$$e_s = (r_s^2\mathbf{\Lambda}_{hh^T} + \sigma_{\mathbf{y}}^2\mathbf{I})^{-1}$$

$$\sqrt{\bar{\alpha}_s}\mathbf{F}[\mathbf{\Sigma}_0\left(\bar{\alpha}_s\mathbf{\Sigma}_0 + (1-\bar{\alpha}_s)\mathbf{I}\right)^{-1}\mathbf{H}^T(r_s^2\mathbf{H}\mathbf{H}^T + \sigma_{\mathbf{y}}^2\mathbf{I})^{-1})(\mathbf{y} - \mathbf{H}\mathbf{x}_0^*)] = c_s\mathbf{\Lambda}_{h^*}e_s(\mathbf{y}^{\mathcal{F}} - \mathbf{\Lambda}_h\mathbf{F}\mathbf{x}_0^*)] \tag{91}$$

Substituting Equations 91 and 89 in 88 yields:

$$\mathbf{x}_{s-1}^{\mathcal{F}} = a_s\mathbf{x}_s^{\mathcal{F}} + b_s(c_s\mathbf{x}_s^{\mathcal{F}} + d_s\boldsymbol{\mu}_0^{\mathcal{F}}) + c_s\mathbf{\Lambda}_{h^*}e_s(\mathbf{y}^{\mathcal{F}} - \mathbf{\Lambda}_h(c_s\mathbf{x}_s^{\mathcal{F}} + d_s\boldsymbol{\mu}_0^{\mathcal{F}})]$$

$$\mathbf{x}_{s-1}^{\mathcal{F}} = [a_s + b_sc_s - c_s\mathbf{\Lambda}_{h^*}e_s\mathbf{\Lambda}_hc_s]\mathbf{x}_s^{\mathcal{F}} + [b_sd_s - c_s\mathbf{\Lambda}_{h^*}e_s\mathbf{\Lambda}_hd_s]\boldsymbol{\mu}_0^{\mathcal{F}} + [c_s\mathbf{\Lambda}_{h^*}e_s]\mathbf{y}^{\mathcal{F}}$$

We further define:

$$\mathbf{G}_{\Pi\text{GDM}}(s) = [a_s + b_sc_s - c_s\mathbf{\Lambda}_{h^*}e_s\mathbf{\Lambda}_hc_s]$$
$$\mathbf{Q}_{\Pi\text{GDM}}(s) = [c_s\mathbf{\Lambda}_{h^*}e_s]$$
$$\mathbf{M}_{\Pi\text{GDM}}(s) = [b_sd_s - c_s\mathbf{\Lambda}_{h^*}e_s\mathbf{\Lambda}_hd_s]$$

This leads to the following spectral-domain recursion:

$$\mathbf{x}_{s-1}^{\mathcal{F}} = \mathbf{G}_{\Pi\text{GDM}}(s)\mathbf{x}_s^{\mathcal{F}} + \mathbf{Q}_{\Pi\text{GDM}}(s)\mathbf{y}^{\mathcal{F}} + \mathbf{M}_{\Pi\text{GDM}}(s)\boldsymbol{\mu}_0^{\mathcal{F}} \tag{92}$$

By recursively applying Equation (92) and following the same derivation steps and notation as in Appendix E, we arrive at the following closed-form distribution for $\mathbf{x}_{0,\Pi\text{GDM}}^{\mathcal{F}}$.

$$\mathbf{x}_{0,\Pi\text{GDM}}^{\mathcal{F}} \sim \mathcal{N}\big(\mathbf{D}_2\mathbf{y}^{\mathcal{F}} + \mathbf{D}_3\boldsymbol{\mu}_0^{\mathcal{F}}\ ,\ \mathbf{D}_1^2\big) \tag{93}$$

I DiffPIR algorithm

In this section, we present the DiffPIR algorithm [58] from a spectral perspective and show how it can be integrated into our method. DiffPIR formulates posterior sampling through the following optimization problem:

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}} \frac{1}{2\sigma_n^2} ||\mathbf{y} - \mathbf{H}\mathbf{x}||^2 + \lambda \mathcal{P}(\mathbf{x})$$

where the measurements follow the degradation model $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}$, $\mathbf{H}$ is a known linear degradation operator, $\mathbf{n} \sim \mathcal{N}(\mathbf{0}, \sigma_y^2 \mathbf{I})$ denotes additive Gaussian noise, and $\lambda \mathcal{P}(\mathbf{x})$ represents the prior regularization term weighted by the hyperparameter λ.

The main idea behind plug-and-play inverse reconstruction methods is to decouple the data-consistency and prior terms using the Half-Quadratic Splitting (HQS) framework [17]:

$$\mathbf{x}_0^{(t)} = \arg\min_{\mathbf{z}} \frac{1}{2\bar{\sigma}_t^2} \|\mathbf{z} - \mathbf{x}_t\|^2 + \mathcal{P}(\mathbf{z}) \tag{94a}$$

$$\hat{\mathbf{x}}_0^{(t)} = \arg\min_{\mathbf{x}} \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2 + \rho_t \|\mathbf{x} - \mathbf{x}_0^{(t)}\|^2 \tag{94b}$$

Equation(94a) promotes consistency with the prior distribution, while Equation (94b) enforces adherence to the measurements. The weighting parameter ρ_s is defined as $\rho_s \triangleq \lambda \sigma_n^2 / \bar{\sigma}_t^2$, where $\bar{\sigma}_t = \sqrt{\frac{1-\bar{\alpha}_t}{\bar{\alpha}_t}}$ depends on the diffusion noise schedule $\{\bar{\alpha}\}_{t=1}^T$ and determines the effective noise level.

Since the prior term $\mathcal{P}(\mathbf{x})$ is generally unknown, DiffPIR employs the denoising process to approximate samples from the prior distribution through $\mathbf{x}_0^{(t)}$. In addition, the optimization problem in Equation (94b) generally admits the following closed-form solution:

$$\hat{\mathbf{x}}_0^{(t)} = \left(\mathbf{H}^\top \mathbf{H} + \rho_t \mathbf{I}\right)^{-1} \left(\mathbf{H}^\top \mathbf{y} + \rho_t \mathbf{x}_0^{(t)}\right) \tag{95}$$

I.1 The reverse process in the time domain

We now analyze the DiffPIR reverse process under the Gaussian prior assumption $\mathbf{x}_0 \sim \mathcal{N}(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$. We begin from the sampling update introduced in [58]:

$$\mathbf{x}_{t-1} = \sqrt{\bar{\alpha}_{t-1}} \hat{\mathbf{x}}_0^{(t)} + \sqrt{1 - \bar{\alpha}_{t-1}} \left(\sqrt{1-\zeta}\, \boldsymbol{\epsilon}_\theta(\mathbf{x}_t, t) + \sqrt{\zeta}\, \boldsymbol{\epsilon}_t \right)$$

where: $\boldsymbol{\epsilon}_\theta(\mathbf{x}_t, t)$ denotes the effective predicted noise and is given by:

$$\boldsymbol{\epsilon}_\theta(\mathbf{x}_t, t) = \frac{\mathbf{x}_t - \sqrt{\bar{\alpha}_t} \hat{\mathbf{x}}_0^{(t)}}{\sqrt{1 - \bar{\alpha}_t}}. \tag{96}$$

Here, $\boldsymbol{\epsilon}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ and ζ controls the amount of stochastic noise injected at each diffusion step.

To enable comparison with other posterior sampling methods, we focus on the deterministic setting $\zeta = 0$, corresponding to DDIM sampling [34]. Accordingly, we use the indexing $\{\}_{s=0}^{S}$ to denote the accelerated diffusion trajectory.

$$\mathbf{x}_{s-1} = \sqrt{\bar{\alpha}_{s-1}} \hat{\mathbf{x}}_0^{(s)} + \sqrt{1 - \bar{\alpha}_{s-1}}\, \boldsymbol{\epsilon}_\theta(\mathbf{x}_s, s) \tag{97}$$

Substituting the predicted noise expression from Equation (I.1) into the reverse update in (97) yields:

$$\mathbf{x}_{s-1} = \sqrt{\bar{\alpha}_{s-1}}\hat{\mathbf{x}}_0^{(s)} + \sqrt{1-\bar{\alpha}_{s-1}}\left[\frac{\mathbf{x}_s - \sqrt{\bar{\alpha}_s}\hat{\mathbf{x}}_0^{(s)}}{\sqrt{1-\bar{\alpha}_s}}\right]$$

which can be rewritten as:

$$\mathbf{x}_{s-1} = \frac{\sqrt{1-\bar{\alpha}_{s-1}}}{\sqrt{1-\bar{\alpha}_s}}\mathbf{x}_s + \left[\sqrt{\bar{\alpha}_{s-1}} - \frac{\sqrt{\bar{\alpha}_s}\sqrt{1-\bar{\alpha}_{s-1}}}{\sqrt{1-\bar{\alpha}_s}}\right]\hat{\mathbf{x}}_0 \tag{98}$$

We define:

$$a_s = \left[\frac{\sqrt{1-\bar{\alpha}_{s-1}}}{\sqrt{1-\bar{\alpha}_s}}\right]\mathbf{I}$$

$$b_s = \left[\sqrt{\bar{\alpha}_{s-1}} - \frac{\sqrt{\bar{\alpha}_s}\sqrt{1-\bar{\alpha}_{s-1}}}{\sqrt{1-\bar{\alpha}_s}}\right]\mathbf{I}.$$

Substituting these definitions into Equation (98) gives

$$\mathbf{x}_{s-1} = a_s\mathbf{x}_s + b_s\hat{\mathbf{x}}_0^{(s)}.$$

At this stage, DiffPIR incorporates measurement information into the diffusion process through the solution of the optimization problem in Equation (94b). Substituting the closed-form solution for $\hat{\mathbf{x}}_0^{(s)}$ yields:

$$\mathbf{x}_{s-1} = a_s\mathbf{x}_s + b_s\left(\mathbf{H}^\top\mathbf{H} + \rho_t\mathbf{I}\right)^{-1}\left(\mathbf{H}^\top\mathbf{y} + \rho_t\mathbf{x}_0^{(t)}\right)$$

In DiffPIR, the estimate $\mathbf{x}_0^{(t)}$ is obtained using Tweedie's formula adapted to variance-preserving diffusion:

$$\mathbf{x}_0^{(t)} = \frac{1}{\sqrt{\bar{\alpha}_t}}\left(\mathbf{x}_t + (1-\bar{\alpha}_t)s_\theta(\mathbf{x}_t, t)\right)$$

Here, we replace the pretrained denoiser with the prior-optimal denoiser under a Gaussian prior assumption. In this setting, the MMSE estimator coincides with the Wiener filter and is given by [4]:

$$\mathbf{x}_0^{(t)} = \mathbf{x}_0^* = (\bar{\alpha}_s\boldsymbol{\Sigma}_0 + (1-\bar{\alpha}_s)\mathbf{I})^{-1}\left(\sqrt{\bar{\alpha}_s}\boldsymbol{\Sigma}_0\mathbf{x}_s + (1-\bar{\alpha}_s)\boldsymbol{\mu}_0\right)$$

Substituting this expression into the reverse process yields:

$$\boxed{\mathbf{x}_{s-1} = a_s\mathbf{x}_s + b_s\left(\mathbf{H}^\top\mathbf{H} + \rho_t\mathbf{I}\right)^{-1}\left(\mathbf{H}^\top\mathbf{y} + \rho_s\mathbf{x}_0^*\right)} \tag{99}$$

I.2 Spectral form of the update rule

Next, we apply the Discrete Fourier Transform (DFT), denoted by $\mathcal{F}\cdot$, to Equation (99). Assuming that $\boldsymbol{\Sigma}_0$ and $\mathbf{H}$ are shift-invariant, both matrices are diagonalized by the DFT.

$$\mathbf{F}\mathbf{x}_{s-1} = \mathbf{F}[a_s\mathbf{x}_s + b_s\left(\mathbf{H}^\top\mathbf{H} + \rho_t\mathbf{I}\right)^{-1}\left(\mathbf{H}^\top\mathbf{y} + \rho_s\mathbf{x}_0^*\right)]$$

$$\mathbf{x}_{s-1}^{\mathcal{F}} = a_s\mathbf{x}_s^{\mathcal{F}} + b_s\mathbf{F}\left(\mathbf{H}^\top\mathbf{H} + \rho_t\mathbf{I}\right)^{-1}\mathbf{F}^T\mathbf{F}\left(\mathbf{H}^\top\mathbf{y} + \rho_s\mathbf{x}_0^*\right)] \tag{100}$$

Using the diagonalization property:

$$\mathbf{F}\left(\mathbf{H}^\top\mathbf{H} + \rho_t\mathbf{I}\right)^{-1}\mathbf{F}^T = (\boldsymbol{\Lambda}_{h^Th} + \rho_s\mathbf{I})^{-1}$$

and denoting:

$$m_s = (\boldsymbol{\Lambda}_{h^Th} + \rho_s\mathbf{I})^{-1}$$

together with Equation 56, we obtain:

$$\mathbf{F}\mathbf{x}_0^* = c_s \mathbf{x}_s^{\mathcal{F}} + d_s \boldsymbol{\mu}_0^{\mathcal{F}} \tag{101}$$

where

$$c_s = \sqrt{\bar{\alpha}_s}\boldsymbol{\Lambda}_0 \left[\bar{\alpha}_s \boldsymbol{\Lambda}_0 + (1-\bar{\alpha}_s)\mathbf{I}\right]^{-1}$$

and

$$d_s = (1-\bar{\alpha}_s) \left[\bar{\alpha}_s \boldsymbol{\Lambda}_0 + (1-\bar{\alpha}_s)\mathbf{I}\right]^{-1}.$$

Substituting these expressions into Equation Equestion(100) yields:

$$a_s \mathbf{x}_s^{\mathcal{F}} + b_s m_s \boldsymbol{\Lambda}_h \mathbf{y}^{\mathcal{F}} + b_s m_s \rho_s c_s \mathbf{x}_s^{\mathcal{F}} + b_s m_s \rho_s d_s \boldsymbol{\mu}_0^{\mathcal{F}}$$

which can be rearranged as:

$$(a_s + b_s m_s \rho_s c_s)\mathbf{x}_s^{\mathcal{F}} + b_s m_s \boldsymbol{\Lambda}_{h^*} \mathbf{y}^{\mathcal{F}} + b_s m_s \rho_s d_s \boldsymbol{\mu}_0^{\mathcal{F}}$$

We further define

$$\mathbf{G}_{\text{DiffPIR}}(s) = [a_s + b_s m_s \rho_s c_s]$$
$$\mathbf{Q}_{\text{DiffPIR}}(s) = [b_s m_s \boldsymbol{\Lambda}_{h^*}]$$

and

$$\mathbf{M}_{\text{DiffPIR}}(s) = [b_s m_s \rho_s d_s]$$

This leads to the following spectral-domain recursion:

$$\mathbf{x}_{s-1}^{\mathcal{F}} = \mathbf{G}_{\text{DiffPIR}}(s)\mathbf{x}_s^{\mathcal{F}} + \mathbf{Q}_{\text{DiffPIR}}(s)\mathbf{y}^{\mathcal{F}} + \mathbf{M}_{\text{DiffPIR}}(s)\boldsymbol{\mu}_0^{\mathcal{F}} \tag{102}$$

Recursively applying Equation (102), together with the derivation steps used in Appendix E, yields the following closed-form distribution for $\mathbf{x}_{0,\text{DiffPIR}}^{\mathcal{F}}$:

$$\boxed{\mathbf{x}_{0,\text{DiffPIR}}^{\mathcal{F}} \sim \mathcal{N}\big(\mathbf{D}_2\mathbf{y}^{\mathcal{F}} + \mathbf{D}_3\boldsymbol{\mu}_0^{\mathcal{F}}\ ,\ \mathbf{D}_1^2\big)} \tag{103}$$

J Additional results: synthetic Gaussian distribution

Section 6.1 considers a Gaussian prior $\mathbf{x}_0 \sim \mathcal{N}(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$, where $\mathbf{x}_0 \in \mathbb{R}^d$ and $\boldsymbol{\Sigma}_0 \in \mathbb{R}^{d\times d}$ is a shift invariant covariance matrix. The covariance is constructed as $\boldsymbol{\Sigma}_0 = A^\top A$, where A is a shift invariant matrix whose first row is given by $a = [-l, -l + \frac{1}{d-1}, \ldots, l - \frac{1}{d-1}, l]$. The mean vector $\boldsymbol{\mu}_0$ is chosen to be constant, corresponding to a stationary model. Figure 3 illustrates the resulting covariance matrix for $d = 50$ and $l = 0.05$.

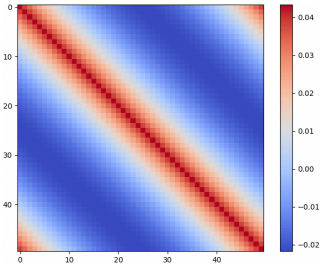


Figure 3: Covariance matrix obtained for $d = 50$ and $l = 0.05$.

J.1 Optimization procedure

Throughout the paper, the optimization problem is solved using the Sequential Least Squares Programming (SLSQP) method [28], as implemented in SciPy.

J.2 DPS algorithm

Using the optimization problem in Equation (16) under the DPS framework, together with the experimental setting described in Section 6.1, we obtain the following spectral recommendation for the hyperparameter vector $\boldsymbol{\zeta}$:

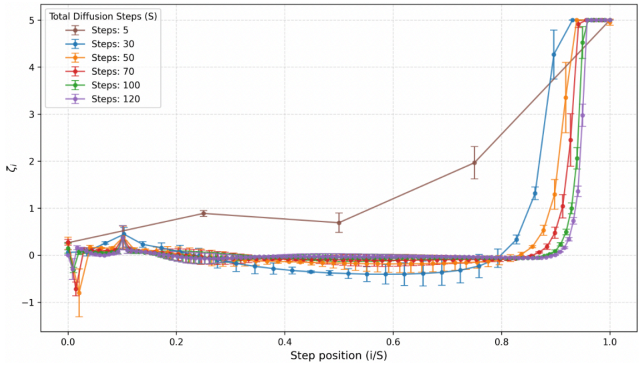


Figure 4: Spectral recommendations for the weighting coefficients $\boldsymbol{\zeta}$ in the DPS framework for different numbers of diffusion steps $S \in [5, 30, 50, 70, 100, 120]$, with variability across realizations.

J.3 ΠGDM algorithm

Here, we derive the spectral recommendations for the ΠGDM algorithm under the setup introduced in Section 6.1. In the original formulation of ΠGDM, the heuristic parameter r_s plays a dual role: it both controls the uncertainty at diffusion step s and serves as a weighting term for the likelihood component in the inference equation (see Eq. 8). These two roles are handled through a single heuristic definition based on the diffusion noise schedule,

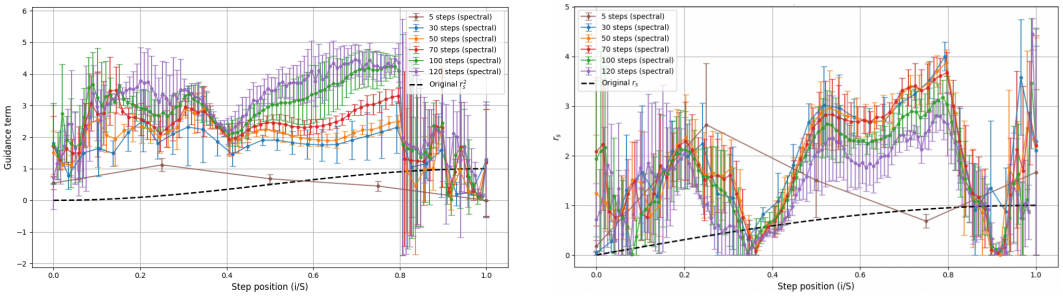
$$r_s = \sqrt{1 - \bar{\alpha}_s}.$$

In our approach, we consider these roles separately, allowing for greater flexibility in balancing their respective influences. Specifically, we introduce a separate guidance parameter g_s while retaining r_s^2 as the uncertainty-related term. The resulting inference update is given by:

$$\begin{aligned}\mathbf{x}_{s-1} = &\sqrt{\bar{\alpha}_{s-1}}\hat{\mathbf{x}}_0 + \sqrt{1-\bar{\alpha}_{s-1}}\,\boldsymbol{\epsilon}_\theta(\mathbf{x}_s, s) \\ &+ g_s \nabla_{\mathbf{x}_s}(\hat{\mathbf{x}}_0)^T \mathbf{H}^T((r_s^2\mathbf{H}\mathbf{H}^T + \sigma_{\mathbf{y}}^2\mathbf{I})^{-1})^T(\mathbf{y} - \mathbf{H}\hat{\mathbf{x}}_0)\end{aligned} \tag{104}$$

Figure 5 presents the resulting spectral recommendations for the weighting coefficients $\{r_s\}_{s=1}^d$ and $\{g_s\}_{s=1}^d$, together with their mean and standard deviation computed over $N = 5$ realizations $\{y_i\}_{i=1}^N$, for different numbers of diffusion steps, as described in Section 6.1. For comparison, we also include the heuristic parameter values used in the original ΠGDM algorithm.

Notably, our recommendations take into account prior information, the observed measurements, and the diffusion dynamics, while the original ΠGDM heuristics define the weighting coefficients solely based on the diffusion noise schedule and decrease over diffusion steps.



(a) Guidance term g_s (b) Uncertainty term r_s

Figure 5: Figures 5a and 5b show the spectral recommendations for the ΠGDM weighting coefficients g_s and r_s, respectively. Results are shown for different diffusion step counts, with the standard deviation across realizations. The original ΠGDM heuristics indicated in black curves

Figure 6 compares the *Wasserstein-2* distance between the reconstructed posterior distributions and the true posterior, for both the heuristic parameter choices in ΠGDM and the proposed spectral recommendations, evaluated across multiple diffusion steps and measurement realizations y_i. We also report the distance obtained from the reconstructed posterior produced by the optimal posterior denoiser derived in (19). The spectral recommendations yield consistently lower Wasserstein distances across diffusion steps and show reduced variability across realizations compared to the ΠGDM heuristics. In addition the exact solution achieves the lowest Wasserstein distance overall.

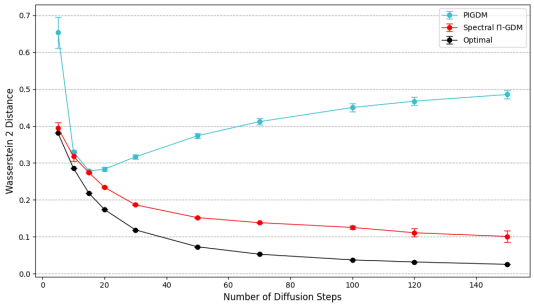


Figure 6: Comparison of the Wasserstein-2 distance for ΠGDM heuristic values, the spectral recommendations applied to ΠGDM (red), and the analytically derived ideal posterior denoiser (black), across different number of diffusion steps $S \in \{5, 10, 15, 20, 30, 50, 70, 100, 120, 150\}$.

J.4 DiffPIR algorithm

Here, we derive the spectral recommendations for the DiffPIR [58] algorithm under the setup introduced in Section 6.1. As discussed in Section I, DiffPIR introduces a single hyperparameter

λ, which controls the balance between the prior and likelihood terms. The remaining weighting parameters are then determined through λ together with the predefined noise level $\bar{\sigma}_t = \sqrt{\frac{1-\bar{\alpha}_t}{\bar{\alpha}_t}}$. In practice, the choice of λ remains heuristic, with no principled rule currently existing for selecting it.

In our approach, we extend the scalar hyperparameter λ to a diffusion-step-dependent vector, $\{\lambda_s\}_{s=1}^{S}$, enabling adaptive weighting across diffusion steps according to the predefined noise schedule.

Figure 7 presents the resulting spectral recommendations for the weighting coefficients $\{\lambda_s\}_{s=1}^{S}$, together with their mean and standard deviation computed over $N = 5$ realizations $\{y_i\}_{i=1}^{N}$, for different numbers of diffusion steps under the setup described in Section 6.1. We note that the original DiffPIR heuristic is not included, since unlike DPS and ΠGDM, DiffPIR relies on a manually predefined scalar parameter that does not explicitly depend on the diffusion schedule or the measurement consistency term. Moreover, the experimental setting considered here was not studied in the original DiffPIR paper, and therefore no corresponding heuristic recommendation is available. Notably, the proposed recommendations jointly account for the prior characteristics, the observed measurements, and the diffusion dynamics, unlike the original DiffPIR heuristics, which are manually predefined.

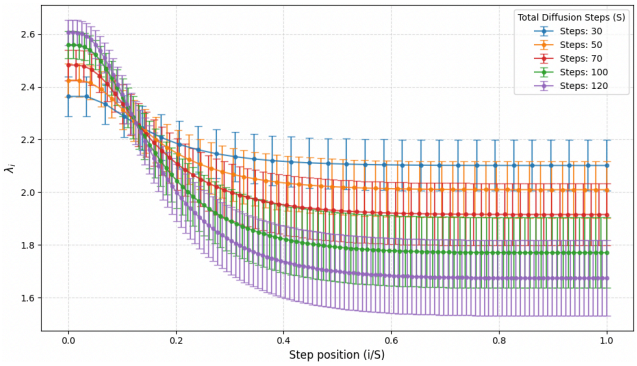


Figure 7: Spectral recommendations for the DiffPIR weighting coefficients $\{\lambda_s\}_{s=1}^{S}$ across different numbers of diffusion steps. The curves include the corresponding standard deviation across realizations.

Figure 8 compares the *Wasserstein-2* distance between the reconstructed posterior distributions and the true posterior for the DiffPIR heuristic parameters with $\lambda = \{0.1, 1, 10, 100\}$ and the proposed spectral recommendations, evaluated across multiple diffusion steps and N measurement realizations $\{y_i\}_{i=1}^{N}$. We also report the distance obtained using the posterior-optimal denoiser derived in (19). The proposed spectral recommendations consistently yield lower Wasserstein distances across diffusion steps and exhibit reduced variability across realizations compared to the original DiffPIR heuristics.

For a small number of diffusion steps, the spectral recommendations slightly outperform the posterior-optimal denoiser in terms of the Wasserstein distance, whereas for larger numbers of diffusion steps the opposite behavior is observed, with a growing gap in favor of the posterior-optimal denoiser. This behavior does not contradict the optimality of the posterior denoiser, since it is derived to minimize the MMSE at each diffusion step rather than directly minimizing the Wasserstein distance between the reconstructed and true posterior distributions.

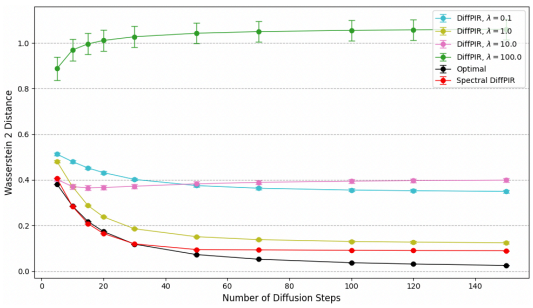


Figure 8: Comparison of the Wasserstein-2 distance for the spectral recommendations applied to DiffPIR (red) and the analytically derived ideal posterior denoiser (black), across different numbers of diffusion steps $S \in \{5, 10, 15, 20, 30, 50, 70, 100, 120, 150\}$.

K Additional results: empirical distribution

K.1 FFHQ dataset

In the following section, we provide a comprehensive comparison of the proposed spectral recommendations for DPS [9] with the heuristic configurations used in DPS and DiffPIR [58], under the same experimental setting described in Section 6.2. For DPS, we evaluate a wide range of heuristic values, $\zeta^{'} = \{0.01, 0.05, 0.1, 0.3, 0.5, 0.7\}$. For DiffPIR, we use the hyperparameter configuration reported in the original paper, ($\lambda = 7$), after verifying it against the neighboring values $\lambda \in \{6, 7, 8\}$. Here, *DiffPIR (D)* denotes the deterministic variant and *DiffPIR (S)* denotes the stochastic variant. The results are presented in terms of FID, LPIPS, SSIM and PSNR.

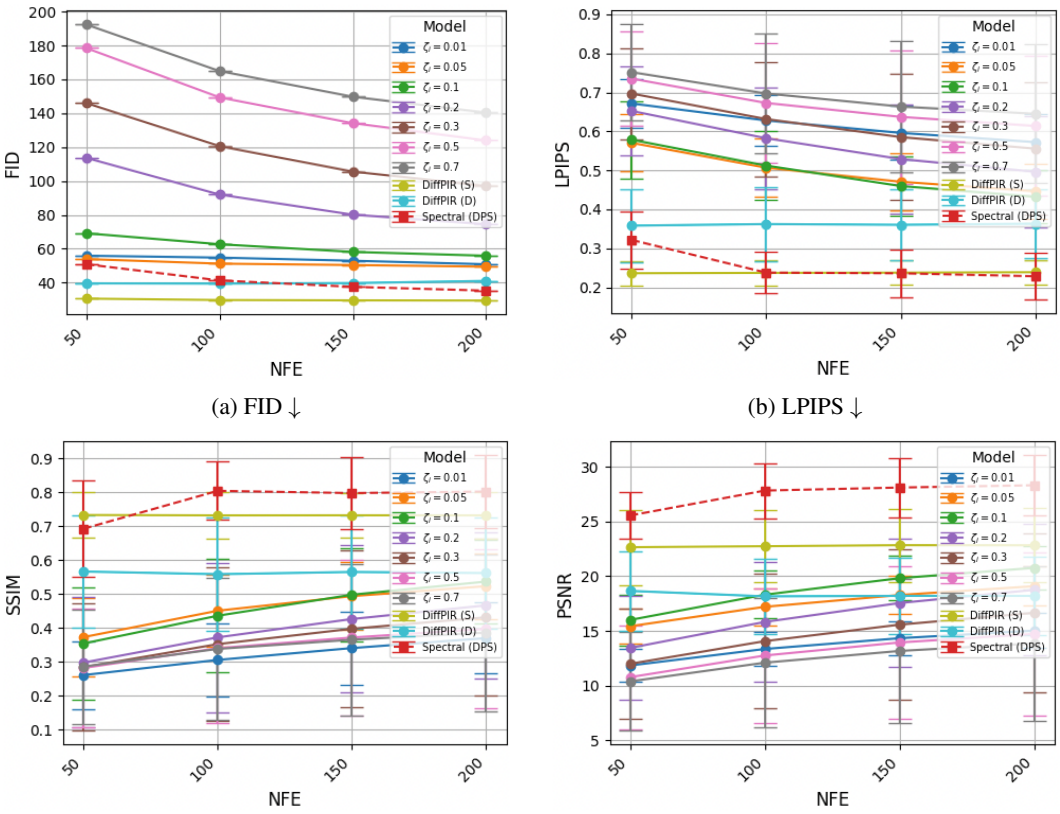
Figure 9 shows that the proposed spectral recommendations consistently outperform all DPS heuristic configurations in terms of reconstruction fidelity, achieving improved distortion metrics and stronger adherence to the underlying signal. In terms of perceptual quality, the spectral recommendations generally achieve lower FID values, although for small numbers of diffusion steps, heuristic configurations with lower guidance weights ζ' attain nearly comparable FID scores. This behavior is consistent with the observation that smaller likelihood guidance weights often favor perceptual quality by placing greater emphasis on the prior distribution, whereas increasing the guidance strength initially improves PSNR before deteriorating for larger values. As discussed in Section 6.2, perceptual quality alone is insufficient in inverse problems, where consistency with the measurements and accurate recovery of the underlying signal remain necessary. Overall, among the evaluated DPS heuristics, $\zeta' = 0.1$ provides the most balanced tradeoff between distortion and perceptual quality.

Figure 9 further compares our approach with DiffPIR heuristics. The proposed spectral recommendations generally achieve higher PSNR values than both the deterministic and stochastic DiffPIR variants. In terms of perceptual metrics such as FID, however, the stochastic DiffPIR process achieves better performance, reflecting the effect of stochastic sampling on perceptual quality. This comparison is intended to position the proposed framework relative to existing posterior sampling approaches. We emphasize that since the spectral recommendations are optimized separately for each posterior sampling method, improved performance for one algorithm does not necessarily imply superior performance relative to alternative methods across all metrics and experimental settings.

K.2 ImageNet dataset

Figure 10 shows the spectral recommendations obtained for the ImageNet 256×256 dataset. Their overall behavior is similar to that observed on FFHQ, with guidance values increasing as the diffusion process progresses; however, the absolute magnitudes differ across the two datasets. By comparison, the DPS heuristic weights remain relatively smaller and exhibit a similar trend across both datasets.

Table 3 presents a comprehensive comparison between the proposed spectral recommendations and the original DPS heuristics on FFHQ-256 and ImageNet-256. We consider several linear degradation settings with parameter pairs $(\mathcal{V}, \sigma_y) \in \{(0.01, 0.1), (0.1, 0.1), (0.1, 0.2)\}$, and evaluate diffusion step counts $S \in \{50, 100, 200, 300, 400\}$, covering regimes that also allow relatively efficient synthesis.



(a) FID ↓ (b) LPIPS ↓

(c) SSIM ↑ (d) PSNR ↑

Figure 9: **LPF degradation.** Comparison between DPS and DiffPIR heuristic configurations and the proposed DPS spectral recommendations under LPF degradation with $\mathcal{V} = 0.1$ and additive Gaussian noise $\sigma_y = 0.1$, evaluated across the number of function evaluations (NFE) on the FFHQ dataset.

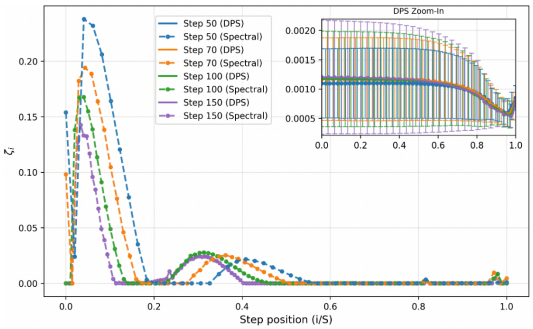


Figure 10: Comparison of spectral recommendations on the ImageNet 256 dataset, with a magnified y-axis view of the DPS heuristics. Results are shown for diffusion steps $S \in \{30, 70, 100, 150\}$.

Selecting the DPS heuristic parameter ζ' is non-trivial, as the recommended values in [9] are tied to particular degradation settings and typically larger numbers of diffusion steps. Following the same methodology, we evaluate a broad range of heuristic values, $\zeta' = \{0.01, 0.05, 0.1, 0.2, 0.3, 0.5, 0.7\}$ across different step counts and selected $\zeta' = 0.1$ for FFHQ (see Section K.1) and $\zeta' = 0.2$ for ImageNet. Both values fall within the range considered in [9]. While smaller values resulted in insufficient measurement fidelity, larger values led to unstable behavior in the final reconstructions. This sensitivity underscores the advantage of the proposed spectral recommendations, which eliminate the need for repeated heuristic tuning.

Table 3 reports results across four evaluation metrics, SSIM, PSNR, LPIPS, and FID, reflecting measurement fidelity and perceptual quality. All experiments were conducted using the implementa-

tion and pretrained models from [9]. Overall, the spectral recommendations tend to achieve a more balanced tradeoff between these two aspects compared to the original DPS heuristics. While some reported values differ from those in [9], this discrepancy is mainly attributable to the use of fewer diffusion steps and different degradation settings.

In some settings with a small number of diffusion steps, the DPS heuristic attains lower FID values; however, this is typically accompanied by reduced performance on measurement fidelity metrics. This behavior may be related to the small guidance step sizes commonly used by DPS, which can yield visually plausible samples that are less strongly constrained by the noisy measurements when using fewer diffusion steps. Representative examples are shown in Figure 11. These observations are also consistent with the findings reported in the original DPS paper [9].

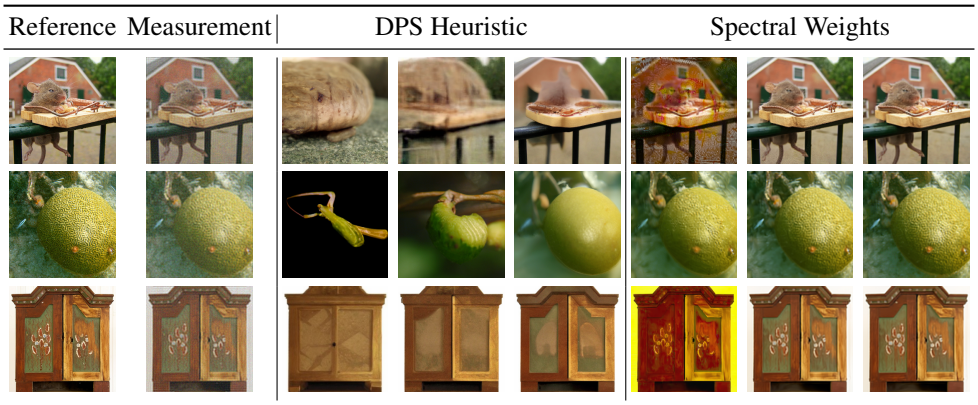


Figure 11: Qualitative comparison of visual results on ImageNet. Each row shows the reference image, the degraded measurement, samples obtained using the DPS heuristic at 100, 200, and 400 diffusion steps (from left to right), and samples obtained using the proposed spectral recommendations at the same step counts.

Figure 11 compares visual results obtained using the DPS heuristic and the proposed spectral recommendations under a degradation with $\mathcal{V} = 0.1$ and $\sigma_y = 0.1$, across different numbers of diffusion steps $S \in \{100, 200, 400\}$ and for the ImageNet dataset.

For a small number of diffusion steps, the DPS heuristic produces visually plausible images; however, these reconstructions often deviate from the reference images in terms of both global structure and fine details. In contrast, the spectral recommendations yield visual results that more closely align with the reference images and the measurements, although they may exhibit minor effects such as localized color variations. This behavior helps explain why the DPS heuristic can attain favorable FID values at low step counts while showing reduced measurements fidelity.

As the number of diffusion steps increases, the DPS heuristic increasingly incorporates information from the measurements, leading to visual results that more closely resemble the reference images, although deviations remain due to the limited heuristic step size. At the same time, the spectral recommendations produce visually natural outputs that remain consistent with the measurements. This convergence in behavior aligns with the quantitative trends reported in Table 11, where the spectral recommendations exhibit a balanced tradeoff between perceptual quality and measurement fidelity.

Figure 12 shows additional visual results on the ImageNet dataset, evaluated at 500 diffusion steps. While the DPS heuristic produces visually natural patterns, some differences with respect to the measurements remain. By comparison, the spectral recommendations yield visually natural samples that are more consistently aligned with the measurements, including at finer levels of detail.

Table 3: Quantitative comparison between the proposed spectral recommendations and the DPS heuristic. Results are reported on FFHQ-256 and ImageNet-256 across multiple degradation settings $(\mathcal{V}, \sigma_y)$ and diffusion step counts, using PSNR, SSIM, LPIPS, and FID.

$\mathcal{V}$	σ_y	METHOD	STEPS	FFHQ 256				IMAGENET 256			
				PSNR↑	SSIM↑	LPIPS↓	FID↓	PSNR↑	SSIM↑	LPIPS↓	FID↓
0.01	0.1	SPECTRAL	50	**11.32**	**0.26**	0.71	197.83	8.10	0.12	0.73	283.29
			100	**13.24**	**0.29**	**0.67**	171.32	11.85	0.24	0.68	184.89
			200	**16.27**	**0.34**	**0.60**	**133.89**	**16.04**	**0.30**	**0.60**	174.88
			300	**17.18**	**0.37**	**0.59**	**132.80**	**14.01**	**0.28**	**0.64**	**137.44**
			400	**18.27**	**0.39**	**0.57**	**121.08**	**15.09**	**0.29**	**0.62**	**144.01**
		DPS	50	4.59	0.24	**0.52**	**169.29**	**12.55**	**0.24**	**0.68**	**132.02**
			100	12.47	0.28	0.68	**155.47**	**12.01**	0.24	**0.67**	**142.94**
			200	13.53	0.30	0.66	142.03	11.83	0.24	0.66	**147.25**
			300	14.56	0.33	0.64	137.69	11.94	0.24	0.66	144.88
			400	14.68	0.33	0.63	137.13	11.94	0.25	0.66	144.08
0.1	0.1	SPECTRAL	50	**25.55**	**0.69**	**0.32**	**50.86**	**15.99**	**0.41**	**0.59**	262.73
			100	**27.82**	**0.80**	**0.23**	**41.31**	10.22	**0.40**	**0.58**	231.42
			200	**28.31**	**0.80**	**0.22**	**35.27**	**15.67**	**0.41**	**0.55**	**120.55**
			300	**28.53**	**0.81**	**0.21**	**31.72**	**14.56**	**0.37**	**0.58**	**139.86**
			400	**28.59**	**0.82**	**0.20**	**29.36**	**13.92**	**0.35**	**0.60**	**54.22**
		DPS	50	16.00	0.35	0.57	69.13	12.73	0.26	0.68	**130.60**
			100	18.31	0.43	0.51	62.68	**12.41**	0.27	0.67	**142.66**
			200	20.76	0.53	0.43	55.88	12.55	0.28	0.66	143.66
			300	22.00	0.57	0.40	54.41	12.92	0.29	0.64	141.28
			400	22.86	0.60	0.38	53.78	13.29	0.31	0.63	131.13
0.1	0.2	SPECTRAL	50	**14.56**	**0.42**	**0.65**	**144.32**	**14.73**	**0.37**	**0.64**	152.59
			100	**15.73**	**0.45**	**0.59**	**120.04**	**14.81**	**0.37**	**0.63**	152.18
			200	**18.72**	**0.50**	**0.50**	**85.81**	**14.83**	**0.38**	**0.62**	145.89
			300	**23.35**	**0.58**	**0.41**	**57.94**	**14.52**	**0.37**	**0.64**	**137.42**
			400	**22.69**	**0.55**	**0.44**	**61.74**	**16.42**	**0.42**	**0.57**	**120.90**
		DPS	50	11.95	0.28	0.69	144.70	12.96	0.26	0.68	**125.55**
			100	13.85	0.34	0.63	120.49	12.58	0.27	0.67	**139.52**
			200	17.08	0.43	0.54	91.55	12.44	0.28	0.66	**141.26**
			300	17.63	0.45	0.53	88.43	12.90	0.29	0.65	137.63
			400	18.64	0.49	0.50	82.39	13.12	0.30	0.64	136.29

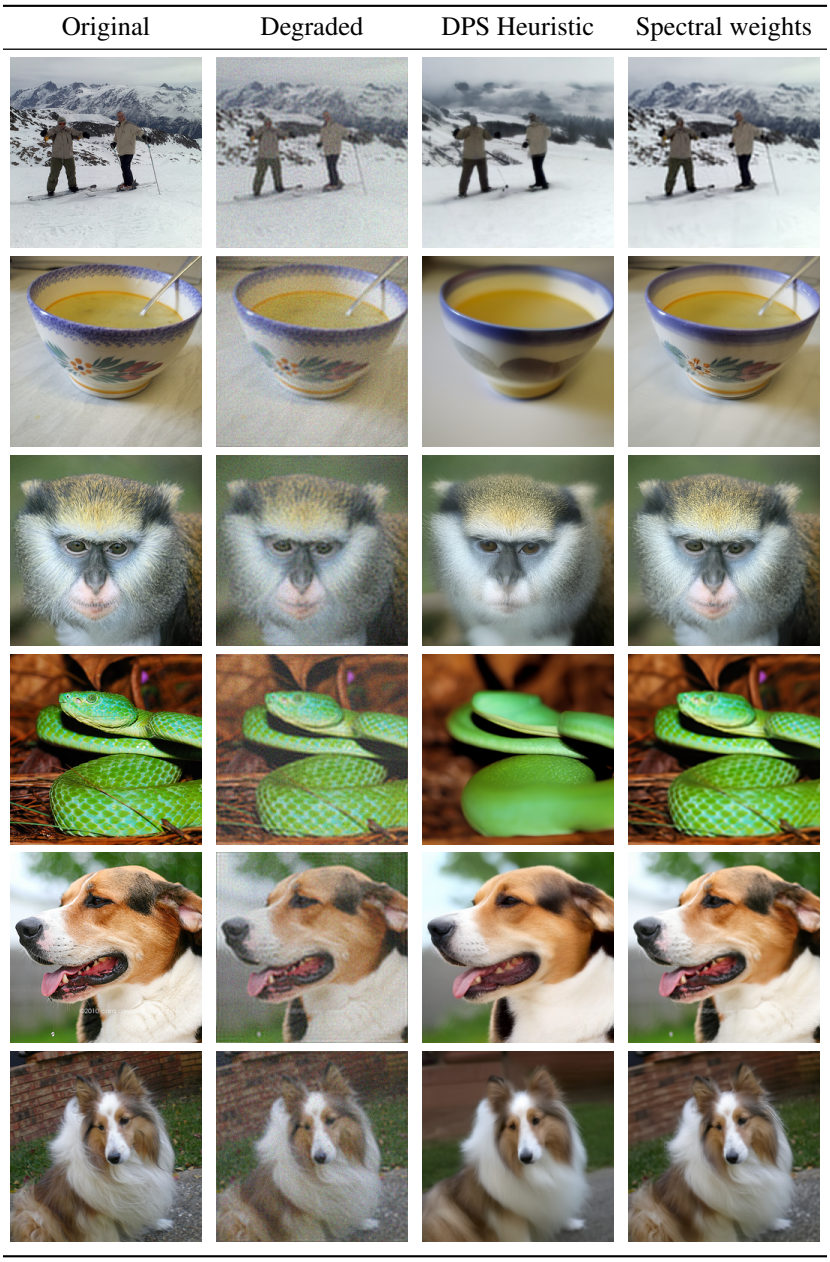


Figure 12: Qualitative comparison of visual results on the ImageNet dataset. Each row shows the reference image, the degraded observation with $\mathcal{V} = 0.1$ and $\sigma_y = 0.1$, and samples obtained using the DPS heuristic and the proposed spectral weighting, both evaluated at 500 diffusion steps.

K.3 General degradation settings

Section K.1 demonstrated that the proposed spectral recommendations extend beyond the Gaussian setting through experiments on FFHQ and ImageNet. Here, we further examine how these recommendations generalize to degradation operators beyond the shift-invariant setting. To this end, we apply the spectral recommendations to two inpainting tasks: random masking with 70% pixel removal and a 128×128 box mask, both corrupted by additive Gaussian noise with standard deviation $\sigma_y = 0.1$. We compare the results with DPS across a range of heuristic guidance weights $\zeta' = \{0.01, 0.05, 0.1, 0.3, 0.5, 0.7, 1\}$, as well as with recent posterior sampling methods, including DAPS [54] and DiffPIR [58]. For DiffPIR, we use the recommended heuristic values, $\lambda = 7$ for box inpainting and $\lambda = 6$ for random inpainting, after verifying their performance against $\lambda \in \{3, 6, 7, 8\}$. For DAPS, we use the heuristic parameters recommended in the original paper and adjust N_{ODE} and N_α to match the number of function evaluations (NFE) used by DPS and DiffPIR.

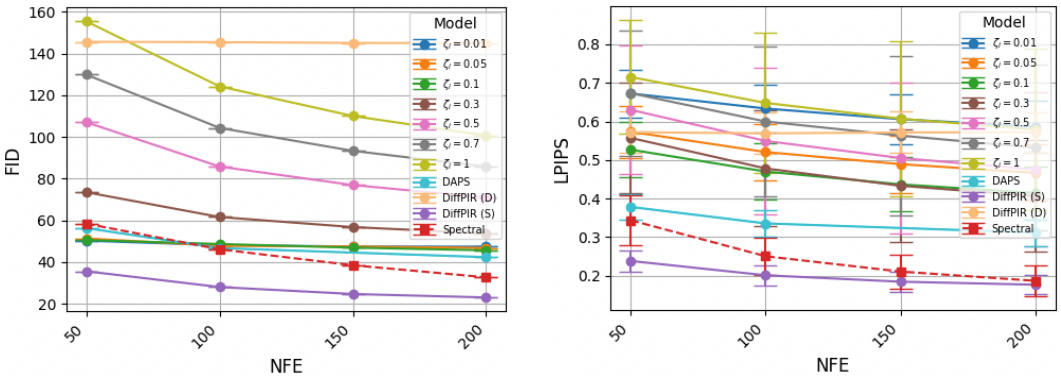
Figure 13 presents the results for the random inpainting task. The proposed spectral recommendations consistently outperform all DPS heuristic configurations in terms of measurement fidelity, achieving improved PSNR values across diffusion steps. In terms of perceptual quality, the spectral recommendations also obtain lower FID values than the heuristic alternatives. However, for small guidance heuristic weights, such as $\zeta' = \{0.01, 0.05\}$, slightly better FID values are observed for a small number of diffusion steps (e.g., 50 steps). As discussed in Section 6.2, perceptual quality alone does not fully characterize performance in inverse problems, where accurate reconstruction of the underlying signal is also required. Figure 14 presents the results for the box inpainting task, showing the same conclusions in terms of both reconstruction fidelity and perceptual quality.

Figures 13 and 14 further compare the spectral recommendations derived for the DPS framework with DAPS and DiffPIR. In both cases, we use the heuristic configurations proposed for the FFHQ inpainting task, including both the deterministic and stochastic DiffPIR variants. Overall, the spectral recommendations outperform both methods across most settings. In the random inpainting task, however, the stochastic DiffPIR variant achieves improved FID values. We emphasize that this comparison is intended to position the proposed framework relative to existing posterior sampling approaches. Since the spectral recommendations are optimized separately for each posterior sampling method, improved performance for one algorithm does not necessarily imply superior performance relative to alternative methods across all metrics and experimental settings. Rather, the results demonstrate clear gains in key aspects. For degradations closer to a circulant LPF (e.g., blur with non-circular padding), generalization is more natural, and the optimization can be performed using a circulant approximation.

K.3.1 Additional visual results

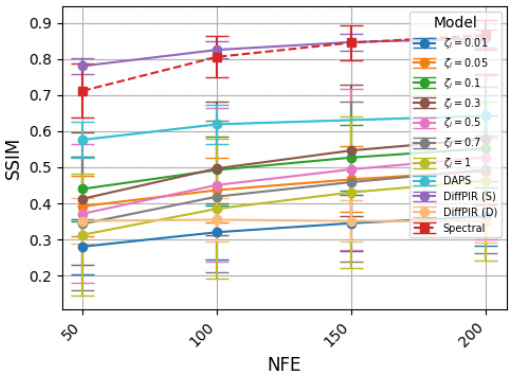
Figures 15 and 16 present representative examples of the DPS algorithm on the FFHQ dataset under random and box inpainting, respectively. In both figures, the first column shows the original image, and the second column shows the degraded input. The next column corresponds to DPS reconstructions using the predefined heuristic for 200 diffusion steps, while the final column presents reconstructions obtained using the spectral recommendations at the same number of steps.

For the DPS heuristic, we show results with $\zeta' = 0.1$, as it provides favorable performance among the tested values of ζ', as observed in Figures 13 and 14. For both degradation settings, the spectral recommendations produce reconstructions that are more consistent with the measurements while preserving a natural visual appearance. In contrast, the DPS heuristic is less faithful to the measurement and tends to introduce new information. For example, in the box inpainting setting, it may also alter regions outside the inpainted area and introduce new content, even when using relatively large guidance weights.

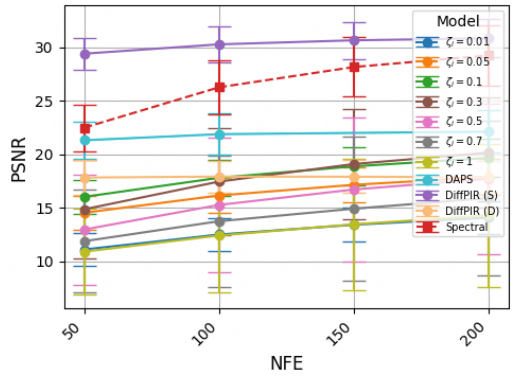


(a) FID ↓

(b) LPIPS ↓

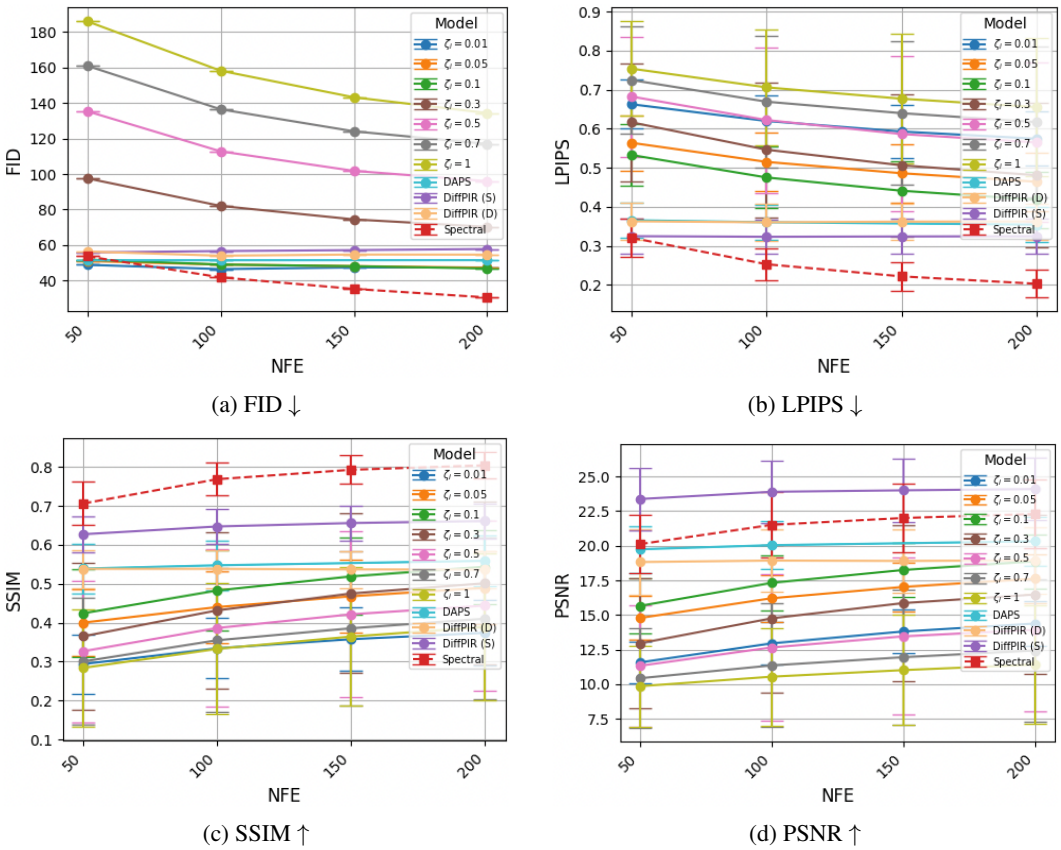


(c) SSIM ↑



(d) PSNR ↑

Figure 13: **Random inpainting degradation.** Comparison between DPS, DiffPIR, and DAPS heuristic configurations and the proposed DPS spectral recommendations under inpainting with 70% randomly removed pixels and additive Gaussian noise $\sigma_y = 0.1$, evaluated across the number of function evaluations (NFE) on the FFHQ dataset.



(a) FID ↓ (b) LPIPS ↓

(c) SSIM ↑ (d) PSNR ↑

Figure 14: **Box inpainting degradation.** Comparison between DPS, DiffPIR, and DAPS heuristic configurations and the proposed DPS spectral recommendations under box inpainting with a 128×128 mask and additive Gaussian noise $\sigma_y = 0.1$, evaluated across the number of function evaluations (NFE) on the FFHQ dataset.

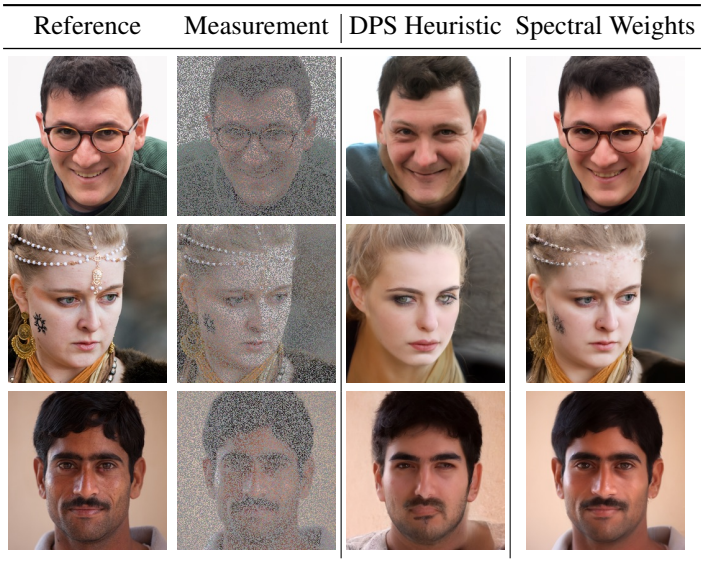


Figure 15: Qualitative comparison of reconstructions on the FFHQ dataset under random inpainting with 70% missing pixels, comparing the DPS heuristic ($\zeta' = 0.1$) and the proposed spectral recommendations across different diffusion step counts.

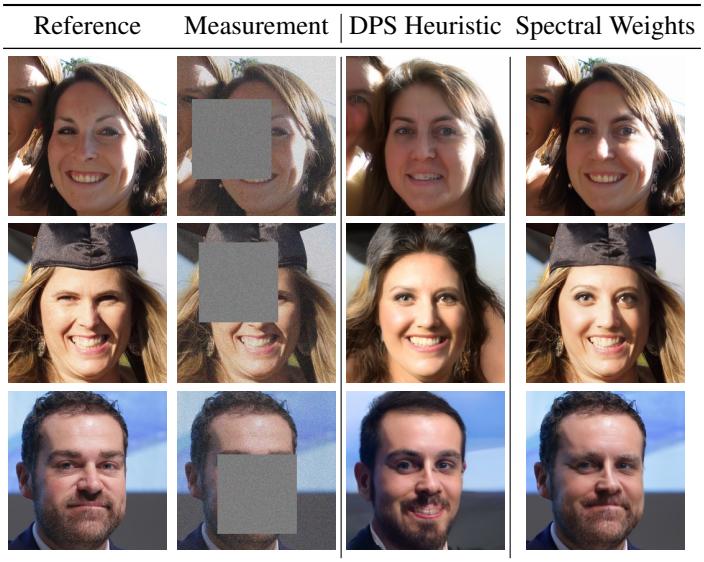


Figure 16: Qualitative comparison of reconstructions on the FFHQ dataset under box inpainting with a 128×128 missing region, comparing the DPS heuristic ($\zeta' = 0.1$) and the proposed spectral recommendations across different diffusion step counts.

K.4 fastMRI dataset

K.4.1 Dataset preparation

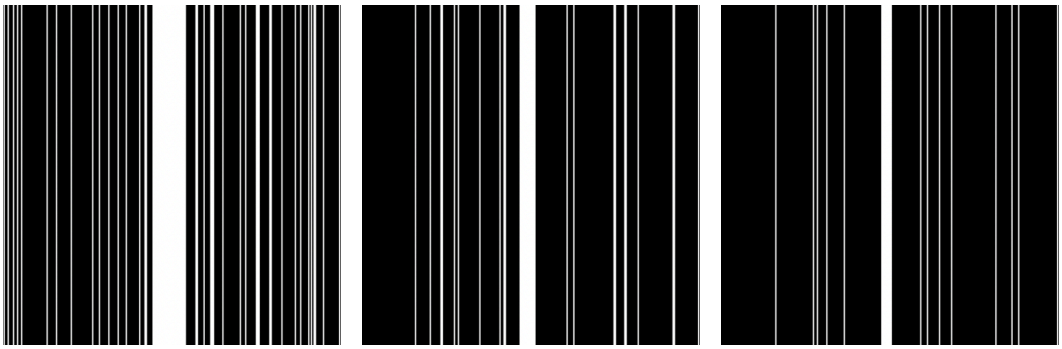
In our work, we consider the single-coil knee MRI scans from the fastMRI dataset [53, 27]. For training, we sample 15 slices at uniformly spaced indices between the one-third and two-third positions of each volume, since slices near the edges typically contain less anatomical information. This results in a training set of $12,546$ images. For validating our spectral recommendations, we use 200 samples extracted from the middle slice of each volume in the validation set. The selected slices are taken from the provided magnitude reconstructions. We center-crop the images to a spatial resolution of 320×320, following [23], and normalize the pixel intensities to the range $[-1, 1]$ before diffusion model training and evaluation.

K.4.2 Degradation operators

Magnetic Resonance Imaging (MRI) is a widely used non-invasive imaging technique that provides clinically important diagnostic information. In MRI systems, measurements are acquired in the Fourier domain, commonly referred to as k-space, through magnetic field encoding. Acquiring the full k-space, however, is often associated with long scan times and increased acquisition costs. As a result, reconstructing high-quality images from incomplete, noisy, or undersampled k-space measurements has become a central inverse problem in accelerated MRI, attracting significant attention in recent years [45, 43].

For the fastMRI single-coil knee MRI slices, we follow the degradation protocol introduced by [23]. In particular, we consider a horizontal frequency undersampling operator in k-space. Given an acceleration factor R, the operator preserves the central $\frac{120}{R}$ low-frequency components and additionally samples $\frac{200}{R}$ frequencies uniformly across the remaining spectrum. Overall, this results in retaining $\frac{320}{R}$ frequency measurements out of the original 320 acquired frequencies.

Notably, the accelerated MRI degradation operator satisfies the shift-invariant property considered in our analysis. In the Fourier domain, its spectral representation $\mathbf{\Lambda}_h$ corresponds to a diagonal sampling mask that selects the acquired k-space frequencies, naturally yielding a diagonal representation of the degradation process. This provides a practically relevant setting in which our framework applies directly to a widely studied real-world inverse problem. Examples of the masks used for acceleration factors $R \in \{4, 8, 12\}$ are shown in Figure 17.



(a) $R = 4$ (b) $R = 8$ (c) $R = 12$

Figure 17: Frequency subsampling masks used for accelerated MRI acquisition with acceleration factors $R \in \{4, 8, 12\}$. As the acceleration factor increases, fewer frequency components are retained, leading to more severe undersampling in k-space.

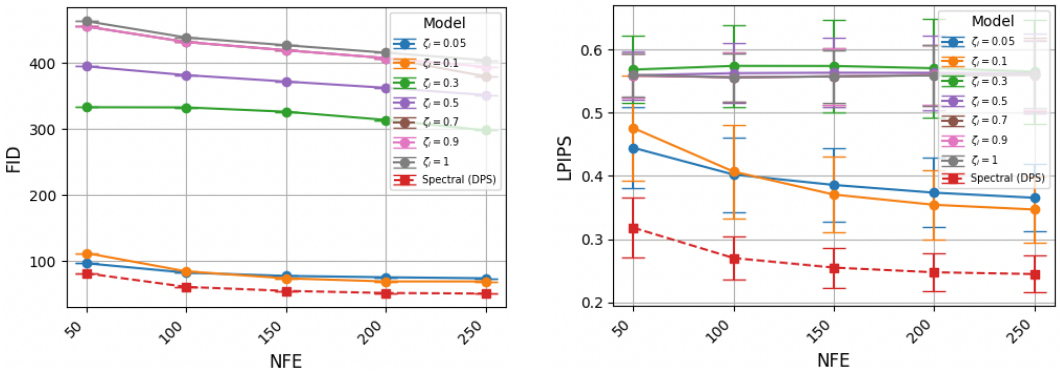
K.4.3 Training details

Since no pretrained prior model is available for the fastMRI dataset, we used the same U-Net architecture employed in [9] for the FFHQ and ImageNet datasets. We adjusted the network input dimensions to match the fastMRI slice dimensions and trained the model using the dataset constructed as described in Section K.4.1. The model was trained for $80{,}000$ steps using eight NVIDIA A40

GPUs with a batch size of 16, using the AdamW optimizer with a learning rate of 10^{-4} and weight decay 0, together with an exponential moving average (EMA) of the model parameters.

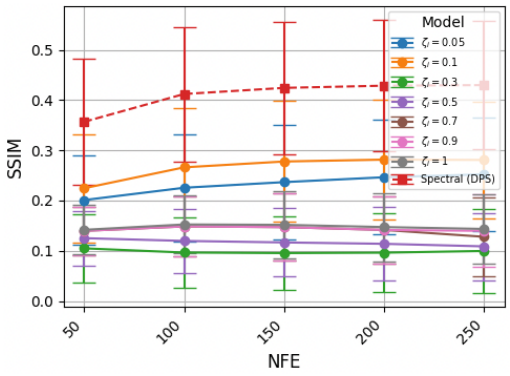
K.4.4 Comprehensive comparison - accelerated sampling

Figure 18 compares the proposed spectral recommendations with multiple DPS heuristic configurations, $\zeta' \in \{0.05, 0.1, 0.3, 0.5, 0.7, 0.9, 1\}$, on the fastMRI validation set described in Section K.4.1, under the LPF degradation setting introduced in Section 6.2. Among the evaluated DPS heuristics, $\zeta' = 0.1$ provides the best tradeoff between perceptual quality and reconstruction fidelity. The spectral recommendations consistently outperform the heuristic configurations across both perceptual and distortion metrics. In particular, noticeable improvements are observed in PSNR, reflecting stronger adherence to the underlying clean scans.

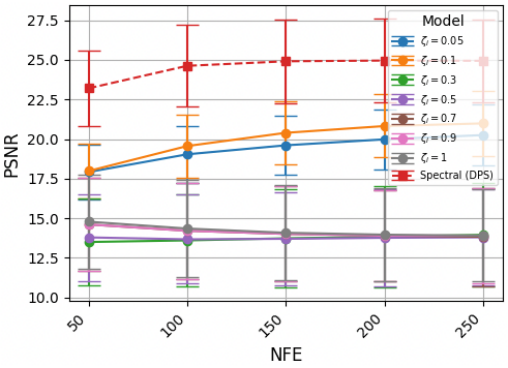


(a) FID ↓

(b) LPIPS ↓



(c) SSIM ↑



(d) PSNR ↑

Figure 18: **LPF degradation.** Comparison between DPS heuristic configurations with $\zeta' = \{0.05, 0.1, 0.3, 0.5, 0.7, 0.9, 1\}$ and the proposed spectral recommendations under LPF degradation with $\mathcal{V} = 0.1$ and additive Gaussian noise $\sigma_y = 0.1$, following the setting in Section 6.2. Results are evaluated across the number of function evaluations (NFE) on the fastMRI dataset.

We now evaluate additional degradation operators. Table 4 presents results on the fastMRI validation set, described in Section K.4.1, under horizontal frequency subsampling degradation with acceleration factors $R \in \{4, 8, 12\}$. For brevity, for each acceleration factor we report only the best-performing DPS heuristic configuration, corresponding to $\zeta' = 0.1$, selected from the range $\zeta' \in \{0.05, 0.1, 0.3, 0.5, 0.7, 0.9, 1\}$.

It can be seen that across all acceleration factors, the spectral recommendations consistently outperform the heuristic configurations in terms of adherence to the original signal $\mathbf{x}_0$. This property is particularly relevant in medical imaging, where, alongside perceptual quality, accurate reconstruction of the underlying scan is required in order to avoid introducing artificial structures that may lead to incorrect interpretation. In terms of perceptual quality, the spectral recommendations also perform well for lower acceleration factors. For the more challenging setting of $R = 12$, however, the selected

heuristic configuration achieves slightly improved FID values for a small number of diffusion steps. This observation further highlights the trade-off between perceptual naturalness and reconstruction fidelity. In particular, under severe degradations, improved PSNR values may be preferable to lower FID scores, since visually plausible scans that deviate from the underlying signal are undesirable in medical reconstruction tasks.

Table 4: Quantitative comparison between the proposed spectral recommendations and the DPS heuristic on the fastMRI dataset under accelerated MRI acquisition with acceleration factors $R \in \{4, 8, 12\}$, using PSNR, SSIM, LPIPS, and FID.

R	METHOD	STEPS	PSNR↑	SSIM↑	LPIPS↓	FID↓
4	SPECTRAL	50	**21.66**	**0.31**	**0.37**	**107.17**
		100	**23.44**	**0.40**	**0.31**	**75.05**
		150	**24.08**	**0.44**	**0.28**	**65.66**
		200	**24.44**	**0.46**	**0.27**	**61.591**
	DPS	50	18.32	0.26	0.46	110.15
		100	19.78	0.27	0.38	75.06
		150	20.29	0.27	0.35	68.59
		200	20.57	0.27	0.34	67.56
8	SPECTRAL	50	**19.21**	**0.23**	**0.44**	143.15
		100	**20.61**	**0.29**	**0.38**	103.48
		150	**21.25**	**0.32**	**0.36**	**86.19**
		200	**21.75**	**0.34**	**0.34**	**81.991**
	DPS	50	17.62	0.23	0.49	**120.27**
		100	19.06	0.27	0.42	**102.31**
		150	19.79	0.27	0.38	88.74
		200	19.99	0.26	0.36	83.13
12	SPECTRAL	50	**17.57**	**0.21**	0.47	142.00
		100	**18.82**	**0.25**	**0.42**	119.50
		150	**19.31**	<u>0.26</u>	**0.40**	**95.15**
		200	**19.61**	<u>0.27</u>	**0.39**	**90.55**
	DPS	50	17.10	0.20	**0.50**	**129.92**
		100	18.26	0.24	0.46	**114.59**
		150	18.96	<u>0.26</u>	0.42	96.63
		200	19.27	<u>0.27</u>	0.40	92.89

K.4.5 Additional visual results

In the following section, we present qualitative comparisons between our spectral guidance recommendations and the DPS heuristic on the fastMRI dataset. Figure 19 considers LPF degradation with $\mathcal{V} = 0.1$ and $\sigma_y = 0.1$, while Figures 20, 21 and 22 show results for accelerated MRI with $R \in \{4,8,12\}$.

In each figure, the first column corresponds to the original image, $\mathbf{x}_0$, and the second to the degraded measurement, $\mathbf{y}$. The following three columns show samples obtained using the DPS heuristic at 50, 100, and 200 diffusion steps (from left to right), while the remaining columns present results obtained using the proposed spectral recommendations at the same step counts.

Since the fastMRI dataset is not considered in the original DPS work [9], no predefined heuristic is available for selecting the guidance parameter. Therefore, for each degradation, we evaluate a wide range of candidate values $\zeta' \in \{0.05, 0.1, 0.3, 0.5, 0.7, 0.9, 1\}$, and select $\zeta' = 0.1$, which consistently provides the best trade-off between perceptual quality and distortion metrics. This setting highlights the practical advantage of our approach. Rather than tuning multiple heuristics, the spectral recommendations are computed once for each degradation and adapt to the prior characteristics, the degradation model, and the diffusion step count.

In Figure 19, the spectral recommendations better preserve the global structure of the original image and the measurement, while the DPS heuristic shows noticeable deviations, especially at low diffusion step counts. As the number of steps increases, the spectral approach progressively refines fine details. For instance, in the second row, a small white spot present in the measurement is recovered using spectral weights, whereas the DPS heuristic fails to reconstruct it.

Although the DPS heuristic often produces visually natural images even at low step counts, as also reflected in the quantitative results, visual quality alone is insufficient in inverse problems, where accurate reconstruction of the underlying signal is required.

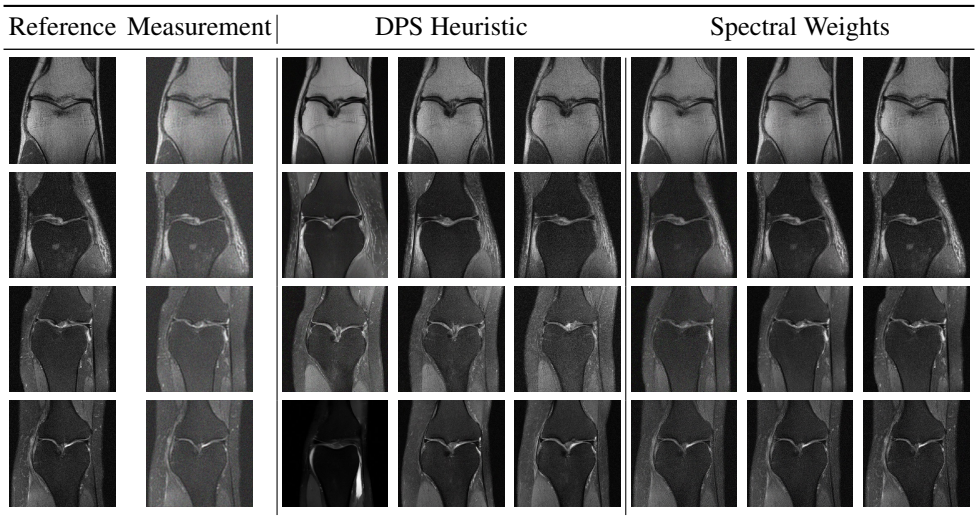


Figure 19: Qualitative comparison of reconstructions on the fastMRI dataset under LPF degradation, comparing the DPS heuristic ($\zeta' = 0.1$) and the proposed spectral recommendations.

Figure 20 demonstrates a similar trend under accelerated sampling using $R = 4$. Across the rows, subtle details such as small white or black spots, which may correspond to clinically relevant structures, are better recovered using the spectral weights, whereas the DPS heuristic typically fails to reconstruct them accurately.

Figures 21 and 22 present comparisons under more severe degradations, using acceleration factors of $R = 8$ and $R = 12$. The same trends persist, with the spectral recovering finer details more accurately. In this regime, the visual naturalness of the DPS heuristic becomes, in some cases, comparable to that of the spectral approach, a behavior also reflected in the FID results in Table 4. This is expected, as relatively small guidance weights such as $\zeta' = 0.1$ tend to favor sampling from the prior, but often struggle to enforce consistency with the measurements. In certain instances, this leads to the

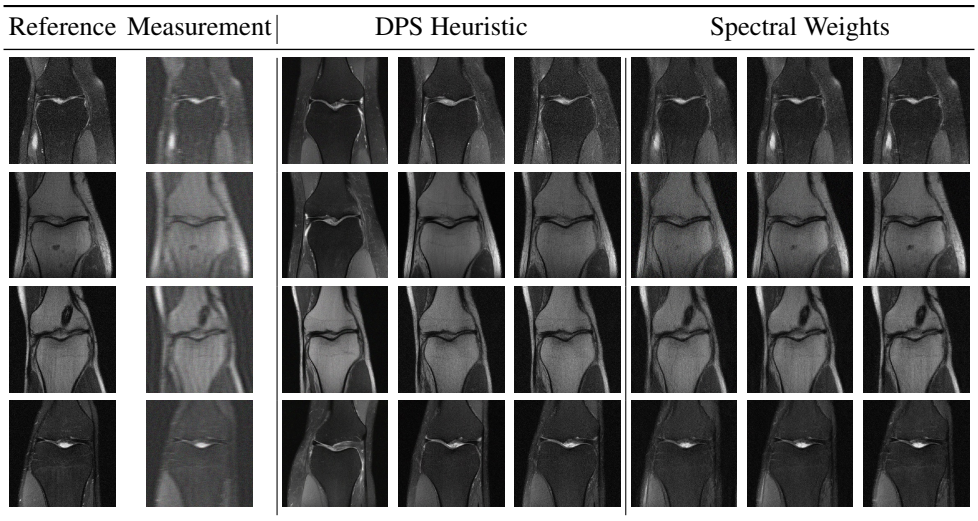


Figure 20: Qualitative comparison of reconstructions on the fastMRI dataset under accelerated MRI acquisition with $R = 4$, comparing the DPS heuristic ($\zeta' = 0.1$) and the proposed spectral recommendations.

introduction of spurious structures. For example, in the second row of the $R = 8$ results, the DPS heuristic produces an artificial dark spot for 50 and 100 diffusion steps that is not present in the measurement.

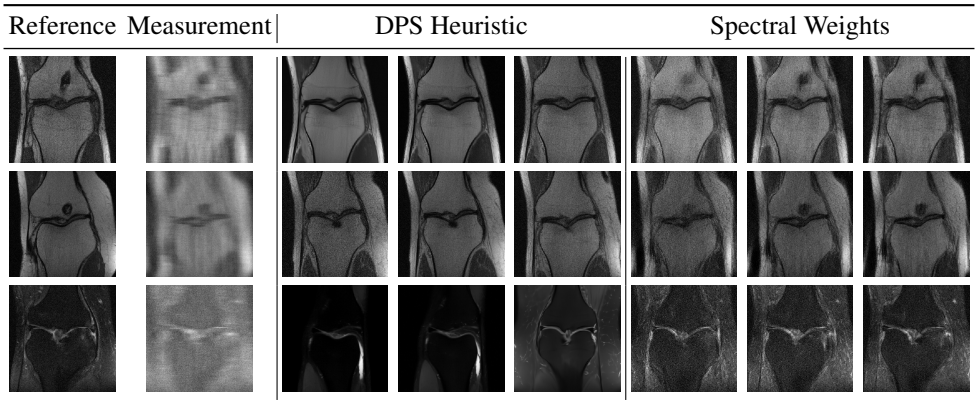


Figure 21: Qualitative comparison of reconstructions on the fastMRI dataset under accelerated MRI acquisition with $R = 8$, comparing the DPS heuristic ($\zeta' = 0.1$) and the proposed spectral recommendations.

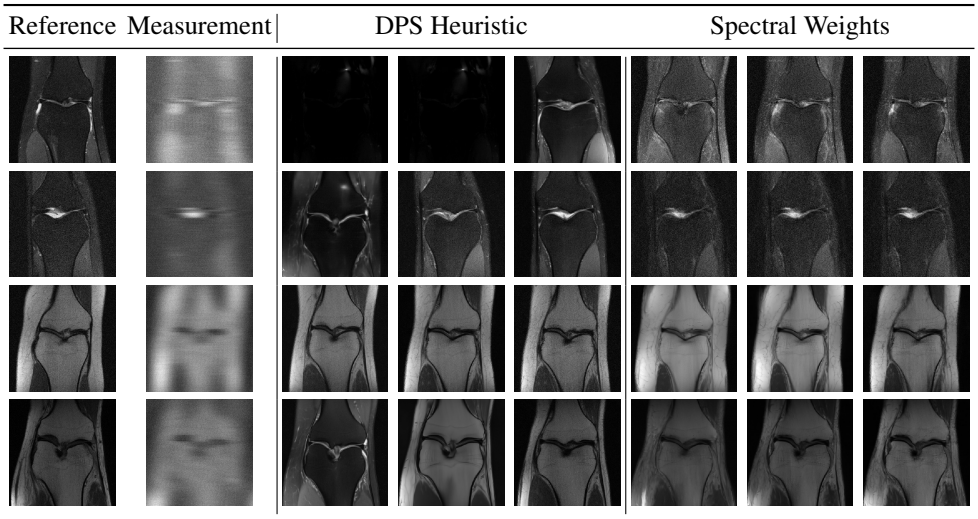


Figure 22: Qualitative comparison of reconstructions on the fastMRI dataset under accelerated MRI acquisition with $R = 12$, comparing the DPS heuristic ($\zeta' = 0.1$) and the proposed spectral recommendations.

K.5 DiffPIR algorithm

K.5.1 spectral recommendations

In this section, we evaluate our framework using the DiffPIR algorithm. As an initial step, Figure 23 presents the spectral recommendations for the guidance parameters $\lambda_{i=1}^{S}$, derived from the deterministic DiffPIR process described in Section I.1, corresponding to the setting $\zeta = 0$. The recommendations are obtained on the FFHQ dataset under the LPF degradation setting described in Section 6.2. Unlike the original DiffPIR heuristic, which employs a constant guidance value independently of the number of diffusion steps, the proposed spectral recommendations vary both across different diffusion step counts and throughout the diffusion process itself.

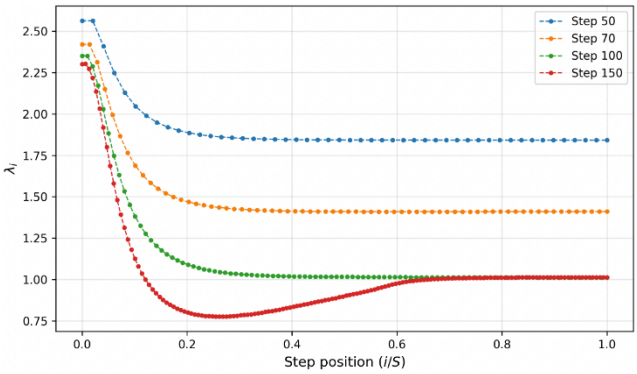


Figure 23: Comparison of the spectral recommendations for the DiffPIR regularization parameter λ_i on the FFHQ 256×256 dataset across selected diffusion step counts $S \in \{50, 70, 100, 150\}$.

K.5.2 Comprehensive comparison

Table 5 compares the heuristic configurations proposed in DiffPIR with the proposed spectral recommendations. We first evaluate both approaches under the LPF degradation setting, and then assess the effectiveness of the spectral recommendations for degradation operators beyond the shift-invariant setting. Specifically, we consider two inpainting tasks: random masking with 70% randomly removed pixels and box inpainting with mask size 128×128. For the inpainting tasks, we report results with $\lambda = 7$ and $\lambda = 6$ for the random and box inpainting settings, respectively, corresponding to the best-performing configurations selected from the heuristic values $\lambda \in 3, 6, 7$, as also recommended in [58]. The spectral recommendations consistently outperform the heuristic configurations under LPF degradation, while a similar trend is observed for the inpainting tasks. In the random inpainting setting, however, the relative performance between the two methods varies across diffusion steps, with a slight advantage shifting between the two methods across diffusion steps.

Table 5: Quantitative comparison between the proposed spectral recommendations and DiffPIR on the FFHQ dataset under LPF degradation, random inpainting with 70% masking, and box inpainting with mask size 128×128, using PSNR, SSIM, LPIPS, and FID.

Degradation	Method	Steps	PSNR↑	SSIM↑	LPIPS↓	FID↓
LPF	Spectral	50	**19.45**	**0.60**	**0.32**	**40.22**
		100	**19.60**	**0.61**	**0.32**	**39.69**
		150	**19.75**	**0.61**	**0.31**	**39.07**
		200	**19.86**	**0.62**	**0.31**	**38.60**
	DiffPIR	50	18.65	0.56	0.36	41.03
		100	18.16	0.55	0.36	39.74
		150	18.20	0.56	0.35	39.41
		200	18.20	0.56	0.35	39.58
Inpainting (random)	Spectral	50	**29.95**	**0.82**	**0.21**	**32.74**
		100	**30.65**	**0.86**	**0.19**	29.03
		150	**30.87**	**0.87**	0.18	25.81
		200	**30.98**	**0.87**	0.18	**23.07**
	DiffPIR	50	29.40	0.78	0.23	35.62
		100	30.27	0.82	0.20	**28.06**
		150	30.65	0.84	0.18	**24.71**
		200	30.83	0.85	0.18	24.25
Inpainting (box)	Spectral	50	**24.01**	**0.70**	**0.27**	**46.52**
		100	**24.56**	**0.73**	**0.26**	**46.39**
		150	**24.71**	**0.75**	**0.26**	**46.33**
		200	**24.76**	**0.76**	**0.26**	**46.25**
	DiffPIR	50	23.37	0.62	0.32	57.73
		100	23.89	0.64	0.32	57.11
		150	24.00	0.65	0.32	56.58
		200	24.09	0.66	0.32	55.81

K.5.3 Additional visual results

Figures 24 and 25 present representative examples of the DiffPIR algorithm on the FFHQ dataset under random and box inpainting, respectively. In both figures, the first column shows the original image, and the second column shows the degraded input. The next two columns correspond to DiffPIR reconstructions using the predefined heuristics, for 50 and 200 diffusion steps (from left to right). The final two columns present the reconstructions obtained using the spectral recommendations, also for 50 and 200 diffusion steps.

For random inpainting (Figure 24), the spectral recommendations produce less noisy reconstructions than the DiffPIR heuristics at lower diffusion step counts. As the number of steps increases, this difference becomes less pronounced, and the reconstructions appear increasingly similar. This behavior is also reflected in Table 5, where the gap between the scores becomes gradually less pronounced across all metrics. For box inpainting (Figure 25), a similar behavior is observed. Under both guidance strategies, the reconstructions appear natural, with a slightly larger gap at lower diffusion step counts in favor of the spectral recommendations. In addition, Compared to the random inpainting setting, the reconstructions appear to be slightly noisier.

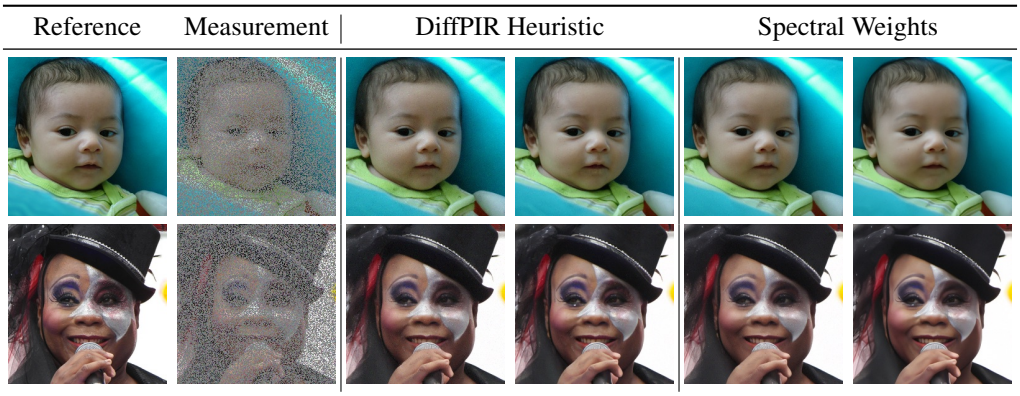


Figure 24: Qualitative comparison of reconstructions on the FFHQ dataset under random inpainting with 70% missing pixels, comparing DiffPIR with predefined heuristics ($\lambda = 7$) and the proposed spectral recommendations across different diffusion step counts.

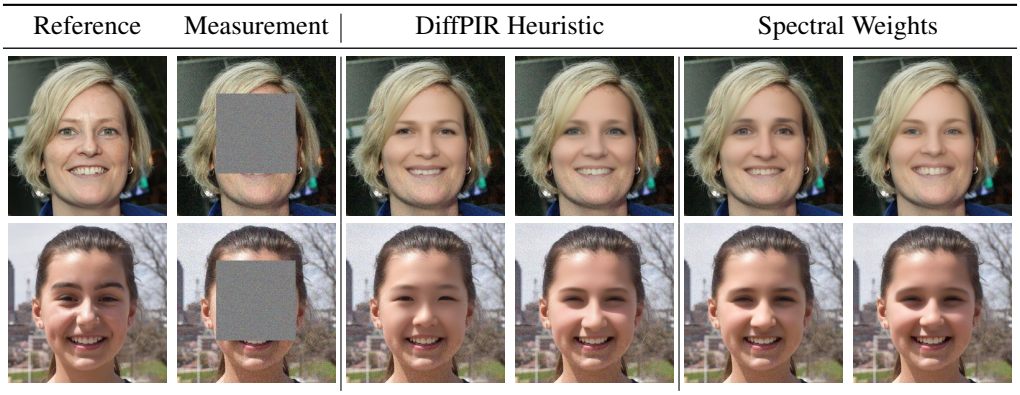


Figure 25: Qualitative comparison of reconstructions on the FFHQ dataset under box inpainting with a 128×128 missing region, comparing DiffPIR with predefined heuristics ($\lambda = 6$) and the proposed spectral recommendations across different diffusion step counts.

L Additional results: realization-specific optimization

This section includes further analysis of posterior behavior under fixed observations. As described in Section 6.2, for three realizations $\mathbf{x}_i$ we generate corresponding degraded observations $\mathbf{y}_i$ and produce $1{,}000$ reconstructions for each case. Evaluation metrics are computed separately for each realization and then averaged.

Table 6 reports the resulting metrics. Here, *Spectral* denotes results obtained using spectral recommendations derived by solving the optimization problem 17, which captures expected posterior behavior across degraded measurements and is solved once for a fixed degradation operator.In contrast, *Spectral* $K{=}1$ corresponds to solving the optimization problem independently for each realization. Overall, both spectral strategies tend to achieve more favorable performance in terms of measurement fidelity and perceptual similarity compared to the DPS heuristic. In addition, solving the optimization problem separately for each realization yields slightly improved results, reflecting the benefit of adapting the weighting to the specific observation.

Table 6: Quantitative results for posterior sampling under fixed observations.

Method	Steps	PSNR↑	SSIM↑	LPIPS↓
DPS	50	12.52	0.27	0.69
	100	14.99	0.35	0.61
	200	18.15	0.44	0.52
	300	19.88	0.49	0.48
	400	21.04	0.53	0.45
	500	21.89	0.55	**0.43**
Spectral	50	**16.63**	**0.48**	**0.60**
	100	17.55	0.51	**0.56**
	200	19.01	0.53	**0.52**
	300	23.05	0.62	**0.43**
	400	23.47	0.61	0.42
	500	22.07	0.61	0.45
Spectral $K{=}1$	50	12.95	0.38	0.66
	100	**18.48**	**0.53**	0.53
	200	**22.24**	**0.59**	0.44
	300	**23.08**	0.62	**0.41**
	400	**24.76**	**0.63**	**0.39**
	500	**22.72**	0.61	0.44

Figures 26 and 27 present qualitative comparisons for two different images. In each figure, the top row shows the reference image on the left and the corresponding degraded observation on the right. Each subsequent row corresponds to a different weighting scheme: DPS (top), the proposed spectral recommendations (middle), and the realization-specific spectral variant with $k{=}1$ (bottom). Within each row, posterior samples obtained with 100, 200, and 400 diffusion steps are shown from left to right. For visualization purposes, each example was selected from the $1,000$ generated samples by choosing the instance that best preserves visual naturalness while remaining consistent with the reference image.

As illustrated in Figures 26 and 27, For a small number of diffusion steps, the DPS heuristic produces samples that reflect the global structure of the reference image, while remaining less constrained by the measurements. In contrast, the spectral recommendations yield samples that adhere more closely to the observed data. We also observe variability across samples produced by all weighting schemes, indicating that the methods do not converge to a single estimate such as the posterior mean.

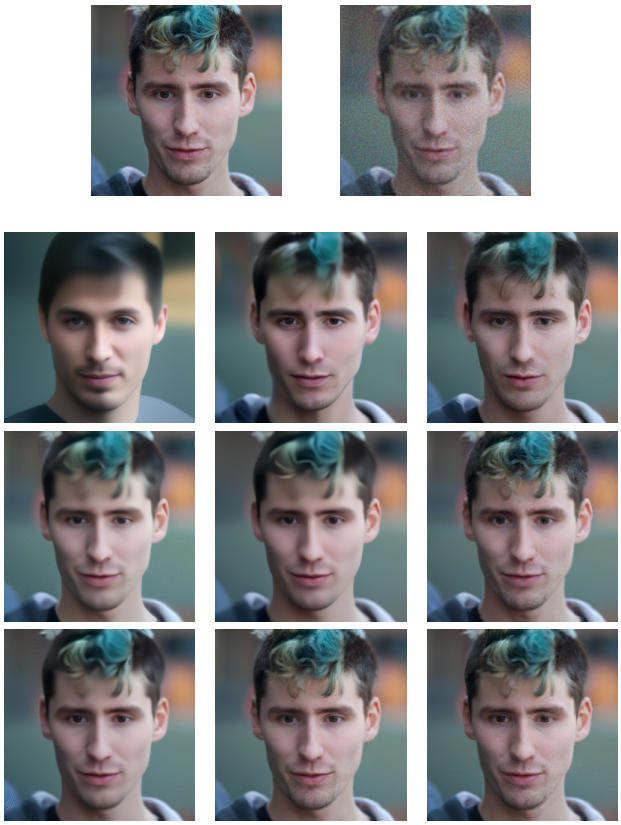


Figure 26: Qualitative comparison of posterior samples under a fixed observation. The top row shows the reference image (left) and the corresponding degraded observation (right). Subsequent rows present reconstructions using different guidance schemes: the DPS heuristic (top), spectral recommendations (middle), and the realization-specific spectral variant with $k = 1$ (bottom). Samples corresponding to 100, 200, and 400 diffusion steps are shown from left to right.

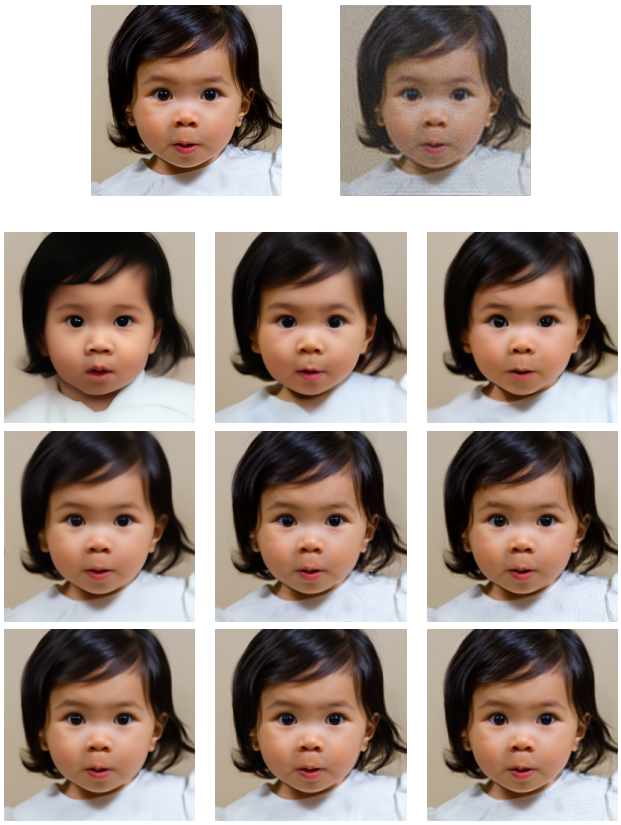


Figure 27: Qualitative comparison of posterior samples under a fixed observation. The top row shows the reference image (left) and the corresponding degraded observation (right). Subsequent rows present reconstructions using different guidance schemes: the DPS heuristic (top), spectral recommendations (middle), and the realization-specific spectral variant with $k = 1$ (bottom). Samples corresponding to 100, 200, and 400 diffusion steps are shown from left to right.

M Covariance matrix estimation

Under stationarity, the covariance matrix of a color image $\mathbf{x}_0 \in \mathbb{R}^{3N}$ has block-circulant with circulant blocks (BCCB) structure:

$$\mathbf{\Sigma}_0 = \begin{pmatrix} \Sigma_{RR} & \Sigma_{RG} & \Sigma_{RB} \\ \Sigma_{GR} & \Sigma_{GG} & \Sigma_{GB} \\ \Sigma_{BR} & \Sigma_{BG} & \Sigma_{BB} \end{pmatrix}.$$

Consequently, each block is diagonalized by the 2D discrete Fourier transform. Let $\mathbf{F}_N$ denote the unitary 2D DFT matrix; then

$$\Sigma_{ij} = \mathbf{F}_N^* \Lambda_{ij} \mathbf{F}_N,$$

where Λ_{ij} are diagonal matrices containing the auto- and cross-channel power spectra.

This yields the Fourier-domain representation of the full covariance

$$\mathbf{\Sigma}_0 = (\mathbf{I}_3 \otimes \mathbf{F}_N^*)\, \mathbf{\Lambda}_0\, (\mathbf{I}_3 \otimes \mathbf{F}_N),$$

where Λ is block-diagonal across spatial frequencies.

This spectral representation allows the optimization problem from Section 3 to be applied, and solved using the acceleration approaches described in Appendix N.

N Optimization time analysis

The computational cost of solving the proposed optimization problem is influenced by several factors, including the signal resolution and the number of diffusion steps. An additional consideration is whether the optimization is performed separately for each measurement realization $\mathbf{y}_i$ (Equation (16)), yielding realization-specific improvements across metrics, or whether a single optimization is solved once for a given degraded system (Equation (17)), and then reused across realizations.

To accelerate the optimization procedure, we adopt two principles introduced in [4]. The first is *iterative optimization*, in which the problem is solved progressively: the process starts with a small number of diffusion steps, allowing fast convergence with relatively low computational requirements. The resulting solution is then interpolated to the appropriate length and used to initialize the optimization with a larger number of steps.

The second principle is *dimensionality reduction*. For high-resolution signals, computational requirements increase significantly; therefore, a low-rank covariance approximation is assumed. Specifically, only the d most significant eigen-directions are retained using principal component analysis (PCA), while components associated with near-zero eigenvalues are discarded.

Table 7 reports the optimization times on the FFHQ dataset at a resolution of $256 \times 256 \times 3$, using iterative initialization and dimensionality reduction, for both the realization-specific problem in Eq. (16) and the degradation-specific problem in Eq. (17). It also reports the average DPS synthesis time per sample as a reference for the computational cost of manual heuristic tuning, which requires repeated synthesis runs and metric evaluation across multiple candidate guidance values. The optimization is performed on a standard CPU using L-BFGS-B via SciPy, with maximum of 2500 iterations, and a function tolerance of 10^{-6}. Synthesis times are measured on an NVIDIA A40 GPU.

Table 7: Optimization and synthesis times as a function of the number of diffusion steps for the FFHQ dataset ($256 \times 256 \times 3$). Optimization results are obtained using iterative initialization and dimensionality reduction. All times are reported in seconds.

Number of diffusion steps	Optimization per degradation (s)	Optimization per realization (s)	Average synthesis time per sample (s)
10	1.77	1.72	1.42
50	13.73	7.84	7.06
100	32.40	45.85	14.11
200	393.30	225.46	28.19
300	189.71	193.98	42.40
400	333.50	330.38	56.51
600	997.22	974.18	84.64
800	1406.56	1280.22	113.00
1000	1902.48	1812.23	125.80

Table 7 compares the optimization times with the average DPS synthesis time. For a fixed degradation and a fixed number of diffusion steps, our optimization problem is solved only once and jointly accounts for the spectral properties of the signal, the measurements, and the diffusion dynamics. In contrast, manual heuristic tuning requires repeated synthesis runs together with the corresponding quality evaluation to identify a suitable guidance parameter. In practice, this limits the search to a finite predefined set of values. In the realization-specific setting, where the optimization is performed separately for each measurement, a comparable heuristic selection likewise requires evaluating candidate guidance parameters separately for each measurement through repeated synthesis and posterior-level evaluation. Overall, the optimization overhead remains modest relative to the synthesis cost and the computational cost associated with heuristic tuning.

NeurIPS Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Justification: The main claims and contributions are clearly stated in the Abstract and reinforced in the Introduction. These are further supported by detailed mathematical derivations in Sections 3 and 4

   Guidelines:

   - The answer [N/A] means that the abstract and introduction do not include the claims made in the paper.
   - The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A [No] or [N/A] answer to this question will not be perceived well by the reviewers.
   - The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
   - It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes]

   Justification: Throughout the paper, we present the underlying assumptions of our approach in both the analytical derivations and the experimental sections. In addition, we discuss the limitations in a dedicated section titled "Limitations" 6.2 and suggest potential directions for future research. We further demonstrate that relaxing these assumptions still yields effective weighting strategies in real-world datasets and degradation operators. We also analyze the optimization time of our method in Appendix N.

   Guidelines:

   - The answer [N/A] means that the paper has no limitation while the answer [No] means that the paper has limitations, but those are not discussed in the paper.
   - The authors are encouraged to create a separate "Limitations" section in their paper.
   - The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
   - The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
   - The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
   - The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
   - If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
   - While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory assumptions and proofs**

   Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

   Answer: [Yes]

   Justification: The Gaussianity and shift-invariant assumptions are explicitly stated, and we further discuss the results when these assumptions are relaxed. All theoretical results are presented in Sections 3 and 4, as well as throughout the appendices.

   Guidelines:

   - The answer [N/A] means that the paper does not include theoretical results.
   - All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
   - All assumptions should be clearly stated or referenced in the statement of any theorems.
   - The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
   - Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
   - Theorems and Lemmas that the proof relies upon should be properly referenced.

4. **Experimental result reproducibility**

   Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

   Answer: [Yes]

   Justification: Throughout the paper, we explicitly describe the optimization problem, including the loss function and the solution method (SLSQP via SciPy J.1). To support reproducibility, the pretrained architectures for the FFHQ and ImageNet datasets are publicly available through the repository of [9]. For the fastMRI dataset, we provide detailed information on the data processing pipeline and the model architecture in Section K.4. We intend to release the corresponding code and pretrained models to further facilitate reproducibility.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - If the paper includes experiments, a [No] answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
   - If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
   - Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general. releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
   - While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
     (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
     (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.

(c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).

(d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

   Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

   Answer: [No]

   Justification: The code for solving the optimization problem is not yet publicly available; However, we provide a detailed description of the optimization procedure in Sections 3.4, 6 and Appendices J, K and L which can be reproduced using standard tools such as the `scipy` package. For the FFHQ and ImageNet datasets, we refer to the publicly available open-source model for reproducibility, where pretrained models and inference code are provided. All datasets used in our experiments are open source and can be accessed as described in Section 6. For the fastMRI dataset, we include detailed information on the data processing pipeline, the model architecture, and the training procedure in Section K.4. The corresponding code will be released shortly.

   Guidelines:

   - The answer [N/A] means that paper does not include experiments requiring code.
   - Please see the NeurIPS code and data submission guidelines (`https://neurips.cc/public/guides/CodeSubmissionPolicy`) for more details.
   - While we encourage the release of code and data, we understand that this might not be possible, so [No] is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
   - The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (`https://neurips.cc/public/guides/CodeSubmissionPolicy`) for more details.
   - The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
   - The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
   - At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
   - Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. **Experimental setting/details**

   Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer) necessary to understand the results?

   Answer: [Yes]

   Justification: Throughout the paper, we provide detailed descriptions of the optimization method, in Section 3 and Appendix J, as well as the datasets used in Section 6. In addition, for the fastMRI dataset, we describe the data processing pipeline together with the training procedure and model architecture in Appendix K.4.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.

- The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
- The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment statistical significance**

   Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

   Answer: [Yes]

   Justification: Throughout all experimental results, we report the mean and standard deviation computed across multiple runs for all evaluation metrics, including LPIPS, FID, PSNR, and SSIM. In addition, we provide uncertainty bars for the spectral recommendations themselves, as well as for the Wasserstein distance evaluations.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The authors should answer [Yes] if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.
   - The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
   - The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
   - The assumptions made should be given (e.g., Normally distributed errors).
   - It should be clear whether the error bar is the standard deviation or the standard error of the mean.
   - It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
   - For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g., negative error rates).
   - If error bars are reported in tables or plots, the authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments compute resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer: [Yes]

   Justification: In section N, we compare execution times for the optimization problem across different covariance matrix sizes and discuss the results, along with specifying the compute resources used.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
   - The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
   - The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. **Code of ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics https://neurips.cc/public/EthicsGuidelines?

Answer: [Yes]

Justification: We have reviewed the NeurIPS Code of Ethics and fully comply with its guidelines.

Guidelines:

- The answer [N/A] means that the authors have not reviewed the NeurIPS Code of Ethics.
- If the authors answer [No], they should explain the special circumstances that require a deviation from the Code of Ethics.
- The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader impacts**

Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

Answer: [N/A]

Justification: We believe this work has limited direct societal impact, as it focuses on the analysis of diffusion-based posterior sampling methods. While our experiments include medical imaging data (fastMRI), the study is purely methodological and does not involve clinical use. Nevertheless, reconstruction errors or artifacts may affect interpretation. More generally, the use of pretrained generative models may introduce biases or misleading outputs, and should be taken into account when applying such methods in practice.

Guidelines:

- The answer [N/A] means that there is no societal impact of the work performed.
- If the authors answer [N/A] or [No], they should explain why their work has no societal impact or why the paper does not address societal impact.
- Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.
- The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate Deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
- The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pre-trained language models, image generators, or scraped datasets)?

Answer: [N/A]

Justification: The paper poses no such risks.

Guidelines:

- The answer [N/A] means that the paper poses no such risks.

- Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
- Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
- We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. **Licenses for existing assets**

    Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

    Answer: [Yes]

    Justification: We acknowledge and properly credit all original sources of the assets used in our work. Specifically, we use the following datasets: FFHQ [24], released under the Creative Commons BY-NC-SA 4.0 license; ImageNet [12], available under its standard non-commercial research use terms; and fastMRI [53, 27], released under the Creative Commons BY-NC-SA 4.0 license. For the pretrained models, we follow the implementation of DPS [9], using its publicly available repository, which is based on the guided-diffusion framework [13]. All licenses and terms of use have been respected and properly cited.

    Guidelines:

    - The answer [N/A] means that the paper does not use existing assets.
    - The authors should cite the original paper that produced the code package or dataset.
    - The authors should state which version of the asset is used and, if possible, include a URL.
    - The name of the license (e.g., CC-BY 4.0) should be included for each asset.
    - For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
    - If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, `paperswithcode.com/datasets` has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
    - For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.
    - If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New assets**

    Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

    Answer: [N/A]

    Justification: The paper does not release new assets.

    Guidelines:

    - The answer [N/A] means that the paper does not release new assets.
    - Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
    - The paper should discuss whether and how consent was obtained from people whose asset is used.
    - At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and research with human subjects**

Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

Answer: [N/A]

Justification: The paper does not involve crowdsourcing nor research with human subjects.

Guidelines:

- The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
- Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
- According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional review board (IRB) approvals or equivalent for research with human subjects**

Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

Answer: [N/A]

Justification: The paper does not involve crowdsourcing nor research with human subjects.

Guidelines:

- The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
- Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
- We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
- For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.

16. **Declaration of LLM usage**

Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research? Note that if the LLM is used only for writing, editing, or formatting purposes and does *not* impact the core methodology, scientific rigor, or originality of the research, declaration is not required.

Answer: [N/A]

Justification: The core method development in this research does not involve LLMs as any important, original, or non-standard components.

Guidelines:

- The answer [N/A] means that the core method development in this research does not involve LLMs as any important, original, or non-standard components.
- Please refer to our LLM policy in the NeurIPS handbook for what should or should not be described.