

Image Restoration via Primal Dual Hybrid Gradient and Flow Generative Model

Ji Li¹, Chao Wang^{2*}

¹Academy for Multidisciplinary Studies, Capital Normal University, Beijing, China

²College of Science, China Agricultural University, Beijing, China
matliji@163.com, wywwwwnx@163.com

Abstract

Regularized optimization has been a classical approach to solving imaging inverse problems, where the regularization term enforces desirable properties of the unknown image. Recently, the integration of flow matching generative models into image restoration has garnered significant attention, owing to their powerful prior modeling capabilities. In this work, we incorporate such generative priors into a Plug-and-Play (PnP) framework based on proximal splitting, where the proximal operator associated with the regularizer is replaced by a time-dependent denoiser derived from the generative model. While existing PnP methods have achieved notable success in inverse problems with smooth squared ℓ_2 data fidelity—typically associated with Gaussian noise—their applicability to more general data fidelity terms remains under-explored. To address this, we propose a general and efficient PnP algorithm inspired by the primal-dual hybrid gradient (PDHG) method. Our approach is computationally efficient, memory-friendly, and accommodates a wide range of fidelity terms. In particular, it supports both ℓ_1 and ℓ_2 norm-based losses, enabling robustness to non-Gaussian noise types such as Poisson and impulse noise. We validate our method on several image restoration tasks, including denoising, super-resolution, deblurring, and inpainting, and demonstrate that ℓ_1 and ℓ_2 fidelity terms outperform the conventional squared ℓ_2 loss in the presence of non-Gaussian noise.

Code — <https://github.com/Chilie/PDHG-Flow>

Extended version — <https://arxiv.org/abs/2511.06748>

Introduction

Many image restoration tasks in image processing and computer vision can be formulated as solving a linear inverse problem:

$$\mathbf{y} = \text{noisy}(\mathbf{A}\mathbf{x}), \quad (1)$$

where $\mathbf{x} \in \mathbb{R}^n$ denotes the unknown clean image, $\mathbf{y} \in \mathbb{R}^m$ and linear matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ describes the degradation model. Image restoration is to recover \mathbf{x} from the noisy observation \mathbf{y} .

In general, problem (1) is ill-posed, and direct inversion is infeasible without incorporating prior knowledge about the

solution. A widely adopted approach is to compute the maximum a posteriori (MAP) estimate of the posterior distribution $p(\mathbf{x}|\mathbf{y})$. The MAP estimate corresponds to the most probable solution and is given by:

$$\hat{\mathbf{x}} = \operatorname{argmax}_{\mathbf{x} \in \mathbb{R}^n} \{\log p(\mathbf{y}|\mathbf{x}) + \log p(\mathbf{x})\}, \quad (2)$$

where $p(\mathbf{y}|\mathbf{x})$ models the noise distribution, and $p(\mathbf{x})$ is the image prior. Since directly computing $\log p(\mathbf{x})$ is often intractable, MAP estimation is commonly reformulated as a regularized optimization problem:

$$\hat{\mathbf{x}} = \operatorname{argmin}_{\mathbf{x} \in \mathbb{R}^n} F(\mathbf{A}\mathbf{x}) + G(\mathbf{x}), \quad (3)$$

where $F(\cdot)$ encodes the data fidelity, measuring consistency between $\mathbf{A}\mathbf{x}$ and \mathbf{y} , while $G(\cdot)$ encodes the regularization, promoting prior knowledge about \mathbf{x} . Problem (3) is typically solved using proximal splitting methods (Parikh, Boyd et al. 2014), especially when $G(\mathbf{x})$ is convex.

In proximal splitting, the proximal step associated with $G(\mathbf{x})$ can be interpreted as a denoising operation. This insight has led to replacing the proximal step with advanced denoisers, such as BM3D (Dabov et al. 2006, 2007), or learned neural network-based denoisers (Zhang et al. 2017, 2020; Ryu et al. 2019). More recently, generative models that learn image priors directly from data have been incorporated into the Plug-and-Play (PnP) framework (Martin et al. 2024; Zhu et al. 2023; Song et al. 2023; Graikos et al. 2022; Hurault, Leclaire, and Papadakis 2022; Kamilov, Mansour, and Wohlberg 2017). Although generative models can in principle provide $\log p(\mathbf{x})$ (Chen et al. 2018), evaluating its gradient via backpropagation is computationally expensive. PnP methods circumvent this issue by directly substituting a denoiser induced by a generative model for the proximal operator. Such approaches have achieved state-of-the-art results in inverse problems using powerful generative priors, including diffusion models (Chung et al. 2022a,b; Song et al. 2022; Kawar et al. 2022) and flow matching models (Pokle et al. 2024; Zhang et al. 2024; Martin et al. 2024).

However, prior PnP methods have predominantly focused on inverse problems with Gaussian noise, where the data fidelity term $F(\cdot)$ is typically the squared ℓ_2 loss. While effective under Gaussian noise, this loss function is suboptimal for non-Gaussian noise models. For example, Poisson noise

*Corresponding author.

is the more appropriate modeling for the shot noise of photons in the observation. Salt-and-pepper noise, also known as impulse noise, is another form of noise in digital images. Although one can still use squared ℓ_2 loss, the mismatch in noise modeling often leads to subpar restoration quality.

In this work, we extend PnP methods based on flow matching generative models to handle non-Gaussian noise by generalizing the data fidelity term $F(\cdot)$ in (3). Specifically, we consider proximity-friendly fidelity functions such as ℓ_1 and ℓ_2 norms, which are better suited for modeling Poisson and impulse noise, respectively. For sparse impulse noise, ℓ_1 -norm fidelity is widely adopted due to its robustness to outliers. While the Poisson likelihood leads naturally to a specific form of F , in practice, ℓ_2 norm fidelity is often preferred for its numerical stability. Our approach unifies various noise models within a general PnP framework, encompassing squared ℓ_2 loss as a special case.

The main contributions of this papers are as follows:

- We propose a generalized PnP framework using the primal-dual hybrid gradient (PDHG) method to solve (3) with non-Gaussian noise, by accommodating flexible fidelity terms $F(\cdot)$.
- The proximal step for $G(\cdot)$ is implemented using an implicit denoiser derived from flow matching models.
- We seamlessly integrate flow matching generative models into the classical PDHG method, achieving an efficient and implementation-friendly solution.
- Extensive experiments on two benchmark datasets across diverse tasks, including denoising, deblurring, inpainting, and super-resolution, demonstrate the superiority of our method under non-Gaussian noise, outperforming existing PnP methods designed for Gaussian noise.

Background

The performance of imaging inverse problems has significantly improved with the advent of generative models that learn expressive image priors. Among these, diffusion (Song et al. 2020; Ho, Jain, and Abbeel 2020) and flow matching models (Lipman et al. 2023; Tong et al.) have attracted considerable attention due to their superior generative capabilities compared to earlier approaches. These models have demonstrated remarkable potential for image restoration when used as priors within optimization frameworks.

Flow Matching Generative Model

Flow matching aims to learn a continuous transformation that maps samples from a simple distribution $p(\mathbf{x}_0)$ (e.g., Gaussian) to samples from a data distribution $p(\mathbf{x}_1)$. Let $\mathcal{P}(\mathbb{R}^n)$ denote the space of probability measures over \mathbb{R}^n , and consider two probability measures P_0 and P_1 representing the latent and data distributions, respectively. Define $\Gamma(P_0, P_1)$ as the set of couplings $\pi \in \mathcal{P}(\mathbb{R}^n \times \mathbb{R}^n)$ with marginals P_0 and P_1 .

Flow matching constructs a continuous probability path $t \mapsto P_t$ for $t \in [0, 1]$, interpolating between P_0 and P_1 . Given a coupling $(\mathbf{x}_0, \mathbf{x}_1)$, one defines the interpolated point

$\mathbf{x}_t = (1 - t)\mathbf{x}_0 + t\mathbf{x}_1$, which induces a trajectory governed by the ordinary differential equation (ODE):

$$\frac{d\mathbf{x}_t}{dt} = \mathbf{v}(\mathbf{x}_t, t), \quad (4)$$

where $\mathbf{v}(\mathbf{x}_t, t)$ is the velocity field. The goal is to learn this velocity field from data. To this end, a neural network $\mathbf{v}_\theta(\mathbf{x}_t, t)$ parameterized by θ is used to approximate $\mathbf{v}(\mathbf{x}_t, t)$ such that the evolution of P_t satisfies the *continuity equation*:

$$\partial_t P_t + \nabla \cdot (P_t \mathbf{v}(\mathbf{x}_t, t)) = 0. \quad (5)$$

To enable the continuous transition from a simple density $p(\mathbf{x}_0)$ to a data density $p(\mathbf{x}_1)$, the core of flow matching is to define the velocity field $\mathbf{v}(\mathbf{x}_t, t)$ in (4).

Similar to normalizing flows (Kingma and Dhariwal 2018), the flow matching model can be trained via maximum likelihood estimation:

$$\mathcal{L}_{\text{ML}}(\theta) = \mathbb{E}_{\mathbf{x}_1 \sim p(\mathbf{x}_1)}[\log p(\mathbf{x}_1)], \quad (6)$$

where the expectation is taken over the dataset. The log-likelihood of $p_1(\mathbf{x})$ can be computed using the differential equation of evolution of the log-likelihood:

$$\frac{\partial \log p(\mathbf{x}_t)}{\partial t} = -\text{trace} \left(\frac{\mathbf{v}_\theta(\mathbf{x}_t, t)}{\partial \mathbf{x}_t} \right). \quad (7)$$

One can additionally obtain the likelihood of the final $\log p(\mathbf{x}_1)$ via integrating (7):

$$\log p(\mathbf{x}_1) = \log p(\mathbf{x}_0) - \int_0^1 \text{trace} \left(\frac{\mathbf{v}_\theta(\mathbf{x}_s, s)}{\partial \mathbf{x}_s} \right) ds. \quad (8)$$

However, evaluating this integral is computationally intensive. To avoid this, simulation-free training methods have been proposed.

The conditional flow matching offers such a simulation-free training approach (Lipman et al. 2023). Its learning objective is

$$\mathcal{L}_{\text{CFM}}(\theta) = \mathbb{E}_t[\mathbb{E}_{(\mathbf{x}_0, \mathbf{x}_1) \sim \pi} \|\mathbf{v}_\theta(\mathbf{x}_t, t) - (\mathbf{x}_1 - \mathbf{x}_0)\|_2^2], \quad (9)$$

where π is a coupling between $p(\mathbf{x}_0)$ and $p(\mathbf{x}_1)$. The minimization of loss (9) is equivalent to the minimization of the direct flow matching loss

$$\mathcal{L}_{\text{FM}}(\theta) := \mathbb{E}_t[\mathbb{E}_{\mathbf{x}_t \sim P_t} \|\mathbf{v}_\theta(\mathbf{x}_t, t) - \mathbf{v}(\mathbf{x}_t, t)\|_2^2]. \quad (10)$$

The choice of coupling π critically affects both training efficiency and generative quality. Optimal coupling can be obtained via standard optimal transport solvers (Tong et al.). Alternatively, rectified flow (Liu, Gong et al. 2023) refines transport maps by iteratively straightening ODE trajectories, improving convergence and model quality.

Denoiser Induced by Flow Model

Let \mathbf{x} be a sample from the random variable $X_t := (1 - t)X_0 + tX_1$, where $X_0 \sim P_0$ and $X_1 \sim P_1$ respectively. For a fixed time $t \in [0, 1]$, the minimizer \mathbf{v}_t^* of the CFM loss (9) over all admissible vector fields is given by:

$$\mathbf{v}_t^*(\mathbf{x}) = \mathbb{E}[X_1 - X_0 | X_t = \mathbf{x}]. \quad (11)$$

Assume we are in the ideal case where $\mathbf{v}_t^\theta = \mathbf{v}_t^*$. Then it follows that, for any \mathbf{x} and $t \in [0, 1]$, the denoising operator

$$D_t(\mathbf{x}) = \mathbf{x} + (1 - t)\mathbf{v}_t^*(\mathbf{x}) \quad (12)$$

satisfies that $D_t(\mathbf{x}) = \mathbb{E}[X_1 | X_t = \mathbf{x}]$, *i.e.*, the conditional mean estimator of the clean image. It implies that one has a natural denoiser operator induced by a flow matching generative model. Thus, flow matching models naturally induce denoisers that can be directly integrated into proximal splitting methods for solving the MAP problem in (3). This framework enables a training-free, zero-shot solution for imaging inverse problems: once trained, the generative model can be applied to a wide range of inverse tasks without retraining or fine-tuning. Compared to methods that directly learn conditional score functions for specific tasks (Whang et al. 2022), our approach provides a universal and versatile framework for solving both linear and nonlinear inverse problems using pre-trained, off-the-shelf generative models. There is no need to fine-tune the model.

Related Works

Measurement-Guided Sampling with Diffusion Model

Prior to the development of flow matching models, diffusion models emerged as a powerful class of generative models and have been widely adopted for image restoration tasks, including both linear and nonlinear inverse problems. The central idea in these approaches is to guide the sampling process using measurement information, typically through a gradient descent step integrated into the unconditional diffusion sampling scheme.

Various methods differ in how they incorporate measurement guidance via fidelity losses. Chung et al. (2022a) introduced a rough fidelity loss based on known likelihoods and Tweedie’s formula. However, this approximation introduces a significant bias, negatively impacting restoration quality. To address this, Song et al. (2022) proposed IIGDM, improving fidelity modeling for Gaussian noise. Nonetheless, IIGDM is tailored to Gaussian noise scenarios and does not generalize well to non-Gaussian or nonlinear problems.

Other notable methods include DDRM (Kawar et al. 2022) and DDNM/DDNM+ (Wang, Yu, and Zhang 2022), which leverage the singular value decomposition (SVD) of the degradation matrix \mathbf{A} for spectral inpainting. However, SVD may be computationally infeasible or unavailable in many practical applications.

Measurement-Guided Linear Inversers via Flows

Building on the efficiency of flow matching models, several works have proposed measurement-guided sampling techniques for image restoration. In particular, Pokle et al. (2024) introduced a training-free flow-based method that incorporates a theoretically justified weighting scheme into the sampling process. This method modifies the standard flow model’s unconditional sampling by introducing a gradient descent step to refine predictions using measurement information. Like diffusion-based counterparts, these methods aim to solve linear inverse problems but with improved efficiency and interpretability.

Plug-and-Play (PnP) Methods

The Plug-and-Play (PnP) framework integrates powerful denoisers into iterative optimization algorithms for image restoration. Zhu et al. (2023) introduced PnP-HQS (also known as DiffPIR), which incorporates a denoiser derived from a pre-trained diffusion model into a half-quadratic splitting (HQS) framework. The algorithm proceeds along a restoration trajectory:

$$\mathbf{x}_T \rightarrow \cdots \rightarrow \mathbf{x}_t \rightarrow \hat{\mathbf{x}}_0(\mathbf{x}_t) \rightarrow \hat{\mathbf{x}}'_0(\mathbf{x}_t) \rightarrow \mathbf{x}_{t-1} \rightarrow \cdots \rightarrow \mathbf{x}_0.$$

At each step, given \mathbf{x}_t , the current estimate \mathbf{x}_t is denoised to obtain $\hat{\mathbf{x}}_0(\mathbf{x}_t)$. A measurement-guided update is then computed by solving the proximal problem:

$$\hat{\mathbf{x}}'_0(\mathbf{x}_t) = \underset{\mathbf{x}}{\operatorname{argmin}} \quad \frac{1}{2} \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2^2 + \frac{\lambda\sigma_t^2}{\alpha_t^2} \|\mathbf{x} - \mathbb{E}[\mathbf{x}_0 | \mathbf{x}_t]\|_2^2.$$

In a related work, Han et al. (2024) employed a generative model to initialize a classical total variation (TV) regularized optimization, solved via the PDHG method, to address limited-angle CT reconstruction. Li and Wang (2024) solves nonlinear Fourier phase retrieval within the framework.

For flow matching models, Martin et al. (2024) proposed PnP-FBS, which integrates flow matching into a forward-backward splitting (FBS) framework. Their algorithm alternates between gradient descent on the data fidelity term and a denoising step using the flow model. Similar to PnP-HQS, a re-noising step is applied before denoising to match the noise level. However, both PnP-HQS and PnP-FBS consider only smooth squared ℓ_2 fidelity terms and are limited to Gaussian noise settings.

Differentiation through Flow ODE

The flow matching model defines a continuous transformation from a latent variable to a data sample via an ODE. Leveraging this structure, Ben-Hamu et al. (2024) formulates D-Flow via an ODE-constrained optimization problem, wherein gradients with respect to the latent variable are computed using adjoint methods. Although this approach allows for exact computation of the negative log-likelihood, it incurs high computational cost due to integration. To alleviate this, Flow-Priors (Zhang et al. 2024) introduced a sequential approximation strategy that sidesteps the need for costly integrals, enabling efficient MAP estimation.

Proposed Method

While existing training-free Plug-and-Play (PnP) methods utilizing flow-based generative models have shown promising results for imaging inverse problems under Gaussian noise, their extension to non-Gaussian noise settings remains underexplored. Previous generative PnP methods, including PnP-HQS and PnP-FBS, primarily address Gaussian noise, resulting in data fidelity terms modeled via the squared ℓ_2 loss. However, in practice, non-Gaussian noise types such as Poisson and impulse noise are common, for which alternative loss functions—specifically the ℓ_1 or ℓ_2 norms—often outperform the squared ℓ_2 loss due to better noise modeling.

To address this gap, we propose a generalized PnP framework that integrates flow-based generative priors and supports both ℓ_1 and ℓ_2 norm-based data fidelity terms. Our key innovation lies in leveraging the Primal-Dual Hybrid Gradient (PDHG) method to efficiently handle non-smooth fidelity losses such as ℓ_1 and ℓ_2 norms, whose proximal operators can be computed in closed form.

Primal Dual Hybrid Gradient (PDHG)

The PDHG method (Chambolle and Pock 2011; Esser, Zhang, and Chan 2010; O’Connor and Vandenberghe 2020) is an effective optimization method designed to solve convex composite problems of the form in (3). The method relies on alternating updates involving the proximal operators of G and the convex conjugate of F , denoted F^* . A more detailed discussion on proximal operators and conjugate functions is provided in Appendix of the arXiv extended version.

Given initializations \mathbf{x}^0 and \mathbf{z}^0 , the PDHG updates proceed as follows:

$$\begin{aligned}\bar{\mathbf{x}}^k &= \text{prox}_{\tau G}(\mathbf{x}^{k-1} - \tau \mathbf{A}^T \mathbf{z}^{k-1}) \\ \bar{\mathbf{z}}^k &= \text{prox}_{\sigma F^*}(\mathbf{z}^{k-1} + \sigma \mathbf{A}(2\bar{\mathbf{x}}^k - \mathbf{x}^{k-1})) \\ \mathbf{x}^k &= \mathbf{x}^{k-1} + \rho_k(\bar{\mathbf{x}}^k - \mathbf{x}^{k-1}) \\ \mathbf{z}^k &= \mathbf{z}^{k-1} + \rho_k(\bar{\mathbf{z}}^k - \mathbf{z}^{k-1}),\end{aligned}\quad (13)$$

where $\tau > 0$ and $\sigma > 0$ are step sizes satisfying $\sigma\tau \|\mathbf{A}\|^2 \leq 1$, where $\|\mathbf{A}\|$ denotes the spectral norm of \mathbf{A} , and $\rho_k \in (0, 2)$ is a relaxation parameter. In our implementation, we set $\rho_k \equiv 1$ for all k .

The generative model induces a time-dependent denoiser $D_t(\cdot)$, which we use to approximate the proximal operator of the regularization term $G(\cdot)$. We take PDHG method (13) to solve (3), and $\rho_k \equiv 1$ in PDHG. The remained thing is to compute the G -relating proximal operator, F^* -relating proximal operator and determine the stepsize. In the following subsections, we detail the implementation of the proximal operators for G and F^* , as well as the selection of step sizes for different noise models.

G -relating Proximal Operator as Denoising Step

Given initializations $\mathbf{x}_0, \mathbf{z}_0$, for per iteration cycle, PDHG algorithm first computes

$$\mathbf{x}_k = \text{prox}_{\tau G}(\mathbf{x}_{k-1} - \tau \mathbf{A}^T \mathbf{z}_{k-1}). \quad (14)$$

The key is to compute the G -relating proximal operator

$$\begin{aligned}\mathbf{v} &= \mathbf{x}_{k-1} - \tau \mathbf{A}^T \mathbf{z}_{k-1} \\ \mathbf{x}_k &= \underset{\mathbf{z}}{\text{argmin}} \left\{ \tau G(\mathbf{z}) + \frac{1}{2} \|\mathbf{z} - \mathbf{v}\|_2^2 \right\}.\end{aligned}\quad (15)$$

Now it’s time to build the connection between the proximal operator and the denoiser from generative models. For the flow matching model, assume that we want to resolve the noiseless \mathbf{x}_k from \mathbf{v} with a time-dependent noise level related to $\tau = \tau_k = t_k^2$. If we assume $G(\mathbf{z}) \simeq -\log p_{\text{data}}(\mathbf{z})$ and the conditional density $p(\mathbf{v}|\mathbf{z}) \simeq \log p_t(\mathbf{v}|\mathbf{z})$, then the proximal operator acts as a denoiser $D_t(\mathbf{v})$ induced from the flow generative model.

However, the input \mathbf{v} does not necessarily have the noise level assumed by the denoiser $D_t(\cdot)$. To address this mismatch, we follow the reprojection technique widely used in the literature (Zhu et al. 2023; Martin et al. 2024). Assuming \mathbf{v} lies approximately on the data manifold (an idealization that works well in practice), we reproject it to the appropriate noise level by adding Gaussian noise.

In summary, the G -relating proximal step can be implemented as follows:

1. **Variable \mathbf{z} update:** which performed $\mathbf{z} = \mathbf{x}_{k-1} - \tau \mathbf{A}^T \mathbf{z}_{k-1}$
2. **Reprojection step:** which added noise to match the assumed noise level. For flow generative models, the renoising steps are as follows

$$\mathbf{x}_k = (1 - t_k) \cdot \epsilon + t_k \cdot \mathbf{z}, \epsilon \sim \mathcal{N}(0, \mathbf{I})$$

3. **Denoising step:** which utilized the generative model to return a denoised version $\mathbf{x}_k = D_t(\mathbf{x}_k)$.

F^* -relating Proximal Operator for Data Fidelity

The next step of PDHG is related to the data fidelity term F . By Moreau’s identity (16):

$$\text{prox}_{\sigma f^*}(\mathbf{y}) = \mathbf{y} - \sigma \text{prox}_{\sigma^{-1} f}(\sigma^{-1} \mathbf{y}), \quad (16)$$

we are required to compute the proximal operator of function $F(\cdot)$. For Gaussian noise measurement, the squared ℓ_2 norm is the standard choice. For sparse noise, the ℓ_1 norm is more suitable over the ℓ_2 norm-based data fidelity. For Poisson noise, ℓ_2 norm-based loss outperforms the maximum likelihood estimation based loss. For sparse impulse noise, ℓ_1 norm is the standard choice. Each of these fidelity terms yields closed-form proximal operators, which are summarized in Table 1.

Adaptive Stepsize Selection

To guarantee the convergence of PDHG algorithm, the step sizes τ, σ shall satisfy $\tau\sigma \|\mathbf{A}\|^2 \leq 1$. As previously mentioned, the stepsize τ shall configured with the assumed noise level. For flow matching or rectified flow model, one can set $\tau_k = \gamma(t_k)^s$ ($s > 0$). For the two flow models, when $s = 2$ and $\gamma = 1$, it degenerates to the assumptions we used. The dual step size is set as $\sigma_k = 1/\tau_k$. The choice of s affects the final restoration performance, we tuned it in experiments. Note that other options of the stepsize can be tried. Despite possible variations, the above step size schedule demonstrated strong empirical performance across tasks. The algorithm flowchart is detailed in Alg. 1.

Loss $\lambda F(\mathbf{x})$	Proximal operator $\text{prox}_{\lambda F}(\mathbf{x})$
$\frac{\lambda}{2} \ \mathbf{x} - \mathbf{y}\ _2^2$	$\frac{\mathbf{x} + \lambda \mathbf{y}}{1 + \lambda}$
$\lambda \ \mathbf{x} - \mathbf{y}\ _1$	$\mathbf{y} + \text{sign}(\mathbf{x} - \mathbf{y}) \max(\ \mathbf{x} - \mathbf{y}\ - \lambda, 0)$
$\lambda \ \mathbf{x} - \mathbf{y}\ _2$	$\mathbf{y} + \left(1 - \frac{\lambda}{\max\{\ \mathbf{x} - \mathbf{y}\ _2, \lambda\}}\right) (\mathbf{x} - \mathbf{y})$

Table 1: Three common losses and their proximal operators.

		(a) Poisson								(b) salt-and-pepper							
Data	Method	Denoising		Deblur		SR		Box Inp		Denoising		Deblur		SR		Box Inp	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
FFHQ	Flow-Priors	25.07	0.654	18.44	0.369	20.89	0.442	20.99	0.461	26.83	0.677	18.54	0.334	22.17	0.621	22.86	0.638
	OT-ODE	23.66	0.596	22.73	0.567	21.64	0.518	22.48	0.566	22.32	0.478	22.79	0.549	18.98	0.372	22.79	0.575
	PnP-HQS	23.90	0.729	23.25	0.703	18.76	0.578	22.97	0.709	23.95	0.633	22.66	0.591	23.31	0.696	22.34	0.548
	PnP-FBS	29.72	0.901	28.45	0.855	23.51	0.590	27.25	0.877	26.27	0.785	25.57	0.786	21.92	0.693	24.90	0.749
	Ours	33.98	0.902	32.87	0.927	29.96	0.836	29.48	0.885	40.27	0.974	37.67	0.974	33.12	0.948	30.57	0.948
AFHQ-Cat	Flow-Priors	24.65	0.670	20.48	0.482	18.80	0.399	20.01	0.390	26.98	0.719	21.37	0.545	14.14	0.336	22.79	0.551
	OT-ODE	23.77	0.540	21.31	0.422	18.65	0.353	20.83	0.496	22.79	0.457	22.09	0.453	20.06	0.389	21.30	0.535
	PnP-HQS	24.57	0.668	23.03	0.608	17.02	0.470	22.81	0.645	25.66	0.693	21.86	0.595	18.86	0.521	22.46	0.565
	PnP-FBS	29.65	0.854	25.57	0.695	23.05	0.558	25.42	0.815	27.21	0.791	24.14	0.648	19.24	0.530	24.44	0.744
	Ours	34.82	0.935	26.28	0.709	26.26	0.747	26.25	0.889	37.94	0.962	28.05	0.791	24.70	0.730	25.43	0.908

Table 2: Quantitative results of different methods for four tasks with Poisson/salt-and-pepper noise.

Algorithm 1: PHGD method with PnP flow generative prior.

Require: Iterations T , Loss F , Denoiser induced by generative model $D_k(\cdot)$, stepsize $\eta_k = \eta\tau_k$.

Ensure: Estimated image \mathbf{x}_0 .

1: Set $\mathbf{x}_0 \sim \mathcal{N}(0, \mathbf{I})$ and $\mathbf{z}_0 = \mathbf{0}$

2: **for** $k = 1 : 1 : T$ **do**

3: \mathbf{z} update: $\mathbf{z} = \mathbf{x}_{k-1} - \eta\tau_{k-1}\mathbf{A}^T\mathbf{z}_{k-1}$

4: Reprojection step:

$$\mathbf{x}_k = (1 - t_k) \cdot \boldsymbol{\epsilon} + t_k \cdot \mathbf{z}, \boldsymbol{\epsilon} \sim \mathcal{N}(0, \mathbf{I})$$

5: Denoising step: $\mathbf{x}_k = D_k(\mathbf{x}_k)$

6: F^* -relating step: $\mathbf{z}_k = \text{prox}_{\tau_{k-1}^{-1}F^*}(\mathbf{z}_{k-1} + \tau_{k-1}^{-1}\mathbf{A}(2\mathbf{x}_k - \mathbf{x}_{k-1}))$.

7: **end for**

Extension to Score-Based Diffusion Model

Although diffusion models are typically formulated using stochastic differential equations (SDEs), they are mathematically equivalent to ordinary differential equations (ODEs) governing the sampling trajectories. This connection enables our proposed PDHG-based plug-and-play (PnP) method to be applied to diffusion models as well, utilizing the denoiser implicitly defined by the diffusion process.

For diffusion model, the mean of $p(\mathbf{x}_0|\mathbf{x}_t)$ is given by the Tweedie’s formula:

$$\mathbb{E}[\mathbf{x}_0|\mathbf{x}_t] = \frac{\mathbf{x}_t + \sigma_t^2 \nabla_{\mathbf{x}_t} \log p(\mathbf{x}_t)}{\alpha_t}, \quad (17)$$

where the score function $\nabla_{\mathbf{x}_t} \log p(\mathbf{x}_t)$ is replaced by the learned generative model in practice. Note that, the mean estimation of $p(\mathbf{x}_0|\mathbf{x}_t)$ just defines a denoiser $D_t(\mathbf{x}_t)$. For variance persevering configuration, *i.e.*, DDPM, we learned the scaled score function, hence the denoiser is given by

$$D_t(\mathbf{x}_t) = \frac{\mathbf{x}_t - \sigma_t \boldsymbol{\epsilon}_\theta(\mathbf{x}_t, t)}{\alpha_t}. \quad (18)$$

For variance exploding model, the corresponding denoiser can be achieved as well.

Numerical Experiments

Our proposed PnP method is model-agnostic and compatible with any ODE-based generative model, including the flow matching model (Lipman et al. 2023), diffusion model (Ho, Jain, and Abbeel 2020), and rectified flow model (Liu, Gong et al. 2023). We have evaluated the algorithm across several applications against these three generative models. For page limitation, we provide the results for flow matching model in the following content. We considered Poisson noise and salt-and-pepper noise in our experiments, as well as Gaussian noise as a baseline for comparison.

Baselines We benchmark our method against four PnP approaches with flow matching imaging prior, including OT-ODE¹, Flow priors, PnP-FBS and PnP-HQS.

Benchmark datasets and publicly available models For easy and fair comparison, we consider the datasets where there are publicly available models. For flow matching, we have the models for CelebA dataset with size 128×128 , and AFHQ-Cat, a subset of the Animal FacesHQ dataset focus on the cat class, with image size 256×256 . All images are normalized to the range $[-1, 1]$. For CelebA, we use the standard training, validation and test splits. For AFHQ-Cat, we follow (Martin et al. 2024) to select random 32 images to create a validation set. For the two datasets, the pretrained model (Martin et al. 2024) is trained using the mini batch OT Flow Matching approach (Tong et al.) for efficiency.

Evaluated imaging problems The considered four image restoration problems are as follows. (a) Denoising; (b) deblurring using a 61×61 Gaussian kernel; (c) super-resolution with average pool downsampling operation ($2 \times$ downsampling for image size 128×128 and $4 \times$ downsampling for image size 256×256); (d) box-inpainting with a centered $s \times s$ mask ($s = 40$ for image size 128×128 and $s = 80$ downsampling for image size 256×256).

Noise types We simulate Poisson noise, salt-and-pepper noise and Gaussian noise for the clean observations.

1. Poisson noise: which is simulated using $\mathbf{y} = \mathbf{z}/\alpha$, $\mathbf{z} \sim$

¹With the unified framework, it is the flow matching based counterpart of diffusion posterior sampling method.

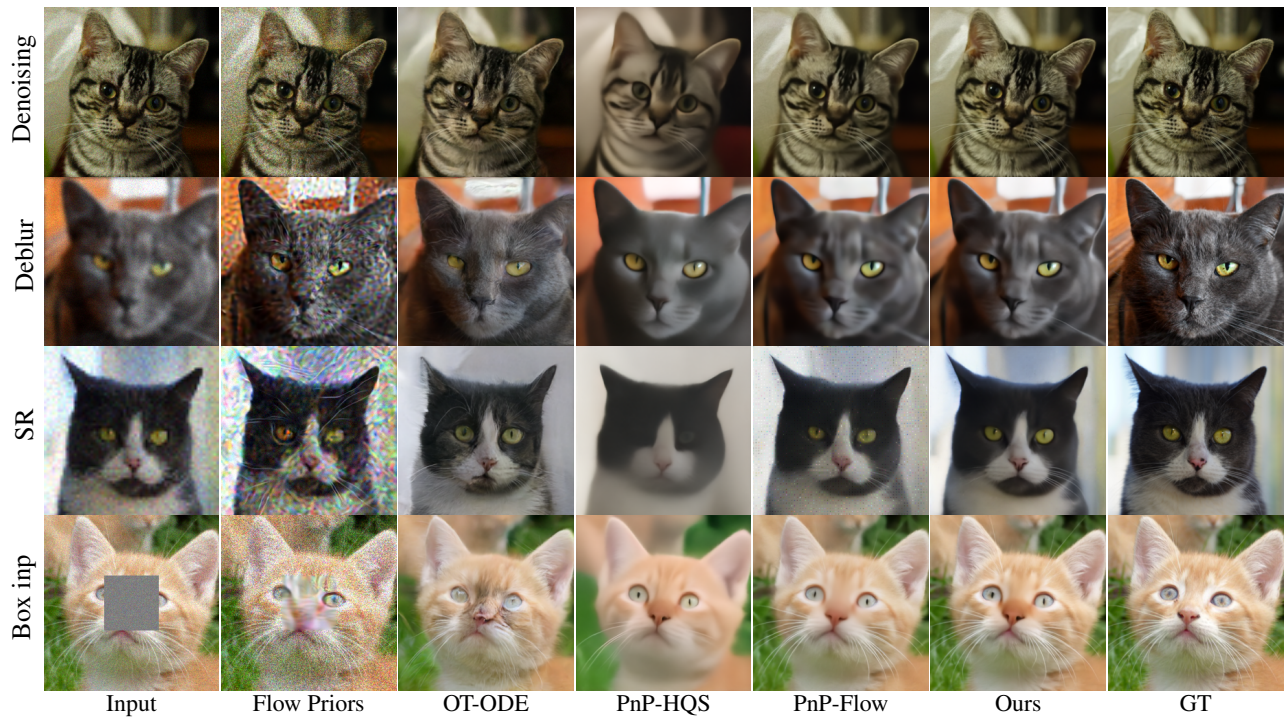


Figure 1: Visualization of different methods on four tasks with Poisson noise for AFHQ-Cat.

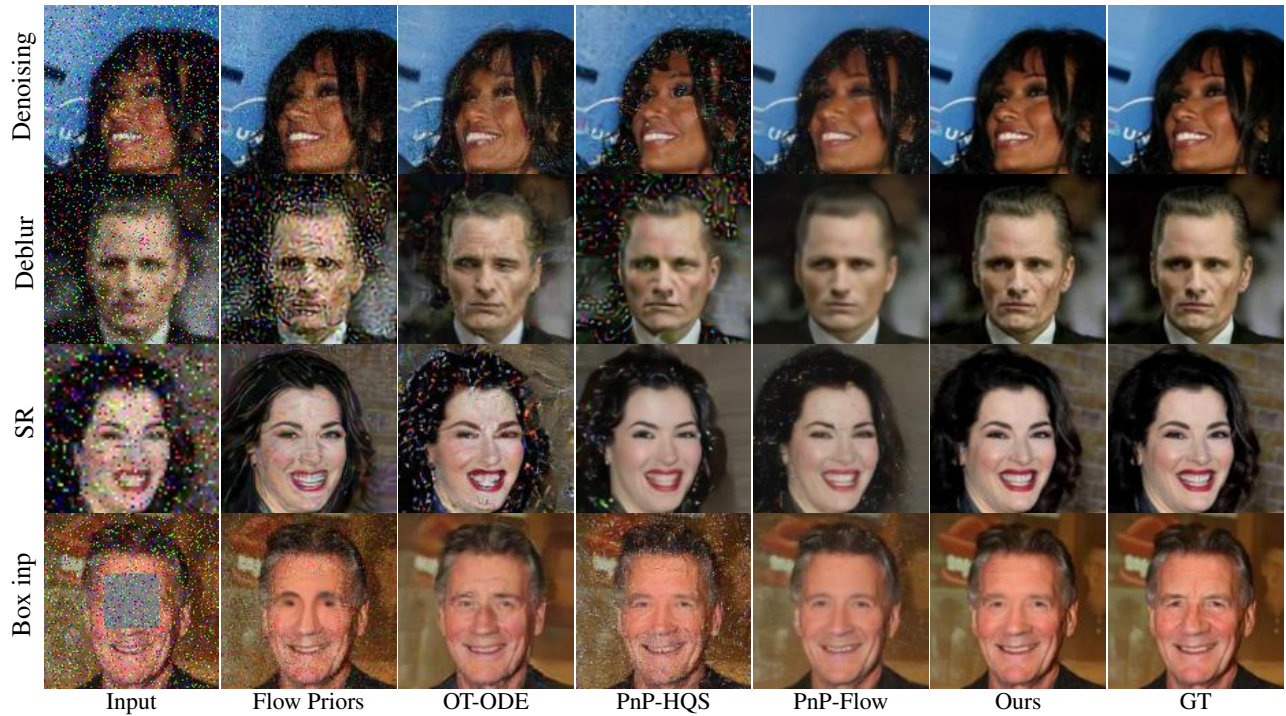


Figure 2: Visualization of different methods on four tasks with impulse noise for CelebA.

Poisson($\mathbf{A}\mathbf{x}_0 \cdot \alpha$), where α denotes the noise level, we consider $\alpha = 1$ in experiment. We set $F = \lambda \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2$ ($\lambda = 200$) for better performance over ℓ_1 norm.

2. Salt-and-pepper noise: which is a sparse noise in digital images. The noise consists of 10% random pixels in image being set to black and white. The data fidelity term is $F = \lambda \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_1$ ($\lambda = 25$).
3. Gaussian noise: Noise level $\sigma = 0.2$ for denoising, $\sigma = 0.05$ for other tasks (except $\sigma = 0.01$ for random inpainting). The data fidelity term is $F = \frac{1}{2\sigma_y^2} \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2^2$, where σ_y is the noise level.

CelebA Method	Denoising		Deblur		SR		Box Inp	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Flow-Priors	29.15	0.771	31.17	0.858	28.17	0.721	29.03	0.861
OT-ODE	26.90	0.681	28.48	0.734	27.70	0.743	26.04	0.728
PnP-HQS	31.91	0.912	20.55	0.617	22.68	0.692	29.84	0.944
PnP-FBS	31.95	0.910	34.18	0.940	30.90	0.905	29.72	0.938
Ours	31.40	0.904	33.90	0.935	30.37	0.898	29.20	0.931

AFHQ-Cat Method	Denoising		Deblur		SR		Box Inp	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Flow-Priors	29.86	0.779	24.51	0.682	22.29	0.499	26.48	0.820
OT-ODE	26.75	0.635	24.71	0.552	23.88	0.568	22.81	0.677
PnP-HQS	31.31	0.880	21.66	0.573	19.18	0.518	26.67	0.917
PnP-FBS	31.03	0.865	27.25	0.746	22.59	0.630	26.83	0.904
Ours	30.73	0.861	27.09	0.742	22.51	0.631	26.31	0.897

Table 3: Quantitative results of different methods for different tasks on the two datasets with Gaussian noise.

Stepsize setting and iterations In all experiments, our method sets a time-dependent stepsize τ_k and $\sigma_k = 1/\tau_k$. For flow model, the stepsize τ_k is of the form $(1 - t_k)^\alpha$ with α tuned for different tasks and datasets. In the case of Poisson and salt-and-pepper noise: we set $\alpha = 0.8$ for all tasks. In the case of Gaussian noise: $\alpha = 0.8$ for denoising; $\alpha = 0.01$ for deblurring; $\alpha = 0.3$ for super-resolution and $\alpha = 0.5$ for box-inpainting. We set $T = 100$. All baseline hyperparameters are tuned following (Martin et al. 2024).

Quantitative results For quantitative comparison, we provide the reference-based PSNR and SSIM metrics. Tables 2 (a) and (b) report PSNR and SSIM on CelebA and AFHQ-Cat for Poisson and impulse noise. For both datasets, our method beats other methods across all the four tasks and the two considered datasets by a large margin. Excluding ours, other methods relies on the squared ℓ_2 norm loss and we tried our best to tune the compared methods. The different performances of PnP-HQS and PnP-FBS are interesting, since they only differ from whether we solve the proximal problem exactly or just perform a gradient step. PnP-FBS is more robust to the mismatch between the noise type and the consistency loss. See Figures 1 and 2 for the visualization of the different methods on the two datasets.

Table 3 report the quantitative results for Gaussian noise. It can see that our method is comparable to PnP-FBS method

and PnP-HQS method for the Gaussian noise scenario. The comparable result is reasonable since all the methods solve the same regularized optimization problem. Hence, the significant improvement is attributed to the exchange of the loss for non-Gaussian noise. **Impact of ℓ_1 vs. ℓ_2 losses for non-**

AFHQ Noise	loss	Denoising		Deblur		SR		Box Inp	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Poisson	ℓ_2	34.82	0.935	26.28	0.709	26.26	0.747	26.25	0.889
	ℓ_1	29.31	0.740	25.70	0.679	24.74	0.717	23.61	0.693
Impulse	ℓ_2	25.81	0.743	24.12	0.627	20.58	0.491	23.08	0.598
	ℓ_1	37.94	0.962	28.05	0.791	24.70	0.730	25.43	0.908

Table 4: Comparison of ℓ_1/ℓ_2 losses for AFHQ-Cat.

Gaussian noise To assess the impact of loss function choice on restoration performance under different noise types, we conduct an ablation study using our method with either ℓ_1 or ℓ_2 data fidelity terms. We report results on the AFHQ-Cat dataset under Poisson and impulse noise. Table 4 presents the quantitative results. For impulse noise, the ℓ_1 loss significantly outperforms the ℓ_2 loss, consistent with the known robustness of ℓ_1 to sparse noise. Conversely, under Poisson noise, the ℓ_2 loss yields better performance than ℓ_1 . These results validate our design choice of using loss functions tailored to noise type and highlight the importance of appropriate fidelity terms in inverse problems.

Efficiency comparison See Table 5 for the efficiency comparison of these methods. Our method is efficient as the existing methods. Flow-priors is slow due to its sequential approximation. PnP-FBS uses 5 samples to implement the denoising step, while ours use only one sample.

Method	Flow-Priors	OT-ODE	PnP-HQS	PnP-FBS	Ours
Time	104	18	12	25	13

Table 5: Computation time (in s) to deblur one image.

Conclusion

The flow generative model can be integrated into the traditional regularized optimization to address the ill-posedness of imaging inverse problems. Existing plug-and-play (PnP) methods established an iterative proximal splitting based method for smooth squared ℓ_2 norm data fidelity term induced from Gaussian noise. However, in many practical applications involving non-Gaussian noise, such as Poisson noise and salt-and-pepper impulse noise, the squared ℓ_2 norm is sub-optimal, and the nonsmooth ℓ_1, ℓ_2 -norm based data fidelities are favorable with better performance. To close the gap, we propose integrating the flow generative model into the traditional primal dual hybrid gradient (PDHG) method to exploit the proximity-friendly structure of the ℓ_1, ℓ_2 -norm-based losses, and the proximal operator corresponding to the regularization term is implicitly implemented via the denoiser induced by the generative models.

Acknowledgments

Ji Li was supported by the National Natural Science Foundation of China (Grant No. 12571472) and supported by the Open Project of Key Laboratory of Mathematics and Information Networks (Beijing University of Posts and Telecommunications), Ministry of Education, China, under Grant No. KF202401). Chao Wang was supported by the National Natural Science Foundation of China (Grant No. 12526572).

References

- Ben-Hamu, H.; Puny, O.; Gat, I.; Karrer, B.; Singer, U.; and Lipman, Y. 2024. D-Flow: Differentiating through Flows for Controlled Generation. In *International Conference on Machine Learning*. PMLR.
- Chambolle, A.; and Pock, T. 2011. A first-order primal-dual algorithm for convex problems with applications to imaging. *Journal of Mathematical Imaging and Vision*, 40: 120–145.
- Chen, R. T.; Rubanova, Y.; Bettencourt, J.; and Duvenaud, D. K. 2018. Neural ordinary differential equations. In *Advances in Neural Information Processing Systems*, volume 31.
- Chung, H.; Kim, J.; Mccann, M. T.; Klasky, M. L.; and Ye, J. C. 2022a. Diffusion Posterior Sampling for General Noisy Inverse Problems. In *The Eleventh International Conference on Learning Representations*.
- Chung, H.; Sim, B.; Ryu, D.; and Ye, J. C. 2022b. Improving diffusion models for inverse problems using manifold constraints. volume 35, 25683–25696.
- Dabov, K.; Foi, A.; Katkovnik, V.; and Egiazarian, K. 2006. Image denoising with block-matching and 3D filtering. In *Image processing: algorithms and systems, neural networks, and machine learning*, volume 6064, 354–365. SPIE.
- Dabov, K.; Foi, A.; Katkovnik, V.; and Egiazarian, K. 2007. Image denoising by sparse 3-D transform-domain collaborative filtering. *IEEE Transactions on Image Processing*, 16(8): 2080–2095.
- Esser, E.; Zhang, X.; and Chan, T. F. 2010. A general framework for a class of first order primal-dual algorithms for convex optimization in imaging science. *SIAM Journal on Imaging Sciences*, 3(4): 1015–1046.
- Graikos, A.; Malkin, N.; Jovic, N.; and Samaras, D. 2022. Diffusion models as plug-and-play priors. volume 35, 14715–14728.
- Han, S.; Xu, Y.; Wang, D.; Morovati, B.; Zhou, L.; Maltz, J. S.; Wang, G.; and Yu, H. 2024. Physics-informed Score-based Diffusion Model for Limited-angle Reconstruction of Cardiac Computed Tomography. *IEEE Transactions on Medical Imaging*, 1–1.
- Ho, J.; Jain, A.; and Abbeel, P. 2020. Denoising diffusion probabilistic models. volume 33, 6840–6851.
- Hurault, S.; Leclaire, A.; and Papadakis, N. 2022. Proximal denoiser for convergent plug-and-play optimization with nonconvex regularization. In *International Conference on Machine Learning*, 9483–9505. PMLR.
- Kamilov, U. S.; Mansour, H.; and Wohlberg, B. 2017. A plug-and-play priors approach for solving nonlinear imaging inverse problems. *IEEE Signal Processing Letters*, 24(12): 1872–1876.
- Kawar, B.; Elad, M.; Ermon, S.; and Song, J. 2022. Denoising Diffusion Restoration Models. In *Advances in Neural Information Processing Systems*.
- Kingma, D. P.; and Dhariwal, P. 2018. Glow: Generative flow with invertible 1x1 convolutions. volume 31.
- Li, J.; and Wang, C. 2024. DiffFPR: Diffusion prior for over-sampled Fourier phase retrieval. In *Forty-first International Conference on Machine Learning*.
- Lipman, Y.; Chen, R. T.; Ben-Hamu, H.; Nickel, M.; and Le, M. 2023. Flow Matching for Generative Modeling. In *The Eleventh International Conference on Learning Representations*.
- Liu, X.; Gong, C.; et al. 2023. Flow Straight and Fast: Learning to Generate and Transfer Data with Rectified Flow. In *The Eleventh International Conference on Learning Representations*.
- Martin, S.; Gagneux, A.; Hagemann, P.; and Steidl, G. 2024. PnP-Flow: Plug-and-Play Image Restoration with Flow Matching. *arXiv preprint arXiv:2410.02423*.
- O’Connor, D.; and Vandenberghe, L. 2020. On the equivalence of the primal-dual hybrid gradient method and Douglas–Rachford splitting. *Mathematical Programming*, 179(1): 85–108.
- Parikh, N.; Boyd, S.; et al. 2014. Proximal algorithms. *Foundations and trends® in Optimization*, 1(3): 127–239.
- Pokle, A.; Muckley, M. J.; Chen, R. T.; and Karrer, B. 2024. Training-free linear image inverses via flows. *Transactions on Machine Learning Research*.
- Ryu, E.; Liu, J.; Wang, S.; Chen, X.; Wang, Z.; and Yin, W. 2019. Plug-and-play methods provably converge with properly trained denoisers. In *International Conference on Machine Learning*, 5546–5557. PMLR.
- Song, J.; Vahdat, A.; Mardani, M.; and Kautz, J. 2022. Pseudoinverse-guided diffusion models for inverse problems. In *International Conference on Learning Representations*.
- Song, J.; Zhang, Q.; Yin, H.; Mardani, M.; Liu, M.-Y.; Kautz, J.; Chen, Y.; and Vahdat, A. 2023. Loss-guided diffusion models for plug-and-play controllable generation. In *International Conference on Machine Learning*, 32483–32498. PMLR.
- Song, Y.; Sohl-Dickstein, J.; Kingma, D. P.; Kumar, A.; Ermon, S.; and Poole, B. 2020. Score-Based Generative Modeling through Stochastic Differential Equations. In *International Conference on Learning Representations*.
- Tong, A.; FATRAS, K.; Malkin, N.; Huguet, G.; Zhang, Y.; Rector-Brooks, J.; Wolf, G.; and Bengio, Y. 2024. Improving and generalizing flow-based generative models with minibatch optimal transport. *Transactions on Machine Learning Research*.
- Wang, Y.; Yu, J.; and Zhang, J. 2022. Zero-Shot Image Restoration Using Denoising Diffusion Null-Space Model.

In *The Eleventh International Conference on Learning Representations*.

Whang, J.; Delbracio, M.; Talebi, H.; Saharia, C.; Dimakis, A. G.; and Milanfar, P. 2022. Deblurring via stochastic refinement. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 16293–16303.

Zhang, K.; Li, Y.; Zuo, W.; Zhang, L.; Van Gool, L.; and Timofte, R. 2020. Plug-and-play image restoration with deep denoiser prior. *arXiv preprint arXiv:2008.13751*.

Zhang, K.; Zuo, W.; Chen, Y.; Meng, D.; and Zhang, L. 2017. Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising. *IEEE transactions on image processing*, 26(7): 3142–3155.

Zhang, Y.; Yu, P.; Zhu, Y.; Chang, Y.; Gao, F.; Wu, Y. N.; and Leong, O. 2024. Flow Priors for Linear Inverse Problems via Iterative Corrupted Trajectory Matching. *arXiv preprint arXiv:2405.18816*.

Zhu, Y.; Zhang, K.; Liang, J.; Cao, J.; Wen, B.; Timofte, R.; and Van Gool, L. 2023. Denoising Diffusion Models for Plug-and-Play Image Restoration. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, 1219–1229.