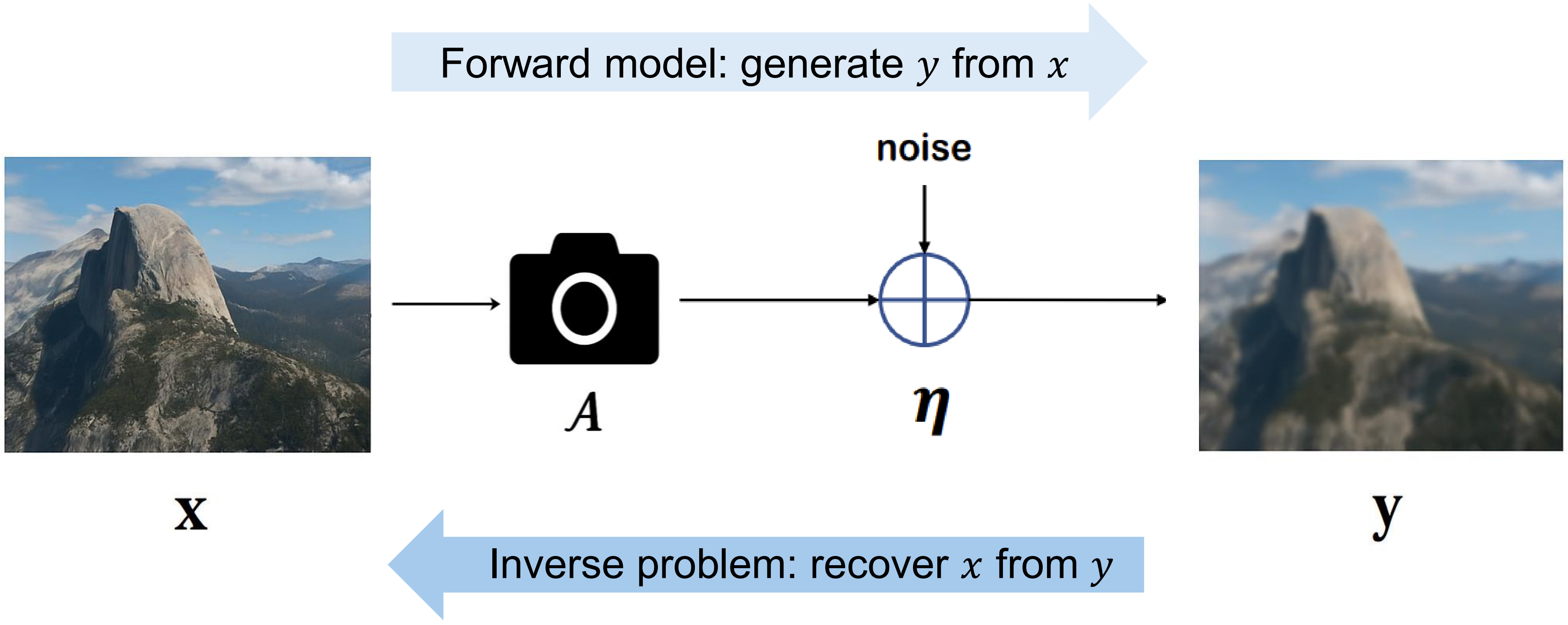
Plug-and-play Diffusion Models for Image Compressive Sensing with Data Consistency Projection

Xiaodong Wang*, Ping Wang*, Zhangyuan Li, Xin Yuan‡

Westlake University & Zhejiang University

@ Sensing and Computational Imaging (SCI) Lab

Inverse problems



Most imaging systems can be formulated as a forward model

- Acquiring the image x is an inverse problem
- Solving an inverse problem is an ill-posed **one-to-many mapping** problem

Inverse problems

Imaging modality	Forward model	Notes
Denoising	$y = Ix$	I is the identity matrix.
Deblur	$y = h * x = Hx$	h is a known blur kernel and $*$ denotes convolution (H is a discrete convolution operator). When h is unknown the reconstruction problem is known as blind deconvolution. [1]
Superresolution	$y = SBx$	S is a subsampling operator (identity matrix with missing rows) and B is a blurring operator cooresponding to convolution with a blur kernel. [2]
Inpainting	$y = Sx$	S is a diagonal matrix where $S_{i,i} = 1$ for the pixels that are sampled and $S_{i,i} = 0$ for the pixels that are not. [3]
Magnetic resonance imaging (MRI)	$y = SFMx$	S is a subsampling operator (identity matrix with missing rows), F is the discrete Fourier transform matrix, and M is a diagonal matrix representing a spatial domain multiplication with the coil sensitivity map (assuming a single coil aquisition with Cartesian sampling in a SENSE framework). [4]
Computed tomography (CT)	$y = Rx$	R is the discrete Radon transform. [5]
Snapshot compressive imaging (SCI)	$y = Mx$	M is the sensing matrix related to the 3D mask. [6]
Single pixel imaging (SPI)	$y = Mx$	M is the sensing matrix related to the 3D mask. [7]
Non-line-of-sight Imaging (NLOS)	$y = R_t^{-1} H R_z x$	The matrix H represents the shift-invariant 3D convolution operation, and the matrices R_t and R_z represent the transformation operations applied to the temporal and spatial dimensions, respectively. [8]
Structured illumination microscopy (SIM)	$y_i = SHM_i x_i$	S is a decimation operator with a downsampling factor of two in each dimension, this is required because SIM aims at doubling the lateral resolution. M_i is a diagonal matrix associated to the ith illumination patterns. H is a discrete convolution operator. [9]
Optical diffraction tomography (ODT)	$y_i = M_i F x$	F is the discrete Fourier transform matrix, M_i is a diagonal matrix associated to the ith illumination patterns. [10]
Phase retrival (PR)	$y = \|Ax\|^2$	$\|\cdot\|$ denotes the absolute value, the square is taken elementwise, and A is a (potentially complex valued) measurement matrix that depends on the application. The measurement

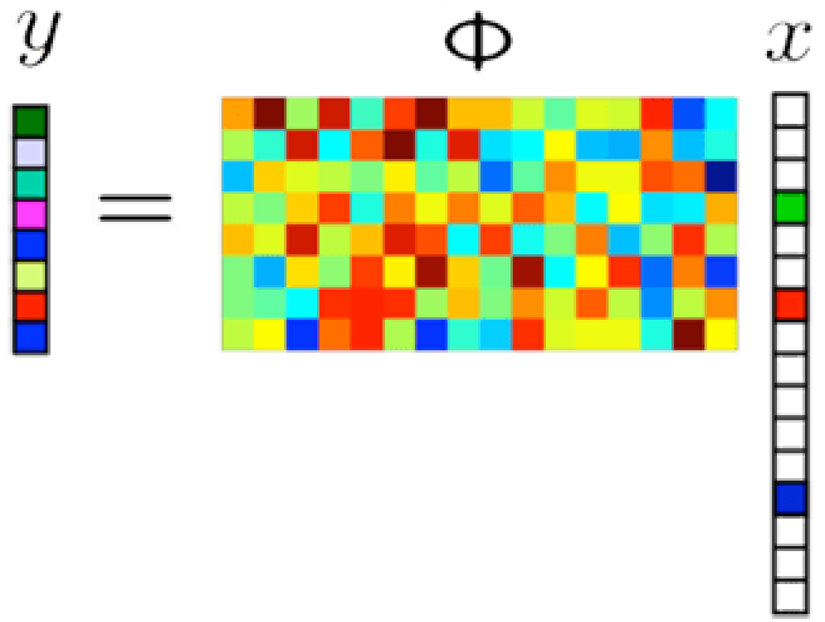


Image Compressive Sensing

Model-based method for inverse problem

Probabilistic formulation of an inverse problem

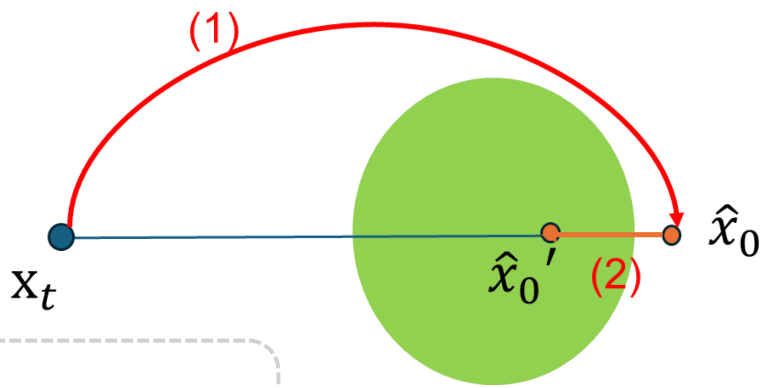
$$y = Ax + \eta \qquad x \sim \mathcal{P}_{\mathrm{x}} \qquad \eta \sim \mathcal{N}(0, \sigma^2 \mathrm{I})$$

Maximum a posteriori probability (MAP) estimator

$$\hat{x} = \underset{x \in \mathcal{R}^n}{\arg\min} \frac{1}{2\sigma^2} \|y - Ax\|_2^2 + h(x), \qquad h(x) = -\log(\mathcal{P}_x(x))$$

Proximal operator is an optimization-based image denoiser

$$\mathrm{prox}_\sigma(z) = \underset{x \in \mathcal{R}^n}{\arg\min} \frac{1}{2\sigma^2} \|z - x\|_2^2 + h(x)$$



$$x^t \leftarrow \mathrm{prox}_\sigma(z^t) \qquad z^t \leftarrow x^t - \gamma\, W\, \nabla_{x^t} \|y - Ax^t\|_2^2$$

Plug-and-play priors (PnP) use pretrained MMSE denoiser as image prior (or proximal operator)

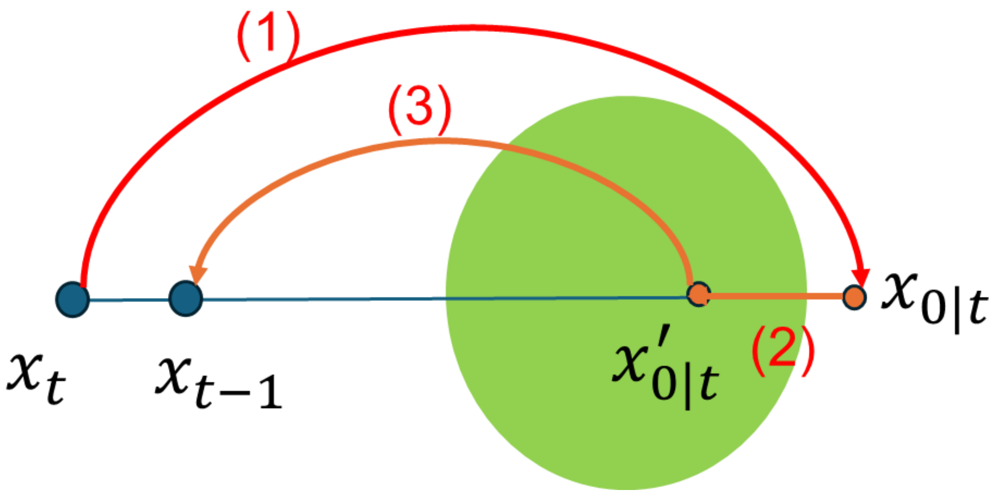
Diffusion model for inverse problem

Forward trajectory of DDIM

$$\mathbf{x}_t = \sqrt{\alpha_t}\mathbf{x}_0 + \sqrt{1-\alpha_t}\,\mathbf{z}, \quad \mathbf{z} \sim \mathcal{N}(0, \mathbf{I}),$$

Backward trajectory of DDIM

$$\mathbf{x}_{t-1} = \sqrt{\alpha_{t-1}}\left(\frac{\mathbf{x}_t + (1-\alpha_t)\nabla_{\mathbf{x}_t}\log p_t(\mathbf{x}_t)}{\sqrt{\alpha_t}}\right) + \sqrt{1-\alpha_{t-1}-\sigma_t^2}\left(-\sqrt{1-\alpha_t}\nabla_{\mathbf{x}_t}\log p_t(\mathbf{x}_t)\right)$$



Measurement-guidance DDIM

$$\nabla_{\mathbf{x}}\log p_t(\mathbf{x}_t \mid \mathbf{y}) = \nabla_{\mathbf{x}_t}\log p_t(\mathbf{x}_t) + \nabla_{\mathbf{x}_t}\log p(\mathbf{y} \mid \mathbf{x}_t),$$

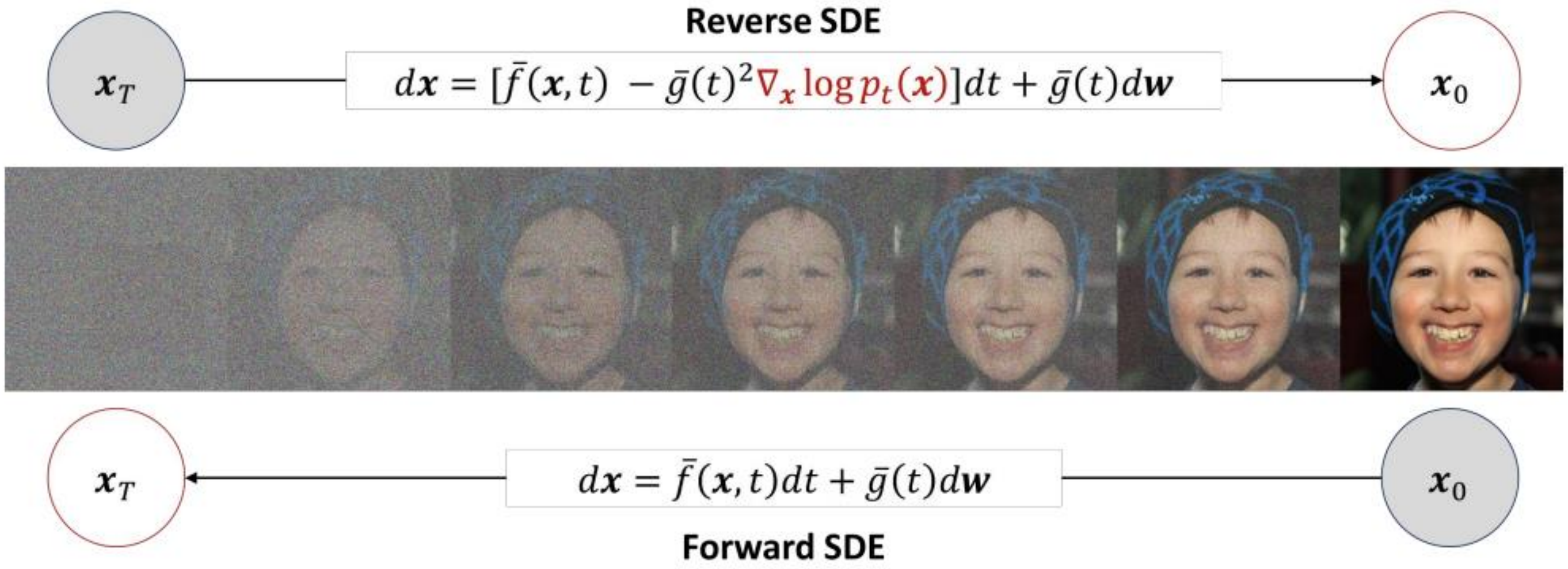
$$\mathbf{x}_{0|t} = \frac{\mathbf{x}_t + (1-\alpha_t)\nabla_{\mathbf{x}_t}\log p_t(\mathbf{x}_t)}{\sqrt{\alpha_t}}, \tag{14a}$$

$$\mathbf{x}'_{0|t} = \mathbf{x}_{0|t} + \mu_t\nabla_{\mathbf{x}_t}\log p(\mathbf{y} \mid \mathbf{x}_t), \tag{14b}$$

$$\mathbf{x}_{t-1} = \sqrt{\alpha_{t-1}}\,\mathbf{x}'_{0|t} - \sqrt{1-\alpha_{t-1}-\sigma_t^2}\left(\sqrt{1-\alpha_t}\nabla_{\mathbf{x}_t}\log p_t(\mathbf{x}_t)\right), \tag{14c}$$

Daras G, Chung H, Lai C H, et al. A survey on diffusion models for inverse problems. 2024[J]. Arxiv

Diffusion model for inverse problem



$$s_{\theta^*}(\mathbf{x}) \approx \nabla_{\mathbf{x}_t} \log p(\mathbf{x}_t)$$

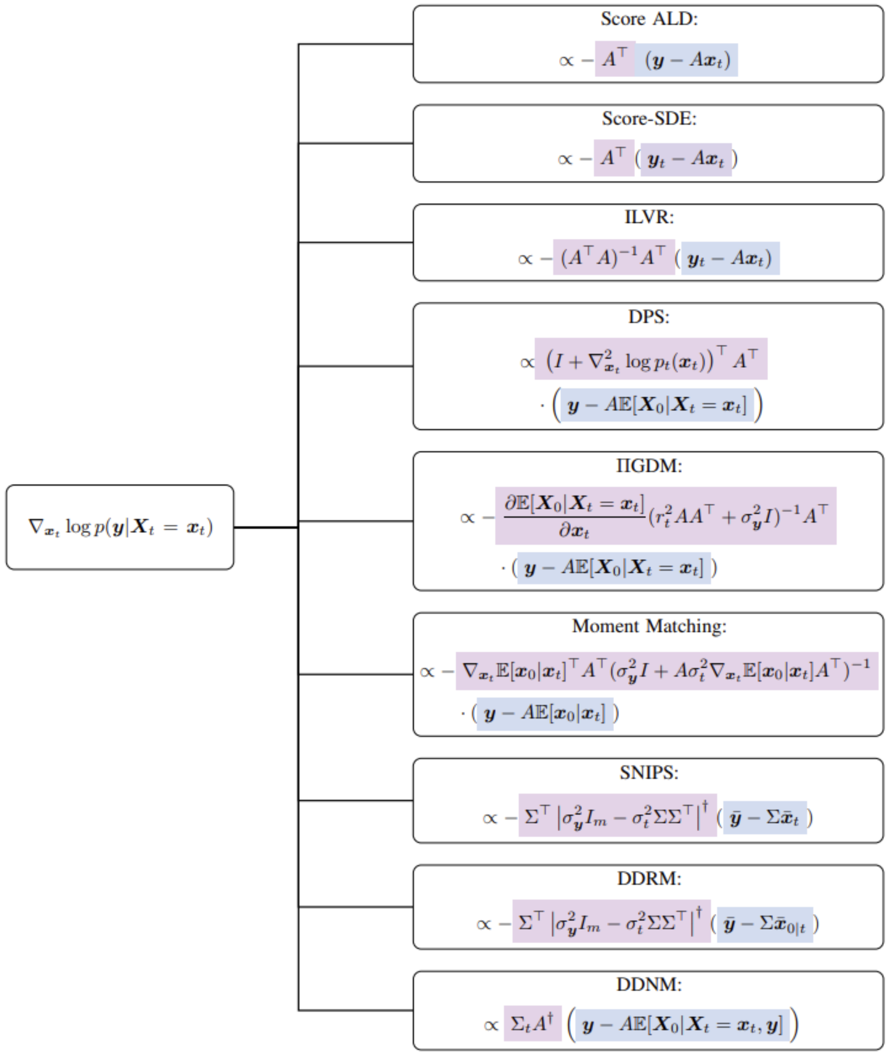
Denoiser

Generative diffusion prior

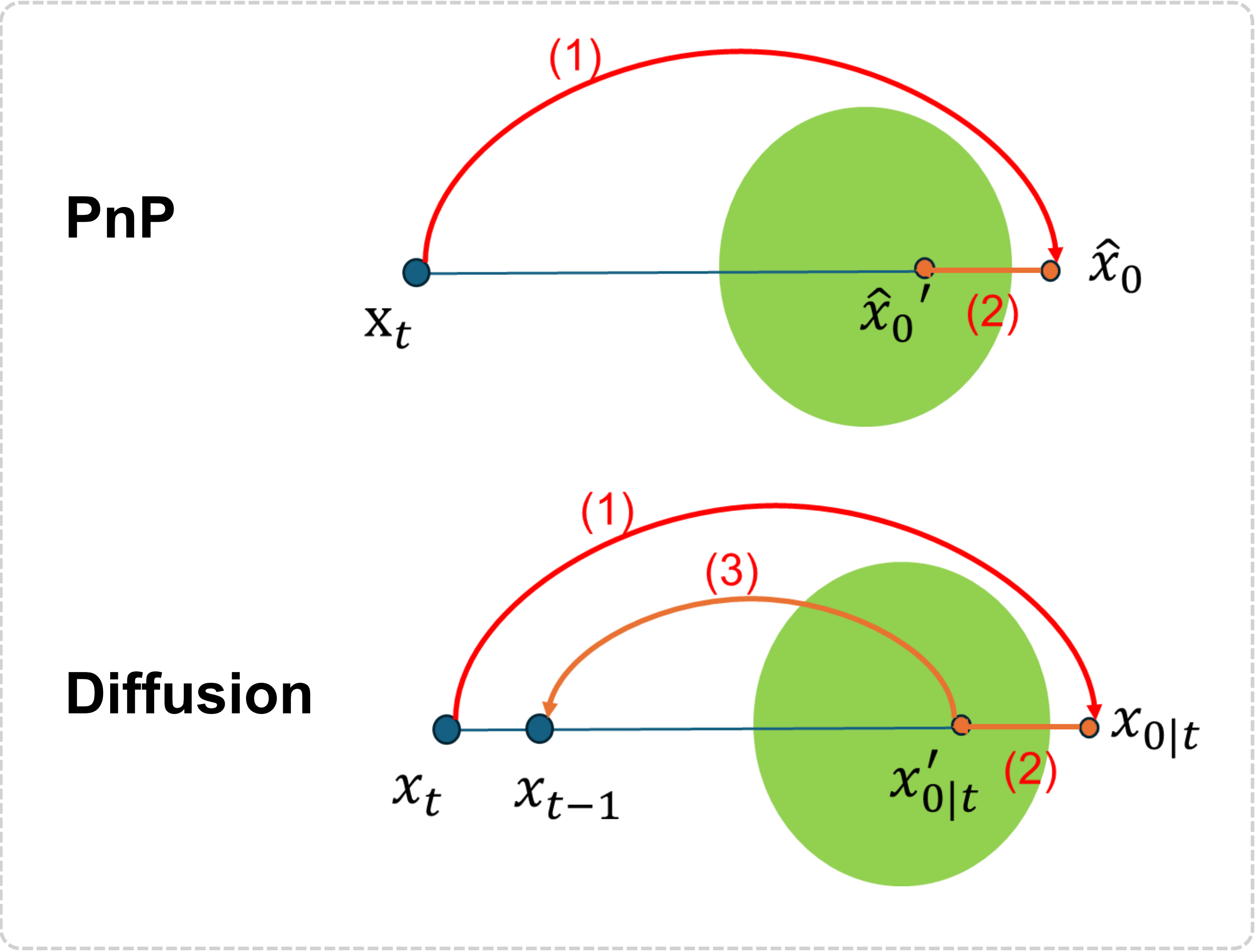
$$\nabla_{\mathbf{x}_t} \log p(\mathbf{x}_t|\mathbf{y}) = \underbrace{\nabla_{\mathbf{x}_t} \log p(\mathbf{y}|\mathbf{x}_t)}_{\text{Measurement process}} + \underbrace{\nabla_{\mathbf{x}_t} \log p(\mathbf{x}_t)}_{\text{Denoiser}}$$

Song, Yang, et al. "Score-based generative modeling through stochastic differential equations." ICLR 2021

Diffusion model for inverse problem



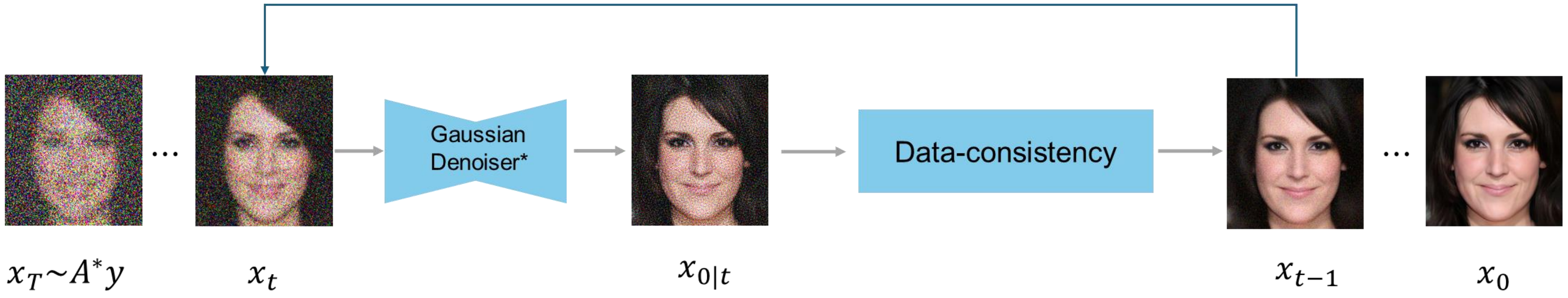
Daras, Giannis, et al. "A survey on diffusion models for inverse problems." arXiv 2024.



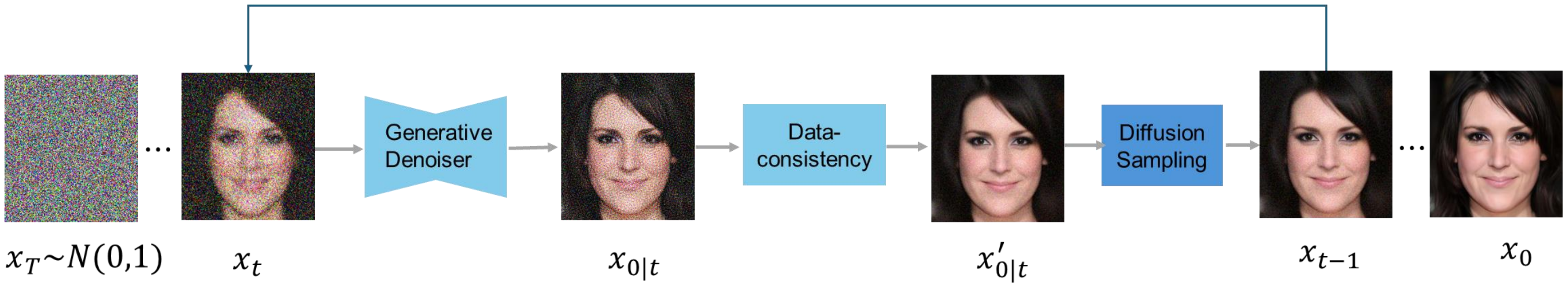
Connection of diffusion and PnP

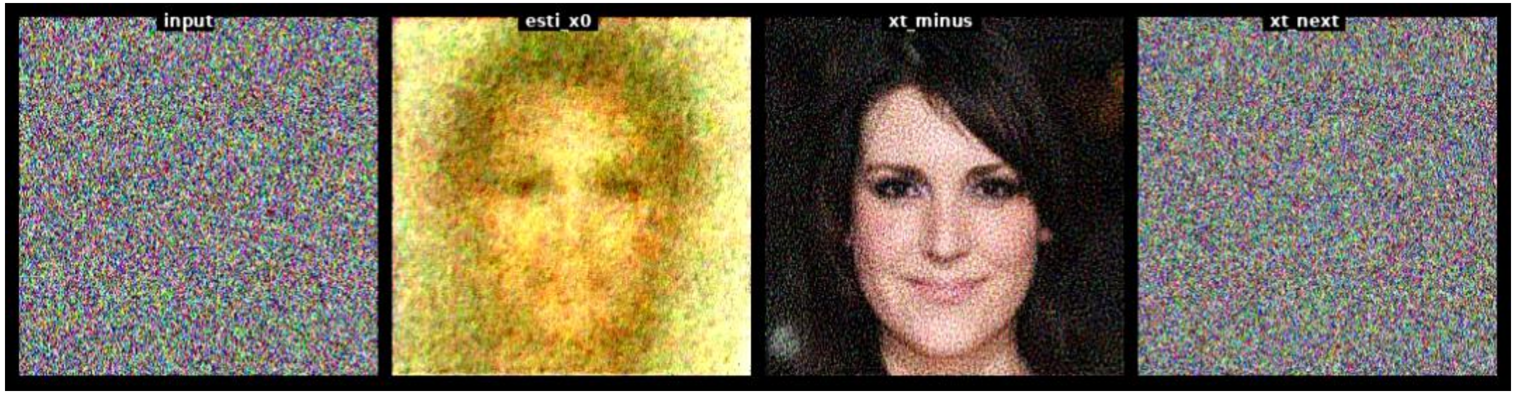
Connection between diffusion model and PnP

(a) Plug-and-play



(b) Diffusion model





Plug-and-play diffusion model for inverse problem

$$x^t \leftarrow \mathrm{prox}_\sigma(z^t)$$

$$z^t \leftarrow x^t - \gamma\, W\, \nabla_{x^t} \|y - Ax^t\|_2^2$$

Solution 1: Half-Quadratic Splitting (HQS)

$$\mathbf{x}'_{0|t} = (\mathbf{H}^T\mathbf{H} + \lambda\mathbf{I})^{-1}(\mathbf{H}^T\mathbf{y} + \lambda\mathbf{x}_{0|t}),$$

Solution 2: Generalized Alternating Projection (GAP)

$$\mathbf{x}'_{0|t} = \mathbf{x}_{0|t} + \mathbf{H}^\dagger(\mathbf{y} - \mathbf{H}\mathbf{x}_{0|t}),$$

Our solution: Fused update of GAP and HQS

$$\mathbf{x}'_{0|t} = (1-\delta_t)(\mathbf{x}_{0|t} + \mathbf{H}^\dagger(\mathbf{y} - \mathbf{H}\mathbf{x}_{0|t})) \\ + \delta_t(\mathbf{H}^\top\mathbf{H} + \lambda\mathbf{I})^{-1}(\mathbf{H}^\top\mathbf{y} + \lambda\mathbf{x}_{0|t}),$$

Algorithm 1 Fused Data Guidance for Diffusion Sampling

Require: Observation $\mathbf{y}$, SPI forward model $\mathbf{H}$, score function, diffusion schedule $\{\sigma_t\}$, fusion weights $\{\delta_t\}$

1: Initialize $\mathbf{x}_T \sim \mathcal{N}(0, \mathbf{I})$
2: **for** $t = T$ to 1 **do**
3: $\quad \mathbf{x}_{0|t} \leftarrow \frac{\mathbf{x}_t + (1-\alpha_t)\nabla_{\mathbf{x}_t}\log p_t(\mathbf{x}_t)}{\sqrt{\alpha_t}}$ // Denoise
4: $\quad \mathbf{g}_{\text{GAP}} \leftarrow \mathbf{x}_{0|t} + \mathbf{H}^\dagger(\mathbf{y} - \mathbf{H}\mathbf{x}_{0|t})$ // GAP term
5: $\quad \mathbf{g}_{\text{HQS}} \leftarrow (\mathbf{H}^T\mathbf{H} + \lambda\mathbf{I})^{-1}(\mathbf{H}^T\mathbf{y} + \lambda\mathbf{x}_{0|t})$ // HQS term
6: $\quad \mathbf{x}'_{0|t} \leftarrow (1-\delta_t)\mathbf{g}_{\text{GAP}} + \delta_t\mathbf{g}_{\text{HQS}}$ // Fused term
7: $\quad \hat{\epsilon}_t \leftarrow \frac{(\mathbf{x}_t - \sqrt{\alpha_t}\,\mathbf{x}'_{0|t})}{\sqrt{1-\alpha_t}}$
8: $\quad \epsilon_t \sim \mathcal{N}(0, \mathbf{I}_n)$
9: $\quad \mathbf{x}_{t-1} = \sqrt{\bar{\alpha}_{t-1}}\,\mathbf{x}'_{0|t} + \sqrt{1-\bar{\alpha}_{t-1}}\left(w_t\sqrt{1-\zeta}\,\hat{\epsilon}_t + \sqrt{\zeta}\,\epsilon_t\right)$
10: $\quad$ // DDIM Sampling
11: **end for**
12: **return** $\mathbf{x}_0$

Fused data-guidance for diffusion model

Results

$\mathbf{H}^{\dagger}\mathbf{y}$ | DDIM + GAP | DDIM + HQS | DDIM + GAP + HQS

Results

Table 1. Reconstruction metrics comparison across methods.

Method	**PSNR (dB)**	**SSIM**	**LPIPS**
$\mathbf{H}^{\dagger}\mathbf{y}$	20.55	0.39	0.54
DDIM+GAP	24.09	0.62	0.33
DDIM+HQS	24.64	0.67	0.38
DDIM+GAP+HQS	24.76	0.68	0.37

Table 2. Reconstruction Results across different CRs.

CR	PSNR	SSIM	LPIPS
1%	21.23	0.47	0.56
5%	24.76	0.68	0.37
10%	25.66	0.70	0.25
20%	27.01	0.78	0.18

Thank you!