Bayesian Optimal Latent Projection for Noisy Image Restoration Supplementary Material

Ziqiang Shi, Rujie Liu
Fujitsu R&D Center
Beijing, China

Jun Takahashi, Takuma Yamamoto
Fujitsu Research
Tokyo, Japan

1. Qualitative Comparison (Mainly for Deblurring)

Due to space limitations in the main text, we present some qualitative results in this supplementary material. Figures 1 and 2 show a qualitative comparison of BOLP and PSLD in removing Gaussian blur. It can be observed that PSLD often generates strange white noise point clouds, whereas BOLP does not have this issue. We believe this is because BOLP can eliminate such unusual white noise by correcting the sampling path used in PSLD. Figure 3 demonstrates that BOLP achieves excellent restoration results across various tasks, including inpainting, deblurring, and super-resolution, and can be applied to any image, whether it is a face, a dog, or a natural scene.

2. Detailed Description of Other Algorithms Compared with BOLP

The notations used in this material are consistent with those in the main text.

In the main text, BOLP is mainly compared with three other algorithms, namely latent DPS (LDPS) [1, 2], PSLD [2], and ReSample [3]. . We summarize the details of these three algorithms in Algorithms 1, 2, and 3 respectively.

3. Tweedie's Formula and Posterior Conditional Sampling Score in Latent Space

In this section, we mainly derive Tweedie's formula and posterior conditional sampling fraction when the observed signal is in pixel domain and the diffusion model is in latent domain.

By Bayes' theorem, we have

$$\nabla_{\mathbf{z}_t} \log p(\mathbf{z}_t|\mathbf{y}) = \nabla_{\mathbf{z}_t} \log p(\mathbf{z}_t) + \nabla_{\mathbf{z}_t} \log p(\mathbf{y}|\mathbf{z}_t). \quad (1)$$

Therefore, the most important thing is to derive $p(\mathbf{y}|\mathbf{z}_t)$.

Algorithm 1 Image restoration with latent DPS

Input and initialization: N, $\mathbf{z}_N \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I})$, $\epsilon_\theta(\mathbf{x}_n, t_n, \mathbf{c})$ the pre-trained latent diffusion generative model.

1: **for** $n = N, N-1, \cdots, 1$

2: # Predict $\mathbf{z}_0^\epsilon$ using ϵ_θ.
$\mathbf{z}_0^\epsilon \leftarrow \frac{1}{\sqrt{\bar{\alpha}_n}}\left[\mathbf{z}_n - \sqrt{1-\bar{\alpha}_n}\epsilon_\theta(\mathbf{z}_n, t_n, \mathbf{c})\right]$.

3: # Calculate the coefficients of $\mathbf{z}_0^\epsilon$ and $\mathbf{z}_n$.
$\mu_n \leftarrow \frac{\sqrt{\alpha_n}(1-\bar{\alpha}_{n-1})}{1-\bar{\alpha}_n}$, and $\mu_0 \leftarrow \frac{\sqrt{\bar{\alpha}_{n-1}}\beta_n}{1-\bar{\alpha}_n}$.

4: # Update $\mathbf{z}$.
$\tilde{\mathbf{z}}_{n-1} \leftarrow \mu_0\mathbf{z}_0^\epsilon + \mu_n\mathbf{z}_n + \sigma_n\mathbf{e}$, where $\mathbf{e} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I})$.

5: # DPS
$\mathbf{z}_{n-1} \leftarrow \tilde{\mathbf{z}}_{n-1} - \eta_n\nabla_{\mathbf{z}_n}\|\mathbf{y} - \mathcal{H}(\mathcal{D}(\mathbf{z}_0^\epsilon)\|$.

Output: $\mathcal{D}(\hat{\mathbf{z}}_0)$.

Algorithm 2 Image restoration with PSLD

Input and initialization: N, $\mathbf{x}_N = \mathbf{y}$, $\epsilon_\theta(\mathbf{x}_n, t_n, \mathbf{c})$ from trained latent models.

1: **for** $n = N, N-1, \cdots, 1$

2: #Predict $\mathbf{z}_0^\epsilon$ using ϵ_θ.
$\mathbf{z}_0^\epsilon \leftarrow \frac{1}{\sqrt{\bar{\alpha}_n}}\left[\mathbf{z}_n - \sqrt{1-\bar{\alpha}_n}\epsilon_\theta(\mathbf{z}_n, t_n, \mathbf{c})\right]$.

3: #Calculate the coefficients of $\mathbf{z}_0^\epsilon$ and $\mathbf{z}_n$.
$\mu_n \leftarrow \frac{\sqrt{\alpha_n}(1-\bar{\alpha}_{n-1})}{1-\bar{\alpha}_n}$, and $\mu_0 \leftarrow \frac{\sqrt{\bar{\alpha}_{n-1}}\beta_n}{1-\bar{\alpha}_n}$.

4: #Update $\mathbf{z}$.
$\tilde{\mathbf{z}}_{n-1} \leftarrow \mu_0\mathbf{z}_0^\epsilon + \mu_n\mathbf{z}_n + \sigma_n\mathbf{e}$,
where $\mathbf{e} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I})$.

5: # DPS vanilla extension,
$\hat{\mathbf{z}}_{n-1} \leftarrow \tilde{\mathbf{z}}_{n-1} - \eta_n\nabla_{\mathbf{z}_n}\|\mathbf{y} - \mathcal{H}(\mathcal{D}(\mathbf{z}_0^\epsilon)\|$.

6: # Gluing of $\mathbf{z}_0$,
$\mathbf{z}_{n-1} \leftarrow \hat{\mathbf{z}}_{n-1} - \gamma_n\nabla_{\mathbf{z}_n}\left\|\mathbf{z}_0^\epsilon - \mathcal{E}(\mathcal{H}^\top\mathbf{y} + (\boldsymbol{I} - \mathcal{H}^\top\mathcal{H})\mathcal{D}(\mathbf{z}_0^\epsilon))\right\|$.

Output: $\mathcal{D}(\hat{\mathbf{z}}_0)$.

From forward diffusion, we know

$$p(\mathbf{z}_n|\mathbf{z}_0) = \frac{1}{(2\pi(1-\bar{\alpha}_n))^{d/2}}\exp\left(-\frac{\left\|\mathbf{z}_n - \sqrt{\bar{\alpha}_n}\mathbf{z}_0\right\|^2}{2(1-\bar{\alpha}_n)}\right). \quad (2)$$

Figure 1. Qualitative comparison of Gaussian deblurring effects using different algorithms on FFHQ. PSLD generates small white noise points that do not exist, while our BOLP does not have this problem. It is best to zoom in and view in color.

So

$$p(\mathbf{z}_n|\mathbf{z}_0) = p_0(\mathbf{z}_n)\exp\big(\mathbf{z}_0^\top T(\mathbf{z}_n) - \phi(\mathbf{z}_0)\big), \quad (3)$$

where

$$p_0(\mathbf{z}_n) = \frac{1}{(2\pi(1-\bar{\alpha}_n))^{d/2}}\exp\left(-\frac{\|\mathbf{z}_n\|^2}{2(1-\bar{\alpha}_n)}\right), \quad (4)$$

$$T(\mathbf{z}_n) = \frac{\sqrt{\bar{\alpha}_n}}{1-\bar{\alpha}_n}\mathbf{z}_n, \quad (5)$$

$$\phi(\mathbf{z}_0) = \frac{\bar{\alpha}_n\|\mathbf{z}_n\|^2}{2(1-\bar{\alpha}_n)}. \quad (6)$$

By integrating $\mathbf{z}_0$, we get $p(\mathbf{z}_n)$

$$\begin{aligned} p(\mathbf{z}_n) &= \int p(\mathbf{z}_n|\mathbf{z}_0)p(\mathbf{z}_0)d\mathbf{z}_0 & (7)\\ &= \int p_0(\mathbf{z}_n)\exp\big(\mathbf{z}_0^\top T(\mathbf{z}_n) - \phi(\mathbf{z}_0)\big)p(\mathbf{z}_0)d\mathbf{z}_0. & (8)\end{aligned}$$

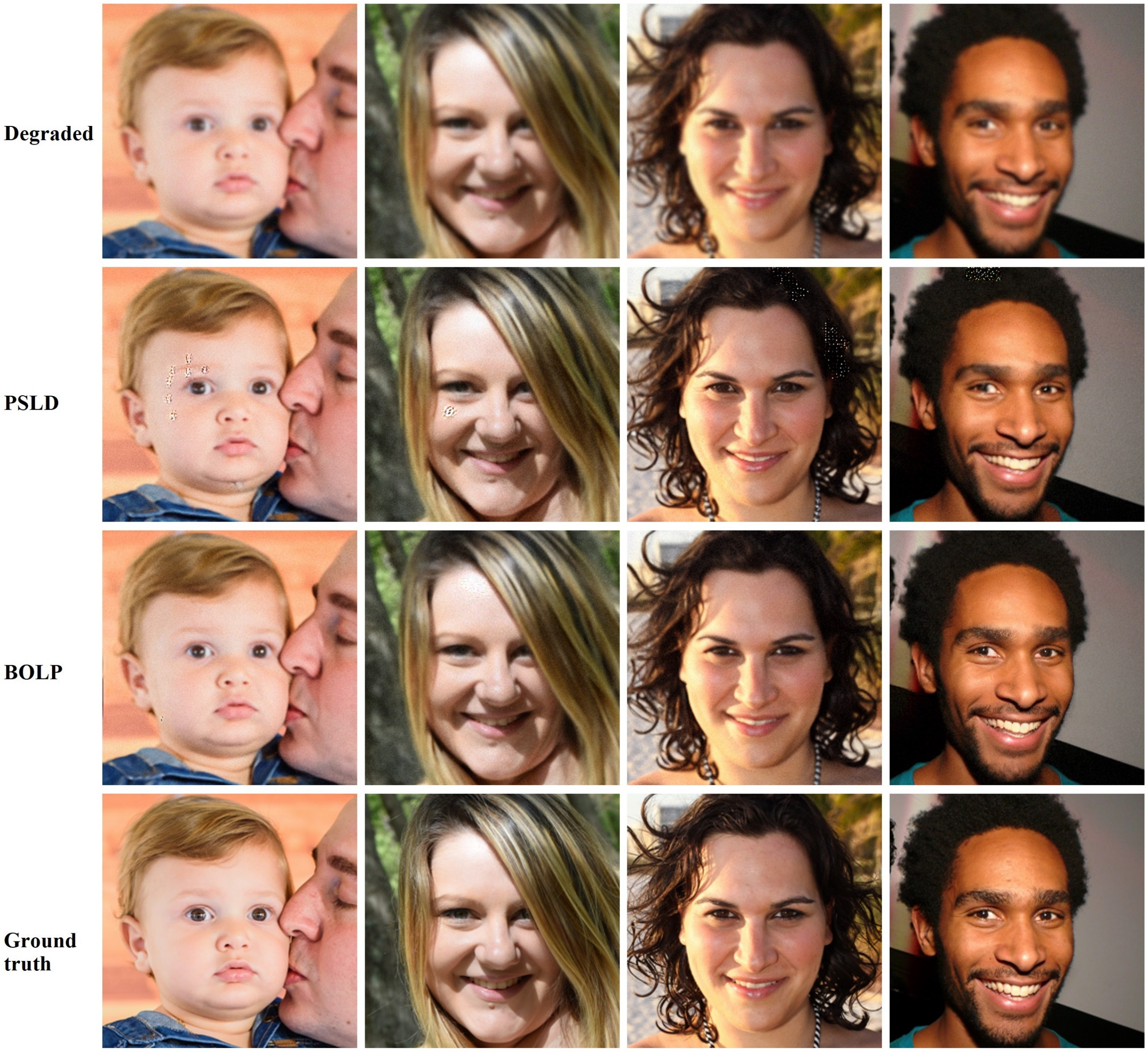

Figure 2. Qualitative comparison of Gaussian deblurring effects using different algorithms on FFHQ. PSLD generates small white noise points that do not exist, while our BOLP does not have this problem. It is best to zoom in and view in color.

Taking the derivative of both sides with respect to $\mathbf{z}_n$,

$$\nabla_{\mathbf{z}_n} p(\mathbf{z}_n) \tag{9}$$

$$= \nabla_{\mathbf{z}_n} p_0(\mathbf{z}_n) \int \exp\big(\mathbf{z}_0^\top T(\mathbf{z}_n) - \phi(\mathbf{z}_0)\big) p(\mathbf{z}_0) d\mathbf{z}_0 \tag{10}$$

$$+ \int (\nabla_{\mathbf{z}_n} T(\mathbf{z}_n))^\top \mathbf{z}_0 p_0(\mathbf{z}_n) \exp\big(\mathbf{z}_0^\top T(\mathbf{z}_n) - \phi(\mathbf{z}_0)\big) p(\mathbf{z}_0) d\mathbf{z}_0 \tag{11}$$

$$= \frac{\nabla_{\mathbf{z}_n} p_0(\mathbf{z}_n)}{p_0(\mathbf{z}_n)} p(\mathbf{z}_n) + (\nabla_{\mathbf{z}_n} T(\mathbf{z}_n))^\top \int \mathbf{z}_0 p(\mathbf{z}_n, \mathbf{z}_0) d\mathbf{z}_0. \tag{12}$$

There

$$\frac{\nabla_{\mathbf{z}_n} p(\mathbf{z}_n)}{p(\mathbf{z}_n)} \tag{13}$$

$$= \frac{\nabla_{\mathbf{z}_n} p_0(\mathbf{z}_n)}{p_0(\mathbf{z}_n)} + (\nabla_{\mathbf{z}_n} T(\mathbf{z}_n))^\top \int \mathbf{z}_0 p(\mathbf{z}_0|\mathbf{z}_n) d\mathbf{z}_0. \tag{14}$$

That is

$$(\nabla_{\mathbf{z}_n} T(\mathbf{z}_n))^\top \mathbb{E}_{\mathbf{z}_0 \sim p(\mathbf{z}_0|\mathbf{z}_n)} [\mathbf{z}_0] = \nabla_{\mathbf{z}_n} \log p(\mathbf{z}_n) - \nabla_{\mathbf{z}_n} \log p_0(\mathbf{z}_n). \tag{15}$$

Simplifying, we obtain the Tweedie's formula in the latent

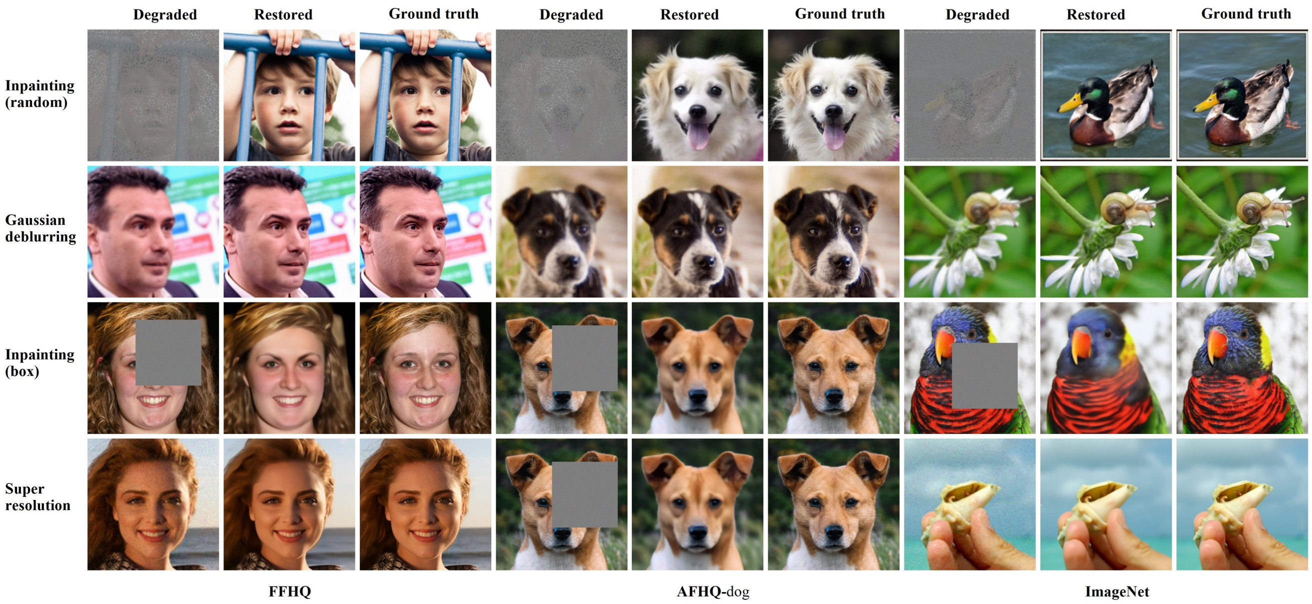

Figure 3. BOLP can be applied to any image restoration under various tasks (inpainting, deblurring, and super resolution), whether it is a human face, a dog, or a natural image.

Algorithm 3 Image restoration with ReSample

Input and initialization: N, $\mathbf{z}_N \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I})$, $\epsilon_\theta(\mathbf{x}_n, t_n, \mathbf{c})$ from trained latent models.

1: **for** $n = N, N-1, \cdots, 1$

2: #Predict $\mathbf{z}_0^\epsilon$ using ϵ_θ.
$\mathbf{z}_0^\epsilon \leftarrow \frac{1}{\sqrt{\bar{\alpha}_n}}\left[\mathbf{z}_n - \sqrt{1-\bar{\alpha}_n}\epsilon_\theta(\mathbf{z}_n, t_n, \mathbf{c})\right]$.

3: #Update $\mathbf{z}$.
$\tilde{\mathbf{z}}_{n-1} \leftarrow \sqrt{\bar{\alpha}_{n-1}}\mathbf{z}_0^\epsilon + \sqrt{1-\bar{\alpha}_{n-1}-\eta\sigma^2}\epsilon_\theta(\mathbf{z}_n, t_n, \mathbf{c}) + \eta\sigma\mathbf{e}$
where $\mathbf{e} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I})$.

4: # ReSample time step,
Solve $\hat{\mathbf{z}}_0(\mathbf{y}) \in \arg\min_{\mathbf{z}} \|\mathbf{y} - \mathcal{H}(\mathcal{D}(\mathbf{z}))\|_2^2$ with initial point $\mathbf{z}_0^\epsilon$.

5: # Stochastic Resample,
#Calculate the coefficients of $\mathbf{z}_0^\epsilon$ and $\tilde{\mathbf{z}}_{n-1}$.
$\mu_n \leftarrow \frac{1-\bar{\alpha}_{n-1}}{\sigma_{n-1}^2+1-\bar{\alpha}_{n-1}}$, $\mu_0 \leftarrow \frac{\sigma_{n-1}^2\sqrt{\bar{\alpha}_{n-1}}}{\sigma_{n-1}^2+1-\bar{\alpha}_{n-1}}$, and $\sigma_e \leftarrow \frac{\sigma_{n-1}^2(1-\bar{\alpha}_{n-1})}{\sigma_{n-1}^2+1-\bar{\alpha}_{n-1}}$.
#Update $\mathbf{z}$.
$\mathbf{z}_{n-1} \leftarrow \mu_0\mathbf{z}_0^\epsilon + \mu_n\tilde{\mathbf{z}}_{n-1} + \sigma_e\mathbf{e}$,
where $\mathbf{e} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I})$.

Output: $\mathcal{D}(\hat{\mathbf{z}}_0)$.

space

$$\hat{\mathbf{z}}_0 = \mathbb{E}_{\mathbf{z}_0 \sim p(\mathbf{z}_0|\mathbf{z}_n)}[\mathbf{z}_0] \tag{16}$$

$$= \mathbb{E}[\mathbf{z}_0|\mathbf{z}_n] = \frac{1}{\sqrt{\bar{\alpha}_n}}(\mathbf{z}_n + (1-\bar{\alpha}_n)\nabla_{\mathbf{z}_n}\log p(\mathbf{z}_n)). \tag{17}$$

Thus we have

$$p(\mathbf{y}|\mathbf{z}_t) = \int p(\mathbf{y}|\mathbf{z}_0)p(\mathbf{z}_0|\mathbf{z}_t)d\mathbf{z}_0 = \mathbb{E}_{\mathbf{z}_0 \sim p(\mathbf{z}_0|\mathbf{z}_t)}[p(\mathbf{y}|\mathbf{z}_0)] \tag{18}$$

$$\geq p(\mathbf{y}|\mathbb{E}_{\mathbf{z}_0 \sim p(\mathbf{z}_0|\mathbf{z}_t)}[\mathbf{z}_0]) = p(\mathbf{y}|\hat{\mathbf{z}}_0) = p(\mathbf{y}|\mathcal{D}(\hat{\mathbf{z}}_0)). \tag{19}$$

That is to say

$$\nabla_{\mathbf{z}_t}\log p(\mathbf{y}|\mathbf{z}_t) \simeq \nabla_{\mathbf{z}_t}\log p(\mathbf{y}|\mathcal{D}(\hat{\mathbf{z}}_0)) \tag{20}$$

$$\simeq -\frac{1}{\sigma_y^2}\nabla_{\mathbf{z}_t}\|\mathbf{y} - \mathcal{H}(\mathcal{D}(\hat{\mathbf{z}}_0))\|^2. \tag{21}$$

This is the posterior score at each moment under the condition of the observed signal that we need.

4. Limitation and Future Work

BOLP does not utilize any images from the target application domain. If a small number of images from the domain (e.g., FFHQ) can be used, then a more accurate estimation of the diffusion sampling trajectory may be achieved. This is the direction we intend to explore in future work.

References

[1] Hyungjin Chung, Jeongsol Kim, Michael Thompson Mccann, Marc Louis Klasky, and Jong Chul Ye. Diffusion posterior sampling for general noisy inverse problems. In *The Eleventh International Conference on Learning Representations*, 2023. 1

[2] Litu Rout, Negin Raoof, Giannis Daras, Constantine Caramanis, Alex Dimakis, and Sanjay Shakkottai. Solving linear inverse problems provably via posterior sampling with latent diffusion models. *Advances in Neural Information Processing Systems*, 36, 2023. 1

[3] Bowen Song, Soo Min Kwon, Zecheng Zhang, Xinyu Hu, Qing Qu, and Liyue Shen. Solving inverse problems with latent diffusion models via hard data consistency. *arXiv preprint arXiv:2307.08123*, 2023. 1