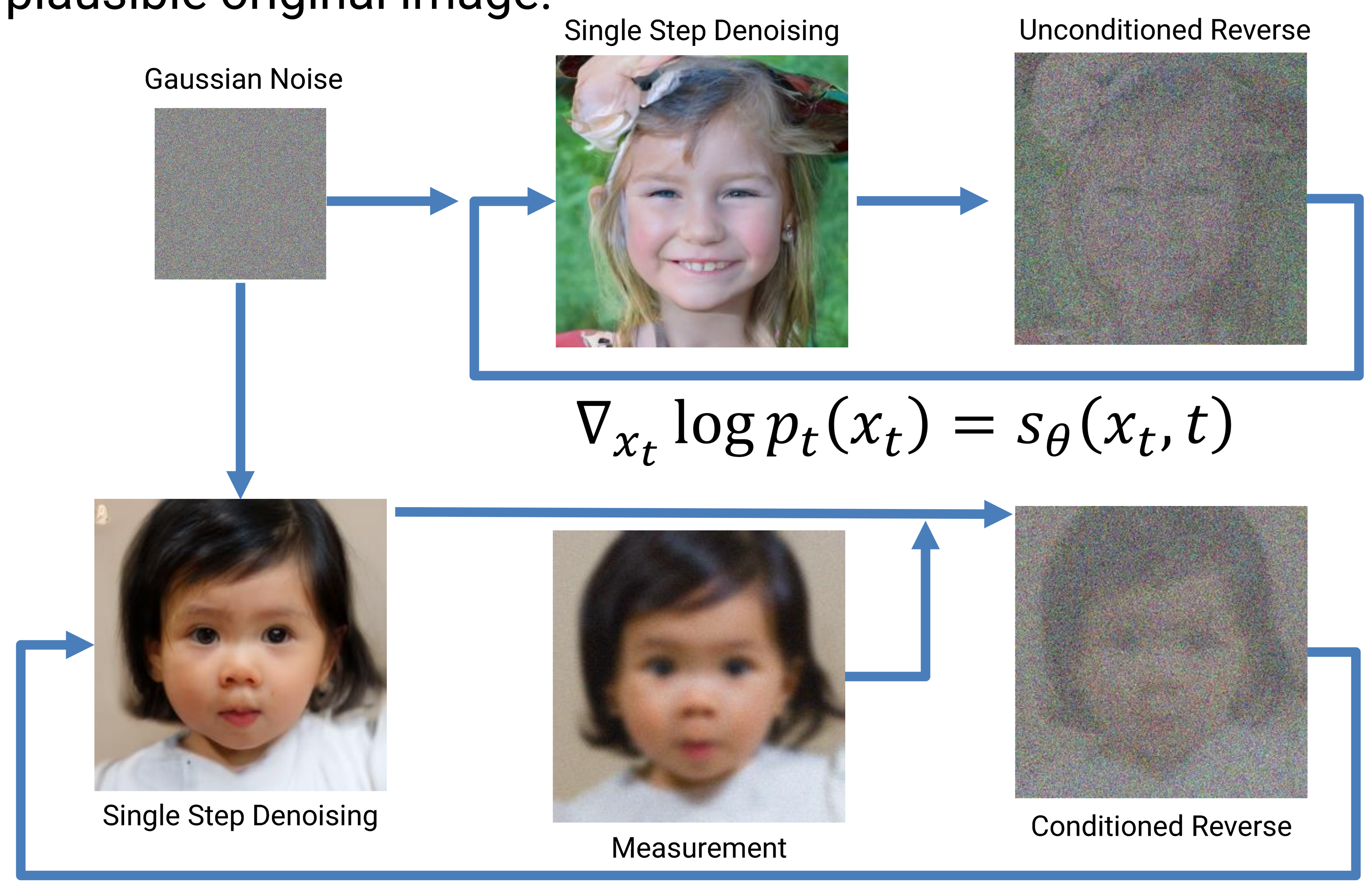
Diffusion-Based Priors for Image Reconstruction

Tim Jacques

Department of Electrical Engineering, Stanford University

Motivation

In modern image processing, diffusion-based generative models have emerged as a capable technique to generate images, often conditioned on text prompts. By harnessing the realistic generation capabilities of these models, we can **better solve "ill-posed" inverse imaging problems**, like deblurring, reconstruction, and inpainting tasks. Each of these methods work by taking a perturbed input image, adding noise, then **performing *conditioned diffusion*** to reconstruct a plausible original image.



$$\nabla_{x_t} \log p_t(x_t) = s_\theta(x_t, t)$$

$$\nabla_{x_t} \log p_t(x_t|y) = s_\theta(x_t, t) + \nabla_{x_t} \log p_t(y|x_t)$$

References

[1] Chung, H., Kim, J., Mccann, M. T., Klasky, M. L., and Ye, J. C. (2023). Diffusion posterior sampling for general noisy inverse problems. In *ICLR*.
[2] Ho, J., Jain, A., and Abbeel, P. (2020). Denoising diffusion probabilistic models. In *NeurIPS*.
[3] Jalal, A., Arvinte, M., Daras, G., Price, E., Dimakis, A. G., and Tamir, J. (2021). Robust compressed sensing mri with deep generative priors.
[4] Meng, C., He, Y., Song, Y., Song, J., Wu, J., Zhu, J.-Y., and Ermon, S. (2022). Sdedit: Guided image synthesis and editing with stochastic differential equations.

Diffusion Conditioning Methods

To condition the reverse diffusion process, we need to estimate the posterior: $\nabla_{x_t} \log p_t(y|x_t)$

Score-Distillation Editing (SDEdit)

Instead of estimating posterior, condition by starting with a noisy version of the measurement at a **partial timestep** in the unconditional reverse diffusion process.

Score Annealed Langevin Dynamics (ScoreALD)

Estimate posterior using:

$$\nabla_{x_t} \log p_t(y|x_t) \approx -\frac{1}{2\zeta_t} \nabla_{x_t} \|A(x_t) - y\|_2^2$$

ζ_t = Annealing Factor

Diffusion Posterior Sampling (DPS)

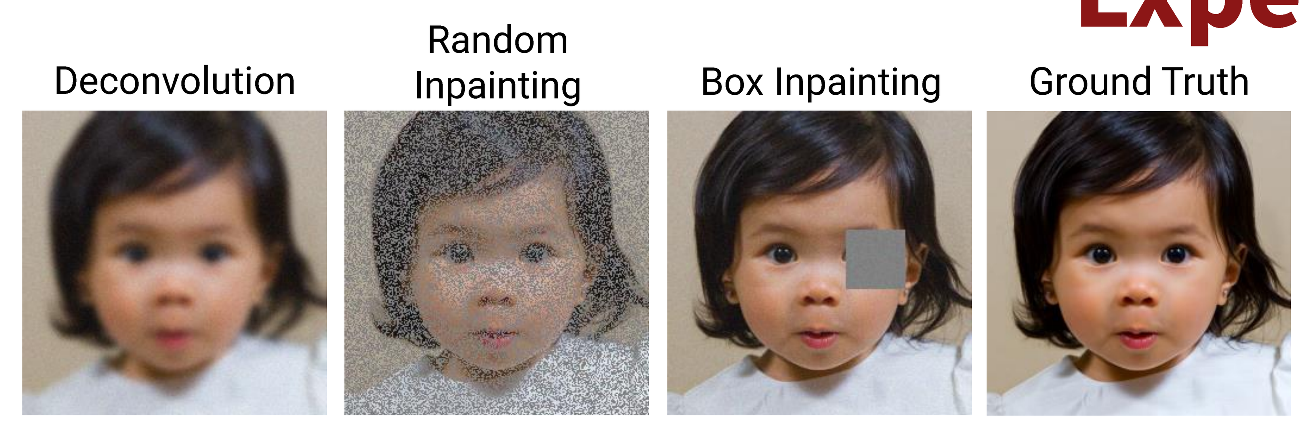
Estimate posterior using:

$$\nabla_{x_t} \log p_t(y|x_t) \approx -\frac{1}{\zeta_t} \nabla_{x_t} \|A(\widehat{x_0}) - y\|_2^2$$

$\zeta_t = \zeta \div \left\|\nabla_{x_t} \|A(\widehat{x_0}) - y\|_2^2\right\|$

ζ = Scale Factor

Experimental Results



- ScoreALD and DPS perform very well in inpainting tasks.
- Deconvolution is best performed by ScoreALD.
- SDEdit needs further tuning for realistic results.

PSNR / LPIPS

Highlighted values = Best result per column