# Local MAP Sampling for Diffusion Models

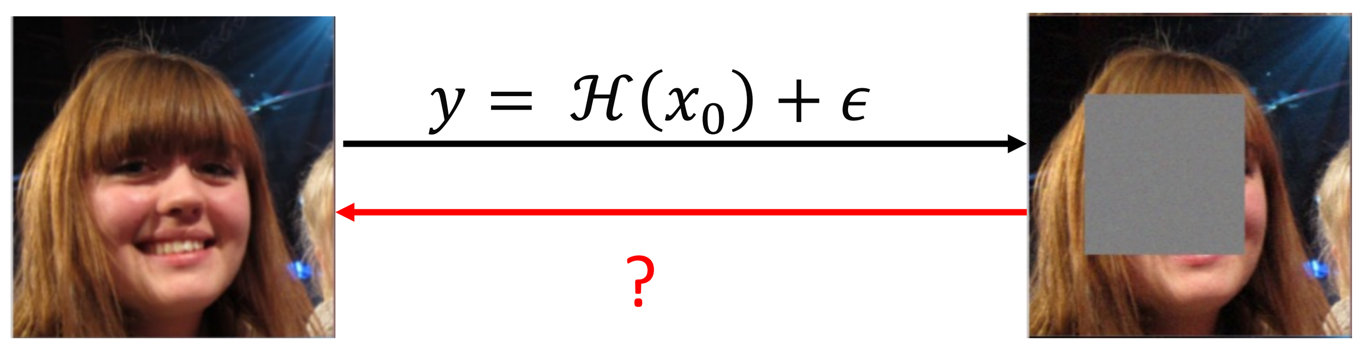
Shaorong Zhang, Rob Brekelmans, Greg Ver Steeg    University of California Riverside

About me    Paper

UC RIVERSIDE University of California Bourns College of Engineering    ICML International Conference On Machine Learning

## Background

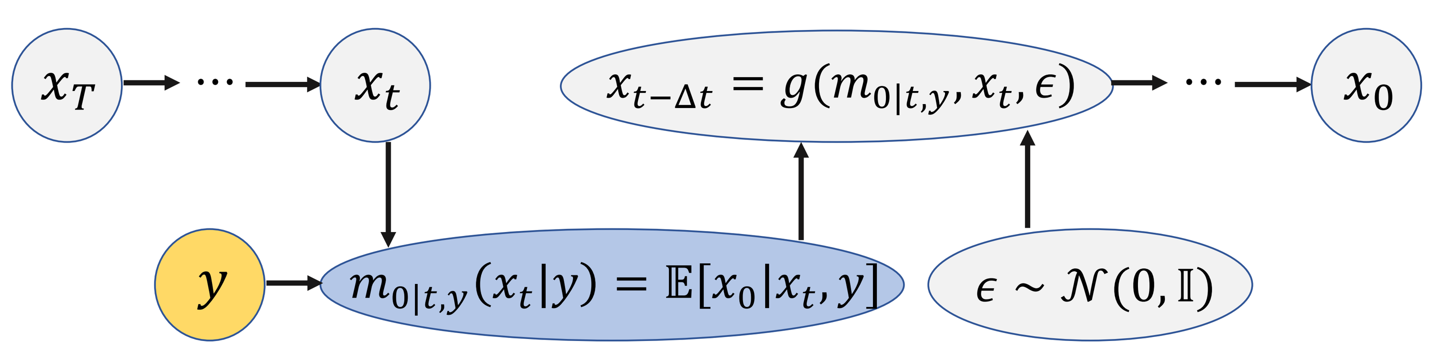
### Problem: Inverse Problem




- $\epsilon \sim \mathcal{N}(0, I)$ represents measurement noise;
- $x_0 \in \mathbb{R}^d$ is drawn from data distribution $\pi_0(x_0)$;
- $\mathcal{H}(\cdot)$ is measurement function;
- $y$ represents the degraded measurement.

## Diffusion Posterior Sampling

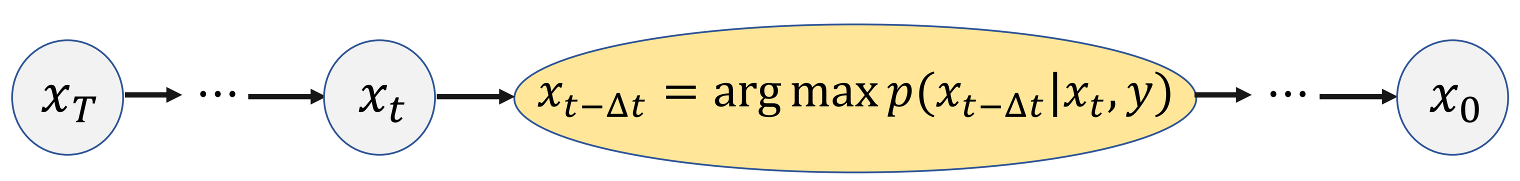
- Adding noise: $x_t = \alpha_t x_0 + \sigma_t\, \epsilon, \epsilon \sim \mathcal{N}(0, I)$



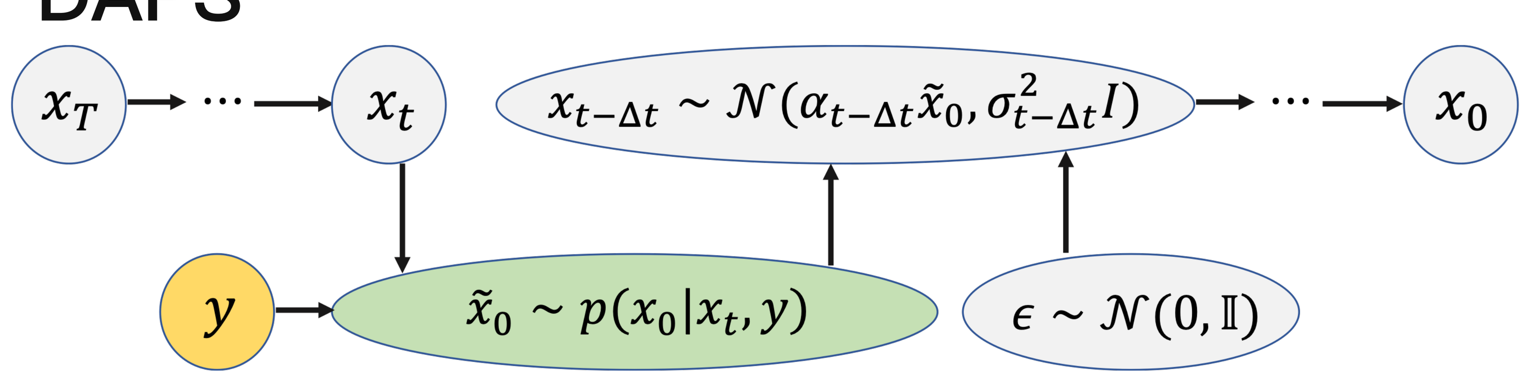

$$g(\xi, x_t, \epsilon) := \alpha_{t-\Delta t}\xi + \sigma_{t-\Delta t}\left(\sqrt{1-\rho_t^2}\,\frac{x_t - \alpha_t \xi}{\sigma_t} + \rho_t \epsilon\right),$$

$$\mathbb{E}[x_0 \mid x_t, y] = m_{0|t} + \frac{\sigma_t^2}{\alpha_t}\nabla_{x_t} \log p(y \mid x_t).$$

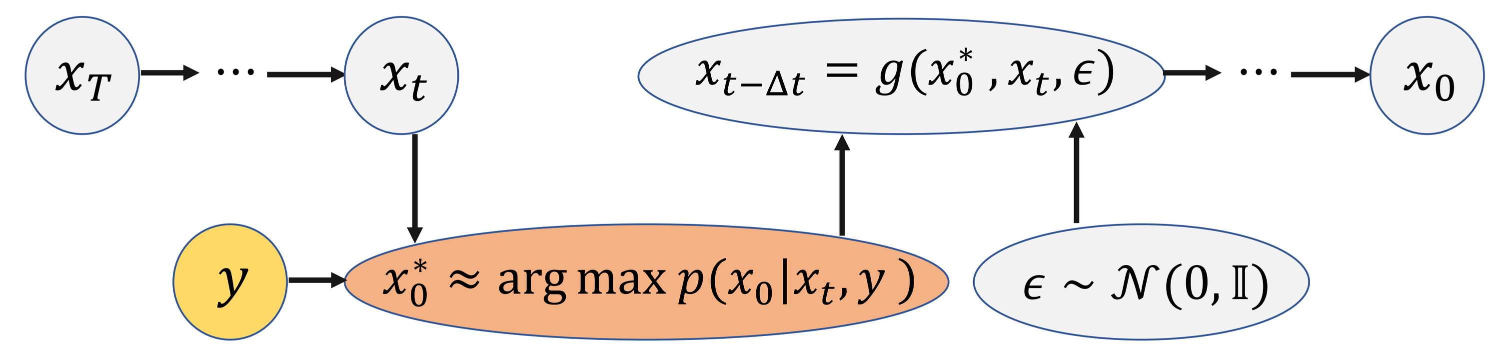
## DMAP

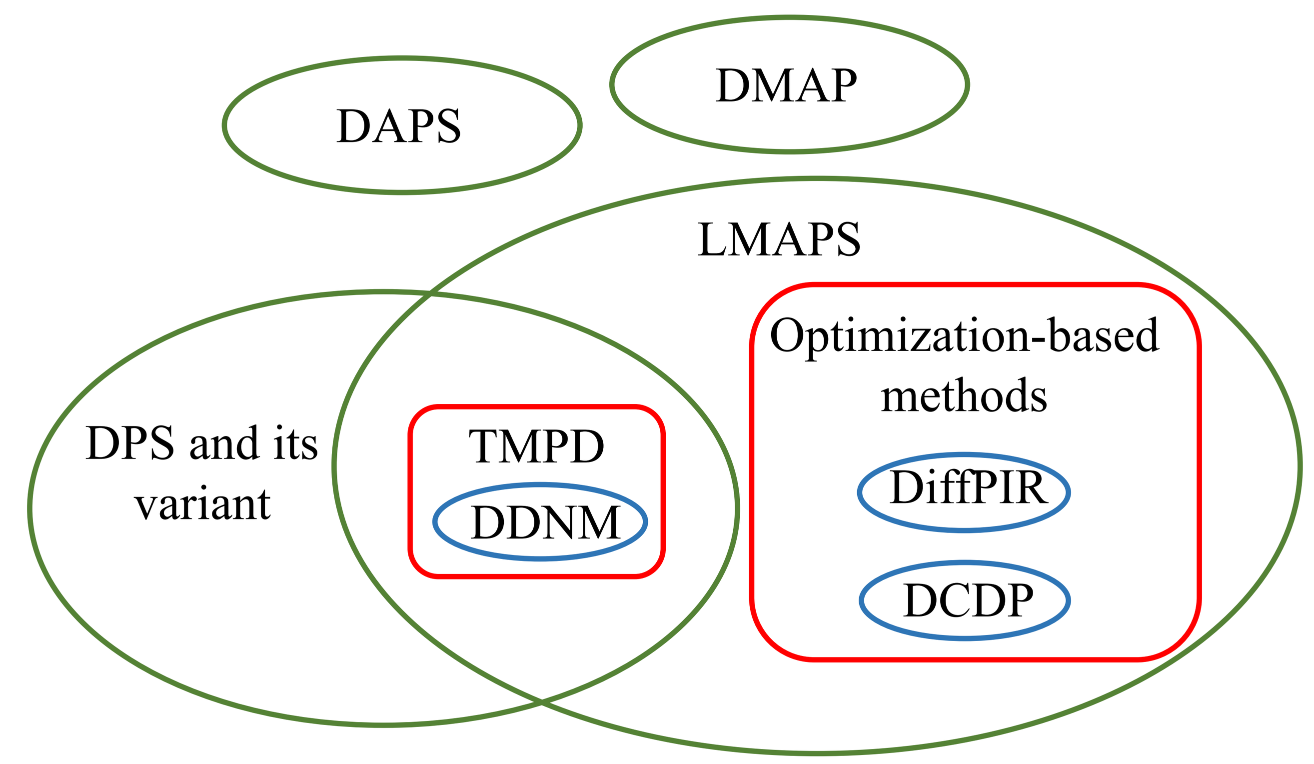



## DAPS




## LMAPS (Local MAP Sampling )




## Solving LMAPS

**Intractable Objective** $x_0^*(t, x_t, y) := \arg\max p(x_0 \mid x_t, y),$

**Equivalent Transformation**

$$x_0^*(t, x_t, y) = \arg\min\{-\log p(x_0 \mid x_t) - \log p(y \mid x_0)\}.$$

**Approximation** $p(x_0 \mid x_t) \approx \mathcal{N}(x_0; m_{0|t}, \Sigma_{0|t})$

$$x_0^* = \arg\min (x_0 - m_{0|t})^T \Sigma_{0|t}^{-1}(x_0 - m_{0|t}) + \frac{1}{\sigma_y^2}\|y - \mathcal{H}(x_0)\|^2.$$

**Approximation** $\Sigma_{0|t} \approx \frac{k}{\mathrm{SNR}}\mathbb{I}$

$$x_0^* = \arg\min\left\{\frac{\mathrm{SNR}}{k}\|x_0 - m_{0|t}\|^2 + \frac{1}{\sigma_y^2}\|y - \mathcal{H}(x_0)\|^2\right\}.$$

**Objective reformulation**

$$x_0^* = \arg\min\left\{\left(1 - \frac{\sigma_t^2}{\sigma_t^2 + k_1^2}\right)\frac{1}{2}\|x_0 - m_{0|t}\|^2 + \frac{\sigma_t^2}{\sigma_t^2 + k_1^2}k_2\|y - \mathcal{H}(x_0)\|^2\right\}$$

## Experimental Results

Image restoration: 10 tasks on FFHQ / ImageNet (100 images each). Scientific tasks: LIS, CS–MRI, Black Hole Imaging.

| 43/60 | 10/10 | 61s |
|---|---|---|
| best image metrics PSNR / SSIM / LPIPS | best scientific PSNR tasks | per image A6000, FFHQ256 |

### Image restoration highlights (FFHQ, PSNR↑/SSIM ↑ /LPIPS↓)

| Task | LMAPS | SITCOM | DAPS |
|---|---|---|---|
| SR 4× | **30.74/0.869/0.165** | 30.55/0.864/0.154 | 29.07/0.818/0.177 |
| Box Inpaint | **25.02/0.876/0.108** | 24.95/0.849/0.131 | 24.07/0.814/0.133 |
| Gaussian Deblur | **30.88/0.867/0.158** | 29.93/0.846/0.172 | 29.19/0.817/0.165 |
| Phase Retrieval | **31.56/0.867/0.126** | – | 30.63/0.851/0.139 |
| Nonlinear Deblur | **29.93/0.855/0.150** | 29.19/0.785/0.190 | 28.29/0.783/0.155 |

### Scientific inverse problems (PSNR↑)

| Tasks | Settings | LMAPS | Δ vs. best |
|---|---|---|---|
| LIS | **NR=360/180/60** | 38.07 / 37.19 / 30.75 | **+1.51 / +1.78 / +1.51** |
| CS–MRI | **4×/8× noisel./raw** | 32.83 / 28.77 / 30.50 / 27.43 | **best in all 4** |
| Black Hole | **100/10/3%** | 26.79 / 24.83 / 24.66 | **+0.72 / +0.26 / +0.41** |

### Sampling efficiency

Gaussian deblurring, FFHQ 256, one NVIDIA A6000 GPU:
LMAPS: 61 s/image (LPIPS 0.158) vs. DAPS: 110 s/image, SITCOM: 73 s/image, DPS: 138 s/image.

**Takeaway: LMAPS keeps MAP–style point estimation in the clean signal space, improving reconstruction fidelity across classical image restoration and scientific inverse problems.**