Linearly Constrained Diffusion Implicit Models

Vivek Jayaram Ira Kemelmacher-Shlizerman Steven M. Seitz John Thickstun

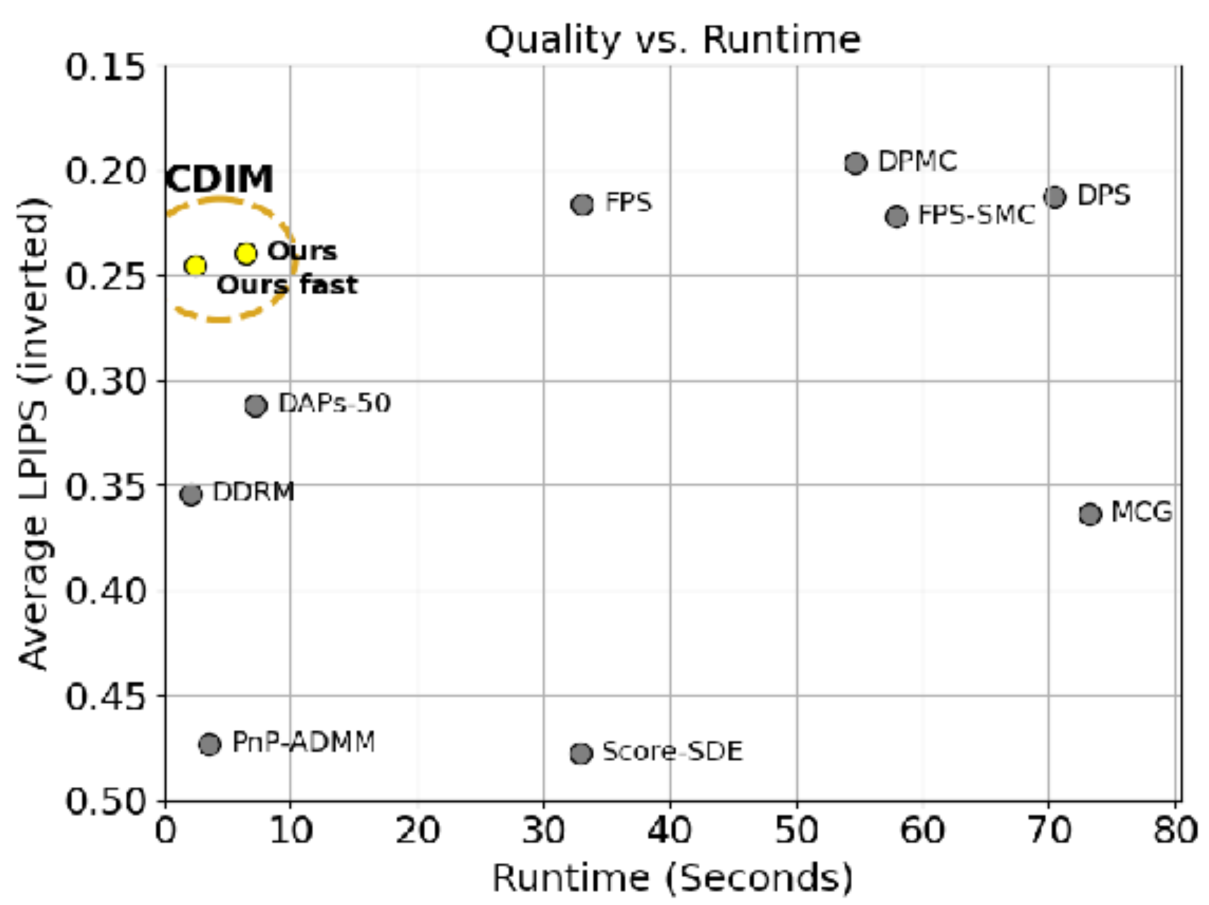
Overview

- We solve noisy linear inverse problems of the form $\mathbf{y} = \mathbf{A}\mathbf{x} + \sigma$.
- Our method is fast and gives exact recovery of the linear constraint $\mathbf{A}$. (e.g. inpainting does not change observed pixels)
- We use a pre-trained diffusion model and alternate unconditional denoising steps with projection steps during the denoising process.
- Our key insight is to use the analytical chi-squared distribution of $\|\mathbf{A}\mathbf{x}_t - \mathbf{y}\|^2$ under the forward noising process to guide the size and number of projection steps. This leads to highly efficient updates.

Example Results

We show pinpointing results for a sparse point cloud reproduction of a movie scene. CDIM is able to inpainting this image where 95% of pixels are missing.

Runtime and Quality Comparisons



Our method (CDIM in top left corner) achieves high performance and extremely fast inference compared to other inverse solvers

Teaser

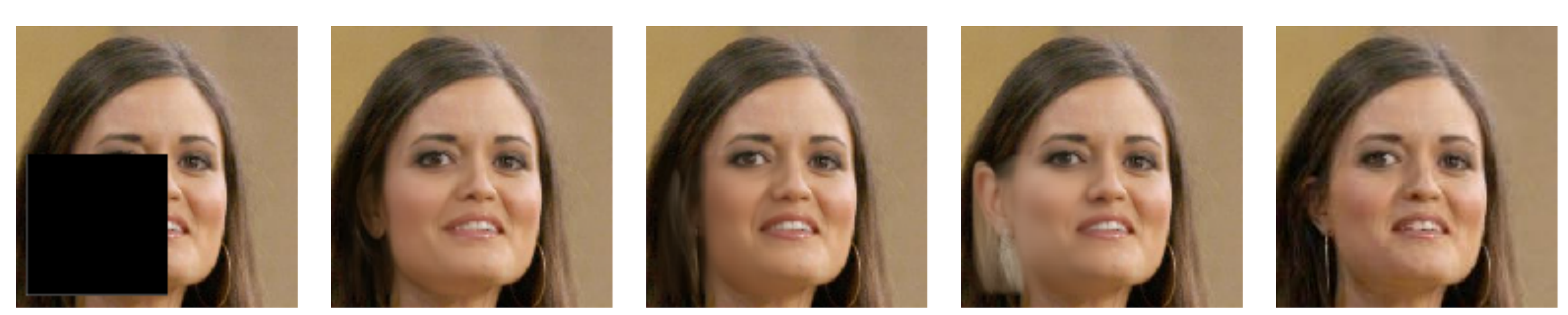
Time-Travel Rephotography | Poisson Noise | Noisy Inpainting | Fast Inference: 3s

Constraint Satisfaction - DPS Comparison

Box inpainting input **y** | DPS + DDIM MAE: 4% | DPS + DDIM (Large η) MAE: 0.3% | CDIM (ours) MAE: 0.05%*

We show comparisons against DPS [1] using DDIM sampling where all algorithms use 30 NFEs. For each output, we report the mean absolute error of the *observed* pixels. DPS either alters the observed pixels or diverges while CDIM achieves strong constraint satisfaction and inpainting results.

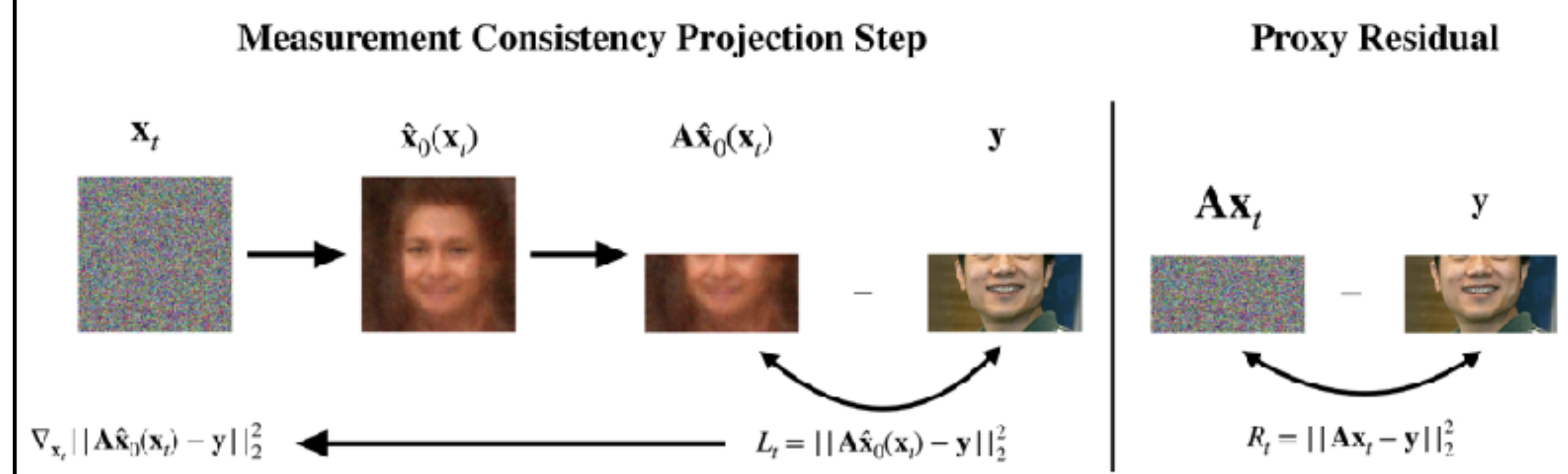
Reducing Inference Steps



Box inpainting input **y** | 25 denoising steps 31 projection steps | 10 denoising steps 15 projection steps | 4 denoising steps 10 projection steps | Ground Truth

We show how results change as we decrease the number of denoising steps. Notably we still produce reasonable results with only 4 denoising steps which leads to a corresponding 10 total projection steps (total of 14 NFEs).

Method Overview



We alternate unconditional diffusion steps with projection steps on our measurement error $\nabla_{\mathbf{x}_t} \|\mathbf{A}\hat{\mathbf{x}}_0(\mathbf{x}_t) - \mathbf{y}\|^2$ (left panel). Our key idea is to use the proxy residual $R_t = \|\mathbf{A}\mathbf{x}_t - \mathbf{y}\|^2$ (right panel), which has analytical chi-squared distribution under the forward noising process, to choose the number and size of projection steps.

Method Details

During the denoising process, we would like to choose projection steps to keep our measurement error $L_t = \|\mathbf{A}\hat{\mathbf{x}}_0(\mathbf{x}_t) - \mathbf{y}\|^2$ close to its expected value during the forward noising process: $\mathbb{E}_{\epsilon_t}[L_t \mid \mathbf{y}]$. Unfortunately this expectation is not tractable. Instead, we show that keeping $R_t = \|\mathbf{A}\mathbf{x}_t - \mathbf{y}\|^2$ near its forward process expectation (tractable chi-squared distribution) is sufficient to guide our projection steps.

CDIM Algorithm

Definitions: $\mu_t(\mathbf{y}) = \mathbb{E}_{\epsilon_t}[R_t \mid \mathbf{y}]$ and $\sigma_t^2(\mathbf{y}) = Var_{\epsilon_t}[R_t \mid \mathbf{y}]$

$\mathbf{x}_T \sim N(0, \mathbf{I})$

for $t = T, T-\delta \ldots 1$ **do**

  $\mathbf{x}_{t-\delta} \leftarrow$ Unconditional DDIM Step

  $\rho_{t-\delta}(\mathbf{y}) \leftarrow \mu_{t-\delta}(\mathbf{y}) + c \cdot \sigma_{t-\delta}(\mathbf{y})$ {Plausible region for $R_{t-\delta}$}

  while $R_{t-\delta} > \rho_{t-\delta}(\mathbf{y})$ {$R_{t-\delta}$ outside plausible region}

    $\hat{\mathbf{x}}_0(\mathbf{x}_{t-\delta}) \leftarrow \mathbb{E}(\mathbf{x}_0 \mid \mathbf{x}_{t-\delta})$ {Compute Tweedie's posterior estimate}

    $\mathbf{g} \leftarrow \nabla_{\mathbf{x}_{t-\delta}} \|\mathbf{A}\hat{\mathbf{x}}_0(\mathbf{x}_{t-\delta}) - \mathbf{y}\|^2$ {Projection step gradient}

    $\eta^* = \text{argmin}_\eta \left| \rho_{t-\delta}(\mathbf{y}) - \|\mathbf{A}(\mathbf{x}_{t-\delta} - \eta\mathbf{g}) - \mathbf{y}\|^2 \right|$ {Best step size}

    $\mathbf{x}_{t-\delta} \leftarrow \mathbf{x}_{t-\delta} - \eta^* \mathbf{g}$ {Projection step

References

[1] *Diffusion Posterior Sampling*, Chung et al. ICLR 2023.