

260 **A Proposition: Controlling the Proxy Residual Controls the True Objective**

261 **Setup and notation.** Fix a reverse-diffusion step  $t \in \{1, \dots, T\}$  with cumulative noise level  $\bar{\alpha}_t \in$   
 262  $(0, 1)$ . Let

$$\text{One-step variance: } b := 1 - \bar{\alpha}_t$$

$$a := \sqrt{\bar{\alpha}_t} - 1,$$

$$\Sigma := \mathbf{A}\mathbf{A}^\top \in \mathbb{R}^{d \times d}.$$

263 Given the current state  $\mathbf{x}_t = \sqrt{\bar{\alpha}_t} \mathbf{x}_0 + \sqrt{b} \boldsymbol{\epsilon}_t$  with  $\boldsymbol{\epsilon}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ , define

$$\text{Proxy residual: } \mathbf{r}_t := \mathbf{A}\mathbf{x}_t - \mathbf{y}, \quad R_t := \|\mathbf{r}_t\|^2,$$

$$\text{Mean of proxy residual: } \mu_t(\mathbf{y}) := \mathbb{E}_{\boldsymbol{\epsilon}_t}[R_t \mid \mathbf{y}]$$

$$\text{Tweedie estimate: } \hat{\mathbf{x}}_0 := \frac{\mathbf{x}_t - \sqrt{b} \boldsymbol{\epsilon}_\theta(\mathbf{x}_t, t)}{\sqrt{\bar{\alpha}_t}}.$$

$$\text{true objective: } L_t := \|\mathbf{A}\hat{\mathbf{x}}_0 - \mathbf{y}\|^2,$$

$$\text{model noise in measurement space: } \boldsymbol{\zeta}_t := \mathbf{A}\boldsymbol{\epsilon}_\theta(\mathbf{x}_t, t).$$

264 Throughout we assume the denoiser is *conditionally unbiased*:  $\mathbb{E}[\boldsymbol{\epsilon}_\theta(\mathbf{x}_t, t) \mid \mathbf{x}_t] = \boldsymbol{\epsilon}_t$ .

265 **Proposition Statement:** Suppose that the proxy residual is close to its expected value:  $|R_t -$   
 266  $\mu_t(\mathbf{y})| \leq \gamma$ . Then with high probability,  $|L_t - \mathbb{E}[L_t \mid \mathbf{y}]| \leq \frac{\gamma}{\bar{\alpha}_t} + \mathcal{O}(\frac{\sqrt{(1-\bar{\alpha}_t)}}{\bar{\alpha}_t})$ , i.e. the target  
 267 residual is close to its expectation.

268 **Step 1: Write  $L_t$  in terms of  $R_t$**

269 Multiply  $\hat{\mathbf{x}}_0$  by  $\mathbf{A}$  and subtract  $\mathbf{y}$ :

$$\mathbf{A}\hat{\mathbf{x}}_0 - \mathbf{y} = \frac{\mathbf{A}\mathbf{x}_t - \sqrt{b}\boldsymbol{\zeta}_t - \sqrt{\bar{\alpha}_t}\mathbf{y}}{\sqrt{\bar{\alpha}_t}} = \frac{\mathbf{r}_t - \sqrt{b}\boldsymbol{\zeta}_t - a\mathbf{y}}{\sqrt{\bar{\alpha}_t}}.$$

270 Squaring the norm yields:

$$L_t = \frac{1}{\bar{\alpha}_t} \left( R_t + b \|\boldsymbol{\zeta}_t\|^2 + a^2 \|\mathbf{y}\|^2 - 2\sqrt{b} \langle \mathbf{r}_t, \boldsymbol{\zeta}_t \rangle - 2a \langle \mathbf{r}_t, \mathbf{y} \rangle + 2a\sqrt{b} \langle \boldsymbol{\zeta}_t, \mathbf{y} \rangle \right). \quad (20)$$

271 **Step 2: Conditional expectation.** Taking  $\mathbb{E}[\cdot \mid \mathbf{y}]$  in Equation (20), using  $\mathbb{E}[\langle \mathbf{r}_t, \boldsymbol{\zeta}_t \rangle \mid \mathbf{y}] =$   
 272  $\mathbb{E}[\langle \boldsymbol{\zeta}_t, \mathbf{y} \rangle \mid \mathbf{y}] = 0$  (given the unbiased-score assumption and independence of  $\boldsymbol{\epsilon}_t$  from  $\mathbf{y}$ ), and using  
 273  $\mathbb{E}[\mathbf{r}_t \mid \mathbf{y}] = \mathbb{E}[\sqrt{\bar{\alpha}_t}\mathbf{A}\mathbf{x}_0 + \sqrt{b}\mathbf{A}\boldsymbol{\epsilon}_t - \mathbf{y}] = a\mathbf{y}$  gives:

$$\begin{aligned} \mathbb{E}[L_t \mid \mathbf{y}] &= \frac{1}{\bar{\alpha}_t} (\mu_t(\mathbf{y}) + b \operatorname{tr} \Sigma + a^2 \|\mathbf{y}\|^2 - 2a^2 \|\mathbf{y}\|^2) \\ &= \frac{1}{\bar{\alpha}_t} (\mu_t(\mathbf{y}) + b \operatorname{tr} \Sigma - a^2 \|\mathbf{y}\|^2). \end{aligned}$$

275 **Step 3: Decompose the deviation.** Now we can calculate how far our true objective deviates from  
 276 its expected value. Write

$$\Delta_1 := \|\boldsymbol{\zeta}_t\|^2 - \operatorname{tr} \Sigma, \quad \Delta_2 := \langle \mathbf{r}_t, \boldsymbol{\zeta}_t \rangle, \quad \Delta_3 := \langle \boldsymbol{\zeta}_t, \mathbf{y} \rangle, \quad \Delta_4 := \langle \mathbf{r}_t, \mathbf{y} \rangle - a\|\mathbf{y}\|^2.$$

277 Then

$$|L_t - \mathbb{E}[L_t | y]| = \frac{1}{\bar{\alpha}_t} \left( |R_t - \mu_t(y)| + b|\Delta_1| + 2\sqrt{b}|\Delta_2| + 2|a|\sqrt{b}|\Delta_3| + 2|a||\Delta_4| \right). \quad (21)$$

278 The first term in Equation (21) is the deviation of the proxy residual from its mean, which satisfies

$$|R_t - \mu_t(\mathbf{y})| \leq \gamma.$$

279 We then use Hanson-Wright and Cauchy-Schwarz to bound the remainder of the terms.

280 **Step 4: Concentration bounds for the  $\Delta_i$ .**

281  $\Delta_1$ : Since  $\zeta_t = A\epsilon_t$  with  $\epsilon_t \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ ,  $\|\zeta_t\|^2$  is a quadratic form in a Gaussian vector. Hanson-  
282 Wright [39] gives

$$|\Delta_1| \leq C_1 \sqrt{\text{tr}(\Sigma^2)}$$

283 with probability at least  $1 - 2e^{-c_1 d}$ . For measurement space dimension  $d$ :  $\mathbf{y} \in \mathbb{R}^d$  and constant  
284  $c_1 > 0$ .

285  $\Delta_2$ : Using Cauchy-Schwarz, a  $\chi^2$  tail for  $\|\zeta_t\|$  and the given bound  $\|r_t\| \leq \sqrt{\mu_t(y) + \gamma}$ ,

$$|\Delta_2| \leq \|r_t\| \|\zeta_t\| \leq \sqrt{\mu_t(y) + \gamma} C_2 \sqrt{\text{tr} \Sigma}.$$

286  $\Delta_3$ : We can use a standard Gaussian tail bound since  $\langle \zeta_t, y \rangle \sim \mathcal{N}(0, y^\top \Sigma y)$ . This yields

$$|\Delta_3| \leq C_2 \|y\| \sqrt{\text{tr} \Sigma}.$$

287  $\Delta_4$ : This is a deterministic bound once  $\|r_t\|$  is bounded. Using  $|a| = \sqrt{\bar{\alpha}_t} - 1 \leq \sqrt{b}$  and  $\|r_t\| \leq$   
288  $\sqrt{\mu_t(y) + \gamma}$ ,

$$|\Delta_4| = |\langle r_t, y \rangle - a\|y\|^2| \leq \|r_t\| \|y\| + |a| \|y\|^2 \leq \sqrt{\mu_t(y) + \gamma} \|y\| + \sqrt{b} \|y\|^2.$$

289 Each of the three genuinely probabilistic bounds occurs with failure probability  $2e^{-cd}$  for measure-  
290 ment space dimension  $d$ :  $\mathbf{y} \in \mathbb{R}^d$  and absorbing constants into  $c > 0$ . In typical proofs involving  
291 Hanson-Wright,  $c_i \approx 10^{-2}$  [40], and since  $d$  is typically much larger than 100, the failure probability  
292 remains exceedingly small.

293 **Step 5: Assemble the pieces.**

294 Insert the bounds for  $\Delta_{1:4}$  into Equation (21), use  $\sqrt{\mu_t(y) + \gamma} \leq \sqrt{\mu_t(y)} + \sqrt{\gamma}$ , and absorb  
295 numerical constants into a universal  $C > 0$ :

$$|L_t - \mathbb{E}[L_t | y]| \leq \frac{\gamma}{\bar{\alpha}_t} + \frac{C\sqrt{b}}{\bar{\alpha}_t} (\sqrt{\gamma} + \|y\|) + \frac{C b}{\bar{\alpha}_t} \|y\|^2.$$

296 **Step 6: Conclusion.**

$$|R_t - \mu_t(\mathbf{y})| \leq \gamma \implies |L_t - \mathbb{E}[L_t | y]| \leq \frac{\gamma}{\bar{\alpha}_t} + \mathcal{O}\left(\frac{\sqrt{1-\bar{\alpha}_t}}{\bar{\alpha}_t} [\sqrt{\gamma} + \|y\|] + \frac{1-\bar{\alpha}_t}{\bar{\alpha}_t} \|y\|^2\right),$$

297 with probability at least  $1 - 6e^{-cd}$ . Because  $1 - \bar{\alpha}_t = b \ll 1$  for all practical timesteps, the additional  
298 terms are dominated by  $\sqrt{b}/\bar{\alpha}_t$ , leaving

$$|L_t - \mathbb{E}[L_t | y]| \leq \frac{\gamma}{\bar{\alpha}_t} + \mathcal{O}\left(\frac{\sqrt{1-\bar{\alpha}_t}}{\bar{\alpha}_t}\right),$$

299 as claimed. Notice that As  $t \rightarrow 0$  ( $\bar{\alpha}_t \rightarrow 1$ ) the second term vanishes, so matching the proxy residual  
300 to its mean immediately controls the true objective with the same statistical precision.

## B Mean and Variance of $\|\mathbf{A}\mathbf{x}_t - \mathbf{y}\|^2$ with Noisy Observations

We derive the mean and variance of our residual energy  $R_t := \|\mathbf{A}\mathbf{x}_t - \mathbf{y}\|^2$  during the forward diffusion process when we have measurement noise:  $\mathbf{y} = \mathbf{A}\mathbf{x}_0 + \boldsymbol{\sigma}_y$  where  $\boldsymbol{\sigma}_y \sim \mathcal{N}(0, \sigma_y^2 \mathbf{I})$ .

**Unbiased substitutions.** Because we never observe the latent projection  $\mathbf{A}\mathbf{x}_0$ , we replace its quadratic forms with statistics that depend only on the noisy measurement  $\mathbf{y}$  and  $m := \dim(\mathbf{y})$ :

$$\|\mathbf{A}\mathbf{x}_0\|^2 = \|\mathbf{y}\|^2 - m\sigma_y^2, \quad (22)$$

$$(\mathbf{A}\mathbf{x}_0)^\top \Sigma (\mathbf{A}\mathbf{x}_0) = \mathbf{y}^\top \Sigma \mathbf{y} - \sigma_y^2 \text{tr} \Sigma, \quad (23)$$

with  $\Sigma := \mathbf{A}\mathbf{A}^\top$ .

**First two moments of the residual energy.** Using the forward-diffusion decomposition  $\mathbf{x}_t = \sqrt{\bar{\alpha}_t} \mathbf{x}_0 + \sqrt{1 - \bar{\alpha}_t} \boldsymbol{\epsilon}_t$ ,  $\boldsymbol{\epsilon}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ , one finds the residual  $\mathbf{A}\mathbf{x}_t - \mathbf{y} = (\sqrt{\bar{\alpha}_t} - 1)\mathbf{A}\mathbf{x}_0 + \sqrt{1 - \bar{\alpha}_t} \mathbf{A}\boldsymbol{\epsilon}_t - \boldsymbol{\sigma}_y$ . After substituting the unbiased identities Equation (22)–Equation (23) and taking expectations over both noise sources  $\boldsymbol{\epsilon}_t$  and  $\boldsymbol{\sigma}_y$ , we obtain closed-form expressions that are fully observable.

**Expectation.**

$$\mathbb{E}[R_t | \mathbf{y}] = \underbrace{(1 - \bar{\alpha}_t) \text{tr}(\Sigma)}_{\text{diffusion noise}} + \underbrace{m\sigma_y^2[1 - (\sqrt{\bar{\alpha}_t} - 1)^2]}_{\text{measurement noise}} + \underbrace{(\sqrt{\bar{\alpha}_t} - 1)^2 \|\mathbf{y}\|^2}_{\text{deterministic bias}}. \quad (24)$$

**Variance.** Writing  $\tilde{\Sigma}_t = (1 - \bar{\alpha}_t)\Sigma + \sigma_y^2 \mathbf{I}$  and  $\tilde{\boldsymbol{\mu}}_t = (\sqrt{\bar{\alpha}_t} - 1)\mathbf{y}$ , the non-central  $\chi^2$  moment formula  $\text{Var}(Q_t) = 2 \text{tr}(\tilde{\Sigma}_t^2) + 4\tilde{\boldsymbol{\mu}}_t^\top \tilde{\Sigma}_t \tilde{\boldsymbol{\mu}}_t$  gives

$$\begin{aligned} \text{Var}[R_t | \mathbf{y}] &= 2 \left[ (1 - \bar{\alpha}_t)^2 \text{tr}(\Sigma^2) + 2(1 - \bar{\alpha}_t) \sigma_y^2 \text{tr} \Sigma + m\sigma_y^4 \right] \\ &\quad + 4(\sqrt{\bar{\alpha}_t} - 1)^2 \left[ (1 - \bar{\alpha}_t) (\mathbf{y}^\top \Sigma \mathbf{y} - \sigma_y^2 \text{tr} \Sigma) + \sigma_y^2 (\|\mathbf{y}\|^2 - m\sigma_y^2) \right]. \end{aligned} \quad (25)$$

## C Convergence of Constraint Satisfaction

We now analyze the convergence properties of the constraint satisfaction procedure in CDIM for the noiseless case. The algorithm alternates between unconditional diffusion updates and projection steps:

1. Unconditional update:  $f_\theta(\mathbf{x}_t) = \text{DDIM.step}(x_t)$
2. Projection:  
 $\mathbf{x}_{t-\delta} = \arg \min_{\mathbf{x}_{t-\delta}} \|\mathbf{x}_{t-\delta} - f_\theta(\mathbf{x}_t)\|^2 \text{ s.t. } \mathbf{A}\hat{\mathbf{x}}_0 = \mathbf{y}$

where  $\hat{\mathbf{x}}_0$  denotes the Tweedie estimate  $E[\mathbf{x}_0 | \mathbf{x}_t]$  at timestep  $t$ . When the constraint is infeasible, we perform gradient descent on  $\|\mathbf{A}\hat{\mathbf{x}}_0 - \mathbf{y}\|^2$ .

We first show that the Tweedie estimate converges to the identity mapping, and the rate of convergence. Given that, we show that as  $t \rightarrow 0$ , finding  $\mathbf{x}_t$  s.t.  $\|\mathbf{A}\hat{\mathbf{x}}_0 - \mathbf{y}\|^2 = 0$  is feasible and satisfiable via the proposed gradient descent algorithm.

**Tweedie Convergence** In this section, we show that as  $t \rightarrow 0$ , the Tweedie estimate converges to the identity mapping:

$$\sup_{\mathbf{x}_t} \|\hat{\mathbf{x}}_0 - \mathbf{x}_t\|_2 \leq \varepsilon(t) \quad (26)$$

where  $\varepsilon(t) \rightarrow 0$  as  $t \rightarrow 0$ .

Consider the forward diffusion process:

$$\mathbf{x}_t = \sqrt{1 - \beta_t} \mathbf{x}_0 + \sqrt{\beta_t} \boldsymbol{\epsilon}, \quad \boldsymbol{\epsilon} \sim \mathcal{N}(0, \mathbf{I}) \quad (27)$$

330 The Tweedie estimate (posterior mean) is given by:

$$\hat{\mathbf{x}}_0 = \mathbb{E}[\mathbf{x}_0|\mathbf{x}_t] = \frac{\mathbf{x}_t}{\sqrt{1-\beta_t}} - \frac{\beta_t}{\sqrt{1-\beta_t}} \nabla_{\mathbf{x}_t} \log p(\mathbf{x}_t) \quad (28)$$

331 For small  $\beta_t$  (as  $t \rightarrow 0$ ), we perform a Taylor expansion:

$$\frac{1}{\sqrt{1-\beta_t}} = 1 + \frac{\beta_t}{2} + O(\beta_t^2) \quad (29)$$

$$\frac{\beta_t}{\sqrt{1-\beta_t}} = \beta_t(1 + \frac{\beta_t}{2}) + O(\beta_t^3) \quad (30)$$

332 Substituting into  $\hat{\mathbf{x}}_0$ :

$$\hat{\mathbf{x}}_0 = \mathbf{x}_t \left(1 + \frac{\beta_t}{2}\right) - \beta_t \left(1 + \frac{\beta_t}{2}\right) \nabla_{\mathbf{x}_t} \log p(\mathbf{x}_t) + O(\beta_t^2) \quad (31)$$

$$= \mathbf{x}_t + \frac{\beta_t}{2} \mathbf{x}_t - \beta_t \nabla_{\mathbf{x}_t} \log p(\mathbf{x}_t) + O(\beta_t^2) \quad (32)$$

333 The deviation from  $\mathbf{x}_t$  is:

$$\hat{\mathbf{x}}_0 - \mathbf{x}_t = \frac{\beta_t}{2} \mathbf{x}_t - \beta_t \nabla_{\mathbf{x}_t} \log p(\mathbf{x}_t) + O(\beta_t^2) \quad (33)$$

334 Taking norms and applying triangle inequality:

$$\|\hat{\mathbf{x}}_0 - \mathbf{x}_t\|_2 \leq \frac{\beta_t}{2} \|\mathbf{x}_t\|_2 + \beta_t \|\nabla_{\mathbf{x}_t} \log p(\mathbf{x}_t)\|_2 + O(\beta_t^2) \quad (34)$$

$$\leq \beta_t \left( \frac{1}{2} \|\mathbf{x}_t\|_2 + \|\nabla_{\mathbf{x}_t} \log p(\mathbf{x}_t)\|_2 \right) + O(\beta_t^2) \quad (35)$$

335 Under standard assumptions for diffusion models:

- 336 1. Bounded data:  $\exists R > 0$  such that  $\|\mathbf{x}_0\|_2 \leq R$  almost surely
- 337 2. Lipschitz score:  $\exists L > 0$  such that  $\|\nabla \log p(\mathbf{x}_t)\|_2 \leq L\|\mathbf{x}_t\|_2 + C$

338 These ensure the terms in parentheses remain bounded. Therefore, for  $0 \leq \beta_t \leq 1$ :

$$\|\hat{\mathbf{x}}_0 - \mathbf{x}_t\|_2 \leq C\beta_t + O(\beta_t^2) = O(\beta_t) \quad (36)$$

339 We show this empirically in Figure 7. This demonstrates that the Tweedie's estimate of the posterior mean converges to the identity mapping in a predictable way.

341 **Convergence of Constraint Satisfaction** Based on the convergence of the tweedie's estimate to  $x_t$ ,  
342 we show that for sufficiently small  $t$ :

- 343 1. The constraint set  $\mathbf{x}_t : \mathbf{A}\hat{\mathbf{x}}_0 = \mathbf{y}$  is non-empty
- 344 2. Gradient descent on  $\|\mathbf{A}\hat{\mathbf{x}}_0 - \mathbf{y}\|^2$  converges to a point satisfying the constraint

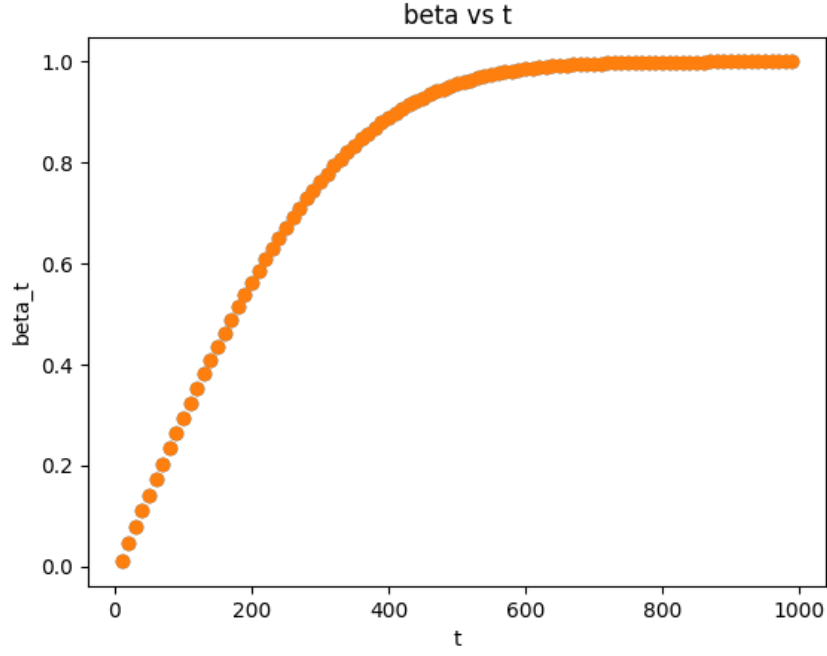
345 First, we show that the optimization landscape becomes increasingly well-behaved as  $t \rightarrow 0$ . Con-  
346 sider the objective:

$$\|\mathbf{A}\hat{\mathbf{x}}_0 - \mathbf{y}\|^2 = \|\mathbf{A}\mathbf{x}_t - \mathbf{y} + \mathbf{A}(\hat{\mathbf{x}}_0 - \mathbf{x}_t)\|^2 \quad (37)$$

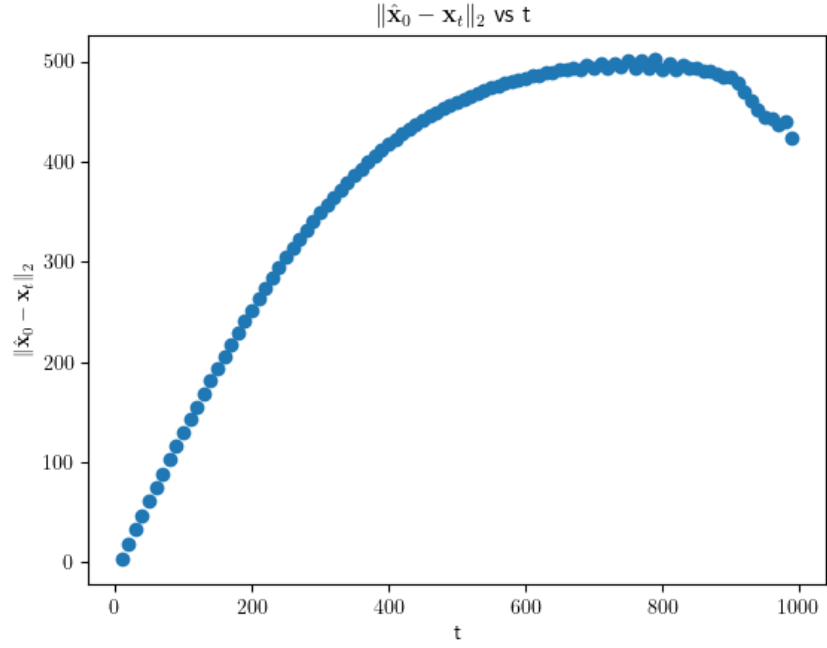
$$\begin{aligned} &= \|\mathbf{A}\mathbf{x}_t - \mathbf{y}\|^2 \\ &\quad + 2\langle \mathbf{A}\mathbf{x}_t - \mathbf{y}, \mathbf{A}(\hat{\mathbf{x}}_0 - \mathbf{x}_t) \rangle \\ &\quad + \|\mathbf{A}(\hat{\mathbf{x}}_0 - \mathbf{x}_t)\|^2 \end{aligned} \quad (38)$$

347 By Tweedie Convergence, the second and third terms are bounded by  $O(\varepsilon(t))$ . Therefore, as  $t \rightarrow 0$ ,  
348 the objective converges to the convex quadratic  $\|\mathbf{A}\mathbf{x}_t - \mathbf{y}\|^2$  which can be optimized with gradient  
349 descent.





(a)



(b)

Figure 7: We show that  $\|\hat{\mathbf{x}}_0 - \mathbf{x}_t\|_2 \leq O(\beta_t)$ , demonstrating that the Tweedie’s estimate of the posterior mean converges to the identity mapping in a predictable way. (a) plots  $\beta_t$  (b) plots  $\|\hat{\mathbf{x}}_0 - \mathbf{x}_t\|_2$  for a Gaussian deblur task. Note that at high values of  $t$ , the diffusion model is a poor estimator of  $\nabla_{\mathbf{x}_t} \log p(\mathbf{x}_t)$  leading to results that don’t follow the convergence pattern.

For feasibility, note that as  $t \rightarrow 0$ , finding  $\mathbf{x}_t$  such that  $\mathbf{A}\hat{\mathbf{x}}_0(\mathbf{x}_t) = \mathbf{y}$  becomes equivalent to finding  $\mathbf{x}_t$  such that  $\mathbf{A}\mathbf{x}_t = \mathbf{y}$  up to an error of  $O(\varepsilon(t))$ . The latter is feasible whenever  $\mathbf{y}$  is in the range of  $\mathbf{A}$ , which is the standard assumption for linear inverse problems.

As  $t \rightarrow 0$ ,  $\varepsilon(t)$  approaches zero, making the constraint feasible. Moreover, since the objective approaches a convex quadratic, gradient descent will converge to the global minimum for sufficiently small  $t$ .

## D Additional Experimental Details

### D.1 Task Details

We describe additional details for each inverse task used in our experiments.

**Super Resolution** Images are downsampled to  $64 \times 64$  using bicubic downsampling with a factor of 4.

**Box Inpainting** A random box of size  $128 \times 128$  is chosen uniformly within the image. Those pixels are masked out affected all three of the RGB channels.

**Gaussian Deblur** A Gaussian Kernel of size  $61 \times 61$  and intensity 3 is applied to the entire image.

**Random Inpainting** Each pixel is masked out with probability 92% affecting all three of the RGB channels

**50% Inpainting** In various figures, we showcase a a 50% inpainting task where the top half of an image is masked out. This task is more challenging than box inpainting and can better illustrate differences between results.

### D.2 Measuring Runtime

To measure wall-clock runtime, we used a single A100 and ran all the inverse problems (super-resolution, box inpainting, gaussian deblur, random inpainting) on the FFHQ dataset. We only consider the runtime of the algorithm, without considering the python initialization time, model loading, or image io. For each task, we measured the runtime on 10 images and averaged the result to produce the final result. We note that the baseline runtimes are taken from [15], where only the box inpainting task was considered. The runtime does not vary much between tasks when using CDIM, so we report our average runtime across tasks as a fair comparison metric.

### D.3 Comparison with DSG

We show a qualitative comparison against DSG [25] on 3 tasks in Figure 8. We used the official code from their github, and generated results with 25 DDIM diffusion steps for both DSG and CDIM (and  $K = 1$  for CDIM). As you can see, the DSG results are blurrier and sometimes contain artifacts

### D.4 ImageNet Results

In Table 5 we report FID and LPIPS for ImageNet.

### D.5 PSNR Results

See Tables 4 and 5

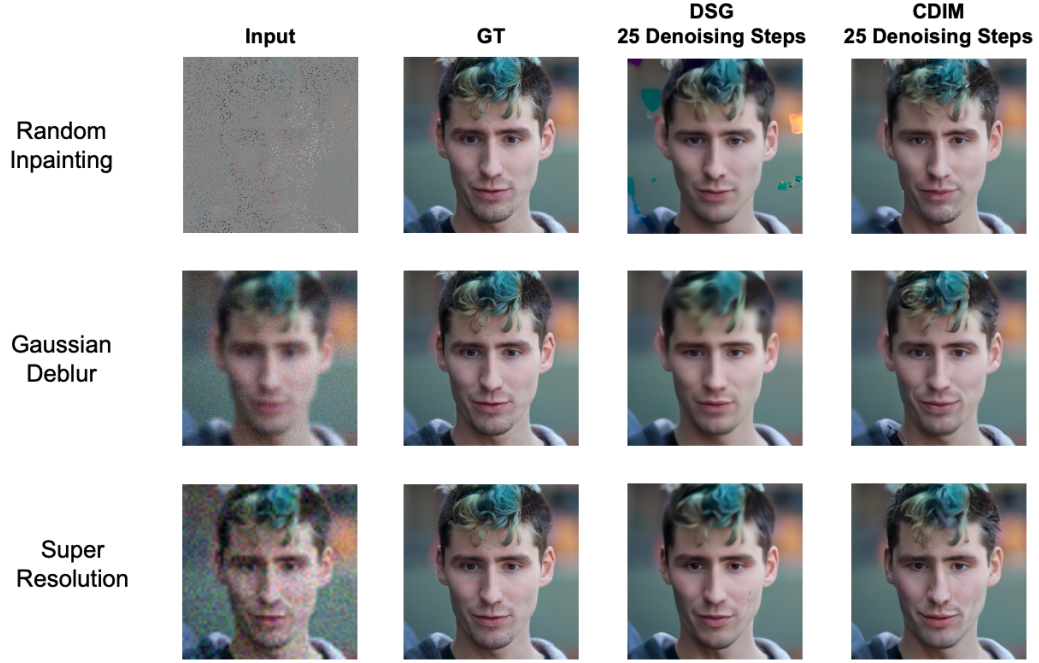


Figure 8: A comparison between DSG [25] and CDIM when both algorithms use 25 DDIM denoising steps. Notice the artifacts in the DSG random inpainting.

Table 3: Quantitative results (FID, LPIPS) of our model and existing models on various linear inverse problems on the Imagenet  $256 \times 256$ -1k validation dataset. (Lower is better)

Imagenet	Super Resolution		Inpainting (box)		Gaussian Deblur		Inpainting (random)	
	FID	LPIPS	FID	LPIPS	FID	LPIPS	FID	LPIPS
Ours - $T^* = 25$	53.70	0.378	52.00	0.267	56.10	0.393	51.96	0.370
Ours - $T^* = 50$	47.45	0.339	50.31	0.251	38.69	0.347	46.20	0.332
FPS-SMC	47.30	0.316	33.24	0.212	54.21	0.403	42.77	0.328
DPS	50.66	0.337	38.82	0.262	62.72	0.444	35.87	0.303
DDRM	59.57	0.339	45.95	0.245	63.02	0.427	114.9	0.665
MCG	144.5	0.637	39.74	0.330	95.04	0.550	39.19	0.414
PnP-ADMM	97.27	0.433	78.24	0.367	100.6	0.519	114.7	0.677
Score-SDE	170.7	0.701	54.07	0.354	120.3	0.667	127.1	0.659
ADMM-TV	130.9	0.523	87.69	0.319	155.7	0.588	189.3	0.510

Table 4: Quantitative results (PSNR) of our model and existing models on various linear inverse problems on the FFHQ 256-1k validation dataset. (Higher is better)

Imagenet	Super Resolution	Inpainting (box)	Gaussian Deblur	Inpainting (random)
	PSNR	PSNR	PSNR	PSNR
Ours - $T^* = 25$	27.08	23.20	26.77	26.49
Ours - $T^* = 50$	27.30	23.47	27.03	27.10
FPS-SMC	28.10	24.70	26.54	27.33
DPS	25.67	22.47	24.25	25.23
DDRM	25.36	22.24	23.36	9.19
MCG	20.05	19.97	6.72	21.57
PnP-ADMM	26.55	11.65	24.93	8.41
Score-SDE	17.62	18.51	7.21	13.52
ADMM-TV	23.86	17.81	22.37	22.03

Table 5: Quantitative results (PSNR) of our model and existing models on various linear inverse problems on the Imagenet  $256 \times 256$ -1k validation dataset. (Higher is better)

<b>Imagenet</b>	<b>Super Resolution</b>	<b>Inpainting (box)</b>	<b>Gaussian Deblur</b>	<b>Inpainting (random)</b>
Methods	PSNR	PSNR	PSNR	PSNR
Ours - $T' = 25$	23.67	19.67	22.78	22.38
Ours - $T' = 50$	23.92	20.06	23.32	22.61
FPS-SMC	24.78	22.03	23.81	24.12
DPS	23.87	18.90	21.97	22.20
DDRM	24.96	18.66	22.73	14.29
MCG	13.39	17.36	16.32	19.03
PnP-ADMM	23.75	12.70	21.81	8.39
Score-SDE	12.25	16.48	15.97	18.62
ADMM-TV	22.17	17.96	19.99	20.96



Figure 9: FFHQ Super-resolution extended results

## 390 D.6 Extended Results





Figure 10: FFHQ Gaussian deblur extended results



Figure 11: FFHQ random inpainting extended results



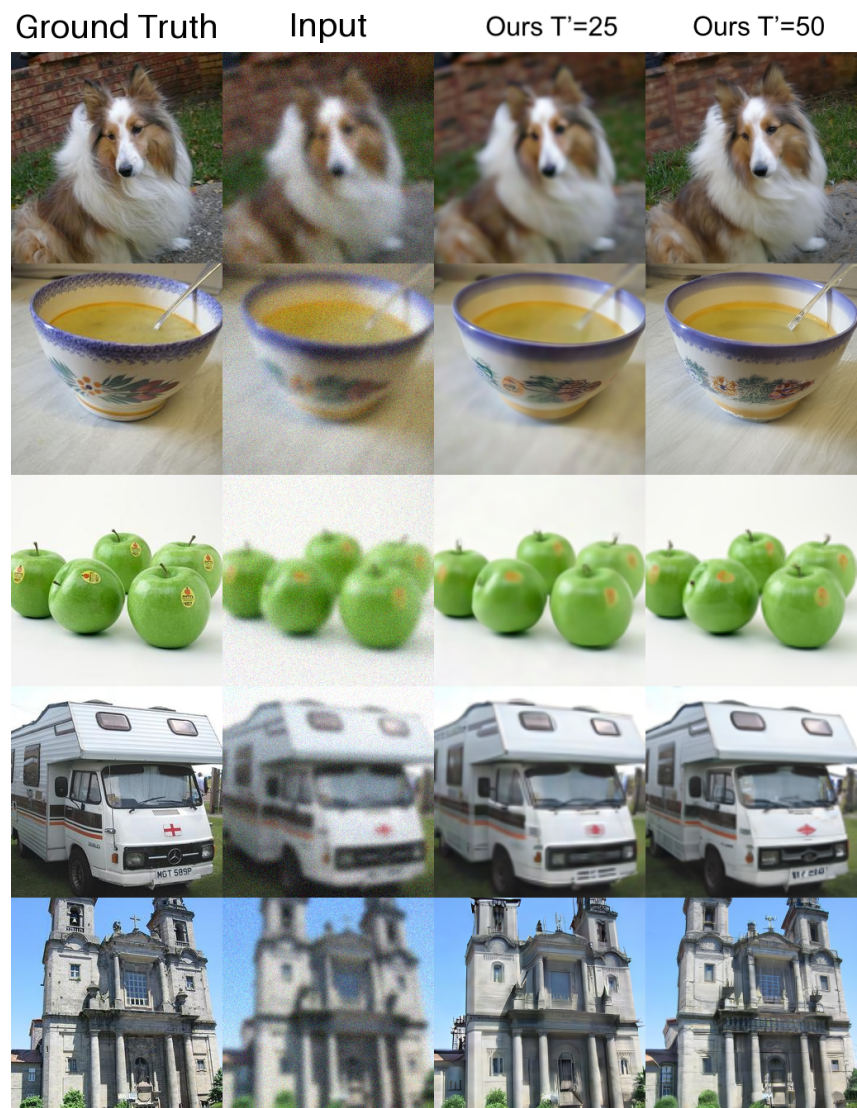


Figure 12: ImageNet Gaussian deblur extended results



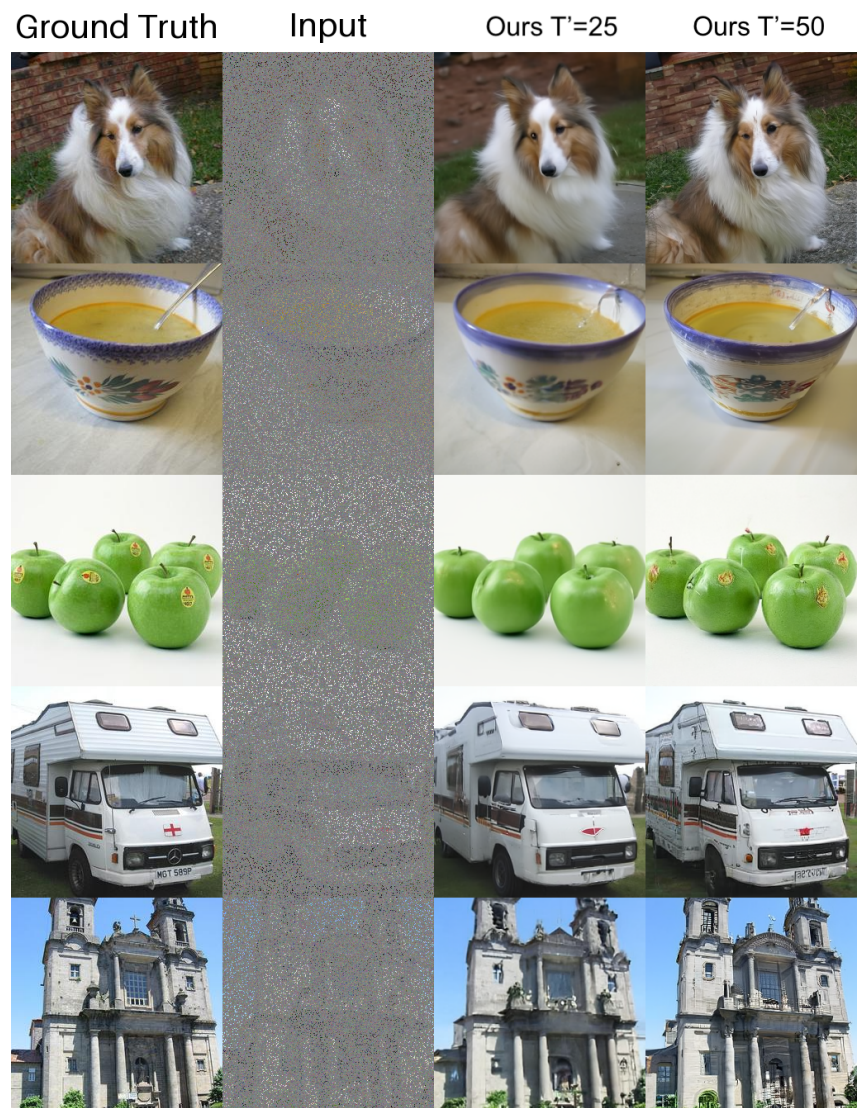


Figure 13: ImageNet random inpainting extended results

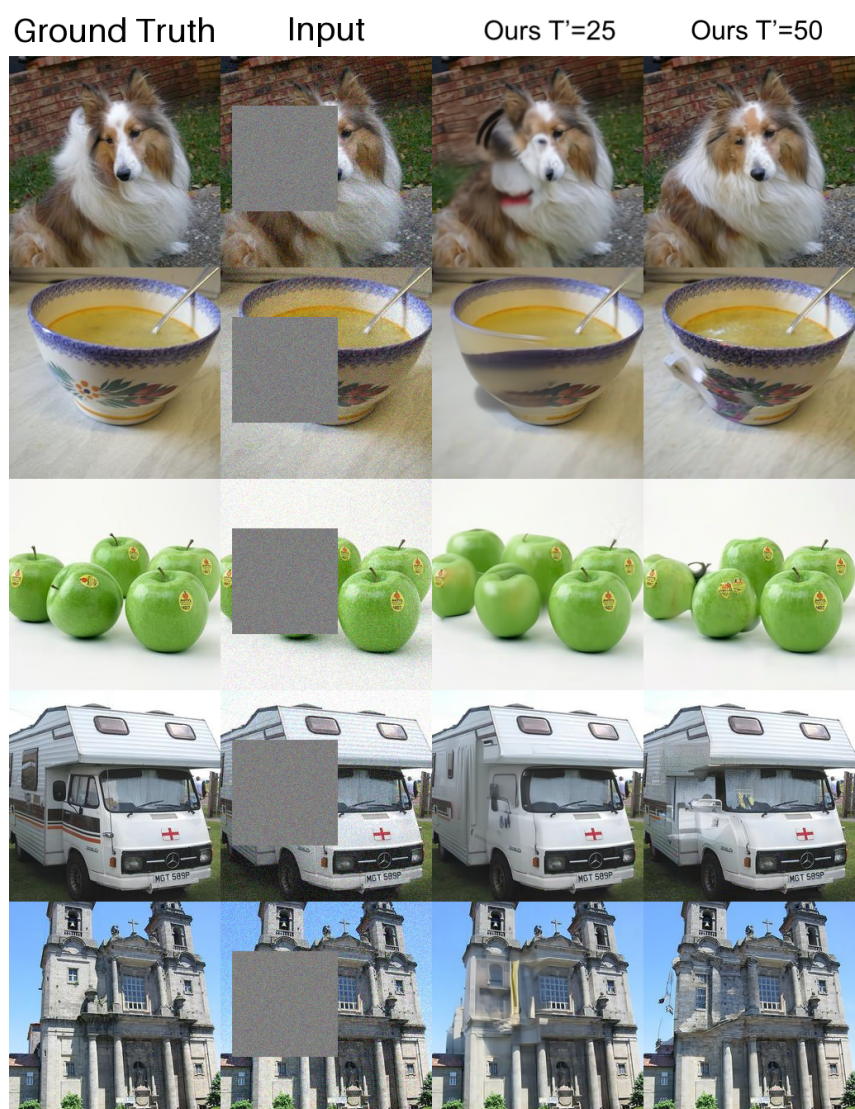


Figure 14: ImageNet box inpainting extended results



Figure 15: Results on inpainting 50% of an image on LSUN Churches dataset.

## References

- [1] Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. *Advances in Neural Information Processing Systems*, 2020. [1](#), [2](#), [3](#), [1](#)
- [2] 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*, 2022. [1](#), [2](#), [4.2](#), [5](#), [5.1](#), [5.2](#)
- [3] Hyungjin Chung, Byeongsu Sim, Dohoon Ryu, and Jong Chul Ye. Improving diffusion models for inverse problems using manifold constraints, 2024. [1](#), [2](#), [4.2](#)
- [4] Yaxuan Zhu, Zehao Dou, Haoxin Zheng, Yasi Zhang, Ying Nian Wu, and Ruiqi Gao. Think twice before you act: Improving inverse problem solving with mcmc, 2024. [1](#), [2](#)
- [5] Jascha Sohl-Dickstein, Eric Weiss, Niru Maheswaranathan, and Surya Ganguli. Deep unsupervised learning using nonequilibrium thermodynamics. In *International conference on machine learning*, 2015. [2](#)
- [6] Shakir Mohamed and Balaji Lakshminarayanan. Learning in implicit generative models. 2017. [2](#)
- [7] Jiaming Song, Chenlin Meng, and Stefano Ermon. Denoising diffusion implicit models. In *International Conference on Learning Representations*, 2021. [2](#), [3](#), [1](#)
- [8] Yang Song, Jascha Sohl-Dickstein, Diederik P Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-based generative modeling through stochastic differential equations. In *International Conference on Learning Representations*, 2020. [2](#)
- [9] Luping Liu, Yi Ren, Zhijie Lin, and Zhou Zhao. Pseudo numerical methods for diffusion models on manifolds. In *International Conference on Learning Representations*, 2021. [2](#)
- [10] Yinhuai Wang, Jiwen Yu, and Jian Zhang. Zero-shot image restoration using denoising diffusion null-space model. In *The Eleventh International Conference on Learning Representations*, 2022. [2](#)
- [11] Bahjat Kawar, Michael Elad, Stefano Ermon, and Jiaming Song. Denoising diffusion restoration models, 2022. [2](#)
- [12] Bahjat Kawar, Gregory Vaksman, and Michael Elad. Snips: Solving noisy inverse problems stochastically, 2021. [2](#)
- [13] Jiaming Song, Arash Vahdat, Morteza Mardani, and Jan Kautz. Pseudoinverse-guided diffusion models for inverse problems. In *International Conference on Learning Representations*, 2023. [2](#)
- [14] Xiangming Meng and Yoshiyuki Kabashima. Diffusion model based posterior sampling for noisy linear inverse problems. *arXiv preprint arXiv:2211.12343*, 2022. [2](#)
- [15] Zehao Dou and Yang Song. Diffusion posterior sampling for linear inverse problem solving: A filtering perspective. In *The Twelfth International Conference on Learning Representations*, 2023. [2](#), [D.2](#)
- [16] Jiaming Song, Qinsheng Zhang, Hongxu Yin, Morteza Mardani, Ming-Yu Liu, Jan Kautz, Yongxin Chen, and Arash Vahdat. Loss-guided diffusion models for plug-and-play controllable generation. In Andreas Krause, Emma Brunskill, Kyunghyun Cho, Barbara Engelhardt, Sivan Sabato, and Jonathan Scarlett, editors, *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pages 32483–32498. PMLR, 23–29 Jul 2023. [2](#)
- [17] Gabriel Cardoso, Yazid Janati El Idrissi, Sylvain Le Corff, and Eric Moulines. Monte carlo guided diffusion for bayesian linear inverse problems, 2023. [2](#)
- [18] Prafulla Dhariwal and Alex Nichol. Diffusion models beat gans on image synthesis, 2021. [2](#), [5.1](#)



- [19] Hyungjin Chung, Byeongsu Sim, and Jong Chul Ye. Come-closer-diffuse-faster: Accelerating conditional diffusion models for inverse problems through stochastic contraction, 2022. [2](#)
- [20] Jooyoung Choi, Sungwon Kim, Yonghyun Jeong, Youngjune Gwon, and Sungroh Yoon. Ilvr: Conditioning method for denoising diffusion probabilistic models, 2021. [2](#)
- [21] Benjamin Boys, Mark Girolami, Jakiw Pidstrigach, Sebastian Reich, Alan Mosca, and O. Deniz Akyildiz. Tweedie moment projected diffusions for inverse problems, 2023. [2](#)
- [22] Stanley H. Chan, Xiran Wang, and Omar A. Elgendy. Plug-and-play admm for image restoration: Fixed point convergence and applications, 2016. [2](#)
- [23] Bo Wahlberg, Stephen Boyd, Mariette Annergren, and Yang Wang. An admm algorithm for a class of total variation regularized estimation problems, 2012. [2](#)
- [24] Hengkang Wang, Xu Zhang, Taihui Li, Yuxiang Wan, Tiancong Chen, and Ju Sun. Dmplug: A plug-in method for solving inverse problems with diffusion models, 2024. [2](#)
- [25] Lingxiao Yang, Shutong Ding, Yifan Cai, Jingyi Yu, Jingya Wang, and Ye Shi. Guidance with spherical gaussian constraint for conditional diffusion, 2024. [2](#) [5.2](#) [D.3](#) [8](#)
- [26] Hyungjin Chung, Jeongsol Kim, Sehui Kim, and Jong Chul Ye. Parallel diffusion models of operator and image for blind inverse problems, 2022. [2](#)
- [27] Charles Laroche, Andrés Almansa, and Eva Coupete. Fast diffusion em: a diffusion model for blind inverse problems with application to deconvolution, 2023. [2](#)
- [28] Bradley Efron. Tweedie’s formula and selection bias. *Journal of the American Statistical Association*, 106(496):1602–1614, 2011. [3](#)
- [29] Giorgio Giannone, Akash Srivastava, Ole Winther, and Faez Ahmed. Aligning optimization trajectories with diffusion models for constrained design generation. In *Thirty-seventh Conference on Neural Information Processing Systems*, 2023. [4.2](#)
- [30] Michael F Hutchinson. A stochastic estimator of the trace of the influence matrix for laplacian smoothing splines. *Communications in Statistics-Simulation and Computation*, 18(3):1059–1076, 1989. [4.2](#)
- [31] Daryl Pregibon. Logistic regression diagnostics. *Annals of Statistics*, 9:705–724, 1981. [4.4](#)
- [32] Donald A. Pierce and Daniel W. Schafer. Residuals in generalized linear models. *Journal of the American Statistical Association*, 81(396):977–986, 1986. [4.4](#)
- [33] Tero Karras, Samuli Laine, and Timo Aila. A style-based generator architecture for generative adversarial networks, 2019. [5.1](#)
- [34] Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, Alexander C. Berg, and Li Fei-Fei. Imagenet large scale visual recognition challenge, 2015. [5.1](#)
- [35] Martin Heusel, Hubert Ramsauer, Thomas Unterthiner, Bernhard Nessler, and Sepp Hochreiter. Gans trained by a two time-scale update rule converge to a local nash equilibrium, 2018. [5.1](#)
- [36] Richard Zhang, Phillip Isola, Alexei A. Efros, Eli Shechtman, and Oliver Wang. The unreasonable effectiveness of deep features as a perceptual metric, 2018. [5.1](#)
- [37] Xuan Luo, Xuaner (Cecilia) Zhang, Paul Yoo, Ricardo Martin-Brualla, Jason Lawrence, and Steven M. Seitz. Time-travel rephotography. *ACM Transactions on Graphics*, 40(6):1–12, December 2021. [5.4](#)
- [38] Johannes L. Schönberger and Jan-Michael Frahm. Structure-from-motion revisited. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 4104–4113, 2016. [5.4](#)

- 482 [39] D. L. Hanson and F. T. Wright. A Bound on Tail Probabilities for Quadratic Forms in Inde-  
483 pendent Random Variables. *The Annals of Mathematical Statistics*, 42(3):1079 – 1083, 1971.  
484 [A](#)
- 485 [40] Mark Rudelson and Roman Vershynin. Hanson-wright inequality and sub-gaussian concentra-  
486 tion, 2013. [A](#)

## NeurIPS Paper Checklist

### 1. Claims

Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

Answer: [\[Yes\]](#)

Justification: Yes, we discuss the accelerated projection algorithm and discuss how it speeds up constrained diffusion compared to existing methods. We focus on our contribution of speedup

Guidelines:

- The answer NA means that the abstract and introduction do not include the claims made in the paper.
- The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A No or NA answer to this question will not be perceived well by the reviewers.
- The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
- It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

### 2. Limitations

Question: Does the paper discuss the limitations of the work performed by the authors?

Answer: [\[Yes\]](#)

Justification: We discuss a few limitations such as not handling latent diffusion. However there could be more discussion of overall limitations

Guidelines:

- The answer NA means that the paper has no limitation while the answer No means that the paper has limitations, but those are not discussed in the paper.
- The authors are encouraged to create a separate "Limitations" section in their paper.
- The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
- The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
- The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
- The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
- If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
- While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

### 3. Theory assumptions and proofs

Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

Answer: [\[Yes\]](#)

Justification: Yes, our main claim (Proposition 1) is fully derived in the appendix. Other claims, such as the chi-squared distribution of residuals are derived in sufficient depth as they are straightforward.

Guidelines:

- The answer NA means that the paper does not include theoretical results.
- All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
- All assumptions should be clearly stated or referenced in the statement of any theorems.
- The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
- Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
- Theorems and Lemmas that the proof relies upon should be properly referenced.

#### 4. Experimental result reproducibility

Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

Answer: [\[Yes\]](#)

Justification: We describe all experimental settings, hardware considerations, hyperparameters, and other requirements for running the experiments. Code will be released.

Guidelines:

- The answer NA means that the paper does not include experiments.
- If the paper includes experiments, a No answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
- If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
- Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general, releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
- While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
  - (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
  - (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
  - (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).
  - (d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

## 5. Open access to data and code

Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

Answer: [Yes]

Justification: Code will be open sourced shortly after the review deadline.

Guidelines:

- The answer NA means that paper does not include experiments requiring code.
- Please see the NeurIPS code and data submission guidelines (<https://nips.cc/public/guides/CodeSubmissionPolicy>) for more details.
- While we encourage the release of code and data, we understand that this might not be possible, so “No” is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
- The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (<https://nips.cc/public/guides/CodeSubmissionPolicy>) for more details.
- The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
- The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
- At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
- Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

## 6. Experimental setting/details

Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer, etc.) necessary to understand the results?

Answer: [Yes]

Justification: Yes, we have discussed the training and test datasets, hyperparameters, stopping criteria, etc.

Guidelines:

- The answer NA means that the paper does not include experiments.
- The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
- The full details can be provided either with the code, in appendix, or as supplemental material.

## 7. Experiment statistical significance

Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

Answer: [Yes]

Justification: We report many metrics like FID, LPIPS, PSNR, etc. These are standard in the community.

Guidelines:

- The answer NA means that the paper does not include experiments.
- The authors should answer “Yes” if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.



- The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
- The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
- The assumptions made should be given (e.g., Normally distributed errors).
- It should be clear whether the error bar is the standard deviation or the standard error of the mean.
- It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
- For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g. negative error rates).
- If error bars are reported in tables or plots, The authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

## 8. Experiments compute resources

Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

Answer: [Yes]

Justification: Yes we describe the hardware used and execution time.

Guidelines:

- The answer NA means that the paper does not include experiments.
- The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
- The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
- The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

## 9. Code of ethics

Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics <https://neurips.cc/public/EthicsGuidelines>?

Answer: [Yes]

Justification: We have done our best to follow all ethical guidelines.

Guidelines:

- The answer NA means that the authors have not reviewed the NeurIPS Code of Ethics.
- If the authors answer No, they should explain the special circumstances that require a deviation from the Code of Ethics.
- The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

## 10. Broader impacts

Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

Answer: [NA]

Justification: There are some negative societal of diffusion posterior sampling but we do not believe it warrants a full broader impact.

Guidelines:

- The answer NA means that there is no societal impact of the work performed.

- If the authors answer NA or No, they should explain why their work has no societal impact or why the paper does not address societal impact.
- Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.
- The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
- The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

## 11. Safeguards

Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pretrained language models, image generators, or scraped datasets)?

Answer: [NA]

Justification: We do not believe safeguards are required for this work.

Guidelines:

- The answer NA means that the paper poses no such risks.
- Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
- Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
- We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

## 12. Licenses for existing assets

Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

Answer: [NA]

Justification: All IP is from standard community datasets.

Guidelines:

- The answer NA means that the paper does not use existing assets.
- The authors should cite the original paper that produced the code package or dataset.
- The authors should state which version of the asset is used and, if possible, include a URL.
- The name of the license (e.g., CC-BY 4.0) should be included for each asset.
- For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.

- If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, [paperswithcode.com/datasets](https://paperswithcode.com/datasets) has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
- For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.
- If this information is not available online, the authors are encouraged to reach out to the asset's creators.

### 13. New assets

Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

Answer: [NA]

Justification: We do not release no assets

Guidelines:

- The answer NA means that the paper does not release new assets.
- Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
- The paper should discuss whether and how consent was obtained from people whose asset is used.
- At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

### 14. Crowdsourcing and research with human subjects

Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

Answer: [NA]

Justification: No crowdsourcing was done

Guidelines:

- The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
- Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
- According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

### 15. Institutional review board (IRB) approvals or equivalent for research with human subjects

Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

Answer: [NA]

Justification: No human subjects were used in experiments.

Guidelines:

- The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
- Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.

- 798           • We recognize that the procedures for this may vary significantly between institutions  
799           and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the  
800           guidelines for their institution.  
801           • For initial submissions, do not include any information that would break anonymity  
802           (if applicable), such as the institution conducting the review.

803 **16. Declaration of LLM usage**

804 Question: Does the paper describe the usage of LLMs if it is an important, original, or  
805 non-standard component of the core methods in this research? Note that if the LLM is used  
806 only for writing, editing, or formatting purposes and does not impact the core methodology,  
807 scientific rigorousness, or originality of the research, declaration is not required.

808 Answer: [NA]

809 Justification: We do not use LLMs outside of helping with basic writing/editing.

810 Guidelines:

- 811           • The answer NA means that the core method development in this research does not  
812           involve LLMs as any important, original, or non-standard components.  
813           • Please refer to our LLM policy (<https://neurips.cc/Conferences/2025/LLM>)  
814           for what should or should not be described.