

000
001
002
003
004
005
006
007
008
009
010
011
012
013
014
015
016
017
018
019
020
021
022
023
024
025
026
027
028
029
030
031
032
033
034
035
036
037
038
039
040
041
042
043
044
045
046
047
048
049
050
051
052
053

ANTITHETIC NOISE IN DIFFUSION MODELS

Anonymous authors

Paper under double-blind review

ABSTRACT

We systematically study antithetic initial noise in diffusion models, discovering that pairing each noise sample with its negation consistently produces strong negative correlation. This universal phenomenon holds across datasets, model architectures, conditional and unconditional sampling, and even other generative models such as VAEs and Normalizing Flows. To explain it, we combine experiments and theory and propose a *symmetry conjecture* that the learned score function is approximately affine antisymmetric (odd symmetry up to a constant shift), supported by empirical evidence. This negative correlation leads to substantially more reliable uncertainty quantification with up to 90% narrower confidence intervals. We demonstrate these gains on tasks including estimating pixel-wise statistics and evaluating diffusion inverse solvers. We also provide extensions with randomized quasi-Monte Carlo noise designs for uncertainty quantification, and explore additional applications of the antithetic noise design to improve image editing and generation diversity. Our framework is training-free, model-agnostic, and adds no runtime overhead. Code is available at <https://anonymous.4open.science/r/Antithetic-Noise-in-Diffusion-Models-8B54>.

1 INTRODUCTION

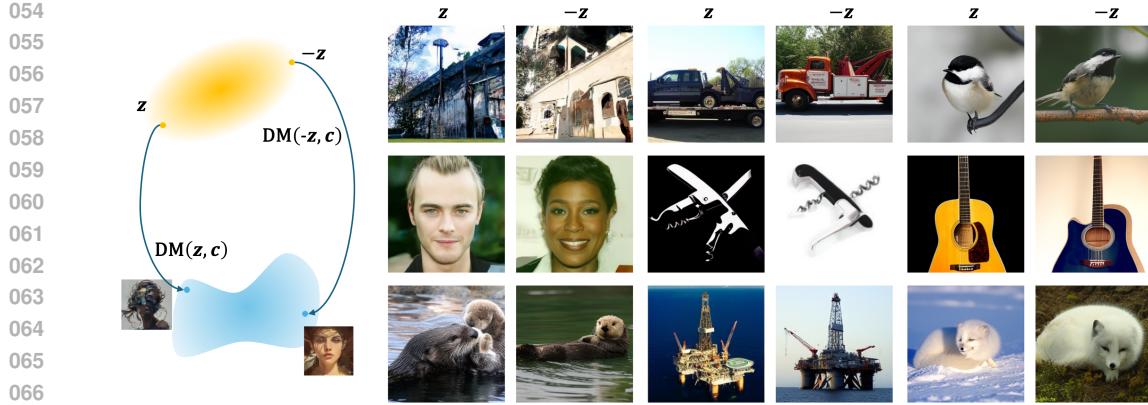
Diffusion models have set the state of the art in photorealistic image synthesis, high-fidelity audio, and video generation (Sohl-Dickstein et al., 2015; Ho et al., 2020; Song et al., 2021b; Kong et al., 2021); they also power applications such as text-to-image generation (Rombach et al., 2022), image editing and restoration (Meng et al., 2022), and inverse problem solving (Song et al., 2022).

For many pretrained diffusion models, sampling relies on three elements: the network weights, the denoising schedule, and the initial Gaussian noise. Once these are fixed, sampling is often deterministic: the sampler transforms the initial noise into an image via successive denoising passes.

Much of the literature improves the first two ingredients and clusters into two strands: (i) architectural and training developments, which improve sample quality and scalability through backbone or objective redesign (e.g., EDM (Karras et al., 2022), Latent Diffusion Models (Rombach et al., 2022), DiT (Peebles & Xie, 2023)), (ii) accelerated sampling, which reduces the number of denoising steps while retaining high-quality generation (e.g., DDIM (Song et al., 2021a), Consistency Models (Song et al., 2023), DPM-Solver++ (Lu et al., 2023), and progressive distillation (Salimans & Ho, 2022)).

However, the third ingredient—the *initial Gaussian noise*—has received comparatively little attention. Prior work has optimized initial noise for generation quality, editing, controllability, or inverse problem solving (Guo et al., 2024; Qi et al., 2024; Zhou et al., 2024; Ban et al., 2025; Chen et al., 2024a; Eyring et al., 2024; Song et al., 2025; Wang et al., 2024; Chihaoui et al., 2024). However, most of these efforts are task-specific. A systematic understanding of how noise itself shapes diffusion model outputs is still missing.

Our perspective is orthogonal to prior work. Our central discovery is both *simple and universal*: pairing every Gaussian noise z with its negation $-z$ —known as *antithetic sampling* (Owen, 2013)—consistently produces samples that are **strongly negatively correlated**. This phenomenon holds regardless of architecture, dataset, sampling schedule, and both conditional and unconditional sampling. It further extends to other generative models such as VAEs and Normalizing Flows. We explain this phenomenon through both experiments and theory. This leads to a symmetry conjecture that the score function is approximately affine antisymmetric, providing a new structural insight supported by empirical evidence.

Figure 1: Use antithetic noise $-z$ and z (with condition c) to generate visually “opposite” images.

This universal property has direct impact on uncertainty quantification and also enables additional applications:

(i) Sharper uncertainty quantification. Antithetic pairs naturally act as control variates, enabling significant variance reduction and thus sharper uncertainty quantification. Our antithetic estimator delivers up to 90% tighter confidence intervals and cuts computation cost by more than 100 times. The efficiency gain immediately leads to huge cost savings in a variety of tasks, including bias detection in generation and diffusion inverse solver evaluation, as we demonstrate in later experiments.

(ii) Other applications. Because each antithetic noise pair drives reverse-diffusion trajectories toward distant regions of the data manifold (see Figure 1), the paired-sampling scheme increases diversity “for free” while preserving high image quality, as confirmed by SSIM and LPIPS in our experiments. Moreover, algorithms that rely on intermediate sampling or approximation steps also benefit from the improved reliability provided by antithetic noise. As an illustration, we present an image editing example in C.2 and show that our antithetic design serves as a *plug-and-play* tool to improve performance at no additional cost.

Building on the antithetic pairs, we generalize the idea to apply quasi-Monte Carlo (QMC) and randomized QMC (RQMC). The resulting RQMC estimator often delivers further variance reduction. Although QMC has been widely used in computer graphics (Keller, 1995; Waechter & Keller, 2011), quantitative finance (Joy et al., 1996; L’Ecuyer, 2009), and Bayesian inference (Buchholz et al., 2018; Liu & Owen, 2021), this is, to our knowledge, its first application to diffusion models.

In summary, we discover a universal property of initial noise, reveal a new symmetry in score networks, and demonstrate concrete benefits in various practical applications. This positions initial noise manipulations as a simple, training-free tool for advancing generative modeling.

The remainder of the paper is organized as follows. Section 2 defines the problem and outlines our motivation. Section 3 presents our central finding that antithetic noise pairs produce strongly negatively correlated outputs. We offer both theoretical and empirical explanations for this phenomenon, and present the *symmetry conjecture* in Section 3.2. Section 4 develops estimators and their confidence intervals via antithetic sampling, and extends the approach to QMC. Section 5 reports experiments on the aforementioned applications. Section 6 discusses our method and outlines future directions. Appendix B, C, and D include proofs, additional experiments and detailed setups, and supplemental visualizations, respectively.

2 SETUP, MOTIVATION, AND RELATED WORKS

Unconditional diffusion model: A diffusion model aims to generate samples from an unknown data distribution p_0 . It first *noises* data towards a standard Gaussian progressively, then learns to reverse the process so that Gaussian noise can be *denoised* step-by-step back into target samples. The forward process simulates a stochastic differential equation (SDE): $d\mathbf{x}_t = \mu(\mathbf{x}_t, t)dt + \sigma_t d\mathbf{w}_t$, where $\{\mathbf{w}_t\}_{t=0}^T$ denotes the standard Brownian motion, and $\mu(\mathbf{x}, t), \sigma_t$ are chosen by the users. Let

108 p_t denote the distribution of \mathbf{x}_t . Song et al. (2021b) states that if we sample $\mathbf{y}_T \sim p_T$ and simulate
 109 the probability-flow ordinary differential equation (PF-ODE) backward from time T to 0 as
 110

$$111 \quad d\mathbf{y}_t = \left(-\mu(\mathbf{y}_t, t) - \frac{1}{2}\sigma_t^2 \nabla \log p(\mathbf{y}_t, t) \right) dt, \quad (1)$$

113 then for every time t the marginal distribution of \mathbf{y}_t coincides with p_t . Thus, in an idealized world,
 114 one can perfectly sample from p_0 by simulating the PF-ODE (1).
 115

116 In practice, the score function $\nabla \log p_t(\mathbf{x}, t)$ is unavailable, and a neural network $\epsilon_\theta^{(t)}(\mathbf{x})$ is trained
 117 to approximate it, where θ denotes its weights. Therefore, one can generate new samples from p_0
 118 by first sampling a Gaussian noise and simulating the PF-ODE (1) through a numerical integrator
 119 from T to 0. For example, DDIM (Song et al., 2021a) has the (discretized) forward process as
 120 $\mathbf{x}_k \mid \mathbf{x}_{k-1} \sim \mathbb{N}(\sqrt{\alpha_k} \mathbf{x}_{k-1}, (1 - \alpha_k)I)$ for $k = 0, \dots, T - 1$, and backward sampling process as

$$122 \quad \mathbf{y}_T \sim \mathbb{N}(0, I), \quad \mathbf{y}_{t-1} = \sqrt{\alpha_{t-1}} \left(\frac{\mathbf{y}_t - \sqrt{1 - \alpha_t} \epsilon_\theta^{(t)}(\mathbf{y}_t)}{\sqrt{\alpha_t}} \right) + \sqrt{1 - \alpha_{t-1}} \epsilon_\theta^{(t)}(\mathbf{y}_t). \quad (2)$$

124 Once θ is fixed, the randomness in DDIM sampling comes solely from the initial Gaussian noise.
 125

126 We remark that samples from p_0 can also be drawn by simulating the backward SDE with randomness
 127 at each step, as in the DDPM sampler (Ho et al., 2020; Song et al., 2021b). Throughout the main text,
 128 we focus on deterministic samplers such as DDIM to explain our idea. In Appendix C.4, we present
 129 additional experiments showing that our findings also extend to stochastic samplers like DDPM.

130 **Text-conditioned latent diffusion:** In Stable Diffusion and its successors SDXL and SDXL Turbo
 131 (Podell et al., 2024; Sauer et al., 2024), a pretrained VAE first compresses each image to a latent
 132 tensor z . A text encoder embeds the prompt as $c \in \mathbb{R}^m$. During training, the network $\epsilon_\theta^{(t)}$ receives
 133 (z_t, c) and learns $\nabla_{z_t} \log p_t(z_t \mid c)$. At generation time we draw latent noise $z_T \sim \mathbb{N}(0, I)$ and run
 134 the reverse diffusion from $t = T$ to 0, yielding a denoised latent z_0 , which the decoder maps back to
 135 pixels. Given a prompt c and an initial Gaussian noise, the sampler produces an image from $p_0(\cdot \mid c)$.

136 **Diffusion posterior sampling:** Diffusion model can also be used as a prior in inverse problems.
 137 Suppose we observe only a partial or corrupted measurement $\mathbf{y}_{\text{obs}} = \mathcal{A}(\mathbf{x}) + \text{noise}$ from an unknown
 138 signal \mathbf{x} . Diffusion posterior sampling aims to sample from the posterior distribution $p(\mathbf{x} \mid \mathbf{y}_{\text{obs}}) \propto$
 139 $p(\mathbf{y}_{\text{obs}} \mid \mathbf{x}) p(\mathbf{x})$, where $p(\mathbf{x})$ is given by a pretrained diffusion model. Given initial noise z and
 140 observed \mathbf{y}_{obs} , diffusion posterior samplers apply a sequence of denoising steps from T to 0, each
 141 step resembling (2) but incorporating \mathbf{y}_{obs} (Chung et al., 2023a; Song et al., 2022; 2024).

142 2.1 MOTIVATION

145 Beyond examining how variations in the initial noise affect the outputs, several technical considerations
 146 motivate our focus on the antithetic noise pair $(z, -z)$ for diffusion models. Let DM denote
 147 certain diffusion model’s mapping from an initial noise vector to a generated image sample.

- 148 • **Preserving quality:** Since $z \sim \mathbb{N}(0, I)$ implies $-z \sim \mathbb{N}(0, I)$, the initial noises z and $-z$ share
 149 the same marginal distribution. Consequently, $\text{DM}(z) \stackrel{d}{=} \text{DM}(-z)$, so all per-sample statistics
 150 remain unchanged. In other words, negating the initial noise does not degrade generation quality.
- 152 • **Maximal separation in the noise space:** High-dimensional Gaussians concentrate on the sphere
 153 of radius $\sqrt{\text{dimension}}$ (Vershynin, 2018), so any draw z and its negation $-z$ lie at opposite poles
 154 of that hypersphere. This antipodal pairing represents the maximum possible perturbation in noise
 155 space, making it a natural extreme test of the sampler’s behavior.
- 156 • **Measuring (non-)linearity:** Recent work has examined how closely score networks approximate
 157 linear operators. Empirical studies across various diffusion settings show that these networks
 158 behave as locally linear maps; this behavior forms the basis for new controllable sampling schemes
 159 (Chen et al., 2024b; Li et al., 2024b; Song et al., 2025). Correlation is a natural choice to measure
 160 linearity: if a score network were exactly linear, feeding it with noises z and $-z$ would yield
 161 a correlation of -1 between $\text{DM}(z)$ and $\text{DM}(-z)$. Thus, the difference between the observed
 correlation and -1 provides a direct measure of the network’s departure from linearity.

162

3 NEGATIVE CORRELATION FROM ANTITHETIC NOISE

163

3.1 PIXEL-WISE CORRELATIONS: ANTITHETIC VS. INDEPENDENT NOISE

164 We compare the similarity between paired images generated under two sampling schemes: PN
 165 (positive vs. negative, z vs. $-z$) and RR (random vs. random, z_1 vs. z_2) under three settings: (1)
 166 *unconditional diffusion models*, (2) *class- or prompt-based conditional diffusion models*, and (3)
 167 *generative models beyond diffusion*.

168 For diffusion models, we evaluate both unconditional and conditional generation using publicly
 169 available pre-trained checkpoints. Notably, we include both traditional U-Net architecture and
 170 transformer-based DiT (Peebles & Xie, 2023). Implementation details are described in Section 5.1.
 171 **We also test on distilled diffusion models (Song et al., 2023) using consistency distillation checkpoints.**
 172 For generative models beyond diffusion, we select two representative baselines: an unconditional
 173 VAE on MNIST and a conditional Glow model (Kingma & Dhariwal, 2018) on CIFAR-10. The
 174 experimental details for consistency models, VAE and Glow are provided in Appendix C.1.

175 To quantify similarity, we use two metrics: the standard Pearson correlation and a centralized Pearson
 176 correlation. Let $x_{i,1}$ and $x_{i,2}$ denote the flattened pixel values of the two images in the i -th generated
 177 pair. The standard Pearson correlation is computed directly between $x_{i,1}$ and $x_{i,2}$. To correct for
 178 dataset-level or class-level bias, we also define a centralized correlation. For a dataset, prompt, or
 179 class with K pairs generated, we compute the mean $\mu_c = \sum_{i=1}^K (x_{i,1} + x_{i,2}) / 2K$. The centralized
 180 correlation of the i -th pair is defined as the standard Pearson correlation between the centralized
 181 images $x_{i,1} - \mu_c$ and $x_{i,2} - \mu_c$. For each comparison, t -test is conducted to assess statistical
 182 significance. In all experiments, the resulting p -values are negligible ($< 10^{-10}$), which confirms
 183 significance; hence, they are omitted from the presentation.

184 Additional metrics such as the Wasserstein distance are presented in Appendix C.4.

185 **Results:** Table 1 summarizes the statistics in all classes of models. PN pairs consistently show
 186 significantly stronger negative correlations than RR pairs, with this contrast also visually evident in
 187 their histograms (Figure 2). In addition, centralization strengthens the negative correlation, since it
 188 removes shared patterns. For example, in CelebA-HQ, centralization removes global facial structure,
 189 while in DiT and Glow, it removes class-specific patterns.

190 The same behavior appears in both DDPM models and diffusion posterior samplers, which we report
 191 in Appendix C.4. These results demonstrate that the negative correlation resulting from antithetic
 192 sampling is a universal phenomenon across diverse architectures and conditioning schemes.

193
 194
 195
 196
 197
 198 Table 1: Correlation results across different models and datasets, shown are means (SD). **Rows**
 199 **1–3** are pretrained **unconditional diffusion models on different datasets**. **Rows 4–5** are **conditional**
 200 **diffusion models**. **Rows 6–8** are pretrained **consistency models on different datasets**. **Rows 9–10** are
 201 **generative models that are not diffusion-based**.

202 203 204 205 206 207 208 209 210 211 212 213 214 215	203 204 205 206 207 208 209 210 211 212 213 214 215	203 204 205 206 207 208 209 210 211 212 213 214 215	203 204 205 206 207 208 209 210 211 212 213 214 215		203 204 205 206 207 208 209 210 211 212 213 214 215	
			203 204 205 206 207 208 209 210 211 212 213 214 215	203 204 205 206 207 208 209 210 211 212 213 214 215	203 204 205 206 207 208 209 210 211 212 213 214 215	203 204 205 206 207 208 209 210 211 212 213 214 215
205 206 207 208 209 209 210 211 212 213 214 215	205 206 207 208 209 209 210 211 212 213 214 215	205 206 207 208 209 209 210 211 212 213 214 215	205 206 207 208 209 209 210 211 212 213 214 215	205 206 207 208 209 209 210 211 212 213 214 215	205 206 207 208 209 209 210 211 212 213 214 215	205 206 207 208 209 209 210 211 212 213 214 215
			205 206 207 208 209 209 210 211 212 213 214 215	205 206 207 208 209 209 210 211 212 213 214 215	205 206 207 208 209 209 210 211 212 213 214 215	205 206 207 208 209 209 210 211 212 213 214 215
			205 206 207 208 209 209 210 211 212 213 214 215	205 206 207 208 209 209 210 211 212 213 214 215	205 206 207 208 209 209 210 211 212 213 214 215	205 206 207 208 209 209 210 211 212 213 214 215
205 206 207 208 209 209 210 211 212 213 214 215	205 206 207 208 209 209 210 211 212 213 214 215	205 206 207 208 209 209 210 211 212 213 214 215	205 206 207 208 209 209 210 211 212 213 214 215	205 206 207 208 209 209 210 211 212 213 214 215	205 206 207 208 209 209 210 211 212 213 214 215	205 206 207 208 209 209 210 211 212 213 214 215
			205 206 207 208 209 209 210 211 212 213 214 215	205 206 207 208 209 209 210 211 212 213 214 215	205 206 207 208 209 209 210 211 212 213 214 215	205 206 207 208 209 209 210 211 212 213 214 215
			205 206 207 208 209 209 210 211 212 213 214 215	205 206 207 208 209 209 210 211 212 213 214 215	205 206 207 208 209 209 210 211 212 213 214 215	205 206 207 208 209 209 210 211 212 213 214 215
205 206 207 208 209 209 210 211 212 213 214 215	205 206 207 208 209 209 210 211 212 213 214 215	205 206 207 208 209 209 210 211 212 213 214 215	205 206 207 208 209 209 210 211 212 213 214 215	205 206 207 208 209 209 210 211 212 213 214 215	205 206 207 208 209 209 210 211 212 213 214 215	205 206 207 208 209 209 210 211 212 213 214 215
			205 206 207 208 209 209 210 211 212 213 214 215	205 206 207 208 209 209 210 211 212 213 214 215	205 206 207 208 209 209 210 211 212 213 214 215	205 206 207 208 209 209 210 211 212 213 214 215
			205 206 207 208 209 209 210 211 212 213 214 215	205 206 207 208 209 209 210 211 212 213 214 215	205 206 207 208 209 209 210 211 212 213 214 215	205 206 207 208 209 209 210 211 212 213 214 215

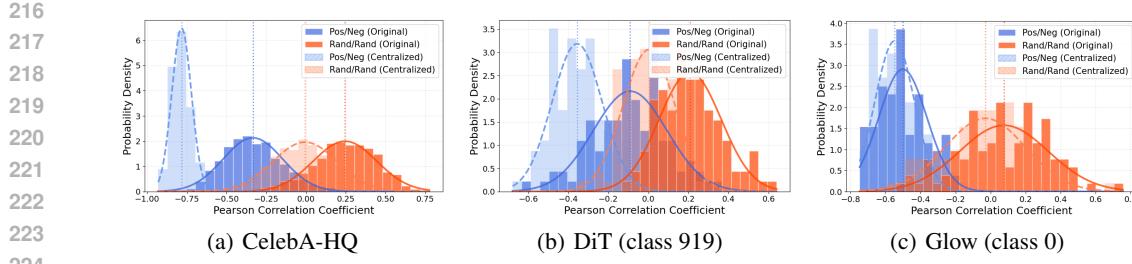


Figure 2: Histograms of standard and centralized Pearson correlation coefficients for CelebA-HQ, DiT class and Glow in single class. Dashed lines indicate the average.

3.2 EXPLANATORY EXPERIMENTS: TEMPORAL CORRELATIONS & SYMMETRY CONJECTURE

We aim to explain the strong negative correlation between $\text{DM}(Z)$ and $\text{DM}(-Z)$ for $Z \sim \mathbb{N}(0, I)$. Because DM performs iterative denoising through the score network $\epsilon_\theta^{(t)}$ (see (2)), we visualize how these correlations evolve over diffusion time-steps in Figure 3.

Throughout the transition from noises to samples, the PN correlation of $\epsilon_\theta^{(t)}$ —indicated by the orange, blue, and green dashed lines in Figure 3—starts at -1 , stays strongly negative, and only climbs slightly in the final steps. This nearly -1 correlation is remarkable, as it suggests that the score network ϵ_θ learns an approximately affine antisymmetric function. To explain, we first state the following results which shows -1 correlation is equivalent to affine antisymmetric, with proof in Appendix B.1.

Lemma 1. *Let $Z \sim \mathbb{N}(0, I_d)$, suppose a map $f : \mathbb{R}^d \rightarrow \mathbb{R}$ satisfies $\text{Corr}(f(Z), f(-Z)) = -1$, then f must be affine antisymmetric at $(0, c)$ for some c , i.e., $f(\mathbf{x}) + f(-\mathbf{x}) = 2c$ for every \mathbf{x} .*

Lemma 1 and Figure 3 leads us to the following conjecture:

Conjecture For each time step t , the score network $\epsilon_\theta^{(t)}$ is approximately affine antisymmetric at $(0, \mathbf{c}_t)$, i.e.,

$$\epsilon_\theta^{(t)}(\mathbf{x}) + \epsilon_\theta^{(t)}(-\mathbf{x}) \approx 2\mathbf{c}_t$$

for some fixed vector \mathbf{c}_t that depends on t .

The conjecture implies that the score network has learned an “almost odd” function up to an additive shift. This is consistent with both theory and observations for large t , where $p_t(\mathbf{x})$ is close to standard Gaussian, thus the score function is almost linear in \mathbf{x} . Nonetheless, our conjecture is far more general: it applies to every timestep t and can accommodate genuinely nonlinear odd behaviors (for example, functions like $\sin x$).

This conjecture combined with the DDIM update rule (2) immediately explains the negative correlation: The one step iteration of F_t in (2) can be written as a linear combination between \mathbf{x} and the output of the score network: $F_t(\mathbf{x}) = a_t \mathbf{x} + b_t \epsilon_\theta^{(t)}(\mathbf{x})$. Given the conjecture, $F_t(-\mathbf{x}) \approx -a_t \mathbf{x} - b_t \epsilon_\theta^{(t)}(\mathbf{x}) + 2b_t \mathbf{c}_t = -F_t(\mathbf{x}) + 2b_t \mathbf{c}_t$, so the one-step DDIM update is affine antisymmetric at $(0, b_t \mathbf{c}_t)$. Thus, beginning with a strongly negatively correlated pair, each DDIM update preserves that strong negative correlation all the way through to the final output.

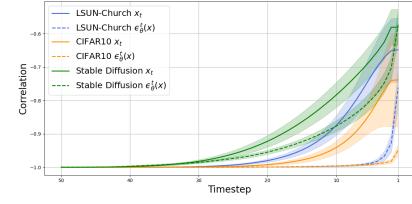


Figure 3: Correlation of x_t (solid) and $\epsilon_\theta^{(t)}$ (dash) between antithetic (PN) pairs. Step 50 is the initial noise and Step 0 is the generated image. Shaded bands show ± 1 std. dev.

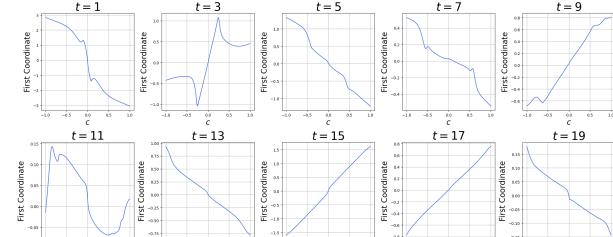


Figure 4: First-coordinate output of the pretrained score network on CIFAR10 as a function of the interpolation scalar c for a 50-step DDIM at $t = 1, 3, \dots, 19$.

To (partly) validate our conjecture, we perform the following experiment. Using a pretrained CIFAR-10 score network with a 50-step DDIM sampler, we picked time steps $t = 1, 3, \dots, 19$ (where smaller t is closer to the image and larger t is closer to pure noise). For each t , we sample $\mathbf{x} \sim \mathbb{N}(0, I_d)$, evaluated the first coordinate of $\epsilon_\theta^{(t)}(c \mathbf{x})$ as c varied from -1 to 1 (interpolating from $-\mathbf{x}$ to \mathbf{x}), and plotted the result in Figure 4. We exclude $t \geq 20$, where the curve is very close to a straight line.

Figure 4 confirms that, at every step t , the first coordinate of the score network is indeed overall affine antisymmetric. Although the curves display nonlinear oscillations at small t , the deviations on either side are approximately mirror images, and for larger t the mapping is almost a straight line. The symmetry center c_t usually lies near zero, but not always—for instance, at $t = 11$, $c_t \approx 0.05$ even though the function spans from -0.05 to 0.15 . In Appendix C.5, we present further validation experiments using alternative coordinates, datasets, and a quantitative metric called the antisymmetry score. Despite the distinct function shapes across coordinates, the conjectured symmetry still persists.

We further evaluate the antisymmetry conjecture across additional coordinates and datasets. On CIFAR-10 and Church, we assess how closely one-dimensional slices of the network outputs behave like affine-antisymmetric functions. To this end, we introduce the *affine antisymmetry score*, which quantifies the degree of antisymmetry. Across datasets, the resulting antisymmetry scores are consistently high (mean values above 0.99 and even the lower quantiles still near 0.9). These results provide strong empirical support for the conjecture; full experimental details and plots are given in Appendix C.5.

Finally, we provide theoretical support for our conjecture in Appendix B.2, B.3, and B.4. Appendix B.2 confirms the conjecture in the large- t regime, Appendix B.3 shows how both the density ratio and the score error converge monotonically as t evolves via a Hermite polynomial expansion, and Appendix B.4 demonstrates that all orthogonal symmetries of the data distribution are preserved throughout the forward process. These results help explain the behavior underlying the conjecture. The forward process preserves every orthogonal symmetry of the initial distribution, including coordinate reflections. Thus, in the ideal case where the data distribution at $t = 0$ is reflection-symmetric, the density remains even and the score remains odd for all t by Proposition 4. Moreover, even if this symmetry holds only approximately for small t , the monotone decay of the density ratio and the score error shows that p_t quickly approaches the Gaussian, so the forward dynamics push the score toward an odd function in a controlled manner.

Both the conjecture and its empirical support could be of independent interest. They imply that the score network learns a function with strong symmetry, even though that symmetry is not explicitly enforced. The conjecture matches the Gaussian linear score approximation in the high-noise regime, which has proven effective (Wang & Vastola, 2024). In the intermediate-to-low noise regime, a single Gaussian approximation performs poorly—this is clear in Figure 4 for $t = 3, 7, 11$, where the score is strongly nonlinear—yet our conjectured approximate symmetry still holds.

Besides explaining the source of negative correlation, we believe this result provides new insight into the structure of diffusion models, which are often treated as black-box functions. Leveraging this finding for algorithms and applications is an important direction for future work.

For readers interested in an additional theoretical perspective, Appendix B.6 presents a discussion that links negative correlation to the FKG inequality, [and provides a detailed analysis of the DDIM sampler using this framework](#).

4 UNCERTAINTY QUANTIFICATION

In uncertainty quantification for a diffusion model DM, the goal is typically to estimate expectations of the form $\mathbb{E}_{z \sim \mathbb{N}(0, I)} [S(\text{DM}(z))]$, where S is a statistic of user's interest. Our goal is to leverage negative correlation to design an estimator with higher accuracy at fixed computation cost of inference.

[Recent studies have examined epistemic uncertainty and bias in diffusion models, including Berry et al. \(2024; 2025\). Our focus is different: we study aleatoric uncertainty from noise sampling and develop a variance-reduction method.](#)

Standard Monte Carlo: To approximate this expectation, the simplest approach is to draw N independent noises $z_1, \dots, z_N \sim \mathbb{N}(0, I)$, calculate $S_i = S(\text{DM}(z_i))$ for each i and form the

standard Monte Carlo (MC) estimator $\hat{\mu}_N^{\text{MC}} := \sum_{i=1}^N S_i/N$ by taking their average. A $(1 - \alpha)$ confidence interval, denoted as $\text{CI}_N^{\text{MC}}(1 - \alpha)$ is then $\hat{\mu}_N^{\text{MC}} \pm z_{1-\alpha/2} \sqrt{(\hat{\sigma}_N^{\text{MC}})^2/N}$, where $\hat{\sigma}_N^2$ is the sample variance of S_1, S_2, \dots, S_N and $z_{1-\alpha/2}$ is the $(1 - \alpha/2)$ -quantile of the standard normal. A formal guarantee of the above construction is given in Proposition 5 in Appendix B.5.

Antithetic Monte Carlo: The observed negative correlation motivates an improved estimator. Let $N = 2K$ be even, users can generate K pairs of antithetic noise $(z_1, -z_1), \dots, (z_K, -z_K)$. Define $S_i^+ = S(\text{DM}(z_i))$, $S_i^- = S(\text{DM}(-z_i))$, and let $\bar{S}_i = 0.5(S_i^+ + S_i^-)$ be their average. Our Antithetic Monte Carlo (AMC) estimator is $\hat{\mu}_N^{\text{AMC}} := \sum_{i=1}^K \bar{S}_i/K$, and confidence interval $\text{CI}_N^{\text{AMC}}(1 - \alpha)$ is $\hat{\mu}_N^{\text{AMC}} \pm z_{1-\alpha/2} \sqrt{2(\hat{\sigma}_N^{\text{AMC}})^2/N}$, where $(\hat{\sigma}_N^{\text{AMC}})^2$ is the sample variance of $\bar{S}_1, \dots, \bar{S}_K$.

The intuition is simple: if the pair (S_i^+, S_i^-) remains negatively correlated, then averaging antithetic pairs can reduce variance by partially canceling out opposing errors—since negative correlation suggests that when one estimate exceeds the true value, the other is likely to fall below it.

The AMC estimator is unbiased, and $\text{CI}_N^{\text{AMC}}(1 - \alpha)$ achieves correct coverage as the sample size increases. Let $\rho := \text{Corr}(S_i^+, S_i^-)$. We can prove that AMC’s standard error and its confidence-interval width are each equal to the Monte Carlo counterparts multiplied by the factor $\sqrt{1 + \rho}$. Thus, when $\rho < 0$, AMC produces **provably lower variance and tighter confidence intervals** than MC, with greater gains as ρ becomes more negative. See Appendix B.5 for a formal statement and proof.

Since both methods have the same computational cost, the variance reduction from negative correlation and the antithetic design yields a direct and cost-free improvement.

K -antithetic noise: We can generalize the antithetic noise pair to a collection of K noise variables, constructed so that every pair has the same negative correlation $-1/(K - 1)$. One way to generate them is as follows. Draw K independent standard Gaussian vectors (w_1, \dots, w_K) and let \bar{w} be their average. For each i , set $z_i = \sqrt{K/(K - 1)}(w_i - \bar{w})$. One can directly check that each z_i is still marginally standard Gaussian, and $\text{Corr}(z_{i,l}, z_{j,l}) = -1/(K - 1)$ for every $i \neq j$ and every pixel l . In particular, when $K = 2$, this construction reduces to our usual antithetic noise pair.

Quasi-Monte Carlo: The idea of using negatively correlated samples has been widely adopted in Monte Carlo methods (Craiu & Meng, 2005) and statistical risk estimation (Liu et al., 2024) for improved performance. This principle of variance reduction can be further extended to QMC methods. As an alternative to Monte Carlo, QMC constructs a deterministic set of N samples that have negative correlation and provide a more balanced coverage of the sample space. QMC samples can also be randomized, resulting in RQMC methods. RQMC maintains the marginal distribution of each sample while preserving the low-discrepancy properties of QMC. Importantly, repeated randomizations allow for empirical error estimation and confidence intervals (L’Ecuyer et al., 2023).

Under regularity conditions (Niederreiter, 1992), QMC and RQMC can improve the error *convergence rate* from $O(N^{-1/2})$ to $O(N^{-1+\epsilon})$ for any $\epsilon > 0$, albeit ϵ may absorb a dimension-dependent factor $(\log N)^d$. For sufficiently smooth functions, RQMC can achieve rates as fast as $O(N^{-3/2+\epsilon})$ (Owen, 1997b;a). Moreover, the space-filling property of QMC may also promote greater sample diversity.

Although the dimension of the Gaussian noise is higher than the typical regime where QMC is expected to be most effective, RQMC still further shrinks our confidence intervals by several-fold compared with standard MC. This hints that the image generator’s effective dimension is much lower than its ambient one, allowing QMC methods to stay useful. Similar gains arise when QMC methods deliberately exploit low-dimensional structure in practical problems (Wang & Fang, 2003; Xiao & Wang, 2019; Liu & Owen, 2023). Developing ways to identify and leverage this structure more systematically is an appealing avenue for future work.

5 EXPERIMENT

Sections 5.1 and 5.2 present two uncertainty quantification applications: estimating pixel-wise statistics and evaluating diffusion inverse problem solvers. The latter considers two popular algorithms across a range of tasks and datasets. For each task, we apply MC, AMC, and, when applicable, QMC estimators as described in Section 4. For each estimator, we report the 95% confidence interval (CI)

378 width and its relative efficiency. The relative efficiency of a new estimator (AMC or QMC) is defined
 379 as the squared ratio of the MC CI width to that of the estimator, $(\text{CI}_{\text{MC}}/\text{CI}_{\text{new estimator}})^2$. It reflects
 380 how many times more MC samples are required to achieve the same accuracy as our new estimator.
 381 Section 5.3 shows that antithetic noise produces more diverse images than independent noise while
 382 preserving quality, and Appendix C.2 explores an image editing example.
 383

385 5.1 PIXEL-WISE STATISTICS

387 We begin by evaluating four pixel-level statistics. These include the (i) pixel value mean, (ii)
 388 perceived brightness, (iii) contrast, and (iv) the image centroid, with definitions in Appendix C.6.
 389 These statistics are actively used in diffusion workflows for diagnosing artifacts and assessing
 390 reliability, with applications in detecting signal leakage (Lin et al., 2024; Everaert et al., 2024),
 391 out-of-distribution detection (Le Bellier & Audebert, 2024), identifying artifact-prone regions (Kou
 392 et al., 2024), and improving the reliability of weather prediction (Li et al., 2024a).

393 **Setup:** We study both unconditional and conditional diffusion models; details are in Appendix C.3:

394 For *unconditional diffusion models*, we evaluate pre-trained models on CIFAR-10 (Krizhevsky et al.,
 395 2009), CelebA-HQ (Xia et al., 2021), and LSUN-Church (Yu et al., 2015). For each dataset, 1,600
 396 image pairs are generated under both PN and RR noise sampling with 50 DDIM steps.
 397

398 For *conditional diffusion models*, we evaluate Stable Diffusion 1.5 on 200 prompts from the Pick-a-Pic
 399 (Kirstain et al., 2023) and DrawBench (Saharia et al., 2022) with classifier-free-guidance (CFG) scale
 400 3.5, and DiT (Peebles & Xie, 2023) on 32 ImageNet classes with CFG scale 4.0. For each class or
 401 prompt, 100 PN and RR pairs are generated with 20 DDIM steps. **We also study the effect of CFG
 402 scale on both models in the Appendix C.4 (Ho & Salimans, 2022).**

403 **Implementation:** To ensure fair comparisons under equal sample budget:

404 For *unconditional diffusion models*: MC uses 3,200 independent samples, AMC ($k = 2$) uses 1,600
 405 antithetic pairs, and **AMC ($k = 8$) uses 400 independent negative correlated batches**. QMC employs
 406 a Sobol' point set of size 64 with 50 independent randomizations. Due to the dimensionality limits,
 407 QMC is restricted to CIFAR-10.

408 For *conditional diffusion models*: for each class or prompt, MC uses 200 independent samples, AMC
 409 uses 100 antithetic pairs, and QMC employs 8 Sobol' points with 25 randomizations.
 410

411 CIs for MC and AMC are constructed as described in Section 4. For QMC, we use a student-*t* interval
 412 to ensure reliable coverage (L'Ecuyer et al., 2023), with details in Appendix C.6.

413 **Results:** Across all statistics, both AMC and QMC have much shorter CIs than MC, with relative
 414 efficiencies ranging from 3.1 to 136. This strongly suggests our estimators can dramatically reduce
 415 cost for uncertainty estimates. AMC and QMC estimators yield comparable results. The performance
 416 of QMC depends on how the total budget is allocated between the size of the QMC point set and the
 417 number of random replicates, a trade-off we explore in further details in Appendix C.6.
 418

420 5.2 EVALUATING DIFFERENT DIFFUSION INVERSE PROBLEM SOLVERS

422 Bayesian inverse problems are important in medical imaging, remote sensing, and astronomy. Re-
 423 cently, many methods have been proposed that leverage diffusion priors to solve inverse problems.
 424 For instance, Zheng et al. (2025) surveys 14 such methods, yet this is still far from exhaustive. As a
 425 result, comprehensive benchmarking is costly. Meanwhile, most existing works on new diffusion
 426 inverse solvers (DIS) largely ignore uncertainty quantification, though it is critical for fair evaluation.
 427 This raises a key challenge: *how to efficiently evaluate posterior sampling algorithms with reliable
 428 estimates of reconstruction quality and uncertainty*. Our approach aims to efficiently address this gap.

429 We show that our antithetic estimator can save substantial computational cost for quantifying un-
 430 certainty by requiring far fewer samples than standard estimators. We focus on two popular DISs,
 431 Diffusion Posterior Sampling (DPS) (Chung et al., 2023a) and Decomposed Diffusion Sampler
 (DDS) (Chung et al., 2023b), and evaluate them on a range of tasks described below.

Table 2: CI lengths and efficiency $(CI_{MC}/CI)^2$ (in parentheses), using MC as baseline.

Dataset		Brightness	Pixel Mean	Contrast	Centroid
CIFAR10	MC	2.00	2.04	1.08	0.14
	QMC	0.35 (32.05)	0.39 (26.94)	0.22 (23.43)	0.04 (9.65)
	AMC ($k = 2$)	0.35 (32.66)	0.39 (27.12)	0.23 (22.05)	0.04 (9.73)
	AMC ($k = 8$)	0.36 (30.35)	0.39 (27.34)	0.24 (20.37)	0.03 (13.94)
CelebA	MC	1.77	1.76	0.60	0.60
	AMC ($k = 2$)	0.26 (47.41)	0.31 (33.15)	0.19 (10.18)	0.20 (7.82)
	AMC ($k = 8$)	0.15 (130.69)	0.16 (115.79)	0.18 (10.90)	0.19 (8.49)
Church	MC	1.66	1.64	1.02	0.82
	AMC ($k = 2$)	0.14 (134.13)	0.16 (103.80)	0.20 (27.16)	0.26 (9.84)
	AMC ($k = 8$)	0.15 (117.90)	0.16 (107.64)	0.19 (27.81)	0.25 (10.87)
Stable Diff.	MC	3.86	3.95	2.18	4.03
	QMC	1.36 (8.07)	1.43 (7.60)	1.03 (4.46)	1.88 (4.54)
	AMC ($k = 2$)	0.90 (18.53)	1.00 (15.53)	0.89 (6.25)	1.62 (6.15)
DiT	MC	7.52	7.74	3.32	3.15
	QMC	3.44 (4.78)	3.49 (4.91)	1.82 (3.31)	1.69 (3.46)
	AMC ($k = 2$)	3.13 (5.79)	3.12 (6.16)	1.72 (3.74)	1.74 (3.30)

5.2.1 HUMAN FACE RECONSTRUCTION

We evaluate DPS across three inverse problems: inpainting, Gaussian deblurring, and super-resolution. Reconstruction error is measured using both PSNR and L_1 distance relative to the ground truth image. For all tasks, we use 200 human face images from the CelebA-HQ dataset. For each corrupted image, we generate 50 DPS reconstructions using both PN and RR noise pairs to compare their estimators. Operator-specific configurations (e.g., kernel sizes, mask ratios) are provided in Appendix C.6.

Results: As shown in Table 3, across all three tasks, AMC achieves substantially shorter confidence intervals than standard MC. This implies AMC yields efficiency gains, ranging from 54% – 84% in inpainting, 41% – 54% in super-resolution, and 34% – 56% in deblurring.

The reconstruction metrics (L_1 , PSNR) differ by less than 0.2% between the two estimators, showing that they produce consistent results. Meanwhile, our method offers a clear advantage in efficiency.

Table 3: Comparison of AMC vs. MC across tasks in DPS with efficiency $(CI_{MC}/CI)^2$ in parentheses

Task	L1		PSNR	
	AMC CI	MC CI	AMC CI	MC CI
Inpainting	0.83 (1.54)	1.03	0.23 (1.84)	0.31
Super-resolution	1.01 (1.41)	1.20	0.24 (1.54)	0.30
Gaussian Deblur	0.89 (1.34)	1.03	0.23 (1.56)	0.29

5.2.2 MEDICAL IMAGE RECONSTRUCTION

We use the DDS (Chung et al., 2023b) as the reconstruction backbone and apply our antithetic noise initialization to estimate reconstruction confidence intervals. Given a single measurement, we generate 100 reconstruction pairs using 50-step DDS sampling. Reconstruction quality is evaluated against ground truth using L_1 distance and PSNR. Dataset details can be found in Appendix C.6.4.

Results: Both MC and AMC estimators give consistent estimates, with estimated L_1 error of 0.0194 and 0.0195, and PSNR values of 31.48 and 31.45, respectively. Similarly, we observe that AMC produces much shorter confidence intervals. The CI lengths decrease with relative efficiency of 1.44 for L_1 and 1.37 for PSNR. This shows that antithetic initialization consistently reduces estimator variance with no extra cost without degrading reconstruction fidelity in large-scale inverse problems.

486
 487 **Conclusion:** These two experiments show that a single antithetic estimator reduces evaluation costs
 488 for diffusion inverse solvers by 34%–84%, a saving that becomes especially valuable given the large
 489 number of solvers in the literature. This makes large-scale benchmarking far more practical.

490 **5.3 DIVERSITY IMPROVEMENT**
 491

492 Many diffusion tools generate a few images from one prompt to let users select a preferred one.
 493 However, these images may look similar with randomly sampled initial noise (Marwood et al., 2023;
 494 Sadat et al., 2024). We show that antithetic noise produces more diverse images than independent
 495 noise, without reducing image quality or increasing computational cost. Diversity is measured using
 496 pairwise SSIM (lower = more diverse) and LPIPS (Zhang et al., 2018) (higher = more diverse).

497 We evaluate the diversity metrics for both unconditional and conditional diffusion on image pairs
 498 generated following the same setup in Section 5.1. As shown in Table 4, antithetic noise pairs
 499 consistently lead to higher diversity, as indicated by lower SSIM and higher LPIPS scores. Importantly,
 500 image quality remains stable for both noise types. We expect our method can be easily integrated into
 501 existing diversity optimization techniques to further improve their performance.

502 **Table 4: Average percentage improvement of PN pairs over RR pairs on SSIM and LPIPS.**

504 Metric	505 Unconditional Diffusion			506 Conditional Diffusion	
	507 CIFAR-10	508 LSUN-Church	509 CelebA-HQ	510 SD1.5	511 DiT
512 SSIM (%)	513 88.78	514 45.69	515 36.78	516 28.32	517 23.99
518 LPIPS (%)	519 6.69	520 3.54	521 15.14	522 5.78	523 10.62

524 **6 DISCUSSION AND FUTURE DIRECTIONS**

525 We find that antithetic initial noise yields negatively correlated samples. This is a robust property that
 526 generalizes across every model that we have tested. The negative correlation is especially useful for
 527 uncertainty quantification, where it brings huge variance reduction and cost savings. This makes our
 528 approach highly effective for evaluation and benchmarking tasks. Meanwhile, this finding can serve
 529 as a simple plug-in to improve related tasks, such as diversity improvement and image editing.

530 A limitation of our work is that while the symmetry conjecture is supported by both empirical and
 531 theoretical evidence, it remains open in full generality. In addition, our method is not a cure-all: the
 532 LPIPS diversity improvements in Table 4 are consistent but modest. While we obtain large gains for
 533 pixel-wise statistics, the improvements for perceptual or text-alignment metrics such as MUSIQ and
 534 CLIP Score are small, since the strong nonlinearity of transformer decoders attenuates most of the
 535 negative correlation present in the diffusion outputs.

536 Looking forward, future directions include systematically studying the symmetry conjecture; integrat-
 537 ing antithetic noise with existing noise-optimization methods; and further leveraging QMC to
 538 improve sampling and estimation.

540 ETHICS STATEMENT
541542 This work does not involve human subjects or personally identifiable data collected by the authors.
543 All datasets used are publicly available and distributed under licenses permitting academic research.
544545 Our medical image reconstruction experiments rely on benchmark datasets that are fully de-identified
546 before release (e.g., publicly available MRI data). No new data collection was performed, and no
547 protected health information is included. The experiments are purely methodological, intended to
548 evaluate reconstruction fidelity and uncertainty quantification techniques, and make no clinical claims
549 or diagnostic recommendations.
550551 Our contributions are methodological, focusing on sampling and uncertainty quantification in gen-
552 erative models. While such methods could be applied broadly, including in sensitive domains, we
553 emphasize that our work aims to improve reliability, transparency, and efficiency in evaluation.
554555 REPRODUCIBILITY STATEMENT
556557 Full experimental details can be found in Section 5 and Appendix C.
558 The code is available at [https://anonymous.4open.science/r/](https://anonymous.4open.science/r/Antithetic-Noise-in-Diffusion-Models-8B54)
559 Antithetic-Noise-in-Diffusion-Models-8B54.560 REFERENCES
561562 Rishabh Agarwal, Levi Melnick, Nicholas Frosst, Xuezhou Zhang, Ben Lengerich, Rich Caruana,
563 and Geoffrey E Hinton. Neural additive models: Interpretable machine learning with neural nets.
564 *Advances in Neural Information Processing Systems*, 34:4699–4711, 2021.565 Dominique Bakry, Ivan Gentil, and Michel Ledoux. *Analysis and geometry of Markov diffusion*
566 *operators*, volume 348. Springer Science & Business Media, 2013.567 Yuanhao Ban, Ruochen Wang, Tianyi Zhou, Boqing Gong, Cho-Jui Hsieh, and Minhao Cheng.
568 The crystal ball hypothesis in diffusion models: Anticipating object positions from initial noise.
569 In *The Thirteenth International Conference on Learning Representations*, 2025. URL <https://openreview.net/forum?id=GpdO9r73xT>.
570571 Lucas Berry, Axel Brando, and David Meger. Shedding light on large generative networks: Estimating
572 epistemic uncertainty in diffusion models. In *The 40th Conference on Uncertainty in Artificial*
573 *Intelligence*, 2024.575 Lucas Berry, Axel Brando, Wei-Di Chang, Juan Camilo Gamboa Higuera, and David Meger. See-
576 ing the unseen: How emoe unveils bias in text-to-image diffusion models. *arXiv preprint*
577 *arXiv:2505.13273*, 2025.578 Alexander Buchholz, Florian Wenzel, and Stephan Mandt. Quasi-Monte Carlo variational inference.
579 In *International Conference on Machine Learning*, pp. 668–677. PMLR, 2018.
580581 Changgu Chen, Libing Yang, Xiaoyan Yang, Lianggangxu Chen, Gaoqi He, Changbo Wang, and
582 Yang Li. FIND: Fine-tuning initial noise distribution with policy optimization for diffusion models.
583 In *Proceedings of the 32nd ACM International Conference on Multimedia*, pp. 6735–6744, 2024a.584 Siyi Chen, Huijie Zhang, Minzhe Guo, Yifu Lu, Peng Wang, and Qing Qu. Exploring low-
585 dimensional subspace in diffusion models for controllable image editing. In *The Thirty-*
586 *eighth Annual Conference on Neural Information Processing Systems*, 2024b. URL <https://openreview.net/forum?id=50aOEfb2km>.
587588 Hamadi Chihaoui, Abdelhak Lemkhter, and Paolo Favaro. Blind image restoration via fast diffusion
589 inversion. *Advances in Neural Information Processing Systems*, 37:34513–34532, 2024.
590591 Hyungjin Chung, Jeongsol Kim, Michael Thompson Mccann, Marc Louis Klasky, and Jong Chul Ye.
592 Diffusion posterior sampling for general noisy inverse problems. In *The Eleventh International*
593 *Conference on Learning Representations*, 2023a. URL <https://openreview.net/forum?id=OnD9zGAGT0k>.
594

594 Hyungjin Chung, Suhyeon Lee, and Jong Chul Ye. Decomposed diffusion sampler for accelerating
 595 large-scale inverse problems. *arXiv preprint arXiv:2303.05754*, 2023b.
 596

597 Radu V Craiu and Xiao-Li Meng. Multiprocess parallel antithetic coupling for backward and forward
 598 Markov chain Monte Carlo. *Annals of Statistics*, pp. 661–697, 2005.

599 Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. Imagenet: A large-scale
 600 hierarchical image database. In *2009 IEEE conference on computer vision and pattern recognition*,
 601 pp. 248–255. Ieee, 2009.

602

603 Martin Nicolas Everaert, Athanasios Fitsios, Marco Bocchio, Sami Arpa, Sabine Süsstrunk, and
 604 Radhakrishna Achanta. Exploiting the signal-leak bias in diffusion models. In *Proceedings of the*
 605 *IEEE/CVF Winter Conference on Applications of Computer Vision*, pp. 4025–4034, 2024.

606 Luca Eyring, Shyamgopal Karthik, Karsten Roth, Alexey Dosovitskiy, and Zeynep Akata. Reno:
 607 Enhancing one-step text-to-image models through reward-based noise optimization. *Advances in*
 608 *Neural Information Processing Systems*, 37:125487–125519, 2024.

609

610 Cees M Fortuin, Pieter W Kasteleyn, and Jean Ginibre. Correlation inequalities on some partially
 611 ordered sets. *Communications in Mathematical Physics*, 22:89–103, 1971.

612 William C. Francis. Variational-autoencoder-for-mnist. <https://github.com/williamcfrancis/Variational-Autoencoder-for-MNIST>, 2022. Pytorch
 613 implementation of a Variational Autoencoder (VAE) that learns from the MNIST dataset and
 614 generates images of altered handwritten digits.

615

616 Xiefan Guo, Jinlin Liu, Miaomiao Cui, Jiankai Li, Hongyu Yang, and Di Huang. Initno: Boosting
 617 text-to-image diffusion models via initial noise optimization. In *Proceedings of the IEEE/CVF*
 618 *Conference on Computer Vision and Pattern Recognition*, pp. 9380–9389, 2024.

619

620 Jonathan Ho and Tim Salimans. Classifier-free diffusion guidance. *arXiv preprint arXiv:2207.12598*,
 621 2022.

622

623 Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. *Advances in*
 624 *neural information processing systems*, 33:6840–6851, 2020.

625

626 Corwin Joy, Phelim P Boyle, and Ken Seng Tan. Quasi-Monte Carlo methods in numerical finance.
 627 *Management science*, 42(6):926–938, 1996.

628

629 Samuel Karlin and Yosef Rinott. Classes of orderings of measures and related correlation inequalities.
 i. multivariate totally positive distributions. *Journal of Multivariate Analysis*, 10(4):467–498, 1980.

630

631 Tero Karras, Miika Aittala, Timo Aila, and Samuli Laine. Elucidating the design space of diffusion-
 632 based generative models. *Advances in Neural Information Processing Systems*, 35:26565–26577,
 633 2022.

634

635 Alexander Keller. A quasi-Monte Carlo algorithm for the global illumination problem in the radiosity
 636 setting. In *Monte Carlo and Quasi-Monte Carlo Methods in Scientific Computing: Proceedings of*
 637 *a conference at the University of Nevada, Las Vegas, Nevada, USA, June 23–25, 1994*, pp. 239–251.
 Springer, 1995.

638

639 Durk P Kingma and Prafulla Dhariwal. Glow: Generative flow with invertible 1x1 convolutions.
 640 *Advances in neural information processing systems*, 31, 2018.

641

642 Yuval Kirstain, Adam Polyak, Uriel Singer, Shahbuland Matiana, Joe Penna, and Omer Levy.
 643 Pick-a-pic: An open dataset of user preferences for text-to-image generation. In *Thirty-seventh*
 644 *Conference on Neural Information Processing Systems*, 2023. URL <https://openreview.net/forum?id=G5RwHpBUv0>.

645

646 Florian Knoll, Jure Zbontar, Anuroop Sriram, Matthew J Muckley, Mary Bruno, Aaron Defazio,
 647 Marc Parente, Krzysztof J Geras, Joe Katsnelson, Hersh Chandarana, et al. fastmri: A publicly
 available raw k-space and dicom dataset of knee images for accelerated mr image reconstruction
 using machine learning. *Radiology: Artificial Intelligence*, 2(1):e190007, 2020.

648 Zhifeng Kong, Wei Ping, Jiaji Huang, Kexin Zhao, and Bryan Catanzaro. Diffwave: A versatile
 649 diffusion model for audio synthesis. In *International Conference on Learning Representations*,
 650 2021. URL <https://openreview.net/forum?id=a-xFK8Ymz5J>.

651 Siqi Kou, Lei Gan, Dequan Wang, Chongxuan Li, and Zhijie Deng. Bayesdiff: Estimating pixel-wise
 652 uncertainty in diffusion via bayesian inference. In *The Twelfth International Conference on Learning
 653 Representations*, 2024. URL <https://openreview.net/forum?id=YcM6ofShwY>.

654 Alex Krizhevsky, Geoffrey Hinton, et al. Learning multiple layers of features from tiny images. 2009.

655 Vladimir Kulikov, Matan Kleiner, Inbar Huberman-Spiegelglas, and Tomer Michaeli. Flowedit:
 656 Inversion-free text-based editing using pre-trained flow models. *arXiv preprint arXiv:2412.08629*,
 657 2024.

658 Georges Le Bellier and Nicolas Audebert. Detecting out-of-distribution earth observation images
 659 with diffusion models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and
 660 Pattern Recognition*, pp. 481–491, 2024.

661 Lizao Li, Robert Carver, Ignacio Lopez-Gomez, Fei Sha, and John Anderson. Generative emulation
 662 of weather forecast ensembles with diffusion models. *Science Advances*, 10(13):eadk4489, 2024a.

663 Xiang Li, Yixiang Dai, and Qing Qu. Understanding generalizability of diffusion models requires
 664 rethinking the hidden gaussian structure. *Advances in neural information processing systems*, 37:
 665 57499–57538, 2024b.

666 Shanchuan Lin, Bingchen Liu, Jiashi Li, and Xiao Yang. Common diffusion noise schedules and
 667 sample steps are flawed. In *Proceedings of the IEEE/CVF winter conference on applications of
 668 computer vision*, pp. 5404–5411, 2024.

669 Sifan Liu and Art B Owen. Quasi-Monte Carlo quasi-Newton in variational Bayes. *Journal of
 670 Machine Learning Research*, 22(243):1–23, 2021.

671 Sifan Liu and Art B Owen. Preintegration via active subspace. *SIAM Journal on Numerical Analysis*,
 672 61(2):495–514, 2023.

673 Sifan Liu, Snigdha Panigrahi, and Jake A. Soloff. Cross-validation with antithetic Gaussian random-
 674 ization. *arXiv preprint arXiv:2412.14423*, 2024.

675 Cheng Lu, Yuhao Zhou, Fan Bao, Jianfei Chen, Chongxuan Li, and Jun Zhu. DPM-solver++:
 676 Fast solver for guided sampling of diffusion probabilistic models, 2023. URL <https://openreview.net/forum?id=4vGwQqviud5>.

677 Pierre L’Ecuyer. Quasi-Monte Carlo methods with applications in finance. *Finance and Stochastics*,
 678 13:307–349, 2009.

679 Pierre L’Ecuyer, Marvin K. Nakayama, Art B. Owen, and Bruno Tuffin. Confidence intervals for
 680 randomized quasi-Monte Carlo estimators. In *2023 Winter Simulation Conference*, pp. 445–446,
 681 2023.

682 David Marwood, Shumeet Baluja, and Yair Alon. Diversity and diffusion: Observations on synthetic
 683 image distributions with stable diffusion. *arXiv preprint arXiv:2311.00056*, 2023.

684 Chenlin Meng, Yutong He, Yang Song, Jiaming Song, Jiajun Wu, Jun-Yan Zhu, and Stefano Ermon.
 685 SDEdit: Guided image synthesis and editing with stochastic differential equations. In *International
 686 Conference on Learning Representations*, 2022. URL https://openreview.net/forum?id=aBscjcpu_tE.

687 H. Niederreiter. *Random Number Generation and Quasi-Monte Carlo Methods*. SIAM, Philadelphia,
 688 PA, 1992.

689 A. B. Owen. Scrambled net variance for integrals of smooth functions. *Annals of Statistics*, 25(4):
 690 1541–1562, 1997a.

691 A. B. Owen. Monte Carlo variance of scrambled net quadrature. *SIAM Journal of Numerical Analysis*,
 692 34(5):1884–1910, 1997b.

702 Art B. Owen. *Monte Carlo Theory, Methods and Examples*. <https://artowen.su.domains/mc/>, 2013.

703

704

705 William Peebles and Saining Xie. Scalable diffusion models with transformers. In *Proceedings of*
 706 *the IEEE/CVF international conference on computer vision*, pp. 4195–4205, 2023.

707

708 Dustin Podell, Zion English, Kyle Lacey, Andreas Blattmann, Tim Dockhorn, Jonas Müller, Joe
 709 Penna, and Robin Rombach. SDXL: Improving latent diffusion models for high-resolution image
 710 synthesis. In *The Twelfth International Conference on Learning Representations*, 2024. URL
<https://openreview.net/forum?id=di52zR8xgf>.

711

712 Zipeng Qi, Lichen Bai, Haoyi Xiong, and Zeke Xie. Not all noises are created equally: Diffusion
 713 noise selection and optimization. *arXiv preprint arXiv:2407.14041*, 2024.

714

715 Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. High-
 716 resolution image synthesis with latent diffusion models. In *Proceedings of the IEEE/CVF confer-
 717 ence on computer vision and pattern recognition*, pp. 10684–10695, 2022.

718

719 Seyedmorteza Sadat, Jakob Buhmann, Derek Bradley, Otmar Hilliges, and Romann M. Weber. CADS:
 720 Unleashing the diversity of diffusion models through condition-annealed sampling. In *The Twelfth
 721 International Conference on Learning Representations*, 2024. URL <https://openreview.net/forum?id=zMoNrajk2X>.

722

723 Chitwan Saharia, William Chan, Saurabh Saxena, Lala Li, Jay Whang, Emily L Denton, Kamyar
 724 Ghasemipour, Raphael Gontijo Lopes, Burcu Karagol Ayan, Tim Salimans, et al. Photorealistic
 725 text-to-image diffusion models with deep language understanding. *Advances in Neural Information
 Processing Systems*, 35:36479–36494, 2022.

726

727 Tim Salimans and Jonathan Ho. Progressive distillation for fast sampling of diffusion models. In
 728 *International Conference on Learning Representations*, 2022. URL <https://openreview.net/forum?id=TIdIXTpzhoI>.

729

730 Axel Sauer, Dominik Lorenz, Andreas Blattmann, and Robin Rombach. Adversarial diffusion
 731 distillation. In *European Conference on Computer Vision*, pp. 87–103. Springer, 2024.

732

733 Jascha Sohl-Dickstein, Eric Weiss, Niru Maheswaranathan, and Surya Ganguli. Deep unsupervised
 734 learning using nonequilibrium thermodynamics. In *International conference on machine learning*,
 pp. 2256–2265. pmlr, 2015.

735

736 Bowen Song, Soo Min Kwon, Zecheng Zhang, Xinyu Hu, Qing Qu, and Liyue Shen. Solving inverse
 737 problems with latent diffusion models via hard data consistency. In *The Twelfth International
 738 Conference on Learning Representations*, 2024. URL <https://openreview.net/forum?id=j8hdRqOUhN>.

739

740 Bowen Song, Zecheng Zhang, Zhaoxu Luo, Jason Hu, Wei Yuan, Jing Jia, Zhengxu Tang, Guanyang
 741 Wang, and Liyue Shen. Ccs: Controllable and constrained sampling with diffusion models via
 742 initial noise perturbation. *arXiv preprint arXiv:2502.04670*, 2025.

743

744 Jiaming Song, Chenlin Meng, and Stefano Ermon. Denoising diffusion implicit models. In *Inter-
 745 national Conference on Learning Representations*, 2021a. URL <https://openreview.net/forum?id=St1giarCHLP>.

746

747 Yang Song, Jascha Sohl-Dickstein, Diederik P Kingma, Abhishek Kumar, Stefano Ermon, and Ben
 748 Poole. Score-based generative modeling through stochastic differential equations. In *International
 749 Conference on Learning Representations*, 2021b. URL <https://openreview.net/forum?id=PxTIG12RRHS>.

750

751 Yang Song, Liyue Shen, Lei Xing, and Stefano Ermon. Solving inverse problems in medical imaging
 752 with score-based generative models. In *International Conference on Learning Representations*,
 753 2022. URL <https://openreview.net/forum?id=vaRCHVj0uGI>.

754

755 Yang Song, Prafulla Dhariwal, Mark Chen, and Ilya Sutskever. Consistency models. In *International
 756 Conference on Machine Learning*, pp. 32211–32252. PMLR, 2023.

756 Roman Vershynin. *High-dimensional probability: An introduction with applications in data science*,
 757 volume 47. Cambridge university press, 2018.
 758

759 Patrick von Platen, Suraj Patil, Anton Lozhkov, Pedro Cuenca, Nathan Lambert, Kashif Rasul,
 760 Mishig Davaadorj, Dhruv Nair, Sayak Paul, William Berman, Yiyi Xu, Steven Liu, and Thomas
 761 Wolf. Diffusers: State-of-the-art diffusion models. <https://github.com/huggingface/diffusers>, 2022.
 762

763 Carsten Waechter and Alexander Keller. Quasi-Monte Carlo light transport simulation by efficient
 764 ray tracing, May 31 2011. US Patent 7,952,583.
 765

766 Bin Xu Wang and John Vastola. The unreasonable effectiveness of gaussian score approximation for
 767 diffusion models and its applications. *Transactions on Machine Learning Research*, 2024. ISSN
 768 2835-8856. URL <https://openreview.net/forum?id=I0uknSHM2j>.
 769

770 Hengkang Wang, Xu Zhang, Taihui Li, Yuxiang Wan, Tiancong Chen, and Ju Sun. Dmplug: A plug-in
 771 method for solving inverse problems with diffusion models. *Advances in Neural Information
 772 Processing Systems*, 37:117881–117916, 2024.
 773

774 Xiaoqun Wang and Kai-Tai Fang. The effective dimension and quasi-Monte Carlo integration.
 775 *Journal of Complexity*, 19:101–124, 2003.
 776

777 Weihao Xia, Yujiu Yang, Jing-Hao Xue, and Baoyuan Wu. Tedigan: Text-guided diverse face image
 778 generation and manipulation. In *IEEE Conference on Computer Vision and Pattern Recognition
 779 (CVPR)*, 2021.

780 Ye Xiao and Xiaoqun Wang. Enhancing quasi-Monte Carlo simulation by minimizing effective
 781 dimension for derivative pricing. *Computational Economics*, 54(1):343–366, 2019.
 782

783 Seungil You, David Ding, Kevin Canini, Jan Pfeifer, and Maya Gupta. Deep lattice networks and
 784 partial monotonic functions. *Advances in neural information processing systems*, 30, 2017.
 785

786 Fisher Yu, Ari Seff, Yinda Zhang, Shuran Song, Thomas Funkhouser, and Jianxiong Xiao. Lsun:
 787 Construction of a large-scale image dataset using deep learning with humans in the loop. *arXiv
 788 preprint arXiv:1506.03365*, 2015.

789 Jure Zbontar, Florian Knoll, Anuroop Sriram, Tullie Murrell, Zhengnan Huang, Matthew J Muckley,
 790 Aaron Defazio, Ruben Stern, Patricia Johnson, Mary Bruno, et al. fastmri: An open dataset and
 791 benchmarks for accelerated mri. *arXiv preprint arXiv:1811.08839*, 2018.
 792

793 Richard Zhang, Phillip Isola, Alexei A Efros, Eli Shechtman, and Oliver Wang. The unreasonable
 794 effectiveness of deep features as a perceptual metric. In *Proceedings of the IEEE conference on
 795 computer vision and pattern recognition*, pp. 586–595, 2018.
 796

797 Hongkai Zheng, Wenda Chu, Bingliang Zhang, Zihui Wu, Austin Wang, Berthy Feng, Caifeng
 798 Zou, Yu Sun, Nikola Borislavov Kovachki, Zachary E Ross, Katherine Bouman, and Yisong
 799 Yue. Inversebench: Benchmarking plug-and-play diffusion priors for inverse problems in physical
 800 sciences. In *The Thirteenth International Conference on Learning Representations*, 2025. URL
 801 <https://openreview.net/forum?id=U3PBIXNG6>.
 802

803 Zikai Zhou, Shitong Shao, Lichen Bai, Zhiqiang Xu, Bo Han, and Zeke Xie. Golden noise for
 804 diffusion models: A learning framework. *arXiv preprint arXiv:2411.09502*, 2024.
 805

806 A USE OF LLMs 807

808 LLMs are used only for writing, editing, or formatting purposes and do not affect the core methodol-
 809 ogy, contributions, and originality of the paper.

810 **B PROOFS**
 811

812 **B.1 PROOFS IN SECTION 3.2**
 813

814 *Proof of Lemma 1.* It suffices to show $\text{Var}(f(Z) + f(-Z)) = 0$. By definition:

815
$$\text{Corr}(f(-Z), f(Z)) = \frac{\text{Cov}(f(-Z), f(Z))}{\sqrt{\text{Var}(f(Z)) \text{Var}(f(-Z))}} = -1.$$

 816
 817

818 Therefore
 819

820
$$\begin{aligned} \text{Var}(f(Z) + f(-Z)) &= \text{Var}(f(Z)) + \text{Var}(f(-Z)) + 2 \text{Cov}(f(Z), f(-Z)) \\ 821 &= \text{Var}(f(Z)) + \text{Var}(f(-Z)) - 2\sqrt{\text{Var}(f(Z)) \text{Var}(f(-Z))} \\ 822 &= \left(\sqrt{\text{Var}(f(Z))} - \sqrt{\text{Var}(f(-Z))} \right)^2. \end{aligned}$$

 823
 824

825 On the other hand, since $-Z \sim \mathbb{N}(0, I)$ given $Z \sim \mathbb{N}(0, 1)$, we have $\text{Var}(f(Z)) = \text{Var}(f(-Z))$
 826 and then $\text{Var}(f(Z) + f(-Z)) = 0$, as claimed. \square
 827

828 **B.2 THEORETICAL GUARANTEE OF THE SYMMETRY CONJECTURE FOR LARGE t**
 829

830 **Background: The Ornstein–Uhlenbeck (OU) process.** The OU process $(X_t)_{t \geq 0}$ in \mathbb{R}^d solves the
 831 SDE

832
$$dX_t = -X_t dt + \sqrt{2} dB_t, \quad X_0 \sim \mu_0,$$

833 where (B_t) is standard Brownian motion in \mathbb{R}^d . This is a Gaussian Markov process with stationary
 834 distribution $\gamma_d = \mathbb{N}(0, I_d)$ and generator

835
$$L f(x) = \Delta f(x) - x \cdot \nabla f(x).$$

 836

837 For any initial law μ_0 , let μ_t denote the law of X_t . Then $\mu_t = \mu_0 P_t$, where $(P_t)_{t \geq 0}$ is the *OU semigroup* defined by

838
$$(P_t f)(x) = \mathbb{E}[f(X_t) \mid X_0 = x].$$

839 Write the *score* of μ_t as $s_t(x) := \nabla \log p_t(x)$, where p_t is Lebesgue density of μ_t , we have
 840

841 **Theorem 1** (Score converges to the Gaussian score). *Under the setup above,*

842
$$\lim_{t \rightarrow \infty} \mathbb{E}_{\mu_t} [\|s_t(X_t) + X_t\|^2] = 0.$$

 843
 844

845 Moreover, the convergence is quantified by
 846

847
$$\mathbb{E}_{\mu_t} [\|s_t(X_t) + X_t\|^2] = \mathcal{I}(\mu_t \mid \gamma_d) \leq e^{-2t} \mathcal{I}(\mu_0 \mid \gamma_d).$$

 848

849 *Proof.* Recall the definition of relative Fisher information
 850

851
$$\mathcal{I}(\mu_t \mid \gamma_d) := \mathbb{E}_{\mu_t} [\|\nabla \log(p_t/\gamma_d)\|^2].$$

 852

853 Since $\nabla \gamma_d(x) = -x$, we have:

854
$$\mathcal{I}(\mu_t \mid \gamma_d) = \mathbb{E}_{\mu_t} [\|s_t(X_t) + X_t\|^2].$$

 855

856 Moreover, it is well known (e.g. Chapter 5 of Bakry et al. (2013)) that $\mathcal{I}(\mu_t \mid \gamma_d) \leq \exp(-2t) \mathcal{I}(\mu_0 \mid \gamma_d)$, this concludes our proof. \square
 857
 858

859 In variance-preserving diffusion models, the forward noising process is exactly the Orn-
 860 stein–Uhlenbeck semigroup. Theorem 1 implies that the true score converges to the Gaussian score
 861 $-x$. Consequently, one-dimensional slices of the score become asymptotically affine antisymmetric,
 862 confirming our symmetry conjecture in the high-noise limit.
 863

The next corollary shows, the 1-step DDIM update has nearly -1 correlation at large t .

864 **Corollary 1** (Correlation of 1-step DDIM). *With all the setup the same as above, assuming we*
 865 *have a score network $\epsilon^{(t)}$ approximating the score function. Let $\eta_t := \max\{\mathbb{E}_{p_t}\|\epsilon^{(t)}(X) -$
 866 $s_t(X)\|^2, \mathbb{E}_{p_t}\|\epsilon^{(t)}(-X) - s_t(-X)\|^2\}$ be the expected (symmetrized) error. Consider one-step*
 867 *DDIM update $F_t(x) = a_t x + b_t \epsilon_\theta^{(t)}(x) := (F_{t,1}(x), \dots, F_{t,d}(x)) \in \mathbb{R}^d$ as defined in (2). Let*

$$869 \quad v_{t,i} := \min\{\text{Var}_{p_t}(F_{t,i}(X)), \text{Var}_{p_t}(F_{t,i}(-X))\}.$$

870 *Then*

$$872 \quad |\text{Corr}(F_{t,i}(X), F_{t,i}(-X)) + 1| \leq \frac{2|b_t|}{\sqrt{v_{t,i}}} \left(\sqrt{\eta_t} + e^{-t} \sqrt{\mathcal{I}(\mu_0 \mid \gamma_d)} \right).$$

875 *Proof.* Theorem 1 shows $\mathcal{I}(\mu_t \mid \gamma_d) \leq \exp(-2t)\mathcal{I}(\mu_0 \mid \gamma_d)$. Write $\epsilon_\theta^{(t)} = s_t + r_t$ and $s_t = -x + \Delta_t$,
 876 so that $\mathbb{E}_{p_t}\|r_t(X)\|^2 \leq \eta_t$ and $\mathbb{E}_{p_t}\|\Delta_t(X)\|^2 = \mathcal{I}(\mu_t \mid \gamma_d)$. Then

$$878 \quad F_t(x) = (a_t - b_t)x + b_t(\Delta_t(x) + r_t(x)), \quad F_t(-x) = -F_t(x) + E_t(x),$$

879 with $E_t(x) := b_t[\Delta_t(-x) + \Delta_t(x) + r_t(-x) + r_t(x)]$. By triangle inequality of $L^2(p_t)$ norm,

$$881 \quad \|E_t\|_{L^2(p_t)} \leq |b_t| (e^{-t} \sqrt{\mathcal{I}(\mu_0 \mid \gamma_d)} + e^{-t} \sqrt{\mathcal{I}(\mu_0 \mid \gamma_d)} + 2\sqrt{\eta_t}).$$

882 Using

$$884 \quad \text{Corr}(F_{t,i}(X), F_{t,i}(-X)) = -1 + \frac{\text{Cov}(F_{t,i}(X), E_t(x))}{\sqrt{\text{Var}(F_{t,i}(X)) \text{Var}(F_{t,i}(-X))}}$$

886 and Cauchy–Schwarz with the variance lower bounds gives

$$888 \quad |\text{Corr}(F_t(X), F_t(-X)) + 1| \leq \frac{\|E_t\|_{L^2(p_t)}}{\sqrt{v_{t,i}}} \leq \frac{2|b_t|}{\sqrt{v_{t,i}}} (\sqrt{\eta_t} + e^{-t} \sqrt{\mathcal{I}(\mu_0 \mid \gamma_d)}),$$

890 as claimed. \square

892 In practice, if $v_{t,i} > 0$, the correlation is close to -1 for large t ; its deviation from -1 is on the order
 893 of the neural network approximation error plus an exponentially decaying term Ce^{-t} .

894 B.3 MONOTONE CONVERGENCE IN t

896 Section B.2 shows that the score converges to the Gaussian score as $t \rightarrow \infty$. We now further
 897 prove that (1) the score error $\mathbb{E}_{\mu_t} [|s_t(X_t) + X_t|^2]$ and (2) the density ratio p_t/γ_d both converge
 898 monotonically in t .

900 B.3.1 DENSITY RATIO CONVERGENCE

902 **Background: Hermite Polynomials in \mathbb{R} :** For $x \in \mathbb{R}$, the (probabilist’s) Hermite polynomials
 903 $\{H_n(x)\}_{n \geq 0}$ form a sequence of polynomials defined by

$$904 \quad H_n(x) = (-1)^n \exp\left(\frac{x^2}{2}\right) \frac{d^n}{dx^n} \exp\left(-\frac{x^2}{2}\right).$$

907 We summarize their known properties here. The first two can be checked by direct calculation. The
 908 latter two can be found on page 105 of Bakry et al. (2013).

910 **Proposition 1.** *Hermite polynomials $\{H_n(x)\}_{n \geq 0}$ satisfy the following*

- 911 • (Recurrence relation) $H_{n+1}(x) = x H_n(x) - H'_n(x), \quad n \geq 0$.
- 912 • (Orthogonality under Gaussian measure) We have:

$$914 \quad \int H_n(x) H_m(x) \gamma_1(dx) = n! \delta_{nm},$$

917 where $\gamma_1(dx) = (2\pi)^{-1/2} e^{-x^2/2} dx$ denotes the one-dimensional standard Gaussian measure, and δ_{nm} is 1 when $n = m$ and 0 otherwise.

- (Completeness) Hermite polynomials $\{H_n(x)\}_{n \geq 0}$ form an orthogonal basis of the Hilbert space $L^2(\gamma) := \{f : \int f^2 \gamma(dx) < \infty\}$ equipped with inner product $\langle f, g \rangle_\gamma := \int f g \gamma(dx)$.
- (Eigenfunction) Let L be a differential operator defined as $L(f) := f'' - x f'$. Hermite polynomials $\{H_n(x)\}_{n \geq 0}$ are eigenfunctions of L :

$$L(H_n) = -n H_n, \quad n \geq 0.$$

Hermite Polynomials in \mathbb{R}^d : One can naturally extend univariate Hermite polynomials to the multivariate setting. For $x = (x_1, \dots, x_d) \in \mathbb{R}^d$, let γ_d denote the d -dimensional standard Gaussian measure,

$$\gamma_d(dx) := (2\pi)^{-d/2} \exp\left(-\frac{\|x\|^2}{2}\right) dx.$$

For a multi-index $\alpha = (\alpha_1, \dots, \alpha_d) \in \mathbb{N}^d$, we define the *multivariate Hermite polynomial*

$$H_\alpha(x) := \prod_{i=1}^d H_{\alpha_i}(x_i),$$

where H_{α_i} is the one-dimensional probabilists' Hermite polynomial. We write $|\alpha| := \alpha_1 + \dots + \alpha_d$ and $\alpha! := \alpha_1! \dots \alpha_d!$. Similar to the one-dimensional case, the family $\{H_\alpha\}_{\alpha \in \mathbb{N}^d}$ forms an orthogonal basis of the Hilbert space $L^2(\gamma_d)$:

$$\int_{\mathbb{R}^d} H_\alpha(x) H_\beta(x) \gamma_d(dx) = \alpha! \delta_{\alpha\beta}.$$

Monotone convergence of the density ratio. Recall from Section B.2 that the generator of the OU process

$$Lf(x) = \Delta f(x) - x \cdot \nabla f(x).$$

and the corresponding Markov semigroup $P_t = \exp(tL)$. We show that the multivariate Hermite polynomials are eigenfunctions of L , and thus also of P_t .

Proposition 2. *We have:*

- $LH_\alpha = -|\alpha| H_\alpha$, $\alpha \in \mathbb{N}^d$,
- $P_t H_\alpha = e^{-|\alpha|t} H_\alpha$, $\alpha \in \mathbb{N}^d, t \geq 0$.
- For every $g \in L^2(\gamma_d)$

$$P_t g(x) = \sum_{\alpha \in \mathbb{N}^d} a_\alpha e^{-t|\alpha|} H_\alpha(x),$$

where $a_\alpha = \langle g, H_\alpha \rangle_{L^2(\gamma_d)}$.

Proof. We can directly calculate

$$\begin{aligned} LH_\alpha(x) &= \Delta H_\alpha(x) - x \cdot \nabla H_\alpha(x) \\ &= \sum_i H''_{\alpha_i}(x_i) H_{\alpha_{-i}}(x_{-i}) - \sum_i x_i H'_{\alpha_i}(x_i) H_{\alpha_{-i}}(x_{-i}) \\ &= \sum_i H_{\alpha_{-i}}(x_{-i}) (H''_{\alpha_i}(x_i) - x_i H'_{\alpha_i}(x_i)) \\ &= \sum_i (-\alpha_i) H_{\alpha_{-i}}(x_{-i}) H_{\alpha_i}(x_i) \\ &= -|\alpha| H_\alpha(x), \end{aligned}$$

where $H_{\alpha_{-i}}(x_{-i}) = \prod_{j \neq i} H_{\alpha_j}(x_j)$. The second to last equality follows from the fourth property of Proposition 1.

Consequently, the OU semigroup $(P_t)_{t \geq 0}$ satisfies

$$P_t H_\alpha = e^{-t|\alpha|} H_\alpha, \quad t \geq 0.$$

972 For every $g \in L^2(\gamma_d)$, we rewrite g according to its Hermite expansion (which is valid according to
 973 properties 2-3 in Proposition 1).

$$975 \quad g(x) = \sum_{\alpha \in \mathbb{N}^d} a_\alpha H_\alpha(x), \quad a_\alpha = \langle g, H_\alpha \rangle_{L^2(\gamma_d)}.$$

977 Therefore

$$979 \quad P_t g(x) = \sum_{\alpha \in \mathbb{N}^d} a_\alpha e^{-t|\alpha|} H_\alpha(x),$$

981 as claimed. \square

984 In particular, let $p_t(x) = f_t(x)\gamma_d(x)$ denote the density at time t of the OU process started from p_0 .
 985 It is known that $f_t = P_t f_0$. Proposition 2 implies

$$987 \quad f_t(x) = \sum_{\alpha \in \mathbb{N}^d} a_\alpha e^{-t|\alpha|} H_\alpha(x) = a_0 + \sum_{|\alpha| > 0} a_\alpha e^{-t|\alpha|} H_\alpha(x),$$

989 where a_α are the Hermite coefficients of f_0 . Since $H_0 \equiv 1$ and

$$991 \quad a_0 = \int_{\mathbb{R}^d} f_0(x) H_0(x) \gamma_d(dx) = \int_{\mathbb{R}^d} f_0(x) \gamma_d(dx) = \int_{\mathbb{R}^d} p_0(x) dx = 1,$$

994 we may rewrite

$$995 \quad f_t(x) = 1 + \sum_{|\alpha| > 0} a_\alpha e^{-t|\alpha|} H_\alpha(x). \quad (3)$$

998 We now show f_t monotonically converges to 1 in $L^2(\gamma_d)$.

1000 **Proposition 3.** *Let $p_t = f_t \gamma_d$ be as above, and assume $f_0 \in L^2(\gamma_d)$. Then $\|f_t - 1\|_{L^2(\gamma_d)}$
 1001 monotonically decays to 0 at the speed of e^{-t} .*

1003 *Proof.* From (3) we have

$$1005 \quad f_t(x) = 1 + \sum_{|\alpha| > 0} a_\alpha e^{-t|\alpha|} H_\alpha(x).$$

1008 Since $\{H_\alpha\}_{\alpha \in \mathbb{N}^d}$ is an orthogonal basis of $L^2(\gamma_d)$, we obtain

$$1009 \quad \|f_t - 1\|_{L^2(\gamma_d)}^2 = \sum_{\alpha \in \mathbb{N}^d} a_\alpha^2 e^{-2t|\alpha|}.$$

1012 Since each term in the sum decreases monotonically to 0, the same holds for $\|f_t - 1\|_{L^2(\gamma_d)}^2$. The
 1013 overall decay rate is governed by the slowest mode, which corresponds to the smallest $|\alpha| = 1$, thus
 1014 the rate of convergence is $\exp(-2t)$. Taking square roots yields the desired result. \square

1017 B.3.2 SCORE ERROR CONVERGENCE

1019 Recall from Section B.2 that $\mathbb{E}_{\mu_t}[\|s_t(X_t) + X_t\|^2] = \mathcal{I}(\mu_t \mid \gamma_d)$. It follows from equation 5.7.2 in
 1020 Bakry et al. (2013) that

$$1021 \quad \frac{d\mathcal{I}(\mu_t \mid \gamma_d)}{dt} \leq -2\mathcal{I}(\mu_t \mid \gamma_d) \leq 0.$$

1024 Thus $\mathcal{I}(\mu_t \mid \gamma_d)$ is monotonically decreasing as a function of t . Further, Grönwall's inequality implies

$$1025 \quad \mathcal{I}(\mu_t \mid \gamma_d) \leq e^{-2t} \mathcal{I}(\mu_0 \mid \gamma_d).$$

1026
1027

B.4 SYMMETRY PRESERVATION UNDER THE FORWARD PROCESS

1028
1029
1030

This section shows the forward process preserves any orthogonal symmetry in the data distribution. More precisely, let G be a subgroup of the orthogonal group $O(d)$ in \mathbb{R}^d (e.g., coordinate sign flips, coordinate permutations, and rotations). We have the following:

1031
1032
1033

Proposition 4. *Assuming p_0 is symmetric about μ under the action of G , i.e., for every $g \in G$, we have $p_0(x) = p_0(g \cdot (x - \mu) + \mu)$ for every x . Then p_t satisfies $p_t(x) = p_t(g \cdot (x - \mu_t) + \mu_t)$ for every x , and $s_t(\mu_t + g \cdot x) = g \cdot s_t(\mu_t + x)$, where $\mu_t = e^{-t}\mu$.*

1034
10351036
1037

Proof. Let $X_0 \sim p_0, Z \sim \mathbb{N}(0, I_d)$. It is known that $X_t := e^{-t}X_0 + \sqrt{1 - e^{-2t}}Z \sim p_t$. For any $g \in G$,

1038
1039
1040
1041

$$\begin{aligned} g \cdot (X_t - \mu_t) + \mu_t &= g \cdot (e^{-t}X_0 + \sqrt{1 - e^{-2t}}Z - \mu_t) + \mu_t \\ &= e^{-t}(g \cdot (X_0 - \mu) + \mu) + \sqrt{1 - e^{-2t}}g \cdot Z \sim p_t \end{aligned}$$

1042
1043
1044

where the last claim uses that $g \cdot (X_0 - \mu) + \mu$ has the same distribution as X_0 (by symmetry) and $g \cdot Z \sim \mathbb{N}(0, I_d)$. This proves the claim on the symmetry of p_t . For the score, since $p_t(\mu_t + g \cdot x) = p_t(\mu_t + x)$, taking derivative on both sides yields:

1045

$$g^\top \cdot \nabla p_t(\mu_t + g \cdot x) = \nabla p_t(\mu_t + x).$$

1046
1047

Since g is an orthogonal matrix, $g^\top = g^{-1}$, we have:

1049
1050

$$\nabla p_t(\mu_t + g \cdot x) = g \cdot \nabla p_t(\mu_t + x).$$

1051
1052
1053

Dividing both sides by $p_t(\mu_t + x)$ proves the claimed result. \square

Sections B.2 (large t limit), B.3 (monotone convergence), and B.4 (symmetry perservation from $t = 0$) together provide insights for the conjecture: The forward OU process preserves all orthogonal symmetries of p_0 , including central reflections. Hence, in the idealized case where p_0 is reflection-symmetric, the density is even and its score is odd. Proposition 4 ensures that this symmetry is preserved for all t . Furthermore, even when the symmetry is only approximate near $t \approx 0$, the monotone convergence of p_t toward the Gaussian is exponentially fast, so the forward dynamics push s_t toward a linear function in a monotone manner.

1061
1062

B.5 PROOFS IN SECTION 4

1063
1064
1065

B.5.1 MONTE CARLO ESTIMATOR:

1066
1067

Proposition 5. Denote $\mathbb{E}_{z \sim \mathbb{N}(0, I)}[S(DM(z))]$ by μ and $\text{Var}_{z \sim \mathbb{N}(0, I)}[S(DM(z))]$ by σ^2 . Assuming $\sigma^2 < \infty$, then we have:

1068
1069

- $\hat{\mu}_N^{MC} \rightarrow \mu$ and $(\hat{\sigma}_N^{MC})^2 \rightarrow \sigma^2$ almost surely, and $\mathbb{P}(\mu \in \text{CI}_N^{MC}(1 - \alpha)) \rightarrow 1 - \alpha$, both as $N \rightarrow \infty$.
- $\mathbb{E}[(\hat{\mu}_N^{MC} - \mu)^2] = \text{Var}[\hat{\mu}_N^{MC}] = \sigma^2/N$.

1070
1071
1072
1073
1074
1075
1076
1077

In words, the above proposition shows: (i) *Correctness*: The standard Monte Carlo estimator $\hat{\mu}_N^{MC}$ converges to the true value, and the coverage probability of $\text{CI}_N^{MC}(1 - \alpha)$ converges to the nominal level $(1 - \alpha)$. (ii) *Reliability*: The expected squared error of $\hat{\mu}_N^{MC}$ equals the variance of a single sample divided by sample size. The confidence interval has width approximately $2\sigma z_{1-\alpha/2}/\sqrt{N}$.

1078
1079

Proof of Proposition 5. Let $S_i := S(DM(z_i))$ for $i = 1, \dots, N$. Because the noises z_i are drawn independently from $\mathbb{N}(0, I)$, the random variables S_1, S_2, \dots are independent and identically distributed with mean μ and variance $\sigma^2 < \infty$.

1080 **Consistency.** By the *strong law of large numbers*,

$$1082 \quad 1083 \quad 1084 \quad \hat{\mu}_N^{\text{MC}} = \frac{1}{N} \sum_{i=1}^N S_i \xrightarrow{\text{a.s.}} \mu.$$

1085 The sample variance estimator

$$1087 \quad 1088 \quad 1089 \quad (\hat{\sigma}_N^{\text{MC}})^2 = \frac{1}{N-1} \sum_{i=1}^N (S_i - \hat{\mu}_N^{\text{MC}})^2 = \frac{1}{N-1} \sum_{i=1}^N S_i^2 - \frac{N}{N-1} (\hat{\sigma}_N^{\text{MC}})^2.$$

1090 The first term converges to $\mathbb{E}[S_1^2]$ almost surely by the law of large numbers, the second term
1091 converges to μ^2 almost surely as shown above. Therefore $(\hat{\sigma}_N^{\text{MC}})^2 \xrightarrow{\text{a.s.}} \sigma^2$.

1093 **Asymptotic normality and coverage.** The classical *central limit theorem* states that

$$1095 \quad 1096 \quad \sqrt{N} \frac{\hat{\mu}_N^{\text{MC}} - \mu}{\sigma} \xrightarrow{d} \mathcal{N}(0, 1).$$

1098 Replacing the unknown σ by the consistent estimator $\hat{\sigma}_N^{\text{MC}}$ and applying *Slutsky's theorem* yields

$$1100 \quad 1101 \quad \sqrt{N} \frac{\hat{\mu}_N^{\text{MC}} - \mu}{\hat{\sigma}_N^{\text{MC}}} \xrightarrow{d} \mathcal{N}(0, 1).$$

1102 Hence

$$1103 \quad 1104 \quad \mathbb{P}(\mu \in \text{CI}_N^{\text{MC}}(1 - \alpha)) = \mathbb{P}\left(\left|\sqrt{N}(\hat{\mu}_N^{\text{MC}} - \mu)/\hat{\sigma}_N^{\text{MC}}\right| \leq z_{1-\alpha/2}\right) \longrightarrow 1 - \alpha.$$

1106 **Mean-squared error.** Because $\hat{\mu}_N^{\text{MC}}$ is the average of N i.i.d. variables,

$$1108 \quad 1109 \quad \text{Var}[\hat{\mu}_N^{\text{MC}}] = \frac{\sigma^2}{N}.$$

1110 Moreover, since $\hat{\mu}_N^{\text{MC}}$ is unbiased ($\mathbb{E}[\hat{\mu}_N^{\text{MC}}] = \mu$), its mean-squared error

$$1112 \quad 1113 \quad \mathbb{E}[(\hat{\mu}_N^{\text{MC}} - \mu)^2] = \mathbb{E}[(\hat{\mu}_N^{\text{MC}} - \mathbb{E}[\hat{\mu}_N^{\text{MC}}])^2] = \frac{\sigma^2}{N}.$$

1114 This completes the proof. □

1116 B.5.2 ANTITHETIC MONTE CARLO ESTIMATOR:

1118 **Proposition 6.** Denote $\mathbb{E}[S(\text{DM}(z))]$ by μ , $\text{Var}[S(\text{DM}(z))]$ by σ^2 , and $\text{Cov}(S_1^+, S_1^-)$ by ρ . Assuming
1119 $\sigma^2 < \infty$, then we have:

- 1121 • $\hat{\mu}_N^{\text{AMC}} \rightarrow \mu$ and $(\hat{\sigma}_N^{\text{AMC}})^2 \rightarrow (1 + \rho)\sigma^2/2$ almost surely as $N \rightarrow \infty$,
- 1123 • $\mathbb{E}[(\hat{\mu}_N^{\text{AMC}} - \mu)^2] = \text{Var}[\hat{\mu}_N^{\text{AMC}}] = \sigma^2(1 + \rho)/N$.
- 1125 • $\mathbb{P}(\mu \in \text{CI}_N^{\text{AMC}}(1 - \alpha)) \rightarrow 1 - \alpha$ as $N \rightarrow \infty$.

1127 *Proof.* Let $N = 2K$ and generate independent antithetic noise pairs $(z_i, -z_i)$ for $i = 1, \dots, K$.
1128 Recall

$$1130 \quad S_i^+ = S(\text{DM}(z_i)), \quad S_i^- = S(\text{DM}(-z_i)), \quad \bar{S}_i = \frac{1}{2}(S_i^+ + S_i^-), \quad i = 1, \dots, K.$$

1132 Because the pairs are independent and identically distributed (i.i.d.), the random variables $\bar{S}_1, \dots, \bar{S}_K$
1133 are i.i.d. with

$$\mathbb{E}[\bar{S}_1] = \mu$$

1134 and

1135
$$\text{Var}[\bar{S}_1] = \text{Var}\left[\frac{1}{2}(S_1^+ + S_1^-)\right] \quad (4)$$

1136
$$= \left(\frac{1}{2}\right)^2 \text{Var}[S_i^+ + S_i^-] \quad (5)$$

1137
$$= \frac{1}{4}(\text{Var}[S_i^+] + \text{Var}[S_i^-] + 2 \text{Cov}(S_i^+, S_i^-)) \quad (6)$$

1138
$$= \frac{1}{4}(\sigma^2 + \sigma^2 + 2\rho\sigma^2) \quad (7)$$

1139
$$= \frac{1+\rho}{2}\sigma^2. \quad (8)$$

1140 where we have written $\text{Cov}(S_i^+, S_i^-) = \rho\sigma^2$ as in the statement. Set $\hat{\mu}_N^{\text{AMC}} = K^{-1} \sum_{i=1}^K \bar{S}_i$ and
1141 let $(\hat{\sigma}_N^{\text{AMC}})^2$ be the sample variance of $\bar{S}_1, \dots, \bar{S}_K$. All three claims in Proposition 6 follow from
1142 Proposition 5. \square

1143

1144

1145 B.6 AN ALTERNATIVE EXPLANATION VIA THE FKG INEQUALITY

1146

1147 We also provide an expository discussion highlighting the FKG connection behind negative correlation.
1148 Consider the univariate case where a scalar Gaussian z is fed through a sequence of one-dimensional
1149 linear maps and ReLU activations. Let F be the resulting composite function. One can show that,
1150 regardless of the number of layers or the linear coefficients used, $\text{Corr}(F(z), F(-z)) \leq 0$. The
1151 proof relies on the univariate FKG inequality (Fortuin et al., 1971), a well-known result in statistical
1152 physics and probability theory.

1153

1154 We generalize this result to higher dimensions via *partial monotonicity*, under which negative
1155 correlation still holds.

1156

1157 We first formally state and prove the claim for the univariate case in Section B.6.

1158 **Proposition 7** (Univariate case). *Let $z \sim \mathbb{N}(0, 1)$ and let $F : \mathbb{R} \rightarrow \mathbb{R}$ be the output of any
1159 one-dimensional feed-forward network obtained by alternating scalar linear maps and ReLU
1160 activations:*

1161
$$h_0(z) = z, \quad h_\ell(z) = \text{ReLU}(w_\ell h_{\ell-1}(z) + b_\ell), \quad \ell = 1, \dots, L, \quad F(z) = h_L(z).$$

1162 For all choices of depths L and coefficients $\{w_\ell, b_\ell\}_{\ell=1}^L$,

1163
$$\text{Corr}(F(z), F(-z)) \leq 0.$$

1164

1165 *Proof.* An important observation is that monotonicity is preserved under composition: combining
1166 one monotonic function with another produces a function that remains monotonic.1167 Each scalar linear map $x \mapsto w_\ell x + b_\ell$ is monotone: it is non-decreasing if $w_\ell \geq 0$ and non-increasing
1168 if $w_\ell < 0$. The ReLU map $x \mapsto \text{ReLU}(x) = \max\{0, x\}$ is non-decreasing. Hence the final function
1169 F is monotone. Without loss of generality, we assume F is non-decreasing.1170 FKG inequality guarantees $\text{Cov}_{Z \sim \mathbb{N}(0, 1)}(f(Z), g(Z)) \geq 0$ provided that f, g are non-decreasing.
1171 Therefore, $\text{Cov}_{Z \sim \mathbb{N}(0, 1)}(F(Z), F(-Z)) = -\text{Cov}_{Z \sim \mathbb{N}(0, 1)}(F(Z), -F(-Z)) \leq 0$. Since

1172
$$\text{Corr}(F(z), F(-z)) = \frac{\text{Cov}_{Z \sim \mathbb{N}(0, 1)}(F(Z), F(-Z))}{\sqrt{\text{Var}(F(z)) \text{Var}(F(-z))}},$$

1173 we have $\text{Corr}(F(z), F(-z)) \leq 0$. \square 1174 To generalize the result to higher dimension, we define a function $f : \mathbb{R}^m \rightarrow \mathbb{R}$ to be
1175 *partially monotone* if for each coordinate j , holding all other inputs fixed, the map $t \mapsto$
1176 $f(x_1, \dots, x_{j-1}, t, x_{j+1}, \dots, x_m)$ is either non-decreasing or non-increasing. Mixed monotonic-
1177 ity is allowed. For example, $f(x, y) = x - y$ is non-decreasing in x and non-increasing in y , yet still
1178 qualifies as partially monotone. We have the following result:

1188 **Proposition 8.** For a diffusion model $DM : \mathbb{R}^d \rightarrow \mathbb{R}^m$ and a summary statistics $S : \mathbb{R}^m \rightarrow \mathbb{R}$, if the
 1189 joint map is partially monotone, then $\text{Corr}(S \circ DM(Z), S \circ DM(-Z)) \leq 0$.
 1190

1191 Now we prove the general case:
 1192

1193 *Proof of Proposition 8.* Let $G := S \circ DM$. For each coordinate $j \in [d]$ fix a sign
 1194

$$s_j = \begin{cases} +1, & \text{if } G \text{ is non-decreasing in } x_j, \\ -1, & \text{if } G \text{ is non-increasing in } x_j. \end{cases}$$

1197 Write $s = (s_1, \dots, s_d) \in \{\pm 1\}^d$ and define, for any $z \in \mathbb{R}^d$,
 1198

$$\tilde{G}(z) = \tilde{G}(z_1, \dots, z_d) := G(s_1 z_1, \dots, s_d z_d).$$

1200 Similarly, define
 1201

$$\tilde{H}(z) := \tilde{G}(-z) = G(-s_1 z_1, \dots, -s_d z_d).$$

1203 By construction, \tilde{G} is coordinate-wise non-decreasing and \tilde{H} is coordinate-wise non-increasing.
 1204

1205 Because each coordinate of $Z \sim \mathbb{N}(0, I_d)$ is symmetric, the random vectors $(s_1 Z_1, \dots, s_d Z_d)$ and
 1206 (Z_1, \dots, Z_d) have the same distribution. Hence

$$\text{Cov}(G(Z), G(-Z)) = \text{Cov}(\tilde{G}(Z), \tilde{H}(Z)).$$

1209 The multivariate FKG inequality (Fortuin et al., 1971) for product Gaussian measures states that,
 1210 when U and V are coordinate-wise non-decreasing, $\text{Cov}(U(Z), V(Z)) \geq 0$. Apply it to the pair
 1211 $(\tilde{G}, -\tilde{H})$: both components are non-decreasing, so
 1212

$$\text{Cov}(\tilde{G}(Z), -\tilde{H}(Z)) \geq 0 \implies \text{Cov}(\tilde{G}(Z), \tilde{H}(Z)) \leq 0 \Leftrightarrow \text{Cov}(G(Z), G(-Z)) \leq 0.$$

1215 Therefore $\text{Corr}(G(Z), G(-Z)) \leq 0$, as claimed. \square
 1216

1217 Proposition 8 relies on checking partial-monotonicity. If S is partially monotone (including any linear
 1218 statistic), then the conditions of Proposition 8 are satisfied by, e.g., Neural Additive Models (Agarwal
 1219 et al., 2021) and Deep Lattice Networks (You et al., 2017). Unfortunately, partial-monotonicity is in
 1220 general hard to verify.

1221 While popular diffusion architectures like DiT and U-Net lack partial monotonicity, we include
 1222 Proposition 8 as an expository attempt to highlight the FKG connection behind negative correlation.
 1223

1224 B.6.1 FKG INEQUALITY FOR DDIM

1226 We first analyze the idealized DDIM process using the FKG inequality. A single step of the idealized
 1227 DDIM is

$$F_t(\mathbf{x}) = a_t \mathbf{x} + c_t s_t(\mathbf{x}), \quad (9)$$

1229 where s_t denotes the score of p_t , and $a_t, c_t \geq 0$ are deterministic coefficients obtained by rearranging
 1230 (2).

1231 The idealized DDIM trajectory is given by the composition:
 1232

$$\text{DDIM}_I = F_1 \circ F_2 \circ F_3 \circ \dots \circ F_T. \quad (10)$$

1235 We present two results that give sufficient conditions under which the local (one-step) DDIM update
 1236 (9) and the global DDIM procedure (10) generate negative correlation.

1237 We need the following definition:

1238 **Definition 1** (MTP₂ with curvature bound). Let $p : \mathbb{R}^d \rightarrow (0, \infty)$ be a C^2 probability density. We
 1239 say that p is multivariate totally positive of order 2 (MTP₂) with curvature bound $\kappa \geq 0$ if the second
 1240 partial derivatives of $\log p$ satisfy, for all $x \in \mathbb{R}^d$,
 1241

$$\partial_{ij}^2 \log p(x) \geq 0 \quad \text{for all } i \neq j, \quad \text{and} \quad \partial_{ii}^2 \log p(x) \geq -\kappa \quad \text{for all } i.$$

1242 MTP₂ distributions are widely studied in statistics and probability Karlin & Rinott (1980). In
 1243 particular, isotropic Gaussian distributions are MTP₂.

1244 **Proposition 9** (One-step DDIM). *With all the notations as above, fix $t > 0$, assume the distribution
 1245 p_t is MTP₂ with curvature bound κ_t , and*

$$1247 \quad a_t \geq c_t \kappa_t,$$

1248 *then F_t is coordinatewise nondecreasing: $\partial_{x_j} F_{t,i}(x) \geq 0$ for all i, j and all x . Moreover, for any
 1249 linear statistics S , we have $\text{Corr}(S \circ F_t(Z)), S \circ F_t(-Z)) \leq 0$.*

1250
 1251 *Proof.* Differentiate componentwise:

$$1253 \quad \frac{\partial F_{t,i}}{\partial x_j}(x) = a_t \delta_{ij} + c_t \partial_{ij}^2 \log p_t(x).$$

1254 For $i \neq j$ this is ≥ 0 by $c_t \geq 0$ and the mixed second derivative condition. For $i = j$,

$$1257 \quad \frac{\partial F_{t,i}}{\partial x_i}(x) \geq a_t + c_t (-\kappa_t) \geq 0.$$

1258 Let $G = S \circ F_t$. Since F_t is coordinate non-decreasing and S is linear, we have that $G = S \circ F_t$ is
 1259 partially monotone. Thus $\text{Corr}(S \circ F_t(Z)), S \circ F_t(-Z)) \leq 0$ by the multivariate FKG inequality
 1260 and Proposition 8. \square

1261 The next result shows that the idealized DDIM trajectory continues to generate negative correlation,
 1262 as long as the curvature bound holds uniformly over all steps.

1263 **Proposition 10** (Global DDIM chain). *With all the notations as above. Assume the distribution p_t is
 1264 MTP₂ with curvature bound κ_t , and*

$$1265 \quad a_t \geq c_t \kappa_t$$

1266 *for every t . Then $\text{DDIM}_I = F_1 \circ F_2 \circ F_3 \circ \dots \circ F_T$ is coordinatewise nondecreasing. Moreover, for
 1267 any linear statistics S , we have $\text{Corr}(S \circ \text{DDIM}_I(Z)), S \circ \text{DDIM}_I(-Z)) \leq 0$.*

1268 *Proof.* Since each F_t is coordinatewise nondecreasing, their composition $F_1 \circ \dots \circ F_T$ is also
 1269 coordinatewise nondecreasing. Therefore, the composition $S \circ \text{DDIM}_I$ is partially monotone, thus
 1270 $\text{Corr}(S \circ \text{DDIM}_I(Z)), S \circ \text{DDIM}_I(-Z)) \leq 0$ as claimed. \square

1271 **Remark 1.** *All these results remain valid if we replace the exact score $\nabla \log p_t$ by the neural network
 1272 s_θ , up to a constant rescaling.*

1273 **Remark 2.** *We note that the condition $a_t \geq c_t \kappa_t$ is satisfied when $\kappa_t = O(1)$ and the probability
 1274 flow ODE is discretized with sufficiently small step size. For the deterministic DDIM update under a
 1275 variance-preserving schedule,*

$$1276 \quad a_t = \sqrt{\frac{\alpha_{t-1}}{\alpha_t}}, \quad c_t = \sqrt{1 - \alpha_{t-1}} - \sqrt{\frac{\alpha_{t-1}}{\alpha_t}} \sqrt{1 - \alpha_t}.$$

1277 *Thus a_t is of unit order, and c_t captures the difference $|\alpha_{t-1} - \alpha_t|$. In the continuous-time limit with
 1278 step size Δt and a smooth function $\alpha(t)$, a Taylor expansion yields $a_t = 1 + O(\Delta t)$ and $c_t = O(\Delta t)$;
 1279 thus $a_t \geq c_t \kappa_t$ holds when Δt is sufficiently small and κ_t stays bounded.*

C ADDITIONAL EXPERIMENTS

C.1 CORRELATION EXPERIMENT SETUP ON OTHER MODELS

C.1.1 CONSISTENCY MODEL

1294 We study both unconditional and conditional consistency models using publicly available check-
 1295 points provided in Hugging Face’s Diffusers library (von Platen et al., 2022): openai/diffusers-
 1296 cd_imagenet64_l2, openai/diffusers-cd_cat256_l2, and openai/diffusers-cd_bedroom256_l2.

1296 For unconditional models, we evaluate pre-trained models on LSUN-Cat and LSUN-Bedroom (Yu
 1297 et al., 2015). For each dataset, 1,600 image pairs are generated under both PN and RR noise sampling
 1298 with 1 DDIM steps.

1299 For conditional models, we evaluate pre-trained models on ImageNet on 32 classes (Deng et al.,
 1300 2009). For each class, 100 PN and RR pairs are generated with 1 steps.

1302 C.1.2 GENERATIVE MODELS BEYOND DIFFUSION

1304 A mentioned in Section 3, we evaluate correlation on a VAE and flow-based models. Unlike the
 1305 diffusion models, which we use public pre-trained checkpoints, both of these models required explicit
 1306 training before evaluation. Here we describe the training setup and generation procedure.

1308 **VAE:** We train the unconditional VAE on MNIST following the publicly available implementation
 1309 provided in Francis (2022). The VAE consists of a simple convolutional encoder–decoder architecture
 1310 with a Gaussian latent prior. After training, we generate 1,600 paired samples under PN and RR
 1311 schemes, respectively.

1312 **Glow:** We train the class-conditional Glow model (Kingma & Dhariwal, 2018) on CIFAR-10 using
 1313 the normflows library. The architecture follows a multiscale normalizing flow design and Glow
 1314 blocks. After training, we generated 100 PN and 100 RR pairs per class.

1316 C.2 IMAGE EDITING

1318 We apply our antithetic initial noise strategy to the image editing algorithm FlowEdit (Kulikov et al.,
 1319 2024), which edits images toward a target text prompt using pre-trained flow models.

1321 In Algorithm 1 of FlowEdit, at each timestep, the algorithm samples n_{avg} random noises $Z \sim \mathcal{N}(0, 1)$
 1322 to create noisy versions of the source image, computes velocity differences between source and target
 1323 conditions, and averages these directions to drive the editing process.

1324 In the $n_{\text{avg}} = 2$ setting, we replace the two independent random samples with antithetic noises: for
 1325 each $Z \sim \mathcal{N}(0, I)$ we also use $-Z$ and average the two velocity updates.

1326 We compare on 76 prompts provided in FlowEdit’s official GitHub repository. For each prompt,
 1327 we generate 10 images using both PN and RR. All other parameters follow the repository defaults.
 1328 We evaluate performance using CLIP(semantic text–image alignment; higher is better) and LPIPS
 1329 (perceptual distance to the source; lower is better), which jointly measure text adherence and structure
 1330 preservation.

1331 As a result, PN improves the mean CLIP score, winning in 56.59% of all pairwise comparisons. It
 1332 also reduces LPIPS, winning in 81.58% of all pairwise comparisons.

1334 C.3 IMPLEMENTATION DETAIL

1336 We use pretrained models from Hugging Face’s Diffusers library (von Platen et al., 2022):
 1337 google/ddpm-church-256, google/ddpm-cifar10-32, and google/ddpm-celebahq-256 for uncondi-
 1338 tional diffusion; Stable Diffusion v1.5 for text-to-image; and the original repository from (Chung
 1339 et al., 2023a) for guided generation. Experiments were run on eight NVIDIA L40 GPUs. The most
 1340 intensive setup—Stable Diffusion—takes about five minutes to generate 100 images for a single
 1341 prompt.

1342 C.4 ADDITIONAL EXPERIMENTS ON PIXEL-WISE SIMILARITY

1344 C.4.1 DDPM

1346 **Antithetic sampling setup:** Unlike DDIM, where the sampling trajectory is deterministic once the
 1347 initial noise is fixed, DDPM adds a random Gaussian noise at every timestep. The update function in
 1348 DDPM is

$$x_{t-1} = \frac{1}{\sqrt{\alpha_t}} (x_t - \frac{1 - \alpha_t}{\sqrt{1 - \alpha_t}} \epsilon_\theta(x_t, t)) + \sigma_t z_t$$

1350 and $z_t \sim \mathcal{N}(0, 1)$ if $t > 1$, else $z = 0$. Therefore, in DDPM, antithetic sampling requires not only
 1351 negating the initial noise but also negating *every noise* z_t added at each iteration.

1353 Table 5 and Table 6 report the standard Pearson correlations and centralized correlation between
 1354 pixel values of image pairs produced using the same pre-trained models under PN and RR noise
 1355 schemes using DDPM (default 1000 steps) across CIFAR10, CelebA-HQ, and LSUN-Church. For
 1356 each dataset, we follow the same setup explained in Section 3. We calculate the pixelwise correction
 1357 with 1600 antithetic noise pairs and 1600 independent noises.

1358 Consistent with the behavior observed in DDIM, PN pairs in DDPM samples also exhibit negative
1359 correlation. Using the standard Pearson correlation, the mean correlation for PN pairs is strongly
1360 negative in all datasets—CIFAR10 (-0.73), Church (-0.45), and in the more identity-consistent
1361 CelebA (-0.18).

1362 The centralized correlation analysis further sharpens this contrast: mean PN correlations are sub-
1363 substantially lower (CIFAR10 -0.80 , CelebA -0.67 , Church -0.65). These results confirm again that
1364 PN noise pairs consistently introduce strong negative dependence, while RR pairs remain close to
1365 uncorrelated or weakly positive.

Table 5: DDPM standard Pearson correlation coefficients for PN and RR pairs

Statistic	CIFAR10		CelebA		Church	
	PN	RR	PN	RR	PN	RR
Mean	-0.73	0.04	-0.18	0.29	-0.45	0.11
Min	-0.95	-0.61	-0.80	-0.50	-0.90	-0.60
25th percentile	-0.81	-0.12	-0.31	0.16	-0.55	-0.02
75th percentile	-0.66	0.20	-0.05	0.44	-0.36	0.23
Max	-0.21	0.80	0.44	0.78	0.07	0.80

Table 6: DDPM centralized Pearson correlation coefficients for PN and RR pairs

Statistic	CIFAR10		CelebA		Church	
	PN	RR	PN	RR	PN	RR
Mean	-0.80	0.01	-0.67	-0.00	-0.65	-0.00
Min	-0.96	-0.61	-0.89	-0.60	-0.93	-0.56
25th percentile	-0.87	-0.15	-0.73	-0.15	-0.72	-0.12
75th percentile	-0.75	0.16	-0.62	0.15	-0.60	0.11
Max	-0.23	0.76	0.14	0.64	-0.32	0.72

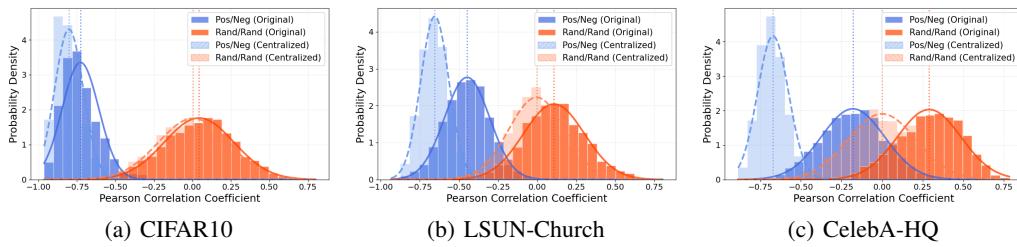


Figure 5: Pearson Correlation histograms for PN and RR pairs across three datasets using DDPM. Dashed lines indicate the mean Pearson correlation for each group.

C.4.2 DIFFUSION POSTERIOR SAMPLING (DPS)

The same pattern persists in the setting of using the diffusion model as a prior for posterior sampling, too, which has been utilized to solve various inverse problems, such as inpainting, super-resolution, and Gaussian deblurring.

1404 Since there is a ground truth image available in the image restoration task, the standard pixel-wise
 1405 correlation is calculated using the difference between reconstructed images and the corresponding
 1406 ground truth, and the centralized correlation is calculated using the same definition described in 3.
 1407 Although the overall standard correlation values are shifted up, due to the deterministic conditioning,
 1408 the posterior nature of sampling—PN pairs still shows significantly lower correlations than RR pairs
 1409 across all tasks.

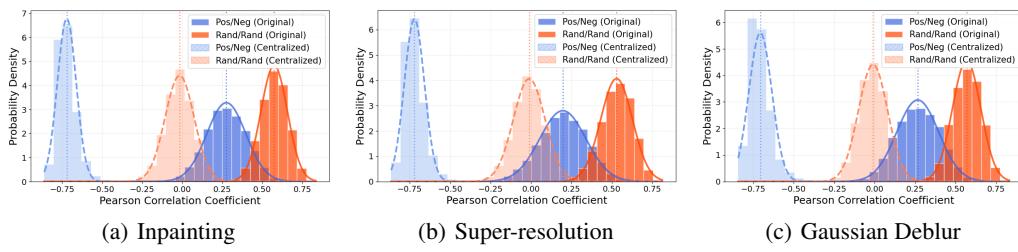
1410 For the standard Pearson correlation, the mean PN correlations range from 0.20 to 0.27, while RR
 1411 correlations consistently lie above 0.50. For the centralized correlation, PN correlations are strongly
 1412 negative across all tasks (means around -0.72). In contrast, RR pairs remain centered near zero
 1413 (mean correlations around -0.01 to -0.02).

1415 Table 7: DPS Pearson correlation coefficients for PN and RR pairs

Statistic	Inpainting		Gaussian Deblur		Super-resolution	
	PN	RR	PN	RR	PN	RR
Mean	0.28	0.57	0.27	0.57	0.20	0.53
Min	-0.14	0.05	-0.22	0.13	-0.34	0.01
25th percentile	0.19	0.52	0.17	0.51	0.10	0.47
75th percentile	0.36	0.63	0.36	0.63	0.30	0.60
Max	0.66	0.83	0.62	0.83	0.64	0.81

1425 Table 8: DPS centralized correlation coefficients for PN and RR pairs

Statistic	Inpainting		Gaussian Blur		Super-resolution	
	PN	RR	PN	RR	PN	RR
Mean	-0.72	-0.02	-0.71	-0.01	-0.72	-0.01
Min	-0.86	-0.43	-0.84	-0.36	-0.87	-0.35
25th percentile	-0.76	-0.07	-0.75	-0.07	-0.76	-0.08
75th percentile	-0.69	0.04	-0.67	0.05	-0.69	0.05
Max	-0.14	0.47	0.01	0.48	-0.14	0.43

1445 Figure 6: Pearson Correlation histograms for PN and RR pairs across three tasks in DPS. Dashed
 1446 lines indicate the mean Pearson correlation for each group.

1449 C.4.3 WASSERSTEIN DISTANCE

1450
 1451 To complement the correlation-based analyses in the main text, we also evaluate similarity using
 1452 the Wasserstein distance, a measure of distributional discrepancy. It quantifies the minimal “effort”
 1453 required to transform one probability distribution into another, which means lower Wasserstein
 1454 values indicate closer alignment between the two distributions, while higher values indicate larger
 1455 differences.1456 To calculate Wasserstein distances, we treat each generated image pair as a sample from a distribution
 1457 under the sampling scheme, PN or RR. As shown in Table 9, PN consistently exhibits larger
 1458 Wasserstein distances than RR across nearly all models, and the differences are statistically significant

1458 at all $p < 10^{-10}$. This implies that antithetic initial noises lead to more divergent distributions than
 1459 random sampling and confirms our results from the correlation analysis.
 1460

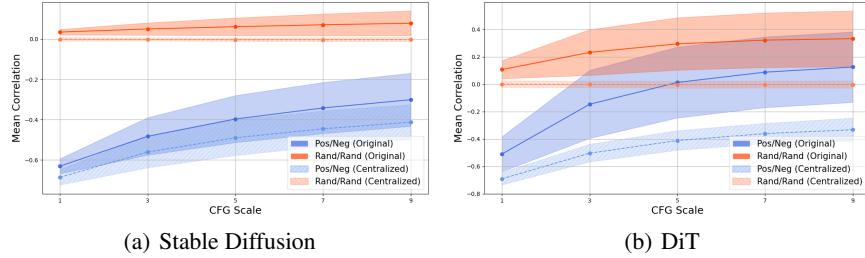
1461
 1462 Table 9: Wasserstein Distance, shown are means (SD) with corresponding t-statistics and p-values.
 1463

Model	Wasserstein Distance		t-stats (p)
	PN	RR	
LSUN-Church	0.16 (0.10)	0.12 (0.07)	$t = 12.09, p = 0$
CIFAR-10	0.19 (0.14)	0.15 (0.09)	$t = 8.30, p = 0$
CelebA-HQ	0.17 (0.11)	0.12 (0.07)	$t = 14.50, p = 0$
SD 1.5	0.10 (0.06)	0.09 (0.04)	$t = 33.11, p = 0$
DiT	0.19 (0.15)	0.14 (0.12)	$t = 14.17, p = 0$
VAE	0.07 (0.05)	0.05 (0.03)	$t = 13.24, p = 0$
Glow	0.15 (0.09)	0.12 (0.08)	$t = 7.92, p = 0$

C.4.4 INFLUENCE OF THE CFG SCALE ON CORRELATION

We extend our pixel-correlation analysis on conditional diffusion models across various Classifier-Free Guidance (CFG) scales. CFG (Ho & Salimans, 2022) is a technique that strengthens conditioning in diffusion models by interpolating between conditional and unconditional score estimates. The guidance scale controls this interpolation strength, with higher values producing samples more aligned with the conditioning signal such as prompt or class.

We conducted experiments on both Stable Diffusion 1.5 (SD1.5) and DiT using CFG scales {1, 3, 5, 7, 9}. For each setting, we generated 100 PN and RR noise pairs across 25 prompts/classes, measuring both standard and centralized correlations between the image samples.



1494 Figure 7: Average pixel correlation versus CFG scale for SD1.5 and DiT. Both standard and central-
 1495 ized correlations are shown. Shaded regions indicate standard deviation of mean correlation across
 1496 25 prompts/classes

1497 As shown in Figure 7, for all CFG scales, negated noise continues to produce strongly negatively
 1498 correlated samples. At the same time, the standard correlation for both PN and RR grows as the CFG
 1499 scale increases. This can be explained as follows: larger CFG values pull samples more strongly
 1500 toward the conditioning signal (prompt/class), effectively shrinking the space of plausible outputs.
 1501 As samples concentrate more tightly around the target distribution, they become more similar to one
 1502 another. Thus, the correlations of both PN and RR pairs increase, while PN remains much more
 1503 negative than RR.

C.4.5 PARTIAL NEGATION OF NOISE VECTORS

We conducted experiments to evaluate how localized antithetic noise influences outputs of unconditional diffusion models trained on LSUN-Church and LSUN-Bedroom. For each dataset, we generated 200 image pairs by sampling $Z_1 \sim \mathbb{N}(0, I)$ and constructing Z_2 by negating the upper half of Z_1 while leaving the lower half unchanged.

We calculate the Pearson standard correlation and the centered correlation between corresponding halves to quantify spatial correspondence induced by the antithetic manipulation. As shown in

1512
 1513
 1514
 1515
 Figure 8, the top halves exhibit sharply negative correlations, and the bottom halves are highly
 1516
 1517
 1518
 1519
 1520
 1521
 1522
 1523
 1524
 1525
 1526
 1527
 1528
 1529
 1530
 1531
 1532
 1533
 1534
 1535
 1536
 1537
 1538
 1539
 1540
 1541
 1542
 1543
 1544
 1545
 1546
 1547
 1548
 1549
 1550
 1551
 1552
 1553
 1554
 1555
 1556
 1557
 1558
 1559
 1560
 1561
 1562
 1563
 1564
 1565
 1566
 1567
 1568
 1569
 1570
 1571
 1572
 1573
 1574
 1575
 1576
 1577
 1578
 1579
 1580
 1581
 1582
 1583
 1584
 1585
 1586
 1587
 1588
 1589
 1590
 1591
 1592
 1593
 1594
 1595
 1596
 1597
 1598
 1599
 1600
 1601
 1602
 1603
 1604
 1605
 1606
 1607
 1608
 1609
 1610
 1611
 1612
 1613
 1614
 1615
 1616
 1617
 1618
 1619
 1620
 1621
 1622
 1623
 1624
 1625
 1626
 1627
 1628
 1629
 1630
 1631
 1632
 1633
 1634
 1635
 1636
 1637
 1638
 1639
 1640
 1641
 1642
 1643
 1644
 1645
 1646
 1647
 1648
 1649
 1650
 1651
 1652
 1653
 1654
 1655
 1656
 1657
 1658
 1659
 1660
 1661
 1662
 1663
 1664
 1665
 1666
 1667
 1668
 1669
 1670
 1671
 1672
 1673
 1674
 1675
 1676
 1677
 1678
 1679
 1680
 1681
 1682
 1683
 1684
 1685
 1686
 1687
 1688
 1689
 1690
 1691
 1692
 1693
 1694
 1695
 1696
 1697
 1698
 1699
 1700
 1701
 1702
 1703
 1704
 1705
 1706
 1707
 1708
 1709
 1710
 1711
 1712
 1713
 1714
 1715
 1716
 1717
 1718
 1719
 1720
 1721
 1722
 1723
 1724
 1725
 1726
 1727
 1728
 1729
 1730
 1731
 1732
 1733
 1734
 1735
 1736
 1737
 1738
 1739
 1740
 1741
 1742
 1743
 1744
 1745
 1746
 1747
 1748
 1749
 1750
 1751
 1752
 1753
 1754
 1755
 1756
 1757
 1758
 1759
 1760
 1761
 1762
 1763
 1764
 1765
 1766
 1767
 1768
 1769
 1770
 1771
 1772
 1773
 1774
 1775
 1776
 1777
 1778
 1779
 1780
 1781
 1782
 1783
 1784
 1785
 1786
 1787
 1788
 1789
 1790
 1791
 1792
 1793
 1794
 1795
 1796
 1797
 1798
 1799
 1800
 1801
 1802
 1803
 1804
 1805
 1806
 1807
 1808
 1809
 1810
 1811
 1812
 1813
 1814
 1815
 1816
 1817
 1818
 1819
 1820
 1821
 1822
 1823
 1824
 1825
 1826
 1827
 1828
 1829
 1830
 1831
 1832
 1833
 1834
 1835
 1836
 1837
 1838
 1839
 1840
 1841
 1842
 1843
 1844
 1845
 1846
 1847
 1848
 1849
 1850
 1851
 1852
 1853
 1854
 1855
 1856
 1857
 1858
 1859
 1860
 1861
 1862
 1863
 1864
 1865
 1866
 1867
 1868
 1869
 1870
 1871
 1872
 1873
 1874
 1875
 1876
 1877
 1878
 1879
 1880
 1881
 1882
 1883
 1884
 1885
 1886
 1887
 1888
 1889
 1890
 1891
 1892
 1893
 1894
 1895
 1896
 1897
 1898
 1899
 1900
 1901
 1902
 1903
 1904
 1905
 1906
 1907
 1908
 1909
 1910
 1911
 1912
 1913
 1914
 1915
 1916
 1917
 1918
 1919
 1920
 1921
 1922
 1923
 1924
 1925
 1926
 1927
 1928
 1929
 1930
 1931
 1932
 1933
 1934
 1935
 1936
 1937
 1938
 1939
 1940
 1941
 1942
 1943
 1944
 1945
 1946
 1947
 1948
 1949
 1950
 1951
 1952
 1953
 1954
 1955
 1956
 1957
 1958
 1959
 1960
 1961
 1962
 1963
 1964
 1965
 1966
 1967
 1968
 1969
 1970
 1971
 1972
 1973
 1974
 1975
 1976
 1977
 1978
 1979
 1980
 1981
 1982
 1983
 1984
 1985
 1986
 1987
 1988
 1989
 1990
 1991
 1992
 1993
 1994
 1995
 1996
 1997
 1998
 1999
 2000
 2001
 2002
 2003
 2004
 2005
 2006
 2007
 2008
 2009
 2010
 2011
 2012
 2013
 2014
 2015
 2016
 2017
 2018
 2019
 2020
 2021
 2022
 2023
 2024
 2025
 2026
 2027
 2028
 2029
 2030
 2031
 2032
 2033
 2034
 2035
 2036
 2037
 2038
 2039
 2040
 2041
 2042
 2043
 2044
 2045
 2046
 2047
 2048
 2049
 2050
 2051
 2052
 2053
 2054
 2055
 2056
 2057
 2058
 2059
 2060
 2061
 2062
 2063
 2064
 2065
 2066
 2067
 2068
 2069
 2070
 2071
 2072
 2073
 2074
 2075
 2076
 2077
 2078
 2079
 2080
 2081
 2082
 2083
 2084
 2085
 2086
 2087
 2088
 2089
 2090
 2091
 2092
 2093
 2094
 2095
 2096
 2097
 2098
 2099
 2100
 2101
 2102
 2103
 2104
 2105
 2106
 2107
 2108
 2109
 2110
 2111
 2112
 2113
 2114
 2115
 2116
 2117
 2118
 2119
 2120
 2121
 2122
 2123
 2124
 2125
 2126
 2127
 2128
 2129
 2130
 2131
 2132
 2133
 2134
 2135
 2136
 2137
 2138
 2139
 2140
 2141
 2142
 2143
 2144
 2145
 2146
 2147
 2148
 2149
 2150
 2151
 2152
 2153
 2154
 2155
 2156
 2157
 2158
 2159
 2160
 2161
 2162
 2163
 2164
 2165
 2166
 2167
 2168
 2169
 2170
 2171
 2172
 2173
 2174
 2175
 2176
 2177
 2178
 2179
 2180
 2181
 2182
 2183
 2184
 2185
 2186
 2187
 2188
 2189
 2190
 2191
 2192
 2193
 2194
 2195
 2196
 2197
 2198
 2199
 2200
 2201
 2202
 2203
 2204
 2205
 2206
 2207
 2208
 2209
 2210
 2211
 2212
 2213
 2214
 2215
 2216
 2217
 2218
 2219
 2220
 2221
 2222
 2223
 2224
 2225
 2226
 2227
 2228
 2229
 2230
 2231
 2232
 2233
 2234
 2235
 2236
 2237
 2238
 2239
 2240
 2241
 2242
 2243
 2244
 2245
 2246
 2247
 2248
 2249
 2250
 2251
 2252
 2253
 2254
 2255
 2256
 2257
 2258
 2259
 2260
 2261
 2262
 2263
 2264
 2265
 2266
 2267
 2268
 2269
 2270
 2271
 2272
 2273
 2274
 2275
 2276
 2277
 2278
 2279
 2280
 2281
 2282
 2283
 2284
 2285
 2286
 2287
 2288
 2289
 2290
 2291
 2292
 2293
 2294
 2295
 2296
 2297
 2298
 2299
 2300
 2301
 2302
 2303
 2304
 2305
 2306
 2307
 2308
 2309
 2310
 2311
 2312
 2313
 2314
 2315
 2316
 2317
 2318
 2319
 2320
 2321
 2322
 2323
 2324
 2325
 2326
 2327
 2328
 2329
 2330
 2331
 2332
 2333
 2334
 2335
 2336
 2337
 2338
 2339
 2340
 2341
 2342
 2343
 2344
 2345
 2346
 2347
 2348
 2349
 2350
 2351
 2352
 2353
 2354
 2355
 2356
 2357
 2358
 2359
 2360
 2361
 2362
 2363
 2364
 2365
 2366
 2367
 2368
 2369
 2370
 2371
 2372
 2373
 2374
 2375
 2376
 2377
 2378
 2379
 2380
 2381
 2382
 2383
 2384
 2385
 2386
 2387
 2388
 2389
 2390
 2391
 2392
 2393
 2394
 2395
 2396
 2397
 2398
 2399
 2400
 2401
 2402
 2403
 2404
 2405
 2406
 2407
 2408
 2409
 2410
 2411
 2412
 2413
 2414
 2415
 2416
 2417
 2418
 2419
 2420
 2421
 2422
 2423
 2424
 2425
 2426
 2427
 2428
 2429
 2430
 2431
 2432
 2433
 2434
 2435
 2436
 2437
 2438
 2439
 2440
 2441
 2442
 2443
 2444
 2445
 2446
 2447
 2448
 2449
 2450
 2451
 2452
 2453
 2454
 2455
 2456
 2457
 2458
 2459
 2460
 2461
 2462
 2463
 2464
 2465
 2466
 2467
 2468
 2469
 2470
 2471
 2472
 2473
 2474
 2475
 2476
 2477
 2478
 2479
 2480
 2481
 2482
 2483
 2484
 2485
 2486
 2487
 2488
 2489
 2490
 2491
 2492
 2493
 2494
 2495
 2496
 2497
 2498
 2499
 2500
 2501
 2502
 2503
 2504
 2505
 2506
 2507
 2508
 2509
 2510
 2511
 2512
 2513
 2514
 2515
 2516
 2517
 2518
 2519
 2520
 2521
 2522
 2523
 2524
 2525
 2526
 2527
 2528
 2529
 2530
 2531
 2532
 2533
 2534
 2535
 2536
 2537
 2538
 2539
 2540
 2541
 2542
 2543
 2544
 2545
 2546
 2547
 2548
 2549
 2550
 2551
 2552
 2553
 2554
 2555
 2556
 2557
 2558
 2559
 2560
 2561
 2562
 2563
 2564
 2565
 2566
 2567
 2568
 2569
 2570
 2571
 2572
 2573
 2574
 2575
 2576
 2577
 2578
 2579
 2580
 2581
 2582
 2583
 2584
 2585
 2586
 2587
 2588
 2589
 2590
 2591
 2592
 2593
 2594
 2595
 2596
 2597
 2598
 2599
 2600
 2601
 2602
 2603
 2604
 2605
 2606
 2607
 2608
 2609
 2610
 2611
 2612
 2613
 2614
 2615
 2616
 2617
 2618
 2619
 2620
 2621
 2622
 2623
 2624
 2625
 2626
 2627
 2628
 2629
 2630
 2631
 2632
 2633
 2634
 2635
 2636
 2637
 2638
 2639
 2640
 2641
 2642
 2643
 2644
 2645
 2646
 2647
 2648
 2649
 2650
 2651
 2652
 2653
 2654
 2655
 2656
 2657
 2658
 2659
 2660
 2661
 2662
 2663
 2664
 2665
 2666
 2667
 2668
 2669
 2670
 2671
 2672
 2673
 2674
 2675
 2676
 2677
 2678
 2679
 2680
 2681
 2682
 2683
 2684
 2685
 2686
 2687
 2688
 2689
 2690
 2691
 2692
 2693
 2694
 2695
 2696
 2697
 2698
 2699
 2700
 2701
 2702
 2703
 2704
 2705
 2706
 2707
 2708
 2709
 2710
 2711
 2712
 2713
 2714
 2715
 2716
 2717
 2718
 2719
 2720
 2721
 2722
 2723
 2724
 2725
 2726
 2727
 2728
 2729
 2730
 2731
 2732
 2733
 2734
 2735
 2736
 2737
 2738
 2739
 2740
 2741
 2742
 2743
 2744
 2745
 2746
 2747
 2748
 2749
 2750
 2751
 2752
 2753
 2754
 2755
 2756
 2757
 2758
 2759
 2760
 2761
 2762
 2763
 2764
 2765
 2766
 2767
 2768
 2769
 2770
 2771
 2772
 2773
 2774
 2775
 2776
 2777
 2778
 2779
 2780
 2781
 2782
 2783
 2784
 2785
 2786
 2787
 2788
 2789
 2790
 2791
 2792
 2793
 2794
 2795
 2796
 2797
 2798
 2799
 2800
 2801
 2802
 2803
 2804
 2805
 2806
 2807
 2808
 2809
 2810
 2811
 2812
 2813
 2814
 2815
 2816
 2817
 2818
 2819
 2820
 2821
 2822
 2823
 2824
 2825
 2826
 2827
 2828
 2829
 2830
 2831
 2832
 2833
 2834
 2835
 2836
 2837
 2838
 2839
 2840
 2841
 2842
 2843
 2844
 2845
 2846
 2847
 2848
 2849
 2850
 2851
 2852
 2853
 2854
 2855
 2856
 2857
 2858
 2859
 2860
 2861
 2862
 2863
 2864
 2865
 2866
 2867
 2868
 2869
 2870
 2871
 2872
 2873
 2874
 2875
 2876
 2877
 2878
 2879
 2880
 2881
 2882
 2883
 2884
 2885
 2886
 2887
 2888
 2889
 2890
 2891
 2892
 2893
 2894
 2895
 2896
 2897
 2898
 2899
 2900
 2901
 2902
 2903
 2904
 2905
 2906
 2907
 2908
 2909
 2910
 2911
 2912
 2913
 2914
 2915
 2916
 2917
 2918
 2919
 2920
 2921
 2922
 2923
 2924

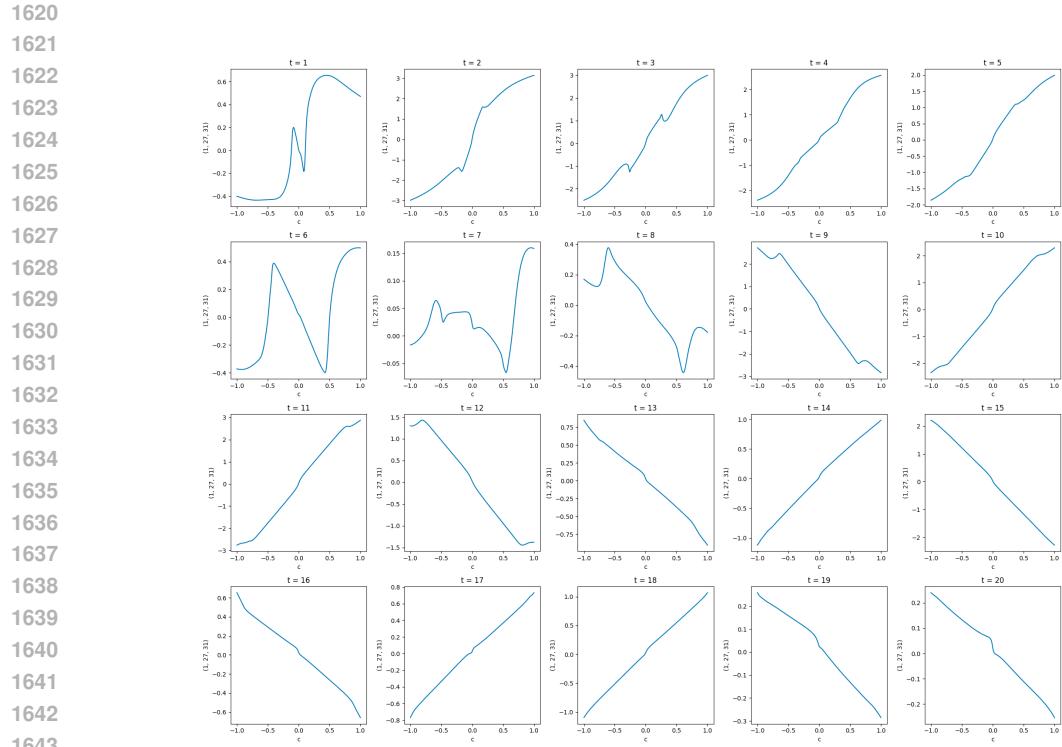


Figure 11: CIFAR10: Coordinate (1, 27, 31)

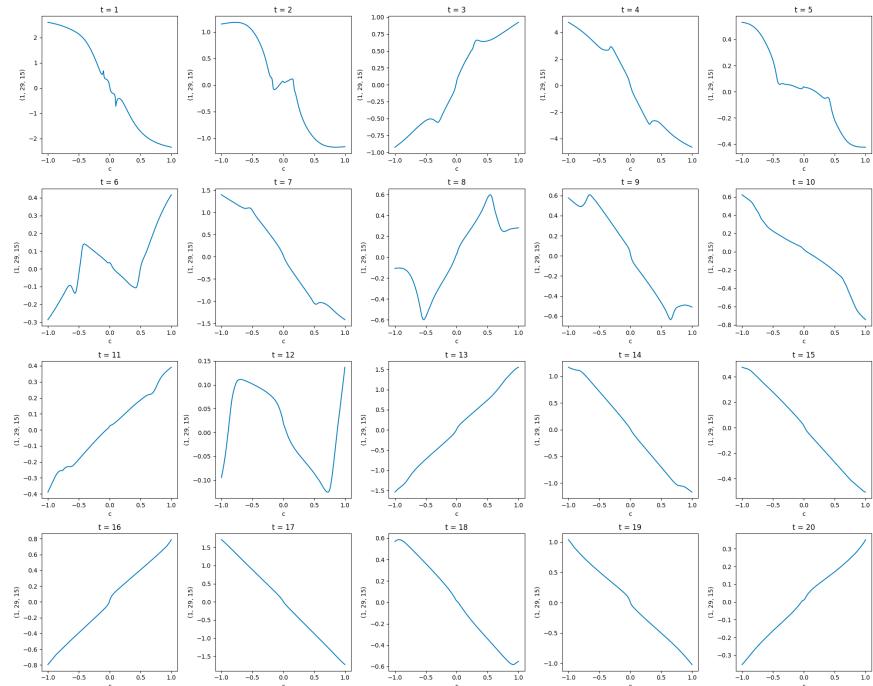


Figure 12: CIFAR10: Coordinate (1, 29, 15)

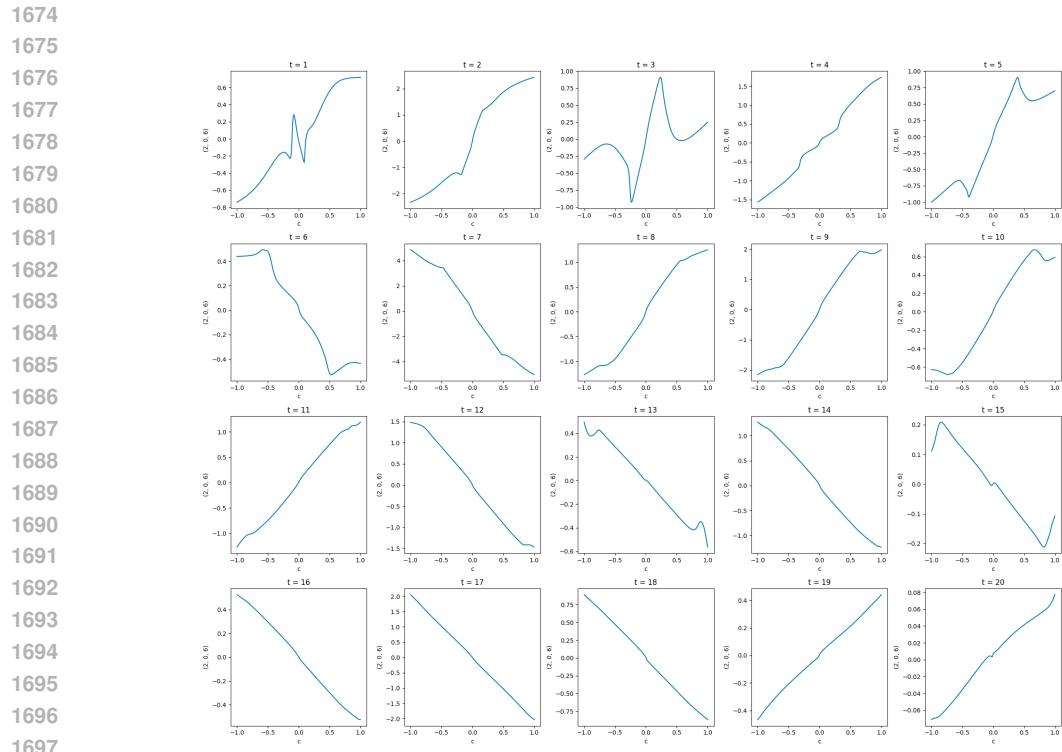


Figure 13: CIFAR10: Coordinate (2, 0, 6)

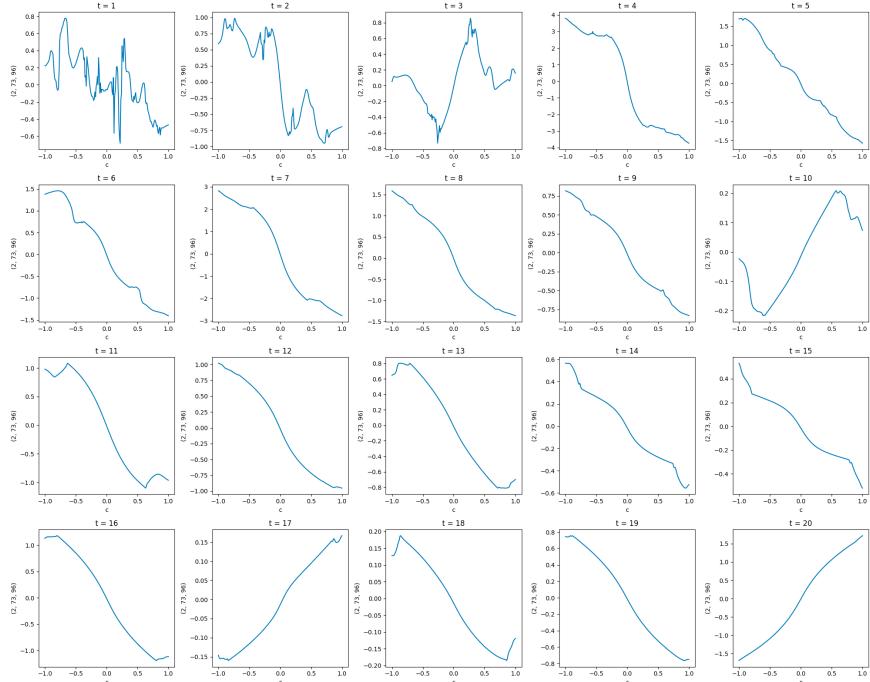
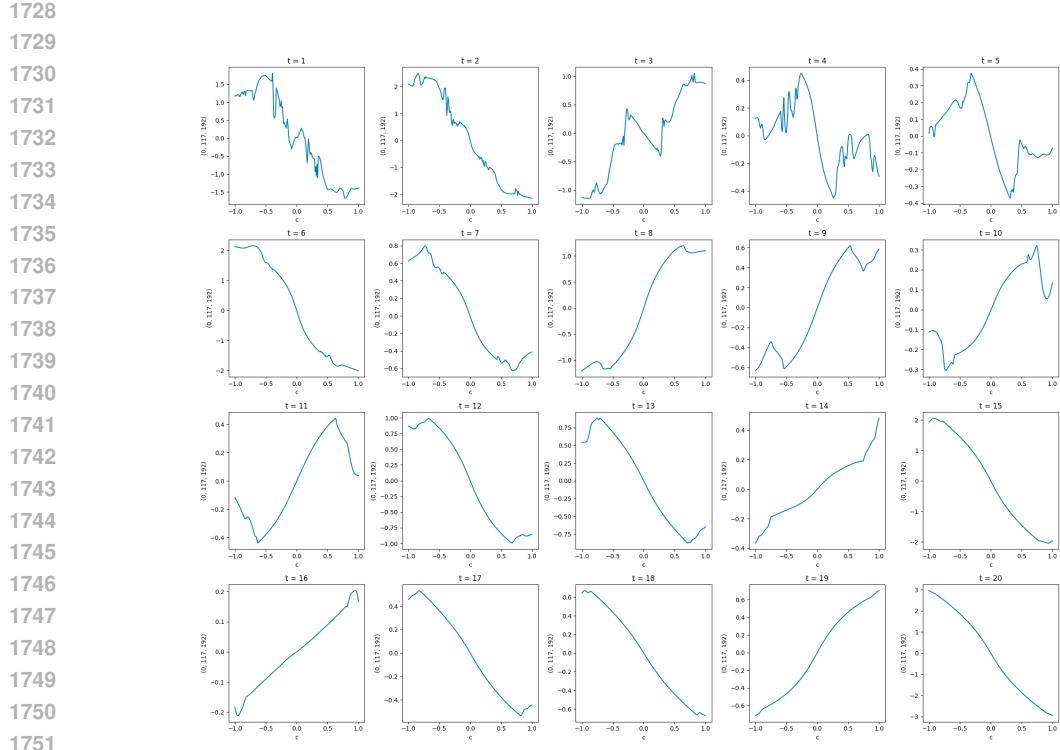
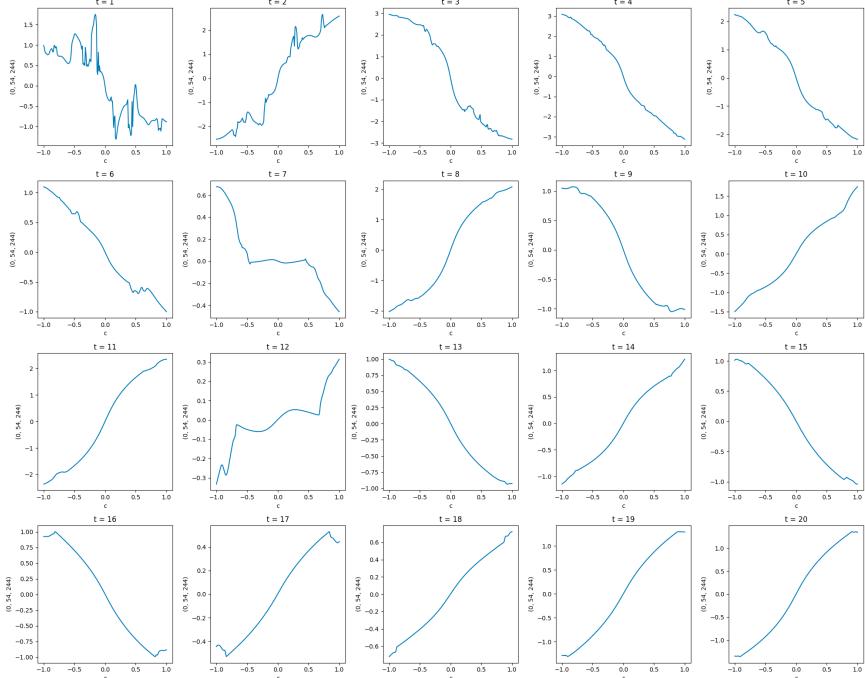


Figure 14: Church: Coordinate (2, 73, 96)

Figure 15: CIFAR10: Coordinate $(0, 117, 192)$ Figure 16: Church: Coordinate $(0, 54, 244)$

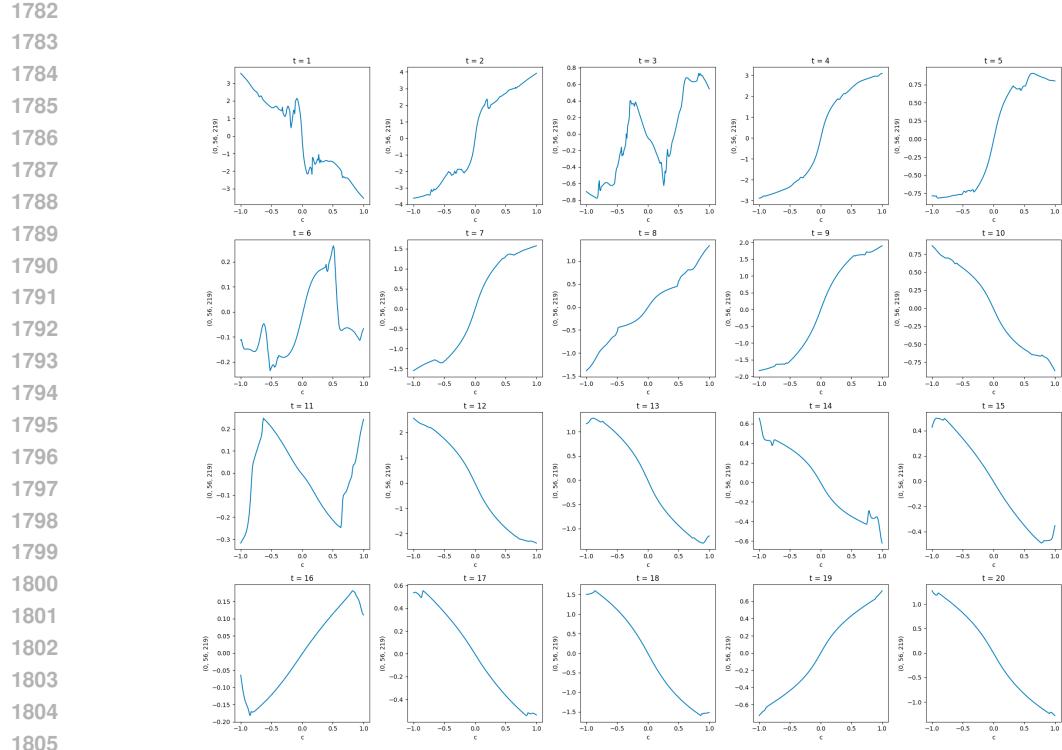
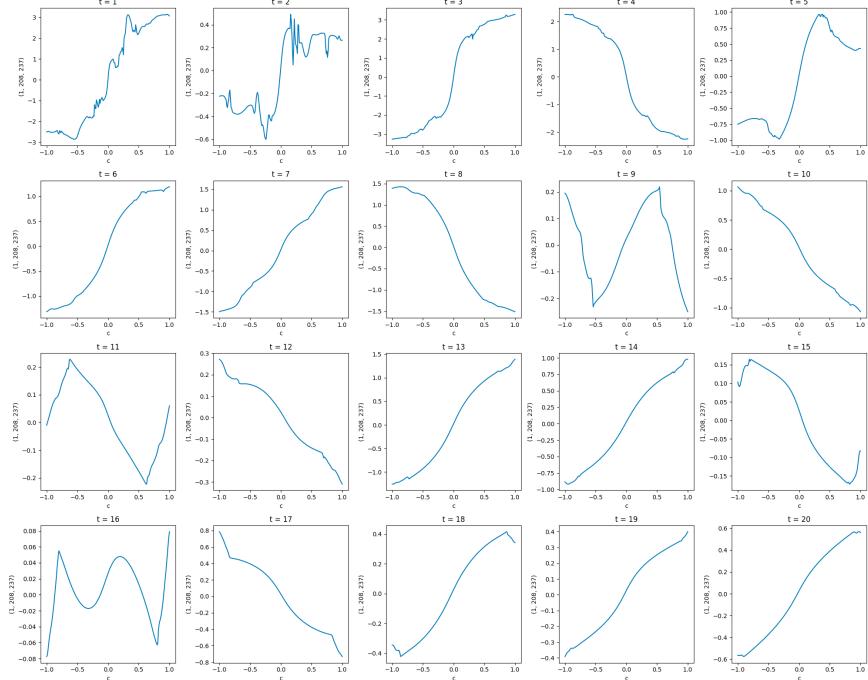
Figure 17: CIFAR10: Coordinate $(0, 56, 219)$ Figure 18: Church: Coordinate $(1, 208, 237)$

Table 10: Affine antisymmetry score

Dataset	CIFAR10	Church
Mean	0.9932	0.9909
Min	0.7733	0.7710
1% quantile	0.9104	0.8673
2% quantile	0.9444	0.8768
5% quantile	0.9690	0.9624
10% quantile	0.9865	0.9750
Median	0.9992	0.9995

1848 C.6 ADDITIONAL EXPERIMENTS FOR UNCERTAINTY QUANTIFICATION

1849 C.6.1 UNCERTAINTY QUANTIFICATION

1850 The image metrics used in uncertainty quantification are used to capture different aspects of pixel
 1851 intensity, color distribution, and perceived brightness.

1852 **Mean pixel value** is defined as the average of all pixel intensities across the image (including all
 1853 channels). Formally, for an image $I \in \mathbb{R}^{C \times H \times W}$, the mean is computed as

$$1854 \mu = \frac{1}{HWC} \sum_{i=1}^C \sum_{j=1}^H \sum_{k=1}^W I_{i,j,k}.$$

1855 **Brightness** is calculated using the standard CIE formula: $0.299 \cdot R + 0.587 \cdot G + 0.114 \cdot B$ to produce
 1856 a grayscale value at *each pixel*, where R, G, B is the red, green, and blue color value of the given
 1857 pixel. It is widely used in video and image compression, and approximates human visual sensitivity
 1858 to color. The brightness of an image is then the average of the grayscale value across all pixels.

1859 **Contrast** is computed as the difference in average pixel intensity between the top and bottom halves of
 1860 an image. Let $x \in [0, 1]^{C \times H \times W}$ be the normalized image, we define contrast as $100 \cdot (\mu_{\text{top}} - \mu_{\text{bottom}})$,
 1861 where μ_{top} and μ_{bottom} are the average intensities over the top and bottom halves, respectively.

1862 **Centroid** measures the coordinate of the brightness-weighted center of mass of the image. For scalar
 1863 comparison, we focus on the vertical component of the centroid to assess spatial uncertainty. After
 1864 converting to grayscale $M \in \mathbb{R}^{H \times W}$ by averaging across channels, we treat the image as a 2D
 1865 intensity distribution and compute the vertical centroid as

$$1866 \frac{1}{\sum_{i=1}^H \sum_{j=1}^W M_{i,j}} \sum_{i=1}^H \sum_{j=1}^W i \cdot M_{i,j},$$

1867 where i denotes the row index.

1868 C.6.2 QMC EXPERIMENTS

1869 **RQMC confidence interval construction:** For the RQMC experiments, we consider R inde-
 1870 pendent randomization of a Sobol' point set of size n , with a fixed budget of $N = Rn$ function
 1871 evaluations. Denote the RQMC point set in the r -th replicate as $\{\mathbf{u}_{r,k}\}_{1 \leq k \leq n} \subset [0, 1]^d$, where
 1872 $\mathbf{u}_{r,k} \sim \text{Unif}([0, 1]^d)$. Applying the Gaussian inverse cumulative distribution function Φ^{-1} to each
 1873 coordinate of $\mathbf{u}_{r,k}$ transforms the uniform samples to standard normal samples. Consequently, the
 1874 estimate in each replicate is given by

$$1875 \hat{\mu}_r^{\text{QMC}} = \frac{1}{n} \sum_{k=1}^n S(\text{DM}(\Phi^{-1}(\mathbf{u}_{r,k}))), \quad r = 1, \dots, R.$$

1890 The overall point estimate is their average
 1891

$$\hat{\mu}_N^{\text{QMC}} = \frac{1}{R} \sum_{r=1}^R \hat{\mu}_r^{\text{QMC}}.$$

1892 Let $(\hat{\sigma}_R^{\text{QMC}})^2$ be the sample variance of $\hat{\mu}_1^{\text{QMC}}, \dots, \hat{\mu}_R^{\text{QMC}}$. The Student t confidence interval is given
 1893 by
 1894

$$\text{CI}_N^{\text{QMC}}(1 - \alpha) = \hat{\mu}_N^{\text{QMC}} \pm t_{R-1, 1-\alpha/2} \frac{\hat{\sigma}_R^{\text{QMC}}}{\sqrt{R}},$$

1895 where $t_{R-1, 1-\alpha/2}$ is the $(1 - \alpha/2)$ -quantile of the t -distribution with $R - 1$ degrees of freedom. If
 1896 the estimates $\hat{\mu}_r^{\text{QMC}}$ are normally distributed, this confidence interval has exact coverage probability
 1897 $1 - \alpha$. In general, the validity of Student t confidence interval is justified by CLT. An extensive
 1898 numerical study by L'Ecuyer et al. (2023) demonstrates that Student t intervals achieve the desired
 1899 coverage empirically.

1900 **Exploring different configurations of R and n** This experiment aims to understand how different
 1901 configurations of R , the number of replicates, and n , the size of the QMC point set, affect the CI
 1902 length of the RQMC method. The total budget of function evaluations is fixed at $Rn = 3200$, to be
 1903 consistent with the AMC and MC experiments. We consider the four image metrics used in Section 4
 1904 and one additional image metric, MUSIQ.

1905 For RQMC, each configuration was repeated five times, and the results were averaged to ensure
 1906 stability. AMC and MC each consist of a single run over 3200 images. All experiments are conducted
 1907 using the CIFAR10 dataset.

1908 The results are shown in Table 11 and underlined values indicate the best CI length among the three
 1909 QMC configurations, while bold values indicate the best CI length across all methods, including MC
 1910 and AMC. For the brightness and pixel mean metrics, the configuration with point set size $n = 64$ and
 1911 number of replicates $R = 50$ reduces CI length the most. In contrast and centroid, the configuration
 1912 with larger n has a better CI length. For MUSIQ, a more complex metric, changes in CI lengths
 1913 across configurations are subtle, and the configuration with the largest R has the shortest CI length.
 1914 While no single configuration consistently advantages, all RQMC and antithetic sampling methods
 1915 outperform plain MC.

1916
 1917
 1918
 1919
 1920
 1921
 1922 Table 11: Average 95% CI length and relative efficiencies (vs MC baseline). The first three rows are
 1923 for RQMC methods, with the configuration of $R \times n$ indicated in the first column.

$R \times n$	Brightness		Mean		Contrast		Centroid		MUSIQ	
	CI	Eff.	CI	Eff.	CI	Eff.	CI	Eff.	CI	Eff.
25 \times 128	0.37	29.81	0.40	25.71	0.22	24.57	0.04	8.17	0.13	1.11
50 \times 64	<u>0.35</u>	<u>32.05</u>	<u>0.39</u>	<u>26.94</u>	0.22	23.43	0.04	7.35	0.13	1.13
200 \times 16	0.47	18.38	0.49	17.30	0.29	14.20	0.05	5.84	<u>0.12</u>	<u>1.21</u>
AMC	0.35	32.66	0.39	27.12	0.23	22.05	0.04	6.96	0.13	1.06
MC	2.00	-	2.04	-	1.08	-	0.11	-	0.13	-

1924 C.6.3 DPS EXPERIMENT IMPLEMENTATION

1925 We evaluate the confidence interval length reduction benefits of antithetic initial noise across three
 1926 common image inverse problems: super-resolution, Gaussian deblurring, and inpainting. For each
 1927 task, the forward measurement operator is applied to the true image, and noisy observations are
 1928 generated using Gaussian noise with $\sigma = 0.05$. We use the official implementation of DPS (Chung
 1929 et al., 2023a) with the following parameters:

1930
 1931
 1932
 1933
 1934
 1935
 1936
 1937
 1938
 1939
 1940
 1941
 1942
 1943

- **Super-resolution** uses the *super_resolution* operator with an input shape of $(1, 3, 256, 256)$ and an
 1943 upsampling scale factor of 2. This models the standard bicubic downsampling scenario followed
 by Gaussian noise corruption.
- **Gaussian deblurring** employs the *gaussian_blur* operator, again with a kernel size of 61 but with
 1944 intensity set to 1, which is intended to test the variance reduction in a simpler inverse scenario.

1944
1945 • **Inpainting** is set up using the *inpainting* operator with a random binary mask applied to the input
1946 image. The missing pixel ratio is drawn from a uniform range between 30% and 70%, and the
1947 image size is fixed at 256×256 . A higher guidance scale (0.5) is used to compensate for the
1948 sparsity of observed pixels.

1949 As shown in Table 3, AMC consistently achieves shorter CI lengths than MC across all tasks and
1950 metrics without sacrificing reconstruction quality, implied by the L1 and PSNR metrics to measure
1951 the difference between reconstructed images and ground truth images.

1952 **C.6.4 DDS DATASET**

1953 Data used in the DDS experiment on uncertainty quantification (Section 4) were obtained from the
1954 NYU fastMRI Initiative database (Knoll et al., 2020; Zbontar et al., 2018). A listing of NYU fastMRI
1955 investigators, subject to updates, can be found at: fastmri.med.nyu.edu. The primary goal of
1956 fastMRI is to test whether machine learning can aid in the reconstruction of medical images.

1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997

1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008

D MORE VISUALIZATIONS

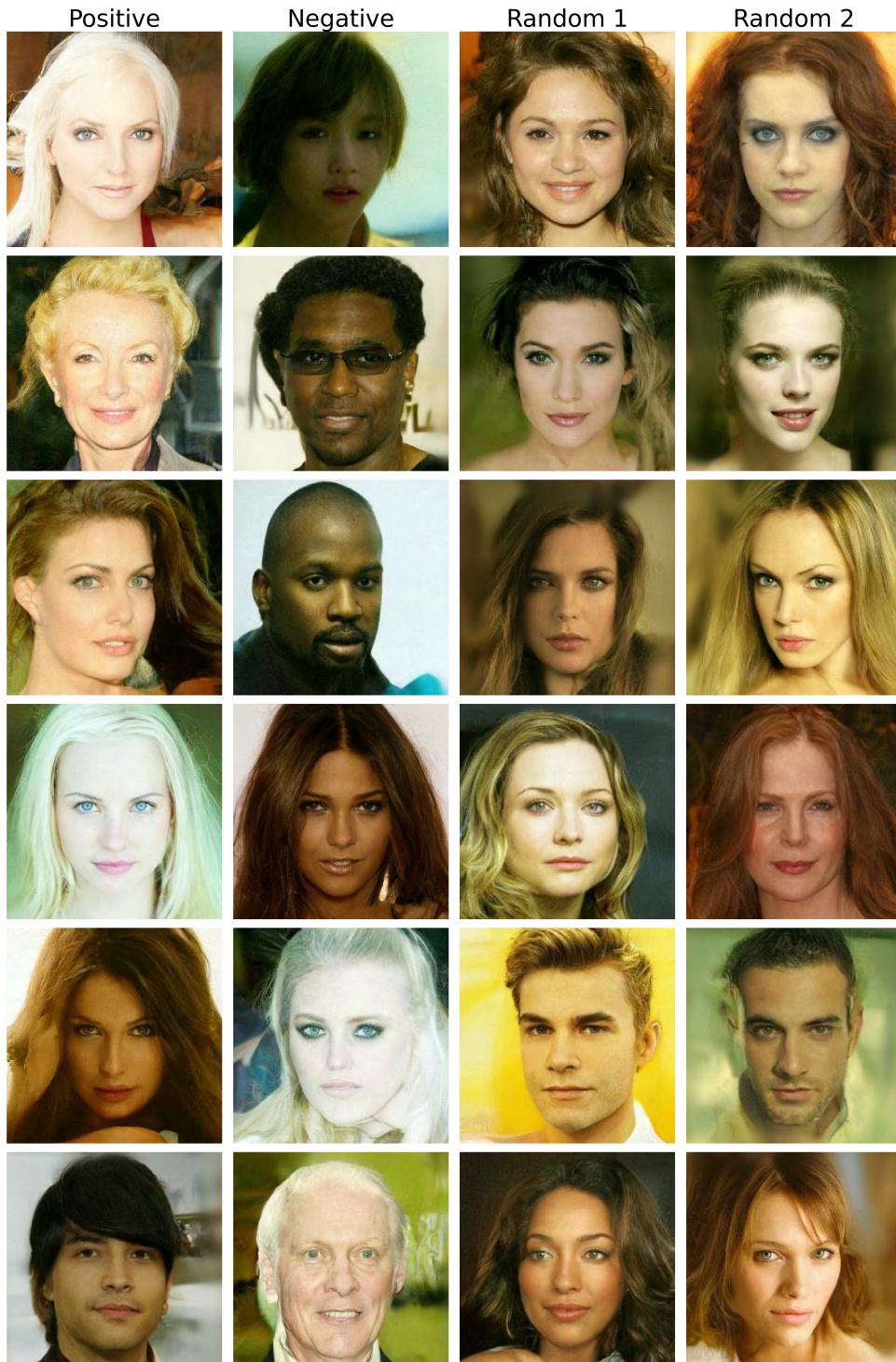
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051

Figure 19: CelebA-HQ Image Generated

2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105

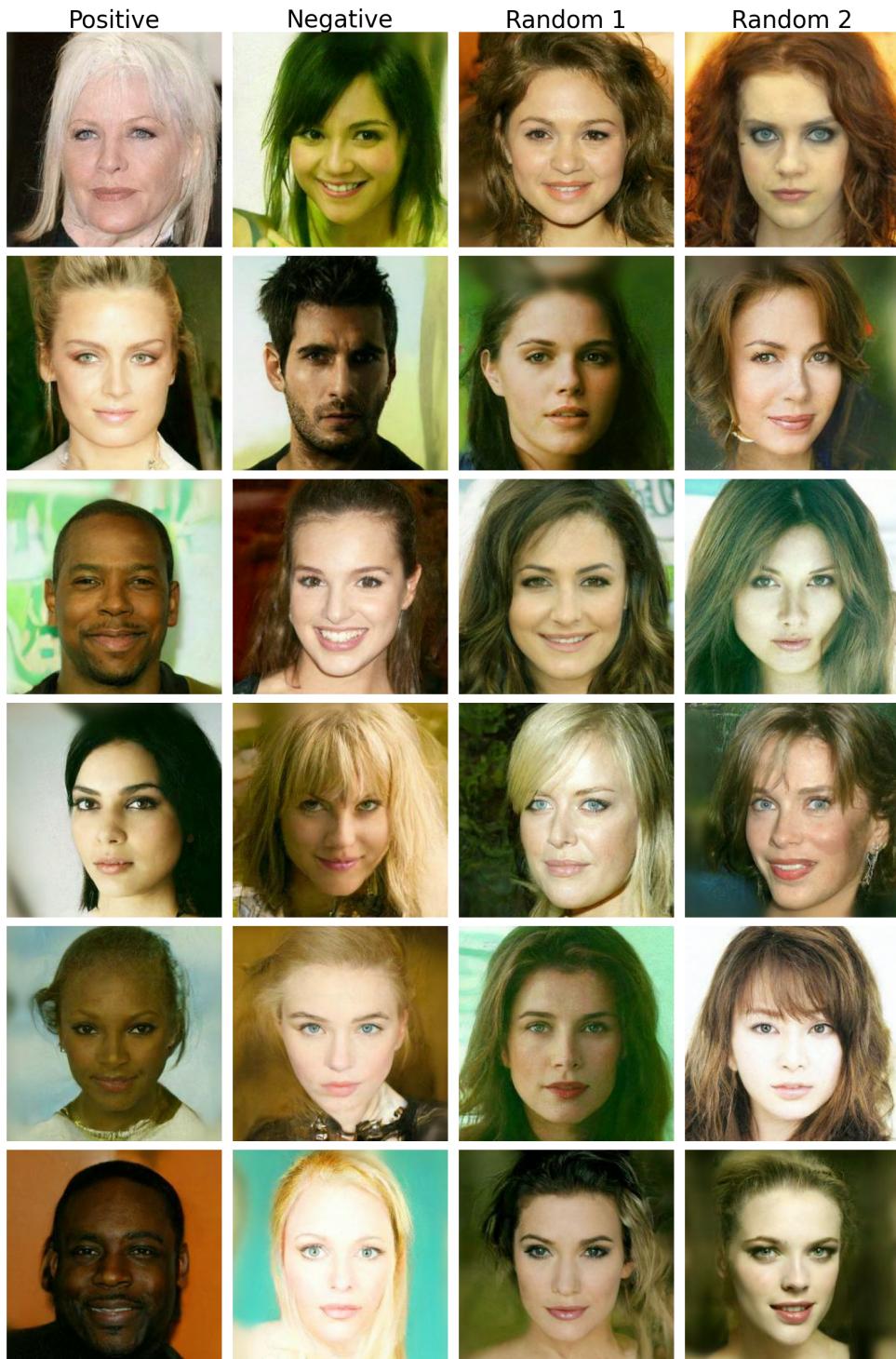


Figure 20: CelebA-HQ Image Generated

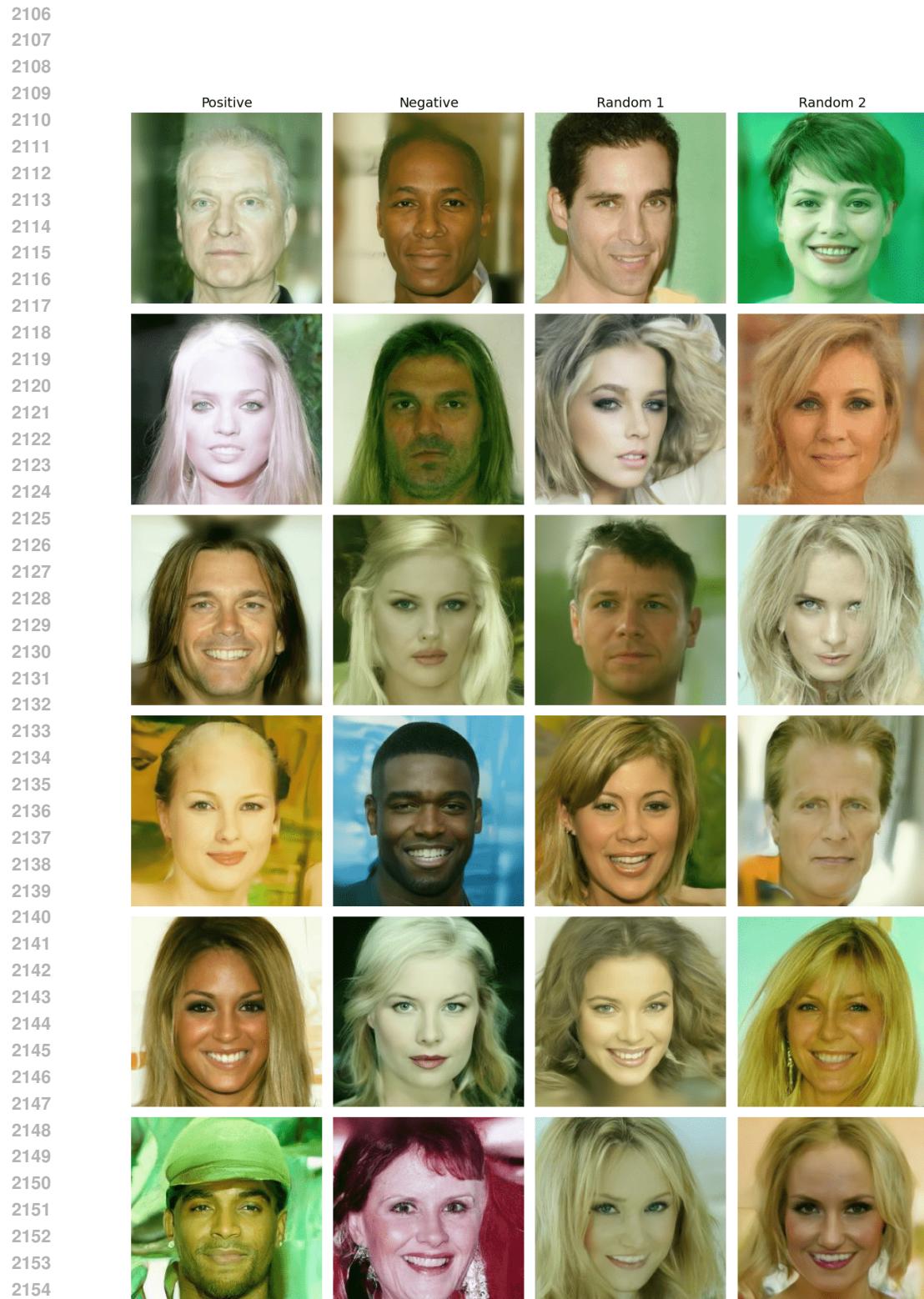


Figure 21: CelebA-HQ Image Generated

2160

2161

2162

2163

2164

2165

2166

2167

2168

2169

2170

2171

2172

2173

2174

2175

2176

2177

2178

2179

2180

2181

2182

2183

2184

2185

2186

2187

2188

2189

2190

2191

2192

2193

2194

2195

2196

2197

2198

2199

2200

2201

2202

2203

2204

2205

2206

2207

2208

2209

2210

2211

2212

2213



Figure 22: DiT Class 974: geyser

2214

2215

2216

2217

2218

2219

2220

2221

2222

2223

2224

2225

2226

2227

2228

2229

2230

2231

2232

2233

2234

2235

2236

2237

2238

2239

2240

2241

2242

2243

2244

2245

2246

2247

2248

2249

2250

2251

2252

2253

2254

2255

2256

2257

2258

2259

2260

2261

2262

2263

2264

2265

2266

2267

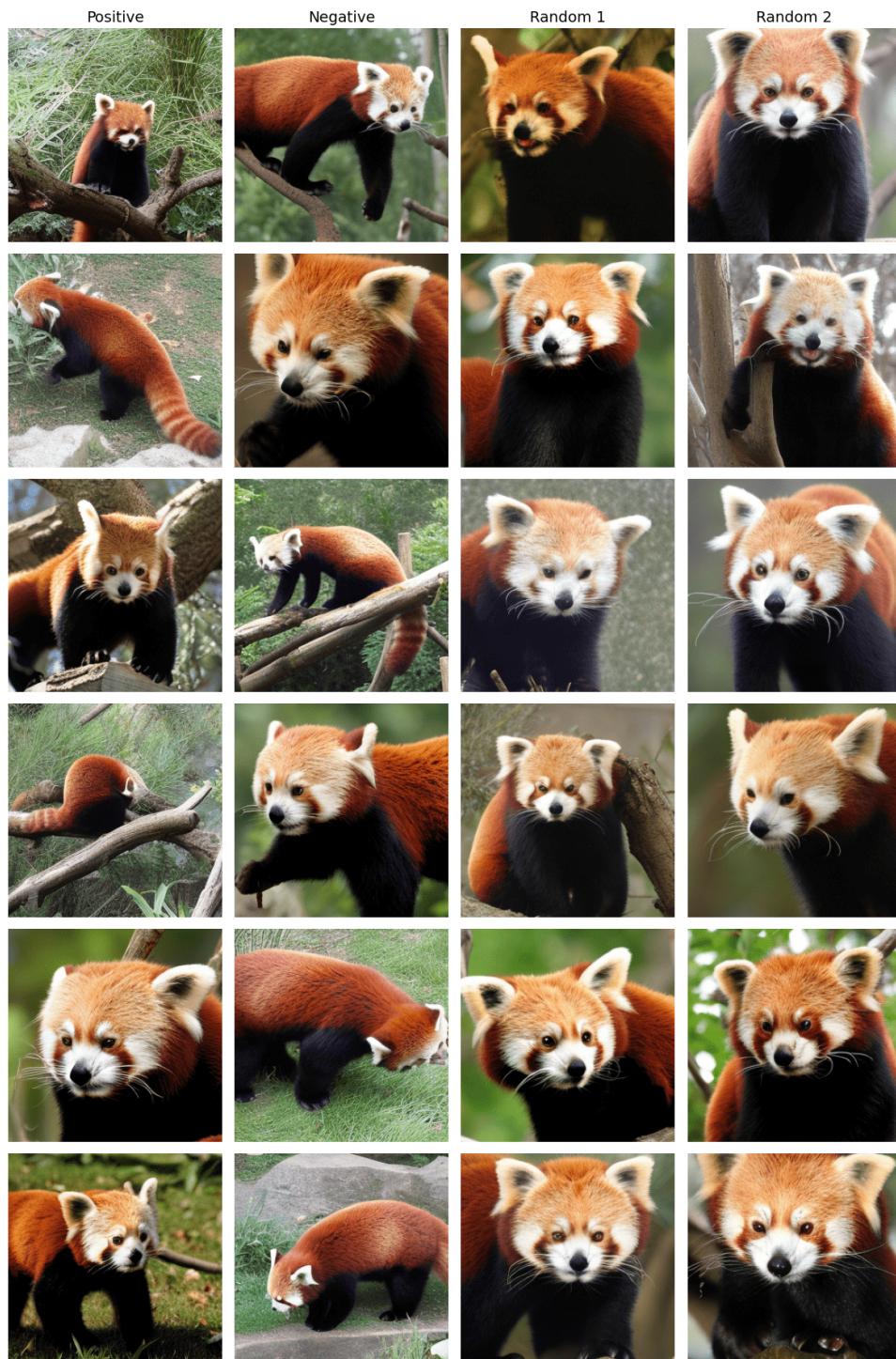


Figure 23: DiT Class 387: lesser panda, red panda

2268

2269

2270

2271

2272

2273

2274

2275

2276

2277

2278

2279

2280

2281

2282

2283

2284

2285

2286

2287

2288

2289

2290

2291

2292

2293

2294

2295

2296

2297

2298

2299

2300

2301

2302

2303

2304

2305

2306

2307

2308

2309

2310

2311

2312

2313

2314

2315

2316

2317

2318

2319

2320

2321

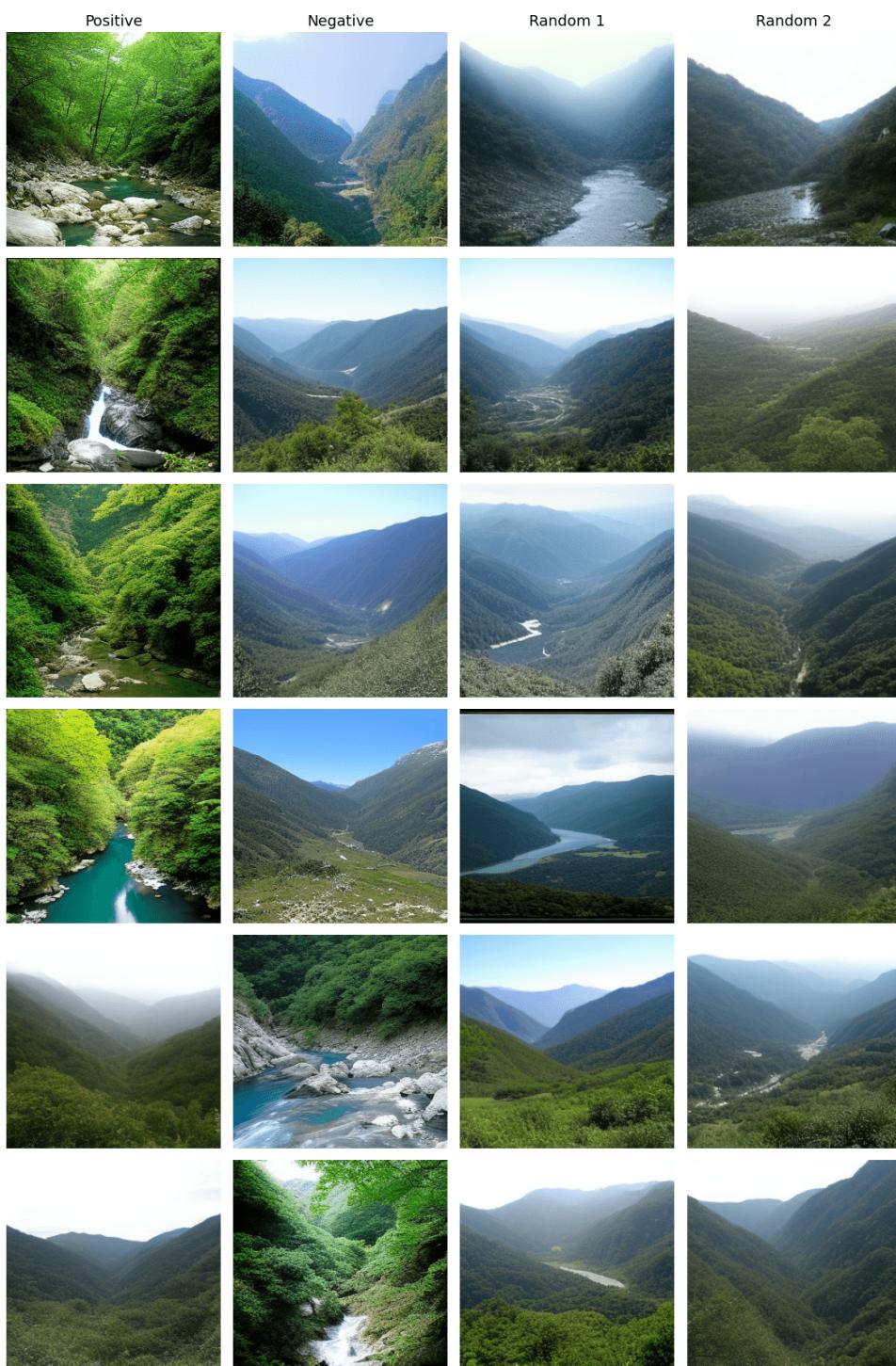


Figure 24: DiT Clas 979: valley, vale

2322

2323

2324

2325

2326

2327

2328

2329

2330

2331

2332

2333

2334

2335

2336

2337

2338

2339

2340

2341

2342

2343

2344

2345

2346

2347

2348

2349

2350

2351

2352

2353

2354

2355

2356

2357

2358

2359

2360

2361

2362

2363

2364

2365

2366

2367

2368

2369

2370

2371

2372

2373

2374

2375

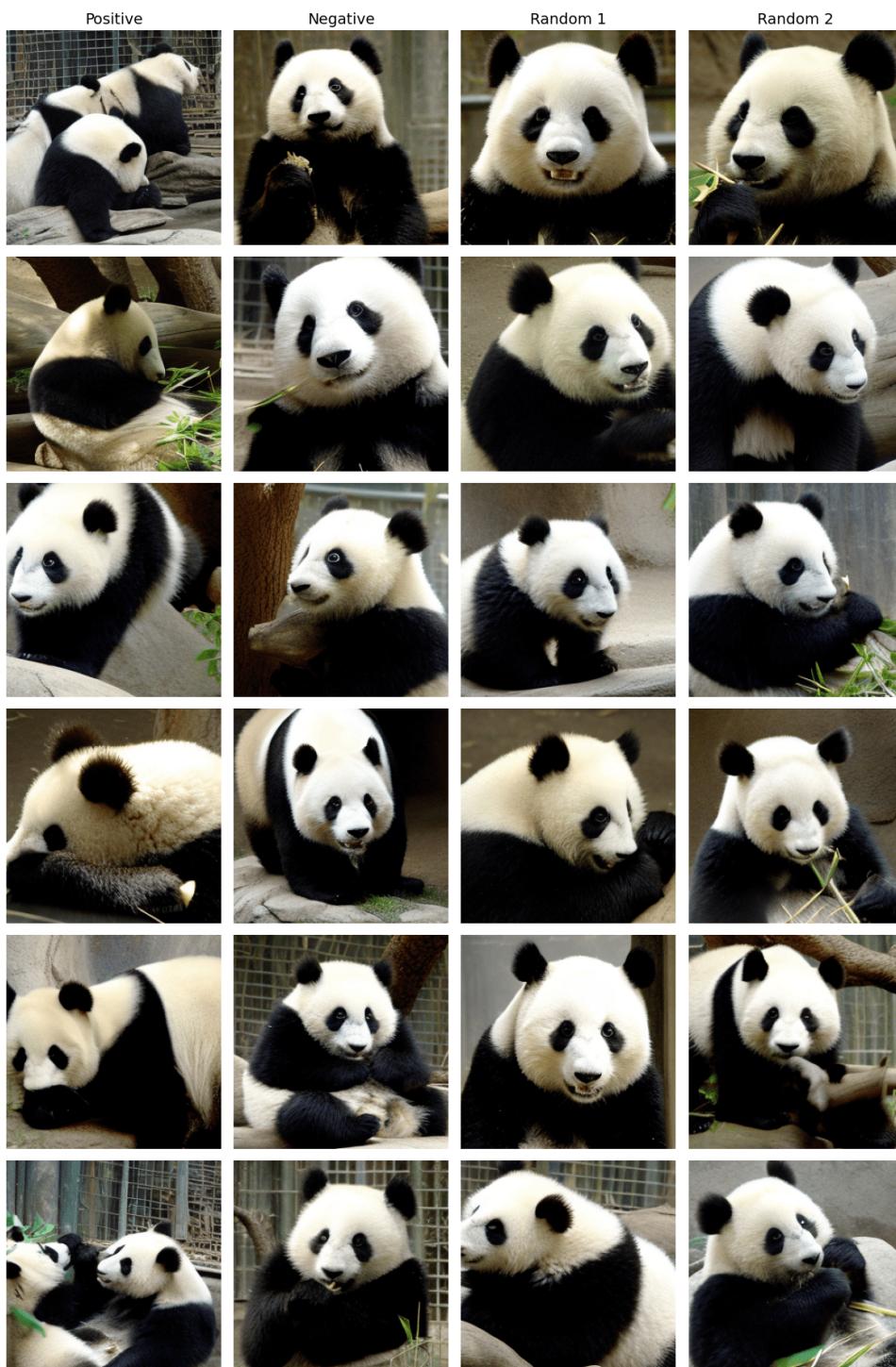


Figure 25: DiT Class 388: giant panda, panda

2376

2377

2378

2379

2380

2381

2382

2383

2384

2385

2386

2387

2388

2389

2390

2391

2392

2393

2394

2395

2396

2397

2398

2399

2400

2401

2402

2403

2404

2405

2406

2407

2408

2409

2410

2411

2412

2413

2414

2415

2416

2417

2418

2419

2420

2421

2422

2423

2424

2425

2426

2427

2428

2429

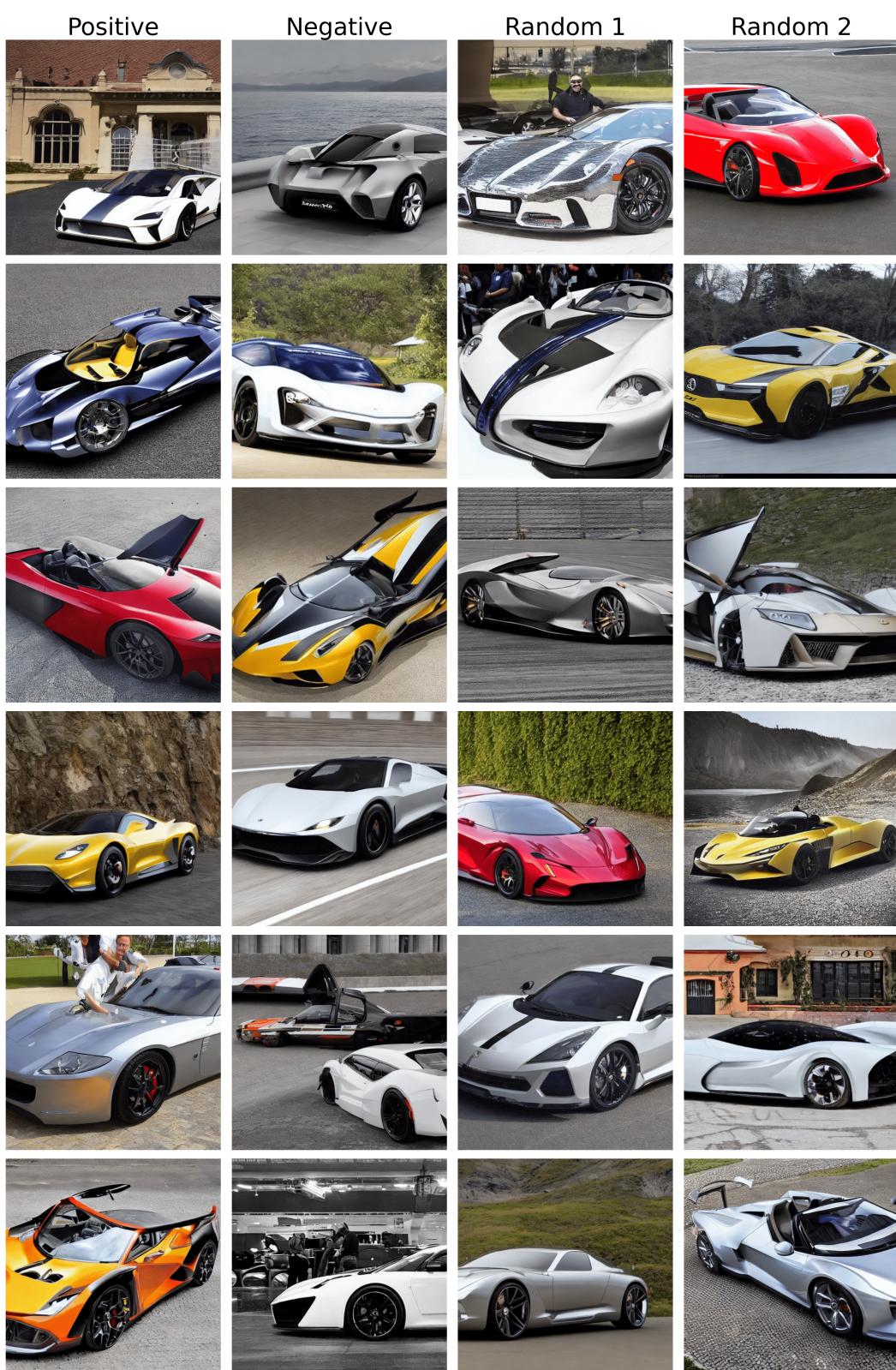


Figure 26: Prompt: “most expensive sports car”

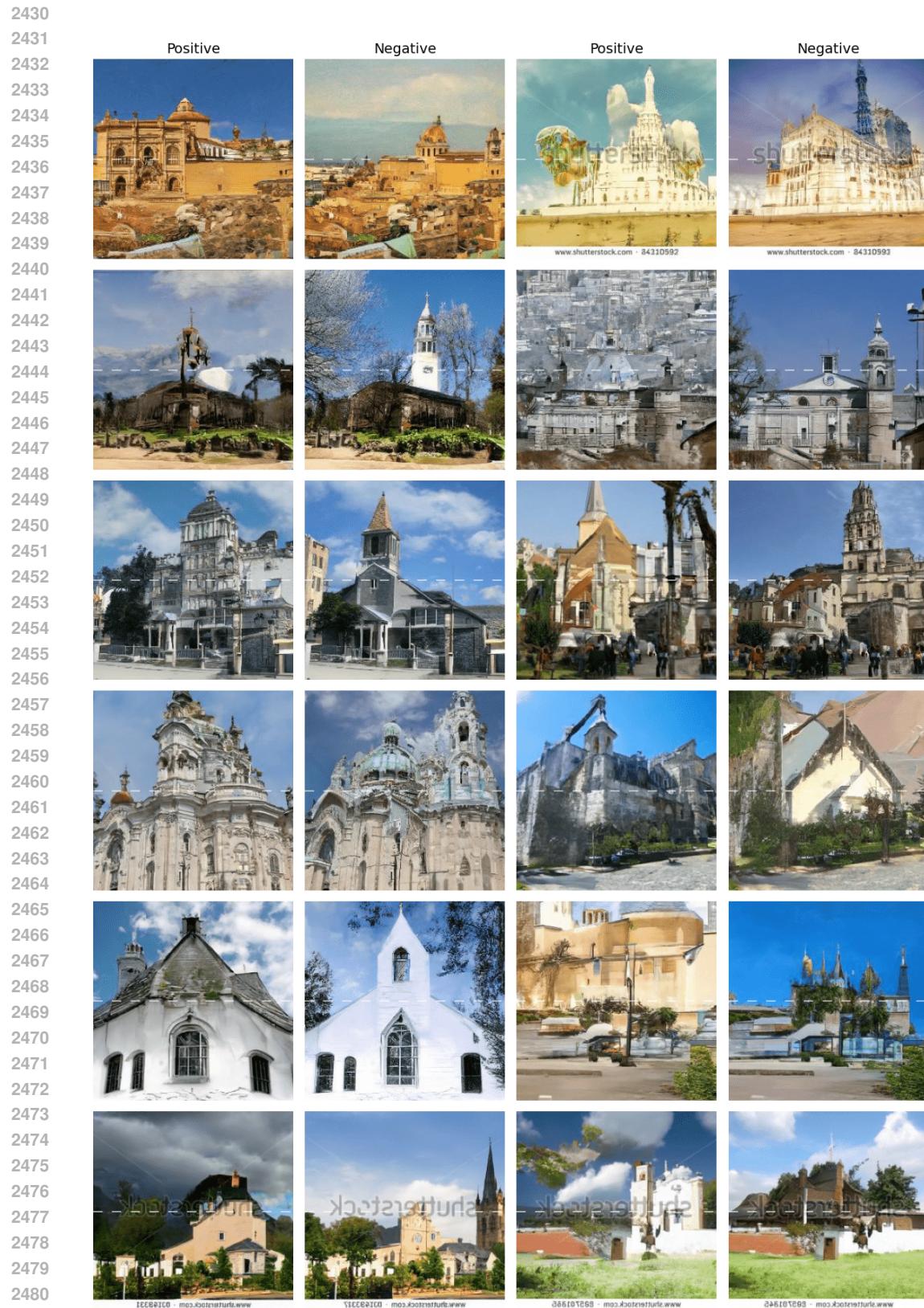


Figure 27: Partial Negation of LSUN-Church

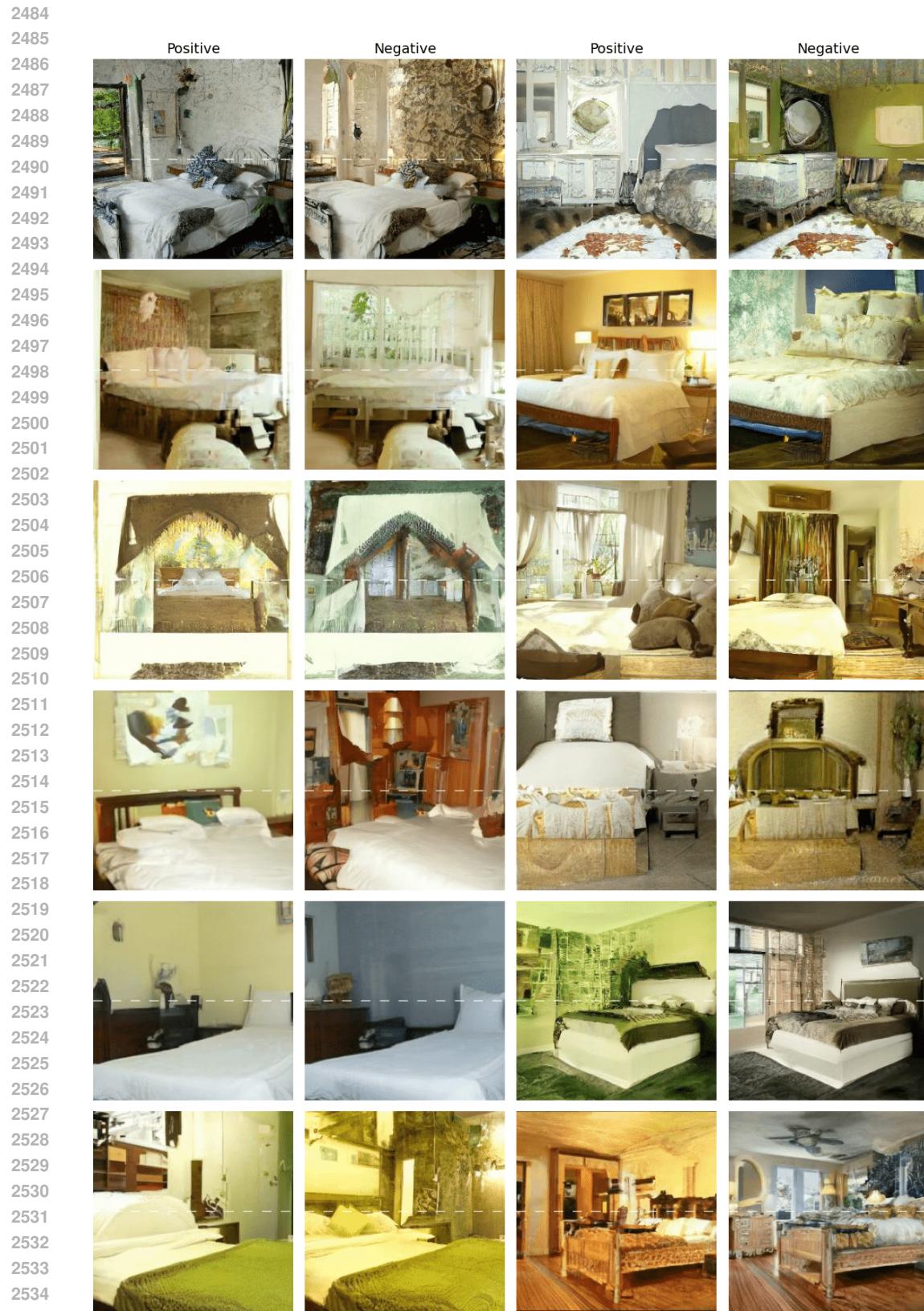


Figure 28: Partial Negation of LSUN-Bedroom