RETHINKING REPARAMETERIZATION OF STOCHASTIC PROCESSES IN GENERATIVE MODELING

Wojciech Kozłowski[1*], **Kamil Adamczewski**[1], **Radosław Kuczbański**[1] **& Maciej Zieba**[12]
[1] WUST, Poland [2] Tooploox, Poland * wojciech.kozlowski@pwr.edu.pl

ABSTRACT

Recent advances in diffusion-based image enhancement have motivated approaches that seek to reuse large pretrained diffusion models across different stochastic processes. In particular, it has been suggested that a diffusion backbone trained under one stochastic process can be repurposed to sample from alternative, task-specific processes through suitable reparameterizations, while relying on the pretrained model to estimate the underlying clean data. This perspective raises the question of whether stochastic processes can, in general, be transformed into one another without retraining the generative backbone. In this work, we study this question from a fundamental standpoint and show that such stochastic process reparameterization is not possible. Specifically, we prove that under commonly satisfied assumptions, any attempt to reparameterize a pretrained stochastic process backbone to represent a different stochastic process necessarily collapses to the original process itself. Our findings imply that a generative backbone trained under a given stochastic process can be used to sample only from that process, and cannot be reused to represent fundamentally different stochastic dynamics through reparameterization alone. We support our theoretical results with empirical ones demonstrating that reparameterization-based methods yield sampling behavior equivalent to standard process with modified sampling method.

1 INTRODUCTION

Diffusion models have emerged as a powerful framework for image generation and restoration, achieving state-of-the-art performance across a wide range of tasks (Sohl-Dickstein et al., 2015; Ho et al., 2020; Song et al., 2020b; Dhariwal & Nichol, 2021; Rombach et al., 2022). Beyond unconditional image synthesis, diffusion-based approaches have been successfully adapted to image enhancement problems such as super-resolution and restoration by conditioning the generative process on a low-quality input image (Saharia et al., 2022b;a). A common strategy concatenates the degraded image with the diffusion model input, yielding effective but task-specific solutions that typically require training a dedicated model for each enhancement setting.

To reduce the cost of training new diffusion backbones, several works have explored reusing pretrained generative models for image enhancement. Guidance-based approaches steer pretrained diffusion models toward high-quality solutions using auxiliary signals (Wang et al., 2022), while lightweight conditioning techniques such as ControlNet (Zhang et al., 2023) and diffusion adapters (Ye et al., 2023; Mou et al., 2024) train small auxiliary modules on top of a frozen diffusion backbone. These methods leverage the strong generative priors encoded in large pretrained models while avoiding full retraining.

In parallel, a different line of work has proposed modifying the *stochastic process itself* to better align diffusion dynamics with the structure of image enhancement tasks. Rather than starting from pure noise, these approaches incorporate the low-quality image directly into the diffusion process. Examples include Ornstein–Uhlenbeck–type processes whose stationary distribution is centered at the degraded image (Delbracio & Milanfar, 2023; Luo et al., 2023; Yue et al., 2023b), as well as diffusion bridge formulations that terminate exactly at the low-quality observation (Li et al., 2023; Zhou et al., 2023; Liu et al., 2023; Yue et al., 2023a; Zhu et al., 2025). While such process definitions are conceptually well aligned with restoration objectives, they typically require training or heavily fine-tuning a dedicated diffusion backbone (Yue et al., 2023b; Liu et al., 2023). As a

result, lightweight conditioning approaches that assume a fixed underlying diffusion process cannot be directly applied.

Recent work has sought to reconcile these two directions by introducing *reparameterization strategies* that aim to connect different stochastic processes during sampling. A representative example is IRBridge (Wang et al., 2025), which proposes switching between diffusion processes such that a pretrained diffusion model can be used to estimate intermediate clean samples while the overall sampling trajectory is intended to follow a different, task-specific stochastic process.

In this work, we study this idea from a more general and fundamental perspective. We ask whether a pretrained diffusion backbone, trained under a fixed stochastic process, can in fact be reused to represent and sample from a *different* stochastic process through reparameterization alone. We show that this is not possible. Specifically, we prove that under broad and commonly satisfied assumptions, any such reparameterization necessarily collapses to the original stochastic process of the pretrained model. While reparameterizations may alter the *parametrization of the sampler*, they cannot modify the underlying stochastic dynamics encoded by the pretrained backbone.

2 METHOD

Preliminaries. Let us consider two d-dimensional stochastic processes, the primary $\mathbf{x} = (\mathbf{x}_t)_{t\in[0,T]}$ and the secondary $\tilde{\mathbf{x}} = (\tilde{\mathbf{x}}_s)_{s\in[0,S]}$ governed by stochastic differential equations of the form $\mathrm{d}\mathbf{z}_t = \mathcal{F}(\mathbf{z}_t, t, \mathbf{y})\,\mathrm{d}t + \mathcal{G}(t)\,\mathrm{d}\mathbf{w}_t$, where $\mathbf{z} \in \{\mathbf{x}, \tilde{\mathbf{x}}\}$, $\mathrm{d}\mathbf{w}_t \in \mathbb{R}^d$ is d-dimensional Wiener process and $\mathbf{y} \in \mathbb{R}^d$ denotes a low-quality (degraded) image used as conditioning information. The drift $\mathcal{F} \in \mathbb{R}^d$ is process-dependent, Lipschitz continuous, and linear in its state variable, while the diffusion coefficient $\mathcal{G}(t) \in \mathbb{R}^+$ depends only on time. We assume that both processes share the same initial condition $\mathbf{x}_0 = \tilde{\mathbf{x}}_0$. Moreover, we assume access to a pretrained generative model for the process $\mathbf{x}$, which provides an approximation of the score function $\nabla_{\mathbf{x}_t} \log p_t(\mathbf{x}_t)$ and, equivalently, allows us to estimate the latent clean image $\hat{\mathbf{x}}_0$ from any state $\mathbf{x}_t$. Solving the associated SDEs would yield the solutions:

$$\mathbf{x}_t = f_t\mathbf{x}_0 + \mathbf{b}_t + \sigma_t\epsilon, \quad (1) \qquad \tilde{\mathbf{x}}_s = \tilde{f}_s\mathbf{x}_0 + \tilde{\mathbf{b}}_s + \tilde{\sigma}_s\epsilon, \quad (2)$$

with $f_t, \tilde{f}_s, \sigma_t, \tilde{\sigma}_s \in \mathbb{R}^+$ and $\mathbf{b}_t, \tilde{\mathbf{b}}_s \in \mathbb{R}^d$ (see Appendix A for a proof). Because processes $\mathbf{x}$ and $\tilde{\mathbf{x}}$ are conditionally independent given $\mathbf{x}_0$ and $\mathbf{y}$, we can introduce the new term that depends on $\tilde{\mathbf{x}}$ into Eq. (1) without changing the distribution of $\mathbf{x}$, since $p(\mathbf{x}_t|\mathbf{x}_0, \mathbf{y}) = p(\mathbf{x}_t|\mathbf{x}_0, \mathbf{y}, \tilde{\mathbf{x}}_s)$. Similarly, we can add $\mathbf{x}_t$ to Eq. (2).

Scope of reparametrizations. In this work, we consider *marginal (state-wise) reparameterization* between stochastic processes, that is, transformations applied independently at each time step to individual states. Concretely, we study reparameterization that map transition kernels from one process to another while preserving their distributions. This class includes reparameterizations of the form Eq. (3)–Eq. (4), which arise naturally for linear SDEs and are used in reparameterization-based methods such as IRBridge Wang et al. (2025).

Reparametrization. Following Wang et al. (2025) we define the *forward* and *reverse* transitions

$$\mathbf{x}_t = \alpha\tilde{\mathbf{x}}_s + (f_t - \alpha\tilde{f}_s)\mathbf{x}_0 + \mathbf{b}_t - \alpha\tilde{\mathbf{b}}_s + \tau\epsilon, \quad (3) \qquad \tilde{\mathbf{x}}_s = \tilde{\alpha}\mathbf{x}_t + (\tilde{f}_s - \tilde{\alpha}f_t)\mathbf{x}_0 + \tilde{\mathbf{b}}_s - \tilde{\alpha}\mathbf{b}_t + \tilde{\tau}\epsilon, \quad (4)$$

where

$$\alpha = \sqrt{\frac{\sigma_t^2 - \tau^2}{\tilde{\sigma}_s^2}}, \quad (5) \qquad \tilde{\alpha} = \sqrt{\frac{\tilde{\sigma}_s^2 - \tilde{\tau}^2}{\sigma_t^2}}, \quad (7)$$

$$\max\left\{0, \sigma_t^2 - \left(\frac{f_t\tilde{\sigma}_s}{\tilde{f}_s}\right)^2\right\} \le \tau^2 \le \sigma_t^2, \quad (6) \qquad \max\left\{0, \tilde{\sigma}_s^2 - \left(\frac{\tilde{f}_s\sigma_t}{f_t}\right)^2\right\} \le \tilde{\tau}^2 \le \tilde{\sigma}_s^2. \quad (8)$$

$\tau, \tilde{\tau}$ are parameters that control the noise level and because all coefficients in Eq. (3) and Eq. (4) have to be non-negative by definition, therefore, Eq. (6) and Eq. (8) have to be met. Note

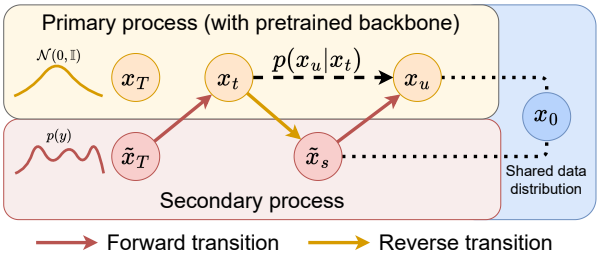

Figure 1: The diagram of the reparametrization scheme. We show that applying *reverse* and *forward* transitions to the primary process (with the pretrained backbone) $\mathbf{x}_t$ is independent of the secondary process $\tilde{\mathbf{x}}$ (from which want to sample) and is equivalent to $p(\mathbf{x}_u|\mathbf{x}_t)$.

that the primary process step t, and the secondary process step s are potentially different timesteps and their relationship is explained in the Appendix C.

The *forward transition* and the *reverse transition* allow us to switch between the two processes during inference, where we can estimate $\hat{\mathbf{x}}_0$ from one process and sample from another. However, during the *forward transition*, we do not have access to the estimate $\hat{\mathbf{x}}_0$, therefore, we have to manipulate t and τ so the coefficient $f_t - \alpha\tilde{f}_s = 0$. Please, see Appendix B for more details.

Reparameterization collapse. We focus on the scenario depicted in Figure 1, where we analyze two subsequent steps: $\mathbf{x}_t \to \tilde{\mathbf{x}}_s \to \mathbf{x}_u$, where $u < t$, and we would like to understand the role of the secondary $\tilde{\mathbf{x}}$ process. For simplicity, we adopt parameters values from Wang et al. (2025). We start by defining $\tilde{\mathbf{x}}_s$ using *reverse transition*

$$\tilde{\mathbf{x}}_s = \frac{\tilde{\sigma}_s}{\sigma_t}\mathbf{x}_t + \left(\tilde{f}_s - \frac{\tilde{\sigma}_s f_t}{\sigma_t}\right)\hat{\mathbf{x}}_0 + \tilde{\mathbf{b}}_s - \frac{\tilde{\sigma}_s \mathbf{b}_t}{\sigma_t} \tag{9}$$

and then use this definition of $\tilde{\mathbf{x}}_s$ and apply it in the *forward transition*, which gives

$$\mathbf{x}_u = \frac{f_u}{\tilde{f}_s}\left(\frac{\tilde{\sigma}_s}{\sigma_t}\mathbf{x}_t + \left(\tilde{f}_s - \frac{\tilde{\sigma}_s f_t}{\sigma_t}\right)\hat{\mathbf{x}}_0 + \tilde{\mathbf{b}}_s - \frac{\tilde{\sigma}_s \mathbf{b}_t}{\sigma_t}\right) + \mathbf{b}_u - \frac{f_u \tilde{\mathbf{b}}_s}{\tilde{f}_s} + \sqrt{\sigma_u^2 - \left(\frac{f_u \tilde{\sigma}_s}{\tilde{f}_s}\right)^2}\,\epsilon. \tag{10}$$

After simplification and defining $\tilde{\phi}_s \equiv \frac{\tilde{\sigma}_s}{\tilde{f}_s}$ we have

$$\mathbf{x}_u = \frac{f_u}{\sigma_t}\left(\tilde{\phi}_s \mathbf{x}_t + \left(\sigma_t - \tilde{\phi}_s f_t\right)\hat{\mathbf{x}}_0 - \tilde{\phi}_s \mathbf{b}_t\right) + \mathbf{b}_u + \sqrt{\sigma_u^2 - \left(\tilde{\phi}_s f_u\right)^2}\,\epsilon, \tag{11}$$

where $\tilde{\phi}_s$ is the only information about the process $\tilde{\mathbf{x}}$. Recall that the *forward transition* needs the coefficient $f_t - \alpha\tilde{f}_s = 0$, while for the *reverse transition*, we want to propagate some information about $\hat{\mathbf{x}}_0$, therefore $\tilde{f}_s - \tilde{\alpha} f_t > 0$. This leads to the following condition

$$\tilde{\sigma}_s^2 - \left(\frac{\tilde{f}_s \sigma_t}{f_t}\right)^2 < 0 \quad \wedge \quad \sigma_u^2 - \left(\frac{f_u \tilde{\sigma}_s}{\tilde{f}_s}\right)^2 \geq 0. \tag{12}$$

Because $\sigma, \tilde{\sigma}, f, \tilde{f}$ are strictly positive (see Appendix A), we can further simplify to

$$\tilde{\phi}_s < \frac{\sigma_t}{f_t} \quad \wedge \quad \tilde{\phi}_s \leq \frac{\sigma_u}{f_u} \quad \wedge \quad \tilde{\phi}_s > 0. \tag{13}$$

Here, for simplicity, we assume that $\mathbf{x}_t$ is a scaled Wiener or Orstein-Uhlenback (OU) process [1] (see Appendix D for a more general solution which also encompasses the family of diffusion bridges). For this group of processes, we have $\frac{\mathrm{d}\sigma_t}{\mathrm{d}t} > 0$ and $\frac{\mathrm{d}f_t}{\mathrm{d}t} \leq 0$ (0 for Variance Exploding (VE) diffusion, and negative otherwise), therefore, the fraction $\frac{\sigma_t}{f_t}$ is strictly increasing. Recall that $u < t$, which implies $\frac{\sigma_u}{f_u} < \frac{\sigma_t}{f_t}$ and leads to the following simplification of Eq. (13)

$$0 < \tilde{\phi}_s \leq \frac{\sigma_u}{f_u}. \tag{14}$$

[1] Note that the diffusion (typical process with pretrained backbone) is the Ornstein-Uhlenbeck process.

Let us define now $\eta_u \in (0, 1]$, so that $\tilde{\phi}_s = \frac{\sigma_u}{f_u}\eta_u$. In Appendix C, we show that, independent of the $\tilde{\mathbf{x}}$ process, the parameter η_u can take any value in $(0, 1]$, with its value determined by how the timesteps of one process are mapped to another. Thus the final definition of $\mathbf{x}_u$

$$\mathbf{x}_u = \frac{1}{\sigma_t}\left(\sigma_u\eta_u\mathbf{x}_t + (f_u\sigma_t - \sigma_u\eta_u f_t)\,\hat{\mathbf{x}}_0 - \sigma_u\eta_u\mathbf{b}_t\right) + \mathbf{b}_u + \sigma_u\sqrt{1-\eta_u^2}\epsilon \tag{15}$$

consists of terms that are all independent of $\tilde{\mathbf{x}}$.

Example. The Eq. (15) shows that the sampling procedure is independent of $\tilde{\mathbf{x}}_s$. To illustrate this result, let us consider the primary process $\mathbf{x}_t$ to be a Variance Preserving (VP) diffusion process. For this specific case, we follow the notation from Ho et al. (2020): $b_t, b_u = 0$, $f_t = \sqrt{\bar{\alpha}_t}$, $f_u = \sqrt{\bar{\alpha}_u}$, and $\sigma_t = \sqrt{1-\bar{\alpha}_t}, \sigma_u = \sqrt{1-\bar{\alpha}_u}$. Then

$$\mathbf{x}_u = \eta_u\sqrt{1-\bar{\alpha}_u}\frac{\mathbf{x}_t - \sqrt{\bar{\alpha}_t}\hat{\mathbf{x}}_0}{\sqrt{1-\bar{\alpha}_t}} + \sqrt{\bar{\alpha}_u}\hat{\mathbf{x}}_0 + \sqrt{1-\bar{\alpha}_u}\sqrt{1-\eta_u^2}\epsilon. \tag{16}$$

Because $\mathbf{x}_t = \sqrt{\bar{\alpha}_t}\hat{\mathbf{x}}_0 + \sqrt{1-\bar{\alpha}_t}\epsilon$, we can define the noise present in $\mathbf{x}_t$ as

$$\epsilon_t = \frac{\mathbf{x}_t - \sqrt{\bar{\alpha}_t}\hat{\mathbf{x}}_0}{\sqrt{1-\bar{\alpha}_t}}. \tag{17}$$

This, together with the definition of $\zeta_u = \sqrt{1-\bar{\alpha}_u}\sqrt{1-\eta_u^2}$ gives us

$$\mathbf{x}_u = \sqrt{\bar{\alpha}_u}\hat{\mathbf{x}}_0 + \sqrt{1-\bar{\alpha}_u-\zeta_u}\epsilon_t + \zeta_u\epsilon, \tag{18}$$

which is equivalent to DDIM method (see Eq. (12) in Song et al. (2020a)). In particular, when for all u we set $\eta_u = 1 \implies \zeta_u = 0$ we got the deterministic version of DDIM $\mathbf{x}_u = \sqrt{\bar{\alpha}_u}\hat{\mathbf{x}}_0 + \sqrt{1-\bar{\alpha}_u}\epsilon_t$.

3 Empirical Results

To validate our theoretical claims at inference time, we consider secondary processes drawn from two families of stochastic processes: Ornstein–Uhlenbeck (OU) processes, represented by IR-SDE and ResShift, and diffusion bridge processes, represented by GOUB and DDBM. As the primary process, we use a DDIM sampler with temperatures ζ_u determined by the timestep mapping. We follow the implementation details of IRBridge (Wang et al., 2025), training ControlNet on the raindrop removal task (Qian et al., 2018). As shown in Table 1, the differences between the two approaches are negligible and can be attributed to floating-point numerical errors rather than to differences in the inference pipeline. For more details, see Appendix E.

Table 1: Mean Squared Error (MSE) between results obtained from the reparametrization of four methods as secondary processes, and the equivalent primary process (in this case, DDIM). Results are shown for different processes and after one step and full inference. Empirically, the difference is negligible.

Method	Difference after	
	1 step	100 steps
IR-SDE (Luo et al., 2023)	1.45e−11	1.85e−05
ResShift (Yue et al., 2023b)	1.09e−11	4.48e−05
GOUB (Yue et al., 2023a)	5.11e−11	3.33e−05
DDBM (Zhou et al., 2023)	4.24e−09	1.26e−05

4 Conclusion

In this work, we showed that a pretrained generative backbone cannot be repurposed to sample from a different stochastic process through reparameterization alone. We established this result across a broad class of stochastic processes, including diffusion, Ornstein–Uhlenbeck, and diffusion bridge formulations, considering both primary and secondary processes. We proved that marginal reparameterizations necessarily collapse back to the original stochastic process, affecting only the parametrization of the sampler rather than the underlying dynamics. This demonstrates that the stochastic process on which a backbone is trained is an intrinsic component of the model itself and cannot be meaningfully altered without retraining, a conclusion supported by our empirical results.

ACKNOWLEDGMENTS

The work conducted by Wojciech Kozłowski, Radosław Kuczbański and Maciej Zieba was supported by the National Centre of Science (Poland) grant no. 2021/43/B/ST6/02853.

REFERENCES

Mauricio Delbracio and Peyman Milanfar. Inversion by direct iteration: An alternative to denoising diffusion for image restoration. *arXiv preprint arXiv:2303.11435*, 2023.

Prafulla Dhariwal and Alexander Nichol. Diffusion models beat gans on image synthesis. *Advances in neural information processing systems*, 34:8780–8794, 2021.

Joseph L Doob. *Classical potential theory and its probabilistic counterpart*, volume 262. Springer, 1984.

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

Diederik P Kingma. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

Bo Li, Kaitao Xue, Bin Liu, and Yu-Kun Lai. Bbdm: Image-to-image translation with brownian bridge diffusion models. In *Proceedings of the IEEE/CVF conference on computer vision and pattern Recognition*, pp. 1952–1961, 2023.

Guan-Horng Liu, Arash Vahdat, De-An Huang, Evangelos A Theodorou, Weili Nie, and Anima Anandkumar. I˙2 sb: Image-to-image schr\" odinger bridge. *arXiv preprint arXiv:2302.05872*, 2023.

Ziwei Luo, Fredrik K Gustafsson, Zheng Zhao, Jens Sjölund, and Thomas B Schön. Image restoration with mean-reverting stochastic differential equations. *arXiv preprint arXiv:2301.11699*, 2023.

Chong Mou, Xintao Wang, Liangbin Xie, Yanze Wu, Jian Zhang, Zhongang Qi, and Ying Shan. T2i-adapter: Learning adapters to dig out more controllable ability for text-to-image diffusion models. In *Proceedings of the AAAI conference on artificial intelligence*, volume 38, pp. 4296–4304, 2024.

Bernt Øksendal. Stochastic differential equations. In *Stochastic differential equations: an introduction with applications*, pp. 38–50. Springer, 2003.

Rui Qian, Robby T Tan, Wenhan Yang, Jiajun Su, and Jiaying Liu. Attentive generative adversarial network for raindrop removal from a single image. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2482–2491, 2018.

Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. High-resolution image synthesis with latent diffusion models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 10684–10695, June 2022.

Chitwan Saharia, William Chan, Huiwen Chang, Chris Lee, Jonathan Ho, Tim Salimans, David Fleet, and Mohammad Norouzi. Palette: Image-to-image diffusion models. In *ACM SIGGRAPH 2022 conference proceedings*, pp. 1–10, 2022a.

Chitwan Saharia, Jonathan Ho, William Chan, Tim Salimans, David J Fleet, and Mohammad Norouzi. Image super-resolution via iterative refinement. *IEEE transactions on pattern analysis and machine intelligence*, 45(4):4713–4726, 2022b.

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

Jiaming Song, Chenlin Meng, and Stefano Ermon. Denoising diffusion implicit models. *arXiv preprint arXiv:2010.02502*, 2020a.

Yang Song, Jascha Sohl-Dickstein, Diederik P Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-based generative modeling through stochastic differential equations. *arXiv preprint arXiv:2011.13456*, 2020b.

Hanting Wang, Tao Jin, Wang Lin, Shulei Wang, Hai Huang, Shengpeng Ji, and Zhou Zhao. IR-Bridge: Solving image restoration bridge with pre-trained generative diffusion models. In *Forty-second International Conference on Machine Learning*, 2025. URL https://openreview.net/forum?id=b3bJR1quJ3.

Yinhuai Wang, Jiwen Yu, and Jian Zhang. Zero-shot image restoration using denoising diffusion null-space model. *arXiv preprint arXiv:2212.00490*, 2022.

Hu Ye, Jun Zhang, Sibo Liu, Xiao Han, and Wei Yang. Ip-adapter: Text compatible image prompt adapter for text-to-image diffusion models. *arXiv preprint arXiv:2308.06721*, 2023.

Conghan Yue, Zhengwei Peng, Junlong Ma, Shiyan Du, Pengxu Wei, and Dongyu Zhang. Image restoration through generalized ornstein-uhlenbeck bridge. *arXiv preprint arXiv:2312.10299*, 2023a.

Zongsheng Yue, Jianyi Wang, and Chen Change Loy. Resshift: Efficient diffusion model for image super-resolution by residual shifting. *Advances in Neural Information Processing Systems*, 36: 13294–13307, 2023b.

Lvmin Zhang, Anyi Rao, and Maneesh Agrawala. Adding conditional control to text-to-image diffusion models. In *Proceedings of the IEEE/CVF international conference on computer vision*, pp. 3836–3847, 2023.

Linqi Zhou, Aaron Lou, Samar Khanna, and Stefano Ermon. Denoising diffusion bridge models. *arXiv preprint arXiv:2309.16948*, 2023.

Kaizhen Zhu, Mokai Pan, Yuexin Ma, Yanwei Fu, Jingyi Yu, Jingya Wang, and Ye Shi. Unidb: A unified diffusion bridge framework via stochastic optimal control. *arXiv preprint arXiv:2502.05749*, 2025.

A SDE SOLUTIONS

We assume $\mathbf{x}_t$ and $\tilde{\mathbf{x}}_s$ satisfy the following stochastic differential equations (SDEs)

$$\mathrm{d}\mathbf{x}_t = \mathcal{F}(\mathbf{x}_t, t, \mathbf{y})\,\mathrm{d}t + \mathcal{G}(t)\,\mathrm{d}\mathbf{w}_t, \tag{19}$$

$$\mathrm{d}\tilde{\mathbf{x}}_s = \tilde{\mathcal{F}}(\tilde{\mathbf{x}}_s, s, \mathbf{y})\,\mathrm{d}s + \tilde{\mathcal{G}}(s)\,\mathrm{d}\tilde{\mathbf{w}}_s. \tag{20}$$

Here, $\mathrm{d}\mathbf{w}_t$ and $\mathrm{d}\tilde{\mathbf{w}}_s$ are independent Wiener processes, and drifts $\mathcal{F}$, $\tilde{\mathcal{F}}$, as well as diffusion coefficients $\mathcal{G}$ and $\tilde{\mathcal{G}}$ meet the conditions for Theory 5.2.1 from Øksendal (2003) to admit a strong unique solution. Because Eq. (19) and Eq. (20) have similar forms and assumptions, we will only derive the solution for $\mathbf{x}$, but the same steps would apply to $\tilde{\mathbf{x}}$.

We start by assuming Eq. (19) to be the *linear* SDE[2], i.e. $\mathcal{F}(\mathbf{x}_t, t, \mathbf{y}) = \mathcal{A}(t)\mathbf{x}_t + \mathcal{B}(t, \mathbf{y})$ therefore, the mild solution would have the following form

$$\mathbf{x}_t = M_{0,t}\mathbf{x}_0 + \int_0^t M_{s,t}\mathcal{B}(s, \mathbf{y})\,\mathrm{d}s + \int_0^t M_{s,t}\mathcal{G}(s)\,\mathrm{d}\mathbf{w}_t, \tag{21}$$

$$\text{with integrating factor: } M_{s,t} = \exp\left(\int_s^t \mathcal{A}(z)\,\mathrm{d}z\right). \tag{22}$$

By Ito isometry,

$$\mathrm{Var}(\mathbf{x}_t) = \int_0^t M_{s,t}^2\mathcal{G}^2(s)\,\mathrm{d}s. \tag{23}$$

Then, the solution can be written as

$$\mathbf{x}_t = M_{0,t}\mathbf{x}_0 + \int_0^t M_{s,t}\mathcal{B}(s, \mathbf{y})\,\mathrm{d}s + \sqrt{\int_0^t M_{s,t}^2\mathcal{G}^2(s)\,\mathrm{d}s}\;\epsilon. \tag{24}$$

Recall the notation from Eq. (1), which can we now expressed as

$$f_t = \exp\left(\int_0^t \mathcal{A}(s)\,\mathrm{d}s\right), \tag{25}$$

$$b_t = \int_0^t M_{s,t}\mathcal{B}(s, \mathbf{y})\,\mathrm{d}s, \tag{26}$$

$$\sigma_t = \sqrt{\int_0^t M_{s,t}^2\mathcal{G}^2(s)\,\mathrm{d}s}. \tag{27}$$

Because for $t \in [0, T)$, $\mathcal{A}, \mathcal{B}, \mathcal{G}$ are finite, it is clear that $f_t, \sigma_t \in \mathbb{R}^+$ and $\mathbf{b}_t \in \mathbb{R}^d$. The same applies to $\tilde{f}_s, \tilde{\mathbf{b}}_s, \tilde{\sigma}_s$ from Eq. (2). Additionally, note that $\mathbf{b}_t$ and $\tilde{\mathbf{b}}_s$ are the only terms that depend on $\mathbf{y}$.

B TRANSITIONS DETAILS

The *forward* and *reverse* transitions defined in Eq. (3) and Eq. (4) are valid whenever Eq. (6) and Eq. (8) hold. These conditions ensure that the coefficients in the reparameterization remain non-negative, and thus the procedure is correct. Since IRBridge assumes access to the pretrained model only for the $\mathbf{x}_t$ process, additional constraints are required. In the *forward transition*, the model is not available and no estimate of $\hat{\mathbf{x}}_0$ can be formed, so no information about $\hat{\mathbf{x}}_0$ should be transferred. In contrast, in the *reverse transition*, the model is available and can be used, which requires that the coefficient multiplying $\hat{\mathbf{x}}_0$ be greater than 0. We now define these conditions separately for each transition.

[2]This might seem as a strong limitation, but in practice all diffusion, flow matching, OU processes, and diffusion bridges fall into this category.

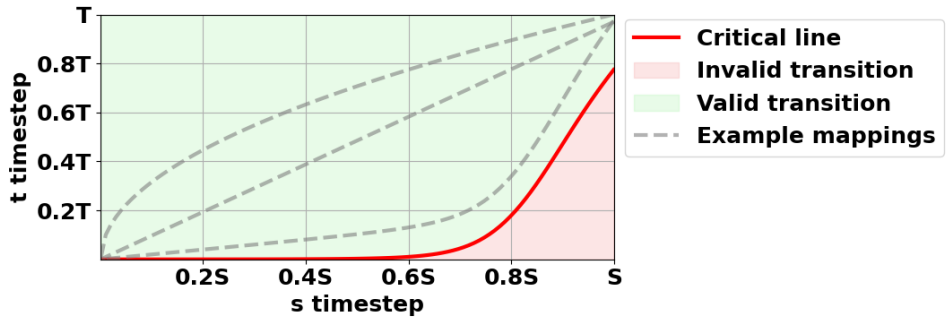

Figure 2: Diagram illustrating the correspondence between the timestep s of IR-SDE Luo et al. (2023) and the timestep t of Stable Diffusion 1.5 Rombach et al. (2022). The green region denotes the set of mappings that satisfy the condition $f_t - \alpha \tilde{f}_s = 0$. Mappings located near the critical line converge toward the deterministic DDIM sampling regime Song et al. (2020a), whereas mappings approaching the upper limit $t = T$ increasingly resemble independent transition kernels during the sampling process.

Forward transition In order to not transfer any information about $\hat{\mathbf{x}}_0$, we need $f_t - \alpha \tilde{f}_s = 0$. Firstly, from LHS of the Eq. (8) we want

$$\sigma_t^2 - \left(\frac{f_t \tilde{\sigma}_s}{\tilde{f}_s} \right)^2 \geq 0, \tag{28}$$

and we need as small τ^2 as possible, which is

$$\tau^2 = \sigma_t^2 - \left(\frac{f_t \tilde{\sigma}_s}{\tilde{f}_s} \right)^2. \tag{29}$$

Such τ applied to Eq. (5) ensures that $f_t - \alpha \tilde{f}_s = 0$. Please note, that there are many possible t that satisfy Eq. (28). If LHS equals 0, we have $\tau = 0$, therefore, it will be a deterministic affine transform. Such t is then called critical timestep. Otherwise if LHS > 0, we can still have $f_t - \alpha \tilde{f}_s = 0$, but there will be some stochasticity in the forward transition. We as well as IRBridge Wang et al. (2025) consider both scenarios.

Reverse transition Because we use $\mathbf{x}$ process to obtain the estimate $\hat{\mathbf{x}}_0$ using a pretrained model, we want use this information in the reverse transition i.e. $\tilde{f}_s - \tilde{\alpha} f_t > 0$. To make this possible, we take LHS of Eq. (8) and set

$$\tilde{\sigma}_s^2 - \left(\frac{\tilde{f}_s \sigma_t}{f_t} \right) < 0. \tag{30}$$

No specific value of $\tilde{\tau}$ is required, however, the authors of IRBride set $\tilde{\tau} = 0$, because it maximizes propagated information about $\hat{\mathbf{x}}_0$. For higher values, we would add noise at the cost of the model's prediction.

C TIMESTEP MAPPING AND CRITICAL LINE

The choice of timesteps matching function from one process to another is bounded by two conditions from Eq. (28) and Eq. (30). For the *forward transition*, LHS can be equal to zero, which is equivalent to the equality of the noise-to-signal ratio of both processes $\frac{\tilde{\sigma}_s}{\tilde{f}_s} = \frac{\sigma_t}{f_t}$ and it is called *critical line*. Figure 2 shows how this equality defines the border of possible timestep mapping, where everything above and on *critical line* is a valid mapping, while the region below the curve indicates too small t. The condition of the *reverse transition* marks the same region, but without the critical line itself.

For our consideration, it is important to recall the definition[3] of $\eta_t \in (0, 1]$, so that $\frac{\tilde{\sigma}_s}{\tilde{f}_s} = \frac{\sigma_t}{f_t} \eta_t$. Therefore, if we select the mapping that lies on the *critical line* , then $\eta_t = 1$, however, when $t \to T$

[3]Note that in the main text, u was used, because t denoted the previous timestep.

then $(f_t \to 0) \implies (\frac{\sigma_t}{f_t} \to \infty)$ which makes $\eta_t \to 0$ counteract the exploding term. Please note that the relative position of the mapping function to the *critical line* is independent of the $\tilde{\mathbf{x}}$ process definition. Similarly, the discretization function that tells which finite subset of $s_i \in [0, S]$ steps is chosen in the inference, is also independent of the definition of the $\tilde{\mathbf{x}}$ process. Therefore, we can select an arbitrary finite set of $t_i \in [0, T]$ time steps and corresponding $\eta_{t_i} \in (0, 1]$ values independent of the $\tilde{\mathbf{x}}$ process.

D Diffusion Bridge as the primary process

In the main body, we focus on the simplified setting in which the primary process x is either an Ornstein–Uhlenbeck process (as in IR-SDE Luo et al. (2023), ResShift Yue et al. (2023b), or the variance-preserving (VP) diffusion process Song et al. (2020b)) or a scaled Wiener process (such as the variance-exploding (VE) diffusion process Song et al. (2020b)). Here, we show that the same conclusions extend to a broader class of stochastic processes known as diffusion bridges Zhou et al. (2023); Liu et al. (2023); Yue et al. (2023a); Li et al. (2023).

The main distinction is that, because σ_t is not necessarily monotonic, we can no longer assume that the ratio $\frac{\sigma_t}{f_t}$ is monotonic, even when f_t itself is decreasing.

The second difference arises from our earlier assumption that $f_t \to 0$ as $t \to T$ implies $\frac{\sigma_t}{f_t} \to \infty$. This implication does not follow directly in the present setting, since for diffusion bridges $\sigma_t \to 0$ as $t \to T$ as well.

In this section, we show that although this is not immediately evident, both assumptions remain valid for diffusion bridges.

Because all of the bridge methods mentioned can be interpreted as the result of the h-transform Doob (1984) applied to a Wiener or OU process, we know that

$$\lim_{t \to T} \mathcal{A}(t) = -\infty, \text{ with asymptotic behavior} \mathcal{A}(t) \sim \frac{-1}{T - t} \tag{31}$$

$$\lim_{t \to T} \mathcal{G}(t) = c, \tag{32}$$

where c is some positive finite constant that depends on the noise scheduler used in the methods. Eq. (31) holds because the h-function is of the form $\mathcal{G}^2 \nabla_{\mathbf{x}_t} \log p_{T|t}(\mathbf{y}|\mathbf{x}_t)$, which for scaled Wiener processes is $\mathcal{G}^2(t) \frac{\mathbf{y} - \mathbf{x}_t}{\int_t^T \mathcal{G}^2(s)\,\mathrm{d}s}$ of which the term next to $\mathbf{x}_t$ has a simple pole at T by the definition. For more details, see Zhou et al. (2023); Liu et al. (2023); Yue et al. (2023a); Li et al. (2023).

We construct the noise-to-signal ratio by

$$\frac{\sigma_t}{f_t} = \frac{\sqrt{\int_0^t M_{s,t}^2 \mathcal{G}^2(s)\,\mathrm{d}s}}{M_{0,t}}. \tag{33}$$

Using exponential properties and linearity of Riemann integral we have

$$M_{s,t} = \exp\left(\int_s^t \mathcal{A}(z)\,\mathrm{d}z\right) = \frac{\exp(\int_0^t \mathcal{A}(z)\,\mathrm{d}z)}{\exp(\int_0^s \mathcal{A}(z)\,\mathrm{d}z)} = \frac{M_{0,t}}{M_{0,s}}. \tag{34}$$

Due to the fact that $M_{0,t} \geq 0$, we have

$$\frac{\sigma_t}{f_t} = \frac{M_{0,t}\sqrt{\int_0^t M_{0,s}^{-2} \mathcal{G}^2(s)\,\mathrm{d}s}}{M_{0,t}} = \sqrt{\int_0^t \frac{\mathcal{G}^2(s)}{M_{0,s}^2}\,\mathrm{d}s}. \tag{35}$$

Monotonicity of $\frac{\sigma_t}{f_t}$. Note that in Eq. (35), integrand is always positive, therefore increasing t would increase the value of $\frac{\sigma_t}{f_t}$.

Limit of $\frac{\sigma_t}{f_t}$ as $t \to T$. We take the limit of Eq. (35)

$$\lim_{t \to T} \sqrt{\int_0^t \frac{\mathcal{G}^2(s)}{M_{0,s}^2}\,\mathrm{d}s}. \tag{36}$$

From Eq. (31) and Eq. (32), we know that when $s \to T$: $\mathcal{G}^2(s) \to c^2$, and $M_{0,s}^2 \to 0$ exponentially. Based on that, the integrand $\frac{\mathcal{G}^2(s)}{M_{0,s}^2} \to \infty$ exponentially which implies that the integral diverges and, therefore, the limit

$$\lim_{t \to T} \frac{\sigma_t}{f_t} = \infty. \tag{37}$$

This ensures that $\eta_t \in (0, 1]$ described in section C is also fulfilled for bridges.

E EXPERIMENTAL DETAILS

The main goal of the experiments we conducted was to validate whether IRBridge collapses to the original process. For this purpose, we could use any data (even random noise) and any backbone (even with random parameters), but for the sake of completeness, we chose to follow the experimental setup from IRBridge.

We used Stable Diffusion 1.5 and trained ControlNet for this model for $100k$ iterations on the Raindrop dataset Qian et al. (2018). The batch size was set to 12, and training was performed using the Adam optimizer Kingma (2014) with a learning rate of $5\mathrm{e}{-5}$.

Once ControlNet achieved satisfactory results, we used the diffusion process with this model as the primary process for IRBridge. We then selected four different methods (IR-SDE Luo et al. (2023), ResShift Yue et al. (2023b), GOUB Yue et al. (2023a), and DDBM (VE version) Zhou et al. (2023)), representing both the Ornstein–Uhlenbeck and diffusion bridge categories, as secondary processes. We ran IRBridge alongside the primary process alone using DDIM with temperatures matching the η_t values from IRBridge and calculated the mean squared error between the outcomes. As shown in Table 1, the difference is negligible and can be attributed to floating-point numerical errors rather than differences in the inference pipeline.