# Image Restoration Driven by Non-Markovian Noise

**Arina Belova**[*]
Fraunhofer HHI

**Gabriel Nobis**[*]
Fraunhofer HHI
Imperial College London

**Maximilian Springenberg**
Fraunhofer HHI

**Tolga Birdal**
Imperial College London

**Wojciech Samek**
Fraunhofer HHI
TU Berlin
BIFOLD

## Abstract

Recent advances in generative diffusion models have explored replacing the standard Brownian driving noise with non-Markovian fractional Brownian motion (fBM) or heavy-tailed Lévy processes. In image restoration, the Generalized OrnsteinUhlenbeck Bridge (GOUB) model maps low-quality to high-quality images by extending the stochastic dynamics of generative bridge models to a time-varying Ornstein-Uhlenbeck (OU) process, while retaining standard Brownian motion as the driving noise. We combine these two directions by investigating how GOUB can be extended to a driving fractional noise, leveraging a Markov approximation of fBM (MA-fBM) for tractable simulation. We show that extending GOUB to fractional noise is inherently difficult: the resulting noise process will depend on the starting value at terminal time, making it infeasible to initialize the backward process during sampling. To nevertheless enable non-Markovian noise trajectories for low-to-high image translation, we employ instead of GOUB the recently proposed Fractional Diffusion Bridge Model (FDBM) for paired data translation and extend its application to image restoration. Empirically, FDBM improves reconstruction quality across a broad range of metrics compared to a Brownian baseline with an identical diffusion constant, but still falls short of purely Brownian-driven GOUB dynamics with a time-dependent diffusion coefficient. Interestingly, FDBM performs best when Brownian motion is modified only minimally toward a non-Markovian process.

## 1 Introduction

Continuous-time generative diffusion models [1] traditionally learn data distributions via stochastic dynamics driven by Brownian motion [2–4], which is a centered Gaussian Markov process with independent and stationary increments. Recent advances in generative diffusion models [5–7] and generative bridge models [8] have replaced the standard Brownian motion with the non-Markovian fractional Brownian motion (fBM) [9, 10], a centered Gaussian process with correlated increments controlled by the Hurst index $H \in (0, 1)$. In contrast to Brownian motion, the trajectories of fBM exhibit distinct regularity depending on the Hurst index $H$. For $H > 0.5$, positively correlated increments lead to smoother, super-diffusive paths, whereas for $H < 0.5$, negative correlations yield rougher, sub-diffusive behavior. Standard Brownian motion is recovered when $H = 0.5$. In the context of image restoration, the Generalized Ornstein-Uhlenbeck Bridge (GOUB) [11] extends stochastic bridge dynamics to a time-varying Ornstein-Uhlenbeck process, while retaining Brownian

[*]Equal contribution

EurIPS 2025 Workshop on Principles of Generative Modeling (PriGM).

motion as the driving noise. The resulting generalized dynamics achieve strong performance across a range of restoration tasks, including inpainting, deraining, and super-resolution. In this work, we investigate whether the GOUB model can be extended to a driving fBM. Following Nobis et al. [8], we note that simulating fractional Brownian bridges directly is intractable. Therefore, we employ a Markov approximation of fBM (MA-fBM) [12, 13] as the driving process. We show that a straightforward extension of GOUB to a driving MA-fBM is not feasible, since the induced noise process depends on the information contained in the clean image at terminal time, thus rendering the initialization of the backward process infeasible. Instead, we adopt Fractional Diffusion Bridge Models (FDBM) [8] and compare the performance of Brownian and fractional noise for image restoration. Our contributions are:

- We derive the backward dynamics of GOUB with a Markovian approximation of fractional Brownian motion and show that it is not suitable for image restoration, as the terminal value of the noise process depends on information from the initial clean image.
- We extend the application of the recently proposed Fractional Diffusion Bridge Models to image restoration - specifically, the image deraining task - thereby enabling non-Markovian interpolating trajectories.
- We show that FDBM achieves substantial improvements over the Brownian baseline with identical diffusion constant, surpassing it on every evaluated metric.

## 2 Generalized Ornstein-Uhlenbeck bridges driven by fractional noise

For ease of notation, we assume one-dimensional data. The proposed theory generalizes to a multidimensional setting, as every data dimension is driven by independent noise. The Generalized Ornstein-Uhlebeck Bridge (GOUB) model of Yue et al. [11] begins with the Generalized Ornstein-Uhlenbeck (GOU) process following for $T > 0$ the dynamics

$$\mathrm{d}X_t = \theta_t(\mu - X_t)dt + g_t \mathrm{d}B_t, \quad t \in [0, T], \tag{1}$$

where $B$ is a standard Brownian motion, $g_t$ is the diffusion function, $\theta_t$ is a drift coefficient that determines how fast $X_t$ approaches the constant $\mu$, and $g_t$ and $\theta_t$ are constrained to satisfy $2\lambda^2 = g_t^2/\theta_t$, for a given constant $\lambda$. The resulting stochastic process $X$ solving eq. (1) is a stationary Gaussian-Markov process, where the mean of the marginal distribution approaches for large $T$ a Gaussian distribution with mean $\mu$ and variance $\lambda^2$ [11]. We replace in the following the standard Brownian motion used in the stochastic dynamics of GOUB by an approximation of

**Definition 2.1** (Type II Fractional Brownian Motion [9])**.** *Let $B = (B_t)_{t\geq 0}$ be a standard Brownian Motion (BM) and $\Gamma$ the Gamma function. The centered Gaussian process*

$$B_t^H = \frac{1}{\Gamma(H + \frac{1}{2})} \int_0^t (t - s)^{H-\frac{1}{2}} \mathrm{d}B_s, \quad t \geq 0, \tag{2}$$

*is called Type II fractional Brownian motion (fBM) with Hurst index $H \in (0, 1)$.*

Since the reverse model of a stochastic process driven by fBM is not available with sufficient structure to train a neural network [6, 14] we use a Markovian-approximation of fBM (MA-fBM) [13, 15], defined by the linear superposition

$$\hat{B}_t^H = \sum_{k=1}^{K} \omega_k Y_t^k, \quad Y_t^k = \int_0^t e^{-\gamma_k(t-s)} \mathrm{d}B_s, \quad t \in [0, T] \tag{3}$$

of the $K \in \mathbb{N}$ Ornstein-Uhlebeck processes $Y^1 \dots Y^K$ with approximation coefficients $\omega_1, \dots, \omega_K$ and geometrically spaced speeds of mean reversion $\gamma_1, ..., \gamma_K$. Importantly, the OU processes $Y_1, ..., Y_K$ are all driven by the same Brownian motion, approximating the time-correlated behavior of fBM. The above approximation is justified by the integral representation of fBM over a family of OU processes by Harms and Stefanovits [12], and we follow Daems et al. [13] in choosing the $L^2$-optimal approximation coefficients $\gamma_k$ and $\omega_k$. We fix $K = 5$ throughout all experiments, as it already provides a good approximation of Type II fBM Daems, Rembert [16, Figures 3.13–3.15]. Nobis et al. [8, Proposition 8] show that $\hat{B}^H$ is non-Markovian and coincides with Brownian motion for $H = 0.5$ only if the mean-reversion speeds of the underlying OU processes are adjusted accordingly. As we fix non-zero $\gamma_1 \neq \gamma_2 \neq \dots \neq \gamma_K$ across all Hurst indices $H \in (0, 1)$, the resulting

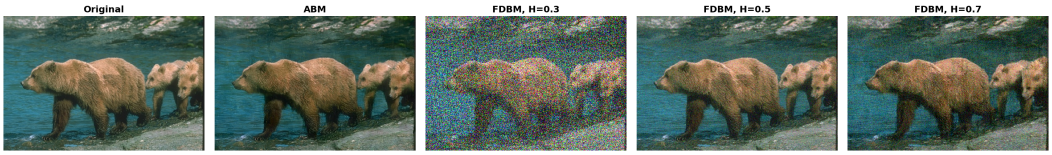



Figure 1: Qualitative results for varying $H$, with $K = 5$, $g = 0.2$, and 10 solver steps for image deraining.

driving noise process remains non-Markovian even for $H = 0.5$ and does not reduce to Brownian motion. Under these fixed $\gamma_1, \dots, \gamma_K$, the case $H = 0.5$ can thus be viewed as an $L^2$-minimal adjustment of Brownian motion to a non-Markovian process, since the approximation coefficients minimize $\int_0^T \mathbb{E}[\hat{B}_t^{0.5} - B_t^{0.5}]\mathrm{d}t$ [13]. Replacing the standard Brownian motion in eq. (1) by MA-fBM we find using the definition of the OU-processes and Ito's formula [17] the forward dynamics

$$\mathrm{d}X_t = \theta_t(\mu - X_t)\mathrm{d}t - g_t \sum_{k=1}^{K} \omega_k \gamma_k Y_t^k \mathrm{d}t + g_t \sum_{k=1}^{K} \omega_k \mathrm{d}B_t, \quad t \in [0, T], \tag{4}$$

with explicit forward representation derived in Section A. Assuming $\int_0^\infty \theta_s ds = \infty$ yields $\mathbb{E}[X_T | X_0] \to \mu$ for $T \to \infty$, implying that the mean-reverting property of Yue et al. [11] carries over to our generalized setting. Considering the augmented OU processes we write $Z = (X, Y^1, ..., Y^K)$ for the augmented forward process. Since $Z$ has a linear drift, we know that $Z$ and $X$ are Gaussian processes [18]. Conditioning $X$ to attain the terminal value $x_T$ we extend the partially pinned process of Nobis et al. [8, Proposition 4], derived for constant diffusion function, to the partially pinned process $Z_{|x_0,x_T} = Z|(X_0 = x_0, X_T = x_T)$ of our setting solving

$$\mathrm{d}Z_{|x_0,x_T}(t) = \left[J(t) + F(t)Z_{|x_0,x_T}(t) + G(t)G(t)^T u(t, Z_{|x_0,x_T}(t))\right] dt + G(t)\mathrm{d}B_t, \tag{5}$$

$$u(t,z) = \begin{pmatrix} \frac{\partial \mu_{T|t}}{\partial x_t}(z) & \frac{\partial \mu_{T|t}}{\partial y_1}(z) & \dots & \frac{\partial \mu_{T|t}}{\partial y_K}(z) \end{pmatrix}^T \frac{x_T - \mu_{T|t}(z)}{\sigma^2_{T|t}}, \tag{6}$$

where $\mu_{T|t}(z)$ and $\sigma^2_{T|t}$ denote the mean and the variance of the conditional terminal $X_T|(Z_t = z)$, respectively. The vector valued function $J(t), G(t) \in \mathbb{R}^{K+1}$ and the matrix valued function $F(t) \in \mathbb{R}^{K+1,K+1}$ are given in Appendix Section B. The above construction generalizes eq. (43) to a driving MA-fBM. To obtain the reverse dynamics, we follow the derivations of Nobis et al. [8, Appendix B.5] in our generalized setting, which yields:

$$\mathrm{d}\bar{Z}_{|x_0,x_T}(t) = \left[J(t) + F(t)\bar{Z}_{|x_0,x_T}(t) - G(t)G^T(t)\nabla_z \log \bar{p}_t(\bar{Z}_{|x_0,x_T}(t)|x_0)\right] \mathrm{d}t + G(t)\mathrm{d}\bar{B}_t. \tag{7}$$

Hence, the reverse dynamics of the partially pinned process coincide with those of the reference process conditioned on $x_0$, and the conditional score function $\nabla \log p_t(\cdot|x_0)$ can serve as a target for training a neural network to approximate the unconditional score function.

**Unfeasible reverse-time simulation of the fractional GOUB dynamics**. As the previous derivations show, the forward-backward scheme of GOUB can be generalized to a driving MA-fBM. However, the resulting backward dynamics are unsuitable for transforming distorted images into clean ones. To illustrate this, we examine the terminal distribution of the augmented OU processes in the partially pinned process. Our forward dynamics reduce for $\theta_t \equiv 0$ and $g_t \equiv g > 0$ to the forward dynamics used in Nobis et al. [8]. We note by the derivations of Nobis et al. [8, Appendix B.4] that we have for $Z_{|x_0,x_T} = (X_{|0,1}, Y^1_{|x_0,x_T}, \dots, Y^K_{|x_0,x_T})$

$$Y^k_{|x_0,x_T}(T) \overset{d}{=} a(x_T - x_0) + \lambda_T \xi, \quad \xi \sim \mathcal{N}(0,1), \tag{8}$$

where $a > 0$ and $\lambda_T > 0$ are known constants and $\xi$ is independent of $(X_0, X_1)$. To see that having only access to samples of $X_T$ does not suffice to initialize the terminal values $Y^k_{X_0,X_T}(T)$ conditionally on $X_T$, we compute

$$\mathbb{E}\big[Y^k_{X_0,X_T}(T) \,\big|\, X_T\big] = \mathbb{E}\big[a(X_T - X_0) + \lambda_T \xi \,\big|\, X_T\big] = a\big(X_T - \mathbb{E}[X_0 \mid X_T]\big), \tag{9}$$

where we use that $\mathbb{E}[\xi \mid X_T] = 0$. Thus, initializing $Y^k_{X_0,X_T}(T)$ given $X_T$ requires knowledge of the conditional law of $X_0 \mid X_T$, in particular of $\mathbb{E}[X_0 \mid X_T]$, which is unknown. Since we have have no access during sampling to $\mathbb{E}[X_0 \mid X_T]$, the fractional GOUB dynamics in eq. (7) are not suitable for training a generative data translation model.

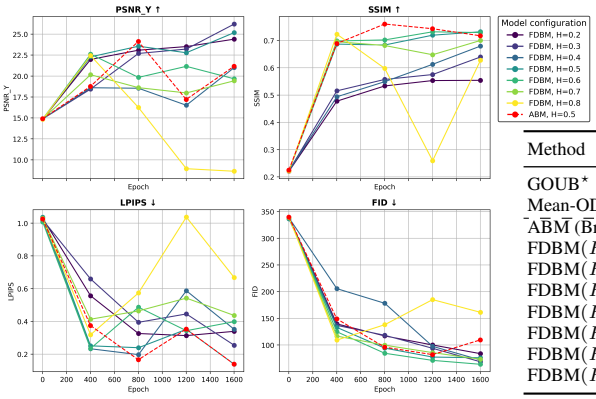



Figure 2: Quantitative results for varying $H$

| Method | PSNR_Y↑ | SSIM↑ | LPIPS↓ | FID↓ |
|---|---|---|---|---|
| GOUB$^\star$ [11] | **31.96** | **0.9** | **0.05** | **18.14** |
| Mean-ODE$^\star$ [11] | **34.56** | **0.9** | **0.08** | **32.83** |
| ABM (Brownian baseline) [19] | 19.76 | 0.71 | 0.38 | 106.74 |
| FDBM($H = 0.2$) | 24.4 | 0.55 | 0.34 | 83.8 |
| FDBM($H = 0.3$) | 26.79 | 0.65 | 0.22 | 63.64 |
| FDBM($H = 0.4$) | 24.28 | 0.7 | 0.22 | 94.43 |
| FDBM($H = 0.5$) | **27.35** | **0.77** | **0.15** | **55.74** |
| FDBM($H = 0.6$) | 22.6 | 0.7 | 0.23 | 124.82 |
| FDBM($H = 0.7$) | 23.66 | 0.68 | 0.2 | 139.67 |
| FDBM($H = 0.8$) | 22.44 | 0.72 | 0.32 | 109.32 |

Table 1: Results for 100 solver steps

## 3 Fractional Diffusion Bridge Models

To incorporate fractional noise into image restoration, we extend the application of Fractional Diffusion Bridge Models (FDBM) [8], which do not require a reverse-time model, to this task. Using eq. (5) within FDBM is generally possible, but it makes sampling from the partially pinned process considerably harder, since the covariances required for sampling are no longer available in closed form Nobis et al. [8, Appendix B.4]. For that reason, we focus in this work on exploring FDBM for image restoration with zero drift $\theta_t \equiv 0$ and a constant diffusion function $g(t) \equiv g > 0$, such that eq. (5) recovers the partially pinned process of FDBM:

$$\mathrm{d}Z_{|x_T,x_0}(t) = FZ_{|x_T,x_0}(t)\mathrm{d}t + GG^T u(t, Z_{|x_T,x_0}(t))\mathrm{d}t + G\mathrm{d}B(t), \quad Z_{|x_T,x_0}(0) = (x_T, 0_K), \tag{10}$$

$$u(t,z) = \begin{pmatrix} 1 & \omega_1\zeta_1(t,T) & \dots & \omega_K\zeta_K(t,T) \end{pmatrix}^T \frac{x_0 - \mu_{T|t}(z)}{\sigma^2_{T|t}}. \tag{11}$$

Here we intentionally swap the roles of $x_0$ and $x_T$, as our goal is to generate the clean image $x_0$ through forward simulation starting from a distorted image $x_T \sim \Pi_T$, while preserving the given coupling $(x_T, x_0) \sim \Pi_{T,0}$ between distorted and clean images. The data transforming process is parameterized as

$$\mathrm{d}Z^\theta_t = FZ^\theta_t\mathrm{d}t + GG^T u^\theta(t, X_0, Z^\theta_t)\mathrm{d}t + G\mathrm{d}B_t, \quad Z^\theta_0 = (X_T, 0, \dots, 0), \tag{12}$$

$$u^\theta(t, x_0, z) = \begin{pmatrix} 1 & \omega_1\zeta_1(t,T) & \dots & \omega_K\zeta_K(t,T) \end{pmatrix}^T \tilde{u}^\theta(t, x_0, \mu_{T|t}(z)), \tag{13}$$

where $\tilde{u}^\theta$ is a time-dependent neural network trained to minimize the KL-divergence $D_{\mathrm{KL}}(\mathbb{P}^\star|\mathbb{P}^\theta)$ of the $\Pi_{T,0}$-coupling-preserving path measure $\mathbb{P}^\star$ of $Z^\star$, constructed in Nobis et al. [8, Proposition 5], and the path measure $\mathbb{P}^\theta$ of $Z^\theta$, via

$$\mathcal{L}^{\text{paired}}_{\text{FDBM}}(\theta) := \int_0^T \mathbb{E}_{\mathbb{P}^\star}\left[\left\| \frac{X^\star_T - \mu_{T|t}(Z^\star_t)}{\sigma^2_{T|t}} - \tilde{u}^\theta(t, X_T, \mu_{T|t}(Z^\star_t))\right\|^2\right] \mathrm{d}t. \tag{14}$$

With the trained model $\tilde{u}^\theta$ at hand, we use the forward dynamics in eq. (12) to generate a clean image using the Euler-Maruyama scheme [17].

## 4 Results

To assess the quality of the generated clean images for the deraining task presented in Yue et al. [11], we evaluate image fidelity with the Peak Signal-to-Noise Ratio (PSNR) [11] and the Fréchet Inception Distance (FID) [20], perceptual quality with the structural similarity index measure (SSIM) [21], and perceptual similarity using the Learned Perceptual Image Patch Similarity (LPIPS) [22]. Following Yue et al. [11], we calculate PSNR, SSIM, and LPIPS on the Y channel YCbCr space, instead of a standard RGB space. We fix $g = 0.2$, as it empirically yields the best results across metrics for the image deraining task. This choice is consistent with the configuration that achieved the best performance in Nobis et al. [8] for predicting conformational changes in proteins. We report qualitative

results in Figure 1, as well as quantitative results in Table 1 and Figure 2. Interestingly, we observe that the setting $H = 0.5$, where Brownian motion is $L^2$-minimally adjusted to a non-Markovian process with respect to the fixed geometrically spaced grid of the approximation, yields the best performance across metrics—significantly outperforming Augmented Bridge Matching (ABM) [19] driven by Brownian motion with the same diffusion constant $g = 0.2$. Comparing our results with the metrics reported in Yue et al. [11] for GOUB shows that FDBM performs worse across all metrics. We emphasize that FDBM was trained for only $1600$ epochs due to computation limits, whereas GOUB was trained for $4000$ epochs. Nonetheless, we acknowledge a significant performance gap. We hypothesize that adapting the forward dynamics of FDBM to a time-dependent diffusion function is necessary to close this gap and to enable a fair comparison between purely Brownian-driven GOUB dynamics with a time-dependent diffusion coefficient and FDBM.

## 5 Conclusion

In this work, we analyzed the feasibility of extending the GOUB model to a driving Markov approximation of fBM and showed that such an extension is conceptually challenging, as the induced noise in the reverse-time dynamics depends on terminal information. In future work, we plan to investigate whether an adjusted Markovian approximation can mitigate this issue or whether it reflects an inherent property of the temporal correlations in fBM itself.

To incorporate fractional noise into image restoration, we employed Fractional Diffusion Bridge Models (FDBM) and demonstrated improved performance on an image-deraining task compared to purely Brownian-driven dynamics with an identical diffusion constant, although FDBM still falls short of Brownian-driven GOUB dynamics with a time-dependent diffusion coefficient. Interestingly, the best-performing noise regime of FDBM arises when Brownian motion is adjusted in an $L^2$-minimal sense to achieve a non-Markovian process.

Our findings highlight the importance of the underlying stochastic dynamics for the performance of generative diffusion bridge models, as evidenced by the performance differences observed across noise regimes. The empirical results further suggest that extending FDBM to a time-dependent diffusion function is a promising direction for future work, and may clarify whether the advantages of non-Markovian noise observed under a constant diffusion function persist when compared to time-dependent, purely Brownian-driven GOUB dynamics.

## Acknowledgments and Disclosure of Funding

T. Birdal acknowledges support from the Engineering and Physical Sciences Research Council [grant EP/X011364/1]. T. Birdal was supported by a UKRI Future Leaders Fellowship [grant number MR/Y018818/1]. This work was also supported by the German Research Foundation (DFG) through research unit DeSBi [KI-FOR 5363] (project ID: 459422098).

## References

[1] Yang Song, Jascha Sohl-Dickstein, Diederik P Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-based generative modeling through stochastic differential equations. In International Conference on Learning Representations, 2021.

[2] Robert Brown. XXVII. A brief account of microscopical observations made in the months of June, July and August 1827, on the particles contained in the pollen of plants; and on the general existence of active molecules in organic and inorganic bodies. The Philosophical Magazine, 4(21):161–173, 1828.

[3] Albert Einstein. Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. Annalen der Physik, pages 549–560, 1905.

[4] Norbert Wiener. Differential-space. Journal of Mathematics and Physics, 2:131–174, 1923.

[5] Anh Tong, Thanh Nguyen-Tang, Toan Tran, and Jaesik Choi. Learning fractional white noises in neural stochastic differential equations. In Advances in Neural Information Processing Systems, volume 35, pages 37660–37675. Curran Associates, Inc., 2022.

[6] Gabriel Nobis, Maximilian Springenberg, Marco Aversa, Michael Detzel, Rembert Daems, Roderick Murray-Smith, Shinichi Nakajima, Sebastian Lapuschkin, Stefano Ermon, Tolga Birdal, Manfred Opper, Christoph Knochenhauer, Luis Oala, and Wojciech Samek. Generative fractional diffusion models. In The Thirty-eighth Annual Conference on Neural Information Processing Systems, 2024. URL https://openreview.net/forum?id=B9qg3wo75g.

[7] Xiao Liang, Wentao Ma, Eric Paquet, Herna Viktor, and Wojtek Michalowski. Prot-gfdm: A generative fractional diffusion model for protein generation. Computational and Structural Biotechnology Journal, 27:3464–3480, 2025. ISSN 2001-0370. doi: https://doi.org/10.1016/j.csbj.2025.07.045. URL https://www.sciencedirect.com/science/article/pii/S2001037025003101.

[8] Gabriel Nobis, Maximilian Springenberg, Arina Belova, Rembert Daems, Christoph Knochenhauer, Manfred Opper, Tolga Birdal, and Wojciech Samek. Fractional diffusion bridge models. In The Thirty-ninth Annual Conference on Neural Information Processing Systems, 2025. URL https://openreview.net/forum?id=Hhc5McwASX.

[9] Paul Lévy. Random functions: general theory with special reference to Laplacian random functions. University of California Publications in Statistics, 1:331–390, 1953.

[10] Benoit B. Mandelbrot and John W. Van Ness. Fractional Brownian Motions, fractional noises and applications. SIAM Review, 10(4):422–437, 1968.

[11] Conghan Yue, Zhengwei Peng, Junlong Ma, Shiyan Du, Pengxu Wei, and Dongyu Zhang. Image restoration through generalized ornstein-uhlenbeck bridge. In Forty-first International Conference on Machine Learning, 2024. URL https://openreview.net/forum?id=oDUJmNCV8D.

[12] Philipp Harms and David Stefanovits. Affine representations of fractional processes with applications in mathematical finance. Stochastic Processes and their Applications, 129(4):1185–1228, 2019. ISSN 0304-4149.

[13] Rembert Daems, Manfred Opper, Guillaume Crevecoeur, and Tolga Birdal. Variational inference for SDEs driven by fractional noise. In The Twelfth International Conference on Learning Representations, 2024.

[14] Sebastien Darses and Bruno Saussereau. Time Reversal for Drifted Fractional Brownian Motion with Hurst Index $H > 1/2$. Electronic Journal of Probability, 12(none):1181 – 1211, 2007. doi: 10.1214/EJP.v12-439.

[15] Philipp Harms. Strong convergence rates for Markovian representations of fractional processes. Discrete and Continuous Dynamical Systems - B, 26(10):5567–5579, 2021. ISSN 1531-3492.

[16] Daems, Rembert. Learning from video in continuous time using physics priors and fractional noise. PhD thesis, Ghent University, 2025.

[17] Samuel N. Cohen and Robert J. Elliott. Stochastic Calculus and Applications. Probability and Its Applications. Birkhäuser, New York, NY, 2st edition, 2015. ISBN 978-1-4939-2866-8.

[18] Simo Särkkä and Arno Solin. Applied Stochastic Differential Equations, volume 10. Cambridge University Press, 2019.

[19] Valentin De Bortoli, Guan-Horng Liu, Tianrong Chen, Evangelos A. Theodorou, and Weilie Nie. Augmented bridge matching, 2023. URL https://arxiv.org/abs/2311.06978.

[20] Martin Heusel, Hubert Ramsauer, Thomas Unterthiner, Bernhard Nessler, and Sepp Hochreiter. GANs trained by a two time-scale update rule converge to a local Nash equilibrium. In Advances in Neural Information Processing Systems, volume 30. Curran Associates, Inc., 2017.

[21] Zhou Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli. Image quality assessment: from error visibility to structural similarity. IEEE Transactions on Image Processing, 13(4):600–612, 2004. doi: 10.1109/TIP.2003.819861.

[22] Richard Zhang, Phillip Isola, Alexei A Efros, Eli Shechtman, and Oliver Wang. The unreasonable effectiveness of deep features as a perceptual metric. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 586–595, 2018.

## A Derivations of the explicit forward representation

We follow the same arguments as in the proof of Nobis et al. [6, Proposition 4.2]. The solution to

$$X_t = -\int_0^t \theta_s X_s ds + \int_0^t [\theta_s \mu - \sum_{k=1}^K \omega_k \gamma_k g(s) Y_s^k] ds + \sum_{k=1}^K \omega_k \int_0^t g_s dB_s \tag{15}$$

is given by [17, Theorem 16.6.1]

$$X_T = c(T) \left[ x_0 + \int_0^T \frac{1}{c(s)} \left( \theta_s \mu - \sum_{k=1}^K \omega_k \gamma_k g(s) Y_s^k \right) ds + \sum_{k=1}^K \omega_k \int_0^T \frac{g(s)}{c(s)} dB_s \right] \tag{16}$$

where $c(t) = \exp(-\int_0^t \theta_s ds)$. Using

$$Y_s^\gamma | Z_t = e^{-\gamma(s-t)} y_t^\gamma + \int_t^s e^{-\gamma(s-u)} dB_u, \quad s > t \tag{17}$$

we find

$$X_T | Z_t = \underbrace{e^{-\int_t^T \theta_s ds} x_t + c(T)\mu \int_t^T e^{\int_0^s \theta_u du} \theta_s ds}_{:= \hat{a}_{T|t}(x_t)} - c(T) \int_t^T \frac{g(s)}{c(s)} \left[ \sum_{k=1}^K \omega_k \gamma_k (Y_s^k | Z_t) \right] ds \tag{18}$$

$$+ c(T) \sum_{k=1}^K \omega_k \int_t^T \frac{g(s)}{c(s)} dB_s \tag{19}$$

$$= \hat{a}_{T|t}(x_t) - c(T) \int_t^T \frac{g(s)}{c(s)} \left( \sum_{k=1}^K \omega_k \gamma_k \left[ e^{-\gamma_k (s-t)} y_t^k + \int_t^s e^{-\gamma_k (s-u)} dB_u \right] \right) ds \tag{20}$$

$$+ c(T) \sum_{k=1}^K \omega_k \int_t^T \frac{g(s)}{c(s)} dB_s \tag{21}$$

$$= \hat{a}_{T|t}(x_t) - c(T) \sum_{k=1}^K \omega_k \gamma_k y_t^k \int_t^T \frac{g(s)}{c(s)} e^{-\gamma_k (s-t)} ds \tag{22}$$

$$- c(T) \sum_{k=1}^K \omega_k \int_t^T \int_t^s \frac{g(s)}{c(s)} e^{-\gamma_k (s-u)} dB_u ds + c(T) \sum_{k=1}^K \omega_k \int_t^T \frac{g(s)}{c(s)} dB_s \tag{23}$$

Using Stochastic Fubini, we have

$$\int_t^T \int_t^s \frac{g(s)}{c(s)} e^{-\gamma_k (s-u)} dB_u ds = \int_t^T \int_u^T \frac{g(s)}{c(s)} e^{-\gamma_k (s-u)} ds dB_u \tag{24}$$

and hence in total

$$X_T | Z_t = e^{-\int_t^T \theta_s ds} x_t + c(T)\mu \int_t^T e^{\int_0^s \theta_u du} \theta_s ds \tag{25}$$

$$e^{-\int_0^T \theta_u du} \sum_{k=1}^K \omega_k \left\{ \int_t^T \underbrace{\left[ \frac{g(u)}{c(u)} - \gamma_k \int_u^T \frac{g(s)}{c(s)} e^{-\gamma_k (s-u)} ds \right]}_{:= \alpha_k(T,u)} dB_u + \gamma_k \int_t^T \frac{g(s)}{c(s)} e^{-\gamma_k (s-t)} ds y_t^k \right\} \tag{26}$$

$$= e^{-\int_t^T \theta_s ds} x_t + c(T)\mu \int_t^T e^{\int_0^s \theta_u du} \theta_s ds + \tag{27}$$

$$e^{-\int_0^T \theta_u du} \sum_{k=1}^K \omega_k \left\{ \int_t^T \alpha_k(T,u) dB_u + \gamma_k \int_t^T \frac{g(s)}{c(s)} e^{-\gamma_k (s-t)} ds y_t^k \right\} \tag{28}$$

Therefore with $c(T) = e^{-\int_0^T \theta_u du}$ we have for the mean

$$\mathbb{E}[X_T|Z_t] = e^{-\int_t^T \theta_s ds} x_t + c(T)\mu \int_t^T e^{\int_0^s \theta_u du} \theta_s ds - c(T) \sum_{k=1}^K \omega_k \gamma_k y_t^k \int_t^T \frac{g(s)}{c(s)} e^{-\gamma_k (s-t)} ds \tag{29}$$

$$= e^{-\int_t^T \theta_s ds} x_t + c(T)\mu [e^{\int_0^T \theta_u du} - e^{\int_0^t \theta_u du}] - c(T) \sum_{k=1}^K \omega_k \gamma_k y_t^k \int_t^T \frac{g(s)}{c(s)} e^{-\gamma_k (s-t)} ds \tag{30}$$

$$= e^{-\int_t^T \theta_s ds} x_t + \mu(1 - e^{-\int_t^T \theta_u du}) - c(T) \sum_{k=1}^K \omega_k \gamma_k y_t^k \int_t^T g(s) e^{\int_0^s \theta_u du} e^{-\gamma_k (s-t)} ds \tag{31}$$

and for the variance

$$\mathbb{V}[X_T|Z_t] = C^2(T) \int_t^T \left[ \sum_{k=1}^K \omega_k \alpha_k(T,u) \right]^2 ds \tag{32}$$

$$= e^{-2\int_0^T \theta_u du} \sum_{i,j}^K \omega_i \omega_j \int_t^T \alpha_i(T,u) \alpha_j(T,u) ds. \tag{33}$$

## B Details of the fractional Generalized Ornstein-Uhlenbeck process

The augmented forward $Z = (X, Y^1 \dots Y^K)$ process follows the dynamics

$$dZ(t) = [J(t) + F(t)Z(t)]\, dt + G(t) \mathrm{d}B(t) \tag{34}$$

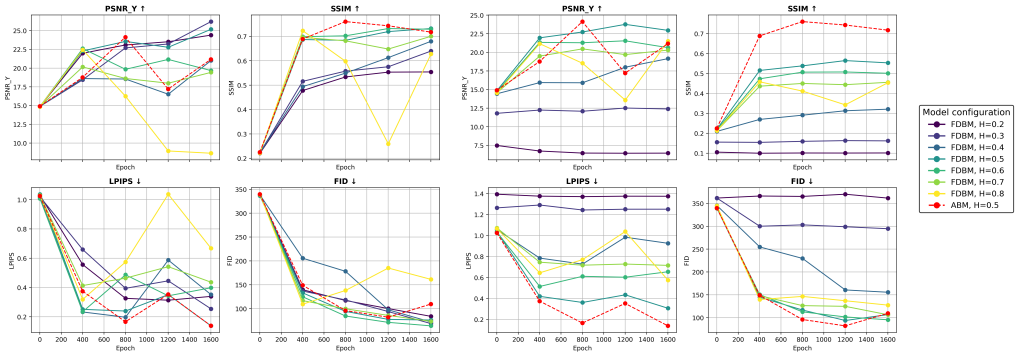
with the vector valued function $J(t), G(t) \in \mathbb{R}^{K+1}$ and the matrix valued function $F(t) \in \mathbb{R}^{K+1,K+1}$ given by

$$F(t) = \begin{pmatrix} -\theta_t & -g_t\omega_1\gamma_1 & -g_t\omega_2\gamma_2 & \dots & -g_t\omega_k\gamma_k \\ 0 & -\gamma_1 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots. & \dots & -\gamma_k \end{pmatrix} \in \mathbb{R}^{(K+1)\mathbf{x}(K+1)} \tag{35}$$

and

$$J(t) = \begin{pmatrix} \theta_t \mu \\ 0 \\ \dots \\ 0 \end{pmatrix} \in \mathbb{R}^{K+1} \quad \textit{and} \quad G(t) = \begin{pmatrix} g_t \sum_{k=1}^K \omega_k \\ 1 \\ \dots \\ 1 \end{pmatrix} \in \mathbb{R}^{K+1} \tag{36}$$

Since $Z$ has a linear drift, we know that $Z$ and $X$ are Gaussian processes [18].




(a) Solver steps = 100 (b) Solver steps = 10

Figure 3: Quantitative results for varying $H$ with different SDE solver steps.

## C Details of the partially pinned dynamics

To derive the control function $u$ in the partially pinned dynamics given in eq. (5) we calculate using $X_T|Z_t \sim \mathcal{N}(\underbrace{\mathbb{E}[X_T|Z_t]}_{m_{T|t}}, \underbrace{\mathbb{V}[X_T|Z_t]}_{\sigma^2_{T|t}})$, we find

$$u(t,z) := \nabla_{z_t} \log q(x_T|z_t) \tag{37}$$

$$= \left( \frac{x_T - m_{T|t}(z_t)}{\sigma^2_{T|t}} \frac{\partial m_{T|t}(z_t)}{\partial x_t}, \frac{x_T - m_{T|t}(z_t)}{\sigma^2_{T|t}} \frac{\partial m_{T|t}(z_t)}{\partial y^1_t}, ..., \frac{x_T - m_{T|t}(z)}{\sigma^2_{T|t}} \frac{\partial m_{T|t}(z_t)}{\partial y^k_t}) \right)^T \tag{38}$$

$$\frac{\partial m_{T|t}(z_t)}{\partial x_t} = \frac{\partial [e^{-\int_t^T \theta_s ds} x_t]}{\partial x_t} = e^{-\int_t^T \theta_s ds} \tag{39}$$

$$\frac{\partial m_{T|t}(z_t)}{\partial y^k_t} = -\frac{\partial [e^{\int_0^T \theta_u du} \sum_{k=1}^K \omega_k \gamma_k y^k_t \int_t^T g(s) e^{\int_0^s \theta_u du} e^{-\gamma_k (s-t)} ds]}{\partial y^k_t} \tag{40}$$

$$= e^{\int_0^T \theta_u du} \omega_k \gamma_k \int_t^T g(s) e^{\int_0^s \theta_u du} e^{-\gamma_k (s-t)} ds. \tag{41}$$

## D Implementation details and additional results

We use the same U-Net architecture as described in GOUB [11]. During training, we do not patch the images; instead, we use the full image size with a batch size of 8. We employ the Adam optimizer with standard parameters $\beta_1 = 0.9$ and $\beta_2 = 0.99$. The total number of training steps is 360000, corresponding to 1600 epochs, given the training dataset size of 1800 images. The learning rate is scheduled using the OneCycleLR policy with a total of 360000 steps and an initial learning rate of 0.0001.

The forward process parameters $F$ and $G$ are described in Appendix B, with the diffusion constant $g$ fixed as a value in $[0, 1]$ for each model configuration. The number of SDE solver steps was selected from $10, 50, 100$.

The effect of varying the number of solver steps can be observed in Figure 3 and Table 2. Increasing the number of SDE solver iterations leads to faster convergence of the metrics toward improved results for almost all model configurations, and also yields superior overall statistics.

## E Generalized Ornstein-Uhlebeck bridges

We summarize in this section the forward-backward dynamics of the Generalized Ornstein-Uhlebeck Bbridge (GOUB) modelYue et al. [11] starting with the Generalized Ornstein-Uhlenbeck (GOU)

(a) Inference steps $N = 100$

| Method / H | PSNR↑ | PSNR_Y↑ | SSIM↑ | LPIPS↓ | FID↓ |
|---|---|---|---|---|---|
| GOUB* [11] | | **31.96** | **0.9** | **0.05** | **20.93** |
| Mean-ODE* [11] | | **34.56** | **0.9** | **0.08** | **32.83** |
| FDBM($H = 0.2$) | 21.01 | 24.4 | 0.55 | 0.34 | 83.8 |
| FDBM($H = 0.3$) | 23.78 | 26.79 | 0.65 | 0.22 | 63.64 |
| FDBM($H = 0.4$) | 22.35 | 24.28 | 0.7 | 0.22 | 94.43 |
| FDBM($H = 0.5$) | **25.35** | **27.35** | **0.77** | **0.15** | **55.74** |
| FDBM($H = 0.6$) | 21.07 | 22.6 | 0.7 | 0.23 | 124.82 |
| FDBM($H = 0.7$) | 21.9 | 23.66 | 0.68 | 0.2 | 139.67 |
| FDBM($H = 0.8$) | 20.94 | 22.44 | 0.72 | 0.32 | 109.32 |
| ABM [19] | 18.36 | 19.76 | 0.71 | 0.38 | 106.74 |
| GOUB [11] | 30.31 | 31.69 | 0.9 | 0.05 | 20.93 |

(b) Inference steps $N = 10$

| Method / H | PSNR↑ | PSNR_Y↑ | SSIM↑ | LPIPS↓ | FID↓ |
|---|---|---|---|---|---|
| ABM [19] | 22.72 | 24.13 | **0.76** | **0.17** | 95.96 |
| FDBM($H = 0.2$) | 2.97 | 7.48 | 0.11 | 1.39 | 361.8 |
| FDBM($H = 0.3$) | 7.76 | 12.52 | 0.16 | 1.25 | 298.99 |
| FDBM($H = 0.4$) | 14.93 | 19.16 | 0.32 | 0.93 | 155.63 |
| FDBM($H = 0.5$) | **21.47** | **24.69** | 0.57 | 0.35 | **84.68** |
| FDBM($H = 0.6$) | 18.91 | 21.64 | 0.5 | 0.53 | 136.5 |
| FDBM($H = 0.7$) | 18.34 | 21.51 | 0.42 | 0.53 | 162.53 |
| FDBM($H = 0.8$) | 19.0 | 22.45 | 0.47 | 0.58 | 106.59 |

Table 2: Comparison of two methods. Results marked with * were recomputed (our baseline).

process following the dynamics

$$\mathrm{d}X_t = \theta_t(\mu - X_t)dt + g_t \mathrm{d}B_t, \tag{42}$$

where $B$ is a standard Brownian motion, $g_t$ is the diffusion function, $\theta_t$ is a drift coefficient, determining how fast $X_t$ approaches the constant $\mu$, and $g_t$ and $\theta_t$ are constrained to satisfy $2\lambda^2 = g_t^2/\theta_t$, for a given constant $\lambda$. The resulting stochastic process $X = (X_t)_{t\in[0,T]}$ solving eq. (1) is a stationary Gaussian-Markov process, where the mean of the marginal distribution approaches for large $T$ a Gaussian distribution with mean $\mu$ and variance $\lambda^2$ [11]. Conditioning $X$ on the terminal value $X_T = x_T$ yields via the $h$-transform [18] the forward dynamics used in GOUB to transform a clean image $x_0$ to a distorted image $x_T$ via [11]

$$\mathrm{d}X_t = \left(\theta_t + g_t^2 \frac{e^{-2\bar{\theta}_{t:T}}}{\bar{\sigma}_{t:T}}\right)(x_T - X_t)dt + g_t \mathrm{d}B_t, \quad X_0 = x_0, \tag{43}$$

where $\tilde{\theta}_{s:t} = \int_s^t \theta_u \mathrm{d}u$, $\tilde{\sigma}_{s:t} = (g_t^2/2\theta)(1 - e^{-2s\bar{\theta}_{s:t}})$ and the Gaussian transition density $p(x|x_0, x_t)$ of $X|(X_0 = x_0, X_T = x_T)$ is available in closed form. We write in the following whenever $X = (X(t))_{t\in[0,1]}$ is a stochastic process and $g$ is a function on $[0, 1]$, we write $\bar{X}(t) = X(1-t)$ for the reverse-time model and $\bar{g}(t) = g(1 - t)$ for the reverse-time function. The reverse-time model to eq. (43) is given by the backward dynamics [11],

$$\mathrm{d}\bar{X}_t = \left[\left(\bar{\theta}_t + \bar{g}_t^2 \frac{e^{-2\tilde{\theta}_{T-t:T}}}{\tilde{\sigma}_{T-t:T}}\right)(x_T - \bar{X}_t) - g_t^2 \nabla_{x_{T-t}} \log \bar{p}_t(\bar{X}_t|x_T)\right] dt + \bar{g}_t \mathrm{d}\bar{B}_t, \; X_T = x_T, \tag{44}$$

The score function $\nabla_{x_t} \log p_t$ is intractable and approximated by a time-dependent neural network $s_\theta$, whose parameters are trained via the maximum-likelihood objective proposed in Proposition 3.3 of Yue et al. [11].