Probabilistic Signature Inversion: Learning Conditional Distributions from Truncated Signatures

Junoh Kang[1] **Kiseop Lee**[2] **Bohyung Han**[1,3]
[1]ECE & [3]IPAI, Seoul National University [2]Department of Statistics, Purdue University
junoh.kang@snu.ac.kr kiseop@purdue.edu bhhan@snu.ac.kr

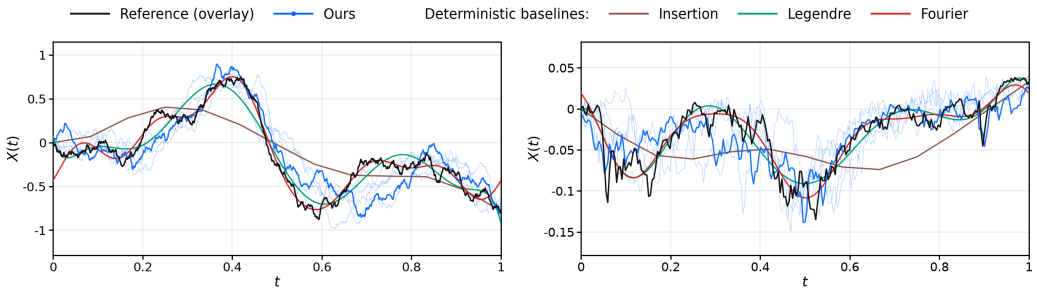
Abstract

The signature transform is a principled feature map for continuous-time paths, valued for its uniqueness and universality. Recovering a path from its truncated signature is, however, structurally ill-posed, because the truncated signature map is not injective. We therefore reframe truncated signature inversion as a probabilistic problem—learning the conditional distribution of a path given its truncated signature—and adopt a signature-conditioned flow matching model as a practical estimator. This formulation elucidates the fundamental difficulty of inversion: posterior-sampling reconstruction error quantifies the irreducible uncertainty remaining after conditioning on a statistic. For Gaussian process families, we derive this error from the tractable posterior under linear-statistics conditioning, obtaining a closed form for log-GBM and numerically tractable formulas for drifted fBM and OU. This oracle error upper-bounds the corresponding error under truncated-signature conditioning, since truncated signatures are more informative than linear statistics. The experimental results are consistent with the theoretical analysis: the learned sampler under linear-statistics conditioning closely matches the theoretically derived posterior in both reconstruction error and distribution, while reconstruction errors decrease under richer truncated-signature conditioning. Moreover, generated paths faithfully recover the conditioning signature and preserve key distributional and temporal structures. Together, these results establish a well-posed probabilistic framework for truncated-signature inversion, with applicability demonstrated on financial data beyond the process families covered by theory. Code is available at https://github.com/Junoh-Kang/ProbSigInv.

1 Introduction

The signature transform represents a continuous-time path as an infinite sequence of its iterated integrals [Lyons et al., 2007, Chevyrev and Kormilitzin, 2016], and owes its power to two foundational properties: uniqueness, which ensures that the full signature identifies a path up to tree-like equivalence [Hambly and Lyons, 2010, Boedihardjo et al., 2016], and universality, which allows any continuous path functional to be approximated by a linear functional of the signature [Chevyrev and Kormilitzin, 2016]. Signatures have thus become useful representations in time series analysis [Gyurkó et al., 2013, Guo et al., 2025] and, more recently, in generative modeling [Ni et al., 2021, Liao et al., 2024]. Uniqueness, in particular, motivates the problem of signature inversion. While a path is provably recoverable from its full signature in the rough-path setting [Geng, 2017], such reconstruction is infeasible in practice, as it requires the infinite-dimensional full signature. Practical methods must instead rely on truncated signatures, employing techniques such as symmetrization-based inversion [Lyons and Xu, 2018], insertion [Chang and Lyons, 2019, Fermanian et al., 2024], or basis-function recovery [Barancikova et al., 2025]. Despite their algorithmic diversity, these approaches predominantly adopt a point-reconstruction viewpoint, aiming to return a single representative path.

40th Conference on Neural Information Processing Systems (NeurIPS 2026).



(a) fBM ($H = 0.7$) (b) S&P 500

Figure 1: **Probabilistic and deterministic signature inversion.** Thin blue curves show five randomly selected paths from our probabilistic signature inversion method, with one draw highlighted in thick blue; deterministic baselines return one path each.

However, this deterministic point reconstruction leaves the non-injectivity induced by finite truncation unaddressed. While the full signature identifies a path up to tree-like equivalence, for any finite depth r, the truncated signature map $X \mapsto S^{\leq r}(X)$ is inherently many-to-one, and thus specifies an entire family of compatible paths rather than a unique inverse. To account for this structural ambiguity, we reframe truncated-signature inversion as the task of learning the prior-induced conditional distribution of paths given their signatures, which we estimate with a signature-conditioned flow matching model.

To quantify the ambiguity induced by finite-depth truncation, we analyze the posterior-sampling reconstruction error, which measures the irreducible uncertainty remaining after conditioning on the available information. For Gaussian process families, we derive this error from the tractable posterior under linear-statistics conditioning—specifically path endpoints and time moments. This yields a closed form for the log-process of geometric Brownian motion (log-GBM) and numerically tractable formulas for drifted fractional Brownian motion (drifted fBM) and the Ornstein–Uhlenbeck (OU) process. Critically, since linear statistics are coarser than truncated signatures, their associated oracle errors provide a computable upper bound on the corresponding error under truncated-signature conditioning.

We evaluate this formulation on three standard process families representing complementary structures in time-series modeling: log-GBM for Brownian diffusion, drifted fBM for long memory, and OU for mean reversion. Across these families, the experimental results are consistent with the theoretical analysis: the learned sampler under linear-statistics conditioning closely matches the theoretically derived posterior in both reconstruction error and distribution, while empirical errors under richer signature conditionings generally follow the predicted σ-algebra hierarchy. Figure 1 illustrates our framework: the blue trajectories form plausible path ensembles around a drifted fBM trajectory and an S&P 500 window. Parameter-estimation and path-level diagnostics further show that generated samples preserve the underlying distributional and temporal structure.

More broadly, probabilistic signature inversion extends the utility of signatures as a path representation, suggesting downstream uses such as probing how signature coordinates affect path behavior and generating signature-conditioned scenarios for financial time series. Our key contributions are as follows:

- We reformulate truncated-signature inversion as probabilistic signature inversion. Since truncation breaks injectivity, we frame the target as a conditional distribution rather than a single path, and we estimate this distribution using signature-conditioned flow matching.
- We quantify the irreducible ambiguity in truncated-signature inversion through the posterior-sampling reconstruction error. For Gaussian processes, we derive a tractable linear-statistics benchmark—providing a closed form for log-GBM and numerically tractable formulas for drifted fBM and OU—which upper-bounds the corresponding errors under richer signature conditionings.
- We empirically validate the framework against tractable theoretical benchmarks and demonstrate that generated samples preserve distributional and temporal structure across Gaussian processes and financial time series.

2 Related work

2.1 Signature inversion

The inverse problem of recovering paths from signatures has been studied from both theoretical and algorithmic perspectives. A theoretical line of work proves reconstruction under appropriate assumptions for several path classes, including Brownian motion [Le Jan and Qian, 2013], Gaussian processes [Boedihardjo and Geng, 2015], and deterministic rough paths [Geng, 2017]. However, these results rely on the full infinite-dimensional signature, limiting their direct practical use.

Explicit reconstruction procedures use truncated signatures to produce algorithmic path approximations under structural restrictions. For example, Lyons and Xu [2018] construct stable approximations for C^1 paths, while the insertion method [Chang and Lyons, 2019, Fermanian et al., 2024] reconstructs piecewise-linear paths through iterative point insertion. Barancikova et al. [2025] impose a different structural restriction, recovering signals by estimating coefficients in a prescribed finite basis, such as a Fourier or polynomial basis. Another line seeks a path whose signature features approximately match a target, using optimization or evolutionary search [Kidger et al., 2019, Bühler et al., 2024]. These approaches ultimately return a single reconstructed path. In contrast, our formulation models the truncation-induced ambiguity under an underlying path distribution, yielding a conditional distribution over compatible paths rather than a unique inverse.

2.2 Time series generation

Recent neural time series generators include adversarial [Yoon et al., 2019], autoregressive [Rasul et al., 2021], and diffusion-based models [Tashiro et al., 2021, Yuan and Qiao, 2024, Tanaka et al., 2025]. However, their conditioning mechanisms serve different objectives: forecasting and imputation models [Rasul et al., 2021, Tashiro et al., 2021] typically condition on partial observations, whereas CoFinDiff [Tanaka et al., 2025] targets population-level summaries like trend and volatility. In contrast, we focus on inversion from an individual path-specific statistic. Note that, due to these fundamental differences in conditioning and task objectives, a direct empirical comparison with forecasting- or population-summary-based generators is not straightforward.

Signatures have also been used in time series generation, but their role differs from ours. One line of work utilizes the signature as a metric for distributional matching, either through adversarial discrimination in the signature space [Ni et al., 2021, Liao et al., 2024] or by employing kernel MMD objectives between path distributions [Lu and Sester, 2025]. A second line of research focuses on generating samples directly within the (log-)signature space, subsequently recovering paths through a separate inversion step [Barancikova et al., 2025]. In contrast, we utilize the signature as conditioning information rather than as a training objective or a target representation to be generated.

3 Preliminaries

3.1 Signature of a path

Definition 1 (Signature)**.** Let the set of all multi-indices be $\mathcal{W} := \bigcup_{k=0}^{\infty} \{1, \dots, d\}^k$. For a continuous path $X : [0, T] \to \mathbb{R}^d$ and a multi-index $I = (i_1, \dots, i_k) \in \mathcal{W}$, the corresponding iterated integral, when well-defined, is

$$S(X)^I_{0,T} := \int_{0<t_1<\cdots<t_k<T} dX^{i_1}_{t_1} \cdots dX^{i_k}_{t_k}. \tag{1}$$

The signature of X over $[0, T]$ is the collection of all iterated integrals, $S(X)_{0,T} = \left(S(X)^I_{0,T}\right)_{I\in\mathcal{W}}$. For bounded-variation paths, these integrals are understood in the Riemann–Stieltjes sense; for continuous semimartingales, in the Stratonovich sense. For fractional Brownian motion with $H \in (1/4, 1)$, they are defined via its canonical geometric rough-path lift.

Theorem 2 (Tree-like equivalence)**.** Let $X, Y : [0, T] \to \mathbb{R}^d$ be continuous paths of bounded variation. Then $S(X)_{0,T} = S(Y)_{0,T}$ if and only if X and Y are equal up to tree-like equivalence.

Informally, two paths are tree-like equivalent if they differ only by segments that retrace themselves and hence cancel out. See Hambly and Lyons [2010] for this classical uniqueness result.

Definition 3 (Signature of observations)**.** Let $X = (x_i)_{i=0}^n$ be observations of a continuous path in $\mathbb{R}^d$ at times $0 = t_0 < \cdots < t_n = T$. Writing X^{lin} for their piecewise-linear interpolation, we define the signature of X as $S(X)_{0,T} := S(X^{\text{lin}})_{0,T}$.

Throughout the paper, $S(X)_{0,T}$ denotes the signature of the continuous-time process in the theoretical analysis and that of the piecewise-linear interpolation of its observations in the experiments.

Definition 4 (Truncated signature)**.** For $r \in \mathbb{N}$, the truncated signature of X at level r, $S(X)_{0,T}^{\leq r}$, is obtained by retaining terms $S(X)_{0,T}^I$ with $|I| \leq r$, where $|I|$ is the length of I.

For finite r, the truncated signature map $X \mapsto S(X)_{0,T}^{\leq r}$ is generally not injective.

3.2 Augmentation

Path augmentation is widely used to enrich the information captured by signatures. Two common choices are time and time lead-lag augmentations, defined below.

Definition 5 (Time augmentation)**.** Let X be observed at times $0 = t_0 < \cdots < t_n = T$. The time augmentation of these observations is a sequence of pairs:

$$\text{TA}(X) := \big((t_0, X_{t_0}), (t_1, X_{t_1}), \ldots, (t_n, X_{t_n})\big). \tag{2}$$

Definition 6 (Time lead-lag augmentation)**.** Let X be observed at times $0 = t_0 < \cdots < t_n = T$. The time lead-lag augmentation of these observations is

$$\text{TLL}(X)_{2i} := (t_i,\, X_{t_i},\, X_{t_i}), \qquad i = 0, \ldots, n, \tag{3}$$

$$\text{TLL}(X)_{2i+1} := (t_{i+1},\, X_{t_{i+1}},\, X_{t_i}), \qquad i = 0, \ldots, n-1, \tag{4}$$

where the second coordinate is the lead and the third is the lag.

The asynchronous updating of lead and lag coordinates encodes the quadratic variation of the original series in the level-2 signature terms [Chevyrev and Kormilitzin, 2016, Flint et al., 2016]. For $A \in \{\text{TA}, \text{TLL}\}$, let $S_A^{\leq r}(X)$ denote the level-r truncated signature of the augmented path $A(X)$.

4 Probabilistic signature inversion

4.1 Problem formulation

The full signature under either augmentation in Section 3.2 determines the path exactly when the starting point is fixed at $X_0 = 0$: the monotone time coordinate strengthens tree-like equivalence in Theorem 2 to exact equivalence [Levin et al., 2013, Morrill et al., 2020]. However, the truncated signature $S_A^{\leq r}(X)$ is a finite-dimensional projection of this infinite object, and multiple distinct paths may share the same truncation. This non-uniqueness motivates our formulation of signature inversion as learning the conditional distribution of paths given their truncated signatures.

Formally, let $X \sim p_{\text{data}}$ be a sample path and write $s = S_A^{\leq r}(X)$ for its truncated signature. Probabilistic signature inversion is defined as the task of learning the conditional distribution

$$P\big(X \mid S_A^{\leq r}(X) = s\big), \tag{5}$$

which is a distribution over paths that are consistent with the signature s and plausible under p_{data}.

4.2 Signature-conditioned flow matching

We employ a flow matching model [Lipman et al., 2023] to approximate this conditional distribution. In the objective below, X denotes the observed sequence at times $0 = t_0 < \cdots < t_n = T$. We train the model to learn a signature-conditioned velocity field that transports samples from a known distribution to data paths by minimizing

$$\mathcal{L}_{\text{FM}}(\phi) = \mathbb{E}_{X^{(1)}, X^{(0)}, \tau}\big[\|u_\phi(X^{(\tau)}; \tau, S_A^{\leq r}(X^{(1)})) - (X^{(1)} - X^{(0)})\|_2^2\big], \tag{6}$$

where $X^{(\tau)} = \tau X^{(1)} + (1-\tau)X^{(0)}$, with $X^{(1)} \sim p_{\text{data}}$, $X^{(0)} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$, and $\tau \sim \mathcal{U}(0,1)$. At inference time, we invert a signature s by solving the following ODE:

$$\frac{dX^{(\tau)}}{d\tau} = u_\phi\big(X^{(\tau)}; \tau, s\big), \quad X^{(0)} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}). \tag{7}$$

5 Posterior-sampling reconstruction error analysis

Probabilistic signature inversion reframes the ambiguity introduced by finite-depth signature truncation as intrinsic, making the inverse problem distributional rather than deterministic. We quantify this ambiguity through the posterior-sampling reconstruction error, which measures the irreducible path-level variation remaining after conditioning on a truncated signature. In this section, we derive tractable oracle errors under linear statistics for Gaussian process families, which upper-bound the corresponding oracle errors under richer signature conditionings.

5.1 Posterior-sampling reconstruction error for general parametric processes

Definition 7 (Posterior-sampling reconstruction error)**.** A prior π_Θ on Θ induces the mixture law $\bar{P}_\Theta := \int_\Theta P_\theta \, \pi(d\theta)$. For a statistic Φ, let $Q_\Theta(\cdot \mid s) = \bar{P}_\Theta(X \in \cdot \mid \Phi(X) = s)$ be the statistic-conditioned distribution. The posterior-sampling reconstruction error under the statistic Φ is

$$\mathcal{E}_\Theta(\Phi) = \mathbb{E}_{X \sim \bar{P}_\Theta,\, \tilde{X} \sim Q_\Theta(\cdot \mid \Phi(X))} \left[\int_0^T |X_t - \tilde{X}_t|^2 \, dt \right]. \tag{8}$$

When $\Phi = S_A^{\leq r}$, Q_Θ is the signature-conditioned distribution.

Proposition 8 (Within–between variance decomposition)**.** For any statistic Φ,

$$\mathcal{E}_\Theta(\Phi) = 2\mathbb{E}_{S \sim \Phi_\# \bar{P}_\Theta} \left[\int_0^T \mathrm{Var}_{Q_\Theta}(X_t \mid S) \, dt \right] \tag{9}$$

$$= 2\mathbb{E}_S \left[\int_0^T \mathbb{E}_{\theta \mid S}\big[\mathrm{Var}_\theta(X_t \mid S)\big] \, dt + \int_0^T \mathrm{Var}_{\theta \mid S}\big(\mathbb{E}_\theta[X_t \mid S]\big) \, dt \right]. \tag{10}$$

Proof. The proof is given in Appendix A.1. □

Definition 9 (Linear statistic)**.** For $r \geq 1$, the linear statistic $L^{(r)} \colon C([0,T];\mathbb{R}) \to \mathbb{R}^r$ is defined by

$$L^{(r)}(X) := \big(\ell_0(X), \ell_1(X), \ldots, \ell_{r-1}(X)\big)^\top, \tag{11}$$

where $\ell_0(X) := X_T$ and $\ell_p(X) := \int_0^T t^{p-1} X_t \, dt$ for $p = 1, \ldots, r-1$.

Proposition 10 (Linear statistic upper bound)**.** For $r \geq 1$,

$$\mathcal{E}_\Theta(S_{\mathrm{TLL}}^{\leq r}) \leq \mathcal{E}_\Theta(S_{\mathrm{TA}}^{\leq r}) \leq \mathcal{E}_\Theta(L^{(r)}). \tag{12}$$

Moreover, $\mathcal{E}_\Theta(S_{\mathrm{TA}}^{\leq r}) = \mathcal{E}_\Theta(L^{(r)})$ for $r \leq 2$.

Proof. Each $\ell_p(X)$ is determined by TA signature coordinates up to level $p+1$, directly for $p = 0$ and by integration by parts for $p \geq 1$. Moreover, $\mathrm{TA}(X)$ is the (time, lead) projection of $\mathrm{TLL}(X)$, up to reparametrization. Hence $\mathcal{F}(L^{(r)}(X)) \subseteq \mathcal{F}(S_{\mathrm{TA}}^{\leq r}(X)) \subseteq \mathcal{F}(S_{\mathrm{TLL}}^{\leq r}(X))$. For $r \leq 2$, the TA signature is also determined by $L^{(r)}(X)$, so their σ-algebras coincide.

For nested σ-algebras $\mathcal{F} \subseteq \mathcal{G}$, the law of total variance and the tower property give $\mathbb{E}[\mathrm{Var}(X_t \mid \mathcal{F})] \geq \mathbb{E}[\mathrm{Var}(X_t \mid \mathcal{G})]$. Under the integrability condition $\mathbb{E}[\int_0^T |X_t|^2 dt] < \infty$, integrating over time, interchanging expectation and integration by Fubini's theorem, and applying Proposition 8 yields the stated inequalities, with equality when the σ-algebras coincide. □

5.2 Oracle posterior-sampling reconstruction error for Gaussian processes

We now focus on Gaussian process families, whose analytical tractability and well-understood distributional properties have made them standard models across many domains. Proposition 8 reduces the oracle error to fixed-parameter conditional mean and variance terms under the observed statistic. For linear statistics, the following proposition gives the Gaussian regression formula for these moments, which will be used for the kernel-based computation in Section 5.4.

Proposition 11 (Gaussian regression under linear statistics)**.** Let $X_t = m(t) + G_t$, where m is deterministic and G_t is a centered Gaussian process with covariance kernel $K(\cdot,\cdot)$. For $r \geq 1$, define

$$k^{(r)}(t) := \mathrm{Cov}\big(G_t, L^{(r)}(G)\big), \qquad \Sigma^{(r)} := \mathrm{Cov}\big(L^{(r)}(G), L^{(r)}(G)\big). \tag{13}$$

Assume $\Sigma^{(r)}$ is invertible. Then $X \mid L^{(r)}(X)$ is a Gaussian process whose conditional mean and pointwise variance are given by

$$\mathbb{E}\big[X_t \mid L^{(r)}(X)\big] = m(t) + k^{(r)}(t)\big(\Sigma^{(r)}\big)^{-1}\,(L^{(r)}(X) - L^{(r)}(m)), \tag{14}$$

$$\mathrm{Var}\big(X_t \mid L^{(r)}(X)\big) = K(t,t) - k^{(r)}(t)\big(\Sigma^{(r)}\big)^{-1}\,k^{(r)}(t)^\top. \tag{15}$$

The conditional variance is independent of the observed value of $L^{(r)}(X)$.

Proof. Since m is deterministic, $\mathcal{F}(L^{(r)}(X)) = \mathcal{F}(L^{(r)}(G))$. For each t, $(G_t, L^{(r)}(G))$ is jointly Gaussian; the standard conditional Gaussian formula gives eqs. (14) and (15). □

Process families. For the rest of the analysis, we consider three classical Gaussian process families chosen to represent complementary regimes in time series modeling: the log-GBM as a canonical diffusion baseline with $\theta=(\mu,\sigma)$, the drifted fBM for long-memory behavior with $\theta=(\mu,\sigma,H)$, and the OU process for mean reversion with $\theta=(\mu,\sigma,\kappa)$. All three are initialized at $X_0 = 0$ and defined by

$$\text{log-GBM:}\quad dX_t = \nu\,dt + \sigma\,dB_t, \qquad \nu := \mu - \tfrac{1}{2}\sigma^2, \tag{16}$$

$$\text{drifted fBM:}\quad X_t = \nu\,t + \sigma B_t^H, \qquad \nu := \mu - \tfrac{1}{2}\sigma^2, \tag{17}$$

$$\text{OU:}\quad dX_t = \kappa(\mu - X_t)\,dt + \sigma\,dB_t, \tag{18}$$

where B is a standard Brownian motion and B^H is a fractional Brownian motion with Hurst parameter $H \in (0,1)$ and covariance $\mathrm{Cov}(B_s^H, B_t^H) = \frac{1}{2}(s^{2H} + t^{2H} - |s-t|^{2H})$.

5.3 Closed-form analysis for log-GBM

For log-GBM, the oracle error under $L^{(r)}$ can be computed in closed form. The key observation is that $L^{(r)}(X)$ can be represented by finitely many Wiener integrals, so the fixed-parameter conditional mean and variance terms in Proposition 8 reduce to an L^2 projection. Proposition 13 carries out this reduction and gives the closed-form oracle posterior-sampling reconstruction error.

Lemma 12 (Gaussian Hilbert-space projection)**.** Let B be a standard Brownian motion and define the Wiener integral $I(f) := \int_0^T f(s)\,dB_s$ for $f \in L^2([0,T])$. Then, I is linear and satisfies the Itô isometry, $\mathbb{E}[I(f)I(g)] = \langle f, g\rangle_{L^2([0,T])}$. For a closed subspace $M \subset L^2([0,T])$ and $f \in L^2([0,T])$,

$$\mathbb{E}\big[I(f) \mid \mathcal{F}(I(g) : g \in M)\big] = I(\Pi_M f), \tag{19}$$

$$\mathrm{Var}\big(I(f) \mid \mathcal{F}(I(g) : g \in M)\big) = \|f - \Pi_M f\|^2_{L^2([0,T])}, \tag{20}$$

where Π_M denotes the orthogonal projection onto M.

Proof. The proof is given in Appendix A.3. □

Proposition 13 (Closed-form oracle posterior-sampling reconstruction error for log-GBM)**.** For log-GBM and $r \geq 1$, there exists a deterministic function $c_r : [0,T] \to \mathbb{R}^r$ such that

$$\mathbb{E}_\theta\big[X_t \mid L^{(r)}(X)\big] = c_r(t)^\top L^{(r)}(X), \qquad t \in [0,T]. \tag{21}$$

In particular, the conditional mean is determined entirely by the observed statistic $L^{(r)}(X)$ and does not depend on $\theta = (\mu,\sigma)$. Thus, the between-parameter conditional-mean variance term in Proposition 8 vanishes. Moreover, the integrated conditional variance is

$$\int_0^T \mathrm{Var}_\theta\big(X_t \mid L^{(r)}(X)\big)\,dt = a_r\sigma^2T^2, \qquad a_r = \frac{r}{2(2r-1)(2r+1)}. \tag{22}$$

Substituting these identities into Proposition 8 yields the oracle error

$$\mathcal{E}_\Theta(L^{(r)}) = 2\,a_r\,T^2\,\mathbb{E}[\sigma^2]. \tag{23}$$

Proof sketch. Let $S := L^{(r)}(X)$. By Fubini, for fixed θ, $S = \nu\beta + \sigma(I(f_0),\dots,I(f_{r-1}))^\top$, where $f_0 \equiv 1$, $f_j(s) = (T^j - s^j)/j$ for $j \geq 1$, and $\beta = (T,\dots,T^r/r)^\top$. Thus, $\mathcal{F}(S) = \mathcal{F}(I(f) : f \in \mathcal{H}_r)$,

where $\mathcal{H}_r := \mathrm{span}\{f_0, \ldots, f_{r-1}\}$. Since $X_t = \nu t + \sigma I(\mathbf{1}_{[0,t]})$, Lemma 12 reduces the conditional mean to projecting $\mathbf{1}_{[0,t]}$ onto $\mathcal{H}_r$. Writing $\Pi_{\mathcal{H}_r}\mathbf{1}_{[0,t]} = c_r(t)^\top(f_0, \ldots, f_{r-1})$,

$$\mathbb{E}_\theta[X_t \mid S] = c_r(t)^\top S + \nu(t - c_r(t)^\top \beta). \tag{24}$$

Since $f_0 \equiv 1 \in \mathcal{H}_r$, $t = \langle \mathbf{1}_{[0,t]}, \mathbf{1}\rangle = \langle \Pi_{\mathcal{H}_r}\mathbf{1}_{[0,t]}, \mathbf{1}\rangle = c_r(t)^\top\beta$, the second term vanishes, yielding (21). Because the resulting conditional mean does not depend on θ, the between-parameter conditional-mean variance term in Proposition 8 vanishes. For the conditional variance,

$$\int_0^T \mathrm{Var}_\theta(X_t \mid S)\, dt = \sigma^2 T^2 \int_0^1 \|\mathbf{1}_{[0,u]} - \Pi_{\mathcal{H}_r}\mathbf{1}_{[0,u]}\|^2\, du = a_r \sigma^2 T^2, \tag{25}$$

where the last equality follows from Lemma 17. The full proof is given in Appendix A.4. □

Corollary 14 (Depth-two oracle error equality for log-GBM)**.** For log-GBM,

$$\mathcal{E}_\Theta(L^{(2)}) = \mathcal{E}_\Theta(S_{\mathrm{TA}}^{\le 2}) = \mathcal{E}_\Theta(S_{\mathrm{TLL}}^{\le 2}). \tag{26}$$

Proof. By Proposition 10, $\mathcal{E}_\Theta(S_{\mathrm{TA}}^{\le 2}) = \mathcal{E}_\Theta(L^{(2)})$. For log-GBM, $S_{\mathrm{TLL}}^{\le 2}$ is equivalent to $(L^{(2)}, \sigma)$, and Proposition 13 shows that the added σ does not change the conditional mean. The conditional-variance decomposition then gives $\mathcal{E}_\Theta(S_{\mathrm{TLL}}^{\le 2}) = \mathcal{E}_\Theta(L^{(2)})$. □

Corollary 15 (Multivariate log-GBM oracle)**.** For multivariate log-GBM with diffusion covariance Σ, the componentwise linear-statistic oracle under squared Euclidean path loss satisfies $\mathcal{E}_\Theta(L_d^{(r)}) = 2a_r T^2 \mathbb{E}_{\pi_\Theta}[\mathrm{tr}(\Sigma)]$ for $r \ge 1$.

Proof. The full statement and proof are given in Appendix A.5. □

5.4 Kernel-based numerical analysis

For Gaussian processes with known deterministic mean and covariance kernel—see Table 4 in Appendix A.2 for the three processes considered in this study—both direct sampling from the linear-statistic posterior and computation of $\mathcal{E}_\Theta(L^{(r)})$ are numerically tractable. At fixed θ, Proposition 11 expresses the conditional mean and pointwise variance terms required by Proposition 8 in terms of $L^{(r)}(m_\theta)$, $k_\theta^{(r)}(t)$, and $\Sigma_\theta^{(r)}$. Lemma 16 in Appendix A.2 writes $k_\theta^{(r)}(t)$ and $\Sigma_\theta^{(r)}$ from the covariance kernel $K_\theta(\cdot,\cdot)$. The two-time extension of eq. (15) gives the conditional covariance kernel needed for path sampling. Finally, writing $S := L^{(r)}(X)$, Bayes' rule gives $p(\theta \mid S) \propto p(S \mid \theta)\pi(\theta)$, where $S \mid \theta \sim \mathcal{N}\big(L^{(r)}(m_\theta), \Sigma_\theta^{(r)}\big)$.

6 Experiment

6.1 Experimental setup

Datasets. We use three Gaussian process families—log-GBM, drifted fBM, and OU—including two-dimensional correlated log-GBM, as controlled benchmarks. We evaluate behavior beyond parametric models using cumulative log returns of the S&P 500 and paired foreign-exchange series (GBP/USD, JPY/USD).

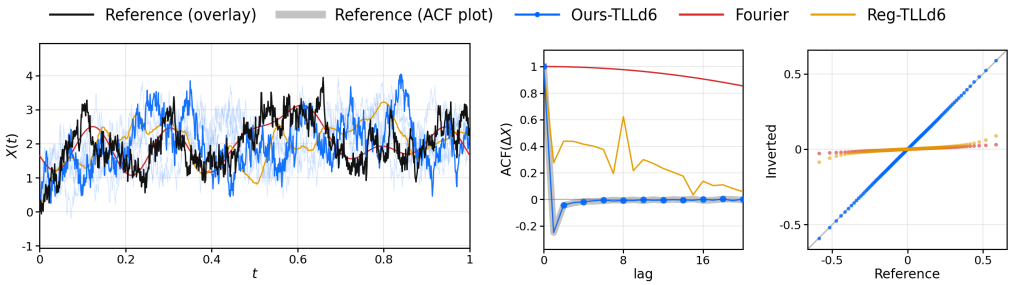
Signature inversion methods. We train a flow-matching model with a Diffusion Transformer [Peebles and Xie, 2023], using linear-statistic (LS), time-augmented (TA), or time lead-lag (TLL) conditioning (see Appendix E.2 for conditional diffusion experiments). Univariate Gaussian-process experiments use depths $r = 1, \ldots, 6$. For 2D log-GBM, we evaluate LS and TA at $r = 2, 4, 6$. S&P 500 and foreign-exchange experiments use TA and TLL at $r = 4$. Deterministic baselines include regression models trained with the ℓ_2 loss using the same conditioning and backbone, `Fourier` and `Legendre` [Barancikova et al., 2025], and `Insertion` [Fermanian et al., 2024].

Evaluation. For each reference path, inversion generates 30 samples using 1000 explicit Euler steps. We assess theoretical alignment using posterior-sampling reconstruction errors. Comparisons with deterministic baselines assess process-parameter recovery for Gaussian processes, and stylized facts, autocorrelation, and cross-series correlation for financial data. We evaluate signature fidelity and provide qualitative comparisons using path overlays, ACF plots, and QQ plots.

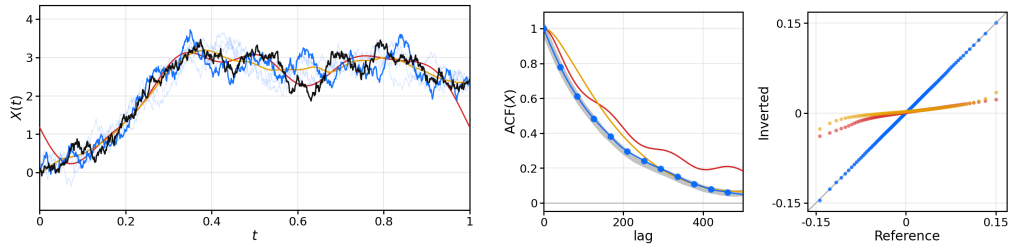
Further experimental details are provided in Appendix B.

Table 1: **Posterior-sampling reconstruction errors.** LS Oracle denotes the LS oracle posterior-sampling reconstruction error from Section 5, computed analytically for log-GBM and estimated numerically for drifted fBM and OU; LS/TA/TLL Emp. are empirical reconstruction-error estimates.

	log-GBM				fBM				OU			
r	LS Oracle	LS Emp.	TA Emp.	TLL Emp.	LS Oracle	LS Emp.	TA Emp.	TLL Emp.	LS Oracle	LS Emp.	TA Emp.	TLL Emp.
1	1.36	1.32	1.32	1.32	1.49	1.49	1.49	1.49	1.16	1.14	1.14	1.14
2	0.54	0.55	0.55	0.55	0.66	0.66	0.66	0.67	0.50	0.51	0.51	0.51
3	0.35	0.35	0.32	0.32	0.45	0.46	0.45	0.44	0.33	0.33	0.31	0.30
4	0.26	0.26	0.20	0.20	0.35	0.36	0.32	0.31	0.25	0.25	0.19	0.20
5	0.21	0.21	0.15	0.15	0.29	0.31	0.25	0.25	0.20	0.20	0.15	0.16
6	0.17	0.17	0.14	0.14	0.25	0.27	0.23	0.25	0.17	0.17	0.14	0.15



(a) fBM ($H = 0.3, \mu = 2.0, \sigma = 2.0$)



(b) OU ($\kappa = 5.0, \mu = 3.0, \sigma = 2.0$)

Figure 2: **Path overlays, ACF plots, and QQ plots for drifted fBM and OU.** ACFs are computed from ΔX for drifted fBM and X for OU. QQ plots are computed from ΔX. Our model uses TLL conditioning at depth $r = 6$. Additional qualitative results are provided in Appendix C.2.

6.2 Theoretical alignment

We empirically validate the theoretical predictions by examining agreement with the LS benchmark and the ordering of posterior-sampling reconstruction errors across conditioning representations. Table 1 shows that under LS conditioning, empirical posterior-sampling reconstruction errors closely track the theory-derived oracle error across process families and truncation depths. The learned sampler also closely matches the numerical LS posterior across most evaluated settings (see Table 9 in Appendix C.1). Although the oracle is derived in continuous time while empirical errors are computed on piecewise-linear paths with $n = 1000$ segments, we expect this discrepancy to be negligible. For log-GBM, where the piecewise-linear oracle is available in closed form, the correction is $O(n^{-1})$ and below 0.002 for all $r \leq 6$ (see Appendix A.6).

Across conditioning representations, empirical errors also decrease with depth and broadly follow the expected LS–TA–TLL ordering, with a few deeper-level TA/TLL reversals that may reflect the harder optimization induced by the larger TLL signature dimension. At depth two for log-GBM, the TLL estimate matches the LS oracle, as predicted by Corollary 14. These findings extend to two-dimensional log-GBM: empirical LS errors agree with the multivariate oracle in Corollary 15, while TA conditioning yields lower reconstruction errors (see Appendix C.3). Furthermore, the truncated signatures of generated paths remain close to the conditioning signatures, indicating consistency with the prescribed conditioning information (see Table 10 in Appendix C.1).

Table 2: **Parameter-estimation errors for Gaussian processes.** For drifted fBM, σ and H errors condition on the other parameter; for OU, σ error conditions on κ, μ. For 2D log-GBM, ρ denotes Brownian correlation. Bold and underlined values indicate the lowest and second-lowest errors in each column, respectively. Our models use depth $r = 6$ for univariate processes and $r = 4$ for 2D log-GBM. Signature depths for each method are given in Appendix B.1.

	log-GBM	fBM		OU	2D log-GBM	
Method	σ err.	σ err.	H err.	σ err.	σ err.	ρ err.
Deterministic						
`Fourier`	1.71	1.69	0.44	1.64	1.75	0.34
`Legendre`	1.76	1.75	0.43	1.74	1.80	0.29
`Insertion`	1.93	1.93	0.46	1.90	–	–
`Reg-TA`	1.73	1.72	0.33	1.72	–	–
`Reg-TLL`	1.69	1.67	0.29	1.69	–	–
Probabilistic						
`Ours-TA`	<u>0.22</u>	<u>0.87</u>	<u>0.05</u>	<u>0.23</u>	<u>0.24</u>	<u>0.24</u>
`Ours-TLL`	**0.06**	**0.10**	**0.01**	**0.06**	**0.07**	**0.08**

6.3 Structural recovery on Gaussian processes

We next evaluate whether probabilistic signature inversion recovers the distributional structure of Gaussian-process paths beyond matching the conditioning signature. Figure 1 (a) and Figure 2 show qualitative comparisons across different path regimes, including smoother paths (fBM with $H = 0.7$), rougher paths (fBM with $H = 0.3$), and strongly mean-reverting paths (OU with $\kappa = 5$). Across these regimes, probabilistic inversions follow the broad shape of the reference path while retaining local fluctuations. Although deterministic baselines also remain close to the conditioning signature (see Table 11 in Appendix C.1), their reconstructions often appear oversmoothed relative to the reference path, missing the characteristic roughness and fluctuation scales of the process. For fBM with $H = 0.3$, our method reproduces the negative short-lag autocorrelation of the reference increments, whereas the two deterministic baselines exhibit positive autocorrelation, failing to capture the antipersistent behavior. For OU, our method closely follows the decay of the reference process-value ACF, whereas the deterministic baselines exhibit slower decay. In both QQ plots, our method closely follows the diagonal, whereas the deterministic baselines compress the reconstructed quantiles toward zero, consistent with their suppressed local fluctuations. Additional qualitative diagnostics are provided in Appendix C.2.

Table 2 adds quantitative parameter-recovery evidence, showing that for univariate processes, TLL gives the lowest estimation errors across all reported parameters, while TA also outperforms the deterministic baselines. This TLL advantage is consistent with time lead-lag augmentation retaining quadratic-variation information, particularly for estimation of σ in log-GBM and OU. TLL also improves estimation of both σ and H for drifted fBM. These findings extend to two-dimensional log-GBM, where probabilistic inversion with TA achieves lower errors in both σ and ρ than the `Fourier` and `Legendre` baselines. Recovery of ρ assesses cross-coordinate dependence in addition to the fluctuation scale of each coordinate.

6.4 Extension to financial data

Figure 1 (b) and Figure 3 show that probabilistic signature inversion extends beyond controlled Gaussian process families to S&P 500 cumulative log-return windows. Probabilistic inversions generate multiple paths around the reference trajectory, following its broad shape while exhibiting diverse local fluctuations. Beyond this qualitative resemblance, the ACF of absolute returns and the QQ plot of returns in Figure 3 show that the generated paths more closely reproduce volatility clustering and the return distribution than deterministic reconstructions. The in-sample results in Table 3 complement these qualitative observations: probabilistic inversions achieve lower median errors than the deterministic baselines in stylized-fact and autocorrelation diagnostics for both S&P 500 and foreign-exchange data, as well as cross-series correlation diagnostics for the latter. Full in-sample and out-of-sample results for both financial datasets are provided in Appendix D.

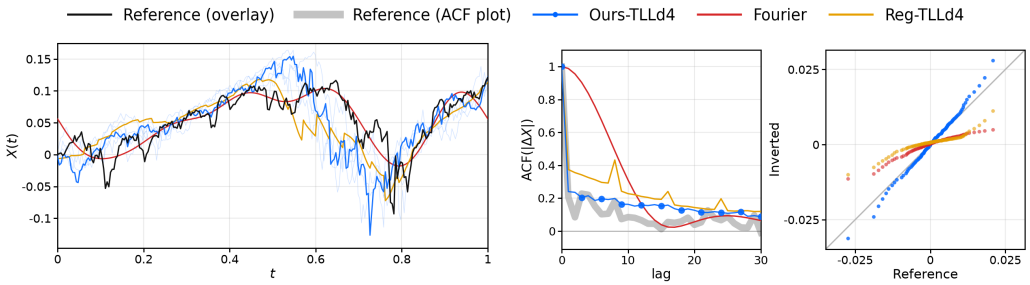


Figure 3: **Path overlay, ACF plot, and QQ plot for out-of-sample S&P 500 data.** The ACF is computed from absolute returns. The QQ plot compares return quantiles between generated and reference paths. Our model uses TLL conditioning at depth $r = 4$. Additional qualitative results are provided in Appendix D.2.

Table 3: **Stylized-fact, autocorrelation, and cross-series correlation errors for financial data (in-sample).** Vol., Skew., and Kurt. denote volatility, skewness, and kurtosis. ACF and ρ denote autocorrelation and cross-series correlation, respectively, for returns r or absolute returns $|r|$. Bold and underlined entries indicate the best and second-best results, respectively, based on unrounded median errors. Signature depths are given in Appendix B.2.

	S&P 500					Foreign exchange				
Method	Vol.	Skew.	Kurt.	ACF(r)	ACF($\lvert r\rvert$)	Vol.	Skew.	Kurt.	ACF($\lvert r\rvert$)	$\rho(\lvert r\rvert)$
Deterministic										
`Fourier`	0.11	0.39	1.89	0.41	0.23	0.07	0.50	2.81	0.35	0.27
`Legendre`	0.12	1.37	5.24	0.31	0.18	0.08	1.38	4.83	0.29	0.22
`Insertion`	0.13	0.50	3.19	0.80	0.50	–	–	–	–	–
`Reg-TAd4`	0.10	1.43	11.33	0.15	0.11	–	–	–	–	–
`Reg-TLLd4`	0.10	1.01	12.04	0.14	0.11	–	–	–	–	–
Probabilistic										
`Ours-TAd4`	<u>0.02</u>	<u>0.30</u>	<u>1.26</u>	<u>0.08</u>	<u>0.08</u>	**0.01**	<u>0.43</u>	<u>2.26</u>	<u>0.07</u>	<u>0.10</u>
`Ours-TLLd4`	**0.02**	**0.29**	**1.19**	**0.08**	**0.08**	<u>0.01</u>	**0.42**	**1.90**	**0.07**	**0.09**

7 Conclusion

We reformulated truncated-signature inversion as a probabilistic problem—learning the conditional distribution of paths given a truncated signature—and quantified the irreducible ambiguity induced by truncation through a posterior-sampling reconstruction error analysis. For Gaussian process families, we derived this error from the tractable posterior under linear-statistics conditioning, obtaining a closed form for log-GBM and numerically tractable expressions for drifted fBM and OU, yielding a concrete theoretical baseline that upper-bounds the corresponding error under richer signature conditionings. Empirically, our signature-conditioned flow matching estimator closely matched the linear-statistics oracle errors in both univariate and bivariate settings and the corresponding posterior distributions across most evaluated univariate settings. It also broadly tracked the expected σ-algebra hierarchy and preserved distributional and temporal structure across Gaussian processes. Experiments on financial data further demonstrated the applicability of probabilistic signature inversion beyond parametric process families.

Beyond inversion itself, the proposed probabilistic signature inversion method may support signature-conditioned scenario generation for financial time series and help interpret the path characteristics captured by individual signature coordinates. Future work could extend the method to conditioning on subsets of signature coordinates to facilitate such interpretation and evaluate its scalability on higher-dimensional paths. On the theoretical side, future work includes deriving tighter bounds or exact characterizations of the reconstruction error under truncated-signature conditioning and extending the analysis beyond Gaussian process families.

Acknowledgments and Disclosure of Funding

This work was partly supported by the Institute of Information & Communications Technology Planning & Evaluation (IITP) grants funded by the Korea government (MSIT) [RS-2022-II220959 (No. 2022-0-00959), (Part 2) Few-Shot Learning of Causal Inference in Vision and Language for Decision Making; RS-2026-25526850, High-Efficiency Neural Networks for Artificial General Intelligence (HERMES-Net); and RS-2021-II211343, Artificial Intelligence Graduate School Program (Seoul National University)].

References

Barbora Barancikova, Zhuoyue Huang, and Cristopher Salvi. Sigdiffusions: Score-based diffusion models for time series via log-signature embeddings. In *ICLR*, 2025.

Horatio Boedihardjo and Xi Geng. The uniqueness of signature problem in the non-markov setting. *Stochastic Processes and Their Applications*, 125(12):4674–4701, 2015.

Horatio Boedihardjo, Xi Geng, Terry Lyons, and Danyu Yang. The signature of a rough path: Uniqueness. *Advances in Mathematics*, 293:720–737, 2016.

Hans Bühler, Blanka Horvath, Terry Lyons, Imanol Perez Arribas, and Ben Wood. A data-driven market simulator for small data environments. In *Modern Topics in Stochastic Analysis and Applications*, pages 273–310, 2024.

Jiawei Chang and Terry J. Lyons. Insertion algorithm for inverting the signature of a path. *arXiv preprint arXiv:1907.08423*, 2019.

Ilya Chevyrev and Andrey Kormilitzin. A primer on the signature method in machine learning. *arXiv preprint arXiv:1603.03788*, 2016.

DLMF. *NIST Digital Library of Mathematical Functions*. `https://dlmf.nist.gov/`, Release 1.2.8 of 2026-09-15. URL `https://dlmf.nist.gov/`. F. W. J. Olver, A. B. Olde Daalhuis, D. W. Lozier, B. I. Schneider, R. F. Boisvert, C. W. Clark, B. R. Miller, B. V. Saunders, H. S. Cohl, and M. A. McClain, eds.

Adeline Fermanian, Jiawei Chang, Terry Lyons, and Gérard Biau. The insertion method to invert the signature of a path. In *Recent Advances in Econometrics and Statistics*, pages 575–595. Springer, 2024.

Guy Flint, Ben Hambly, and Terry Lyons. Discretely sampled signals and the rough Hoff process. *Stochastic Processes and their Applications*, 126(9):2593–2614, 2016.

Xi Geng. Reconstruction for the signature of a rough path. *Proceedings of the London Mathematical Society*, 114(3):495–526, 2017.

Xin Guo, Binnan Wang, Ruixun Zhang, and Chaoyi Zhao. On consistency of signature using Lasso. *Operations Research*, 73(5):2530–2549, 2025.

Lajos Gergely Gyurkó, Terry Lyons, Mark Kontkowski, and Jonathan Field. Extracting information from the signature of a financial data stream. *arXiv preprint arXiv:1307.7244*, 2013.

Ben Hambly and Terry Lyons. Uniqueness for the signature of a path of bounded variation and the reduced path group. *Annals of Mathematics*, 171(1):109–167, 2010.

Jonathan Ho, Ajay N. Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. In *NeurIPS*, 2020.

Patrick Kidger and Terry Lyons. Signatory: Differentiable computations of the signature and logsignature transforms, on both CPU and GPU. In *ICLR*, 2021.

Patrick Kidger, Patric Bonnier, Imanol Perez Arribas, Cristopher Salvi, and Terry Lyons. Deep signature transforms. In *NeurIPS*, 2019.

Wolfram Koepf and Dieter Schmersau. Representations of orthogonal polynomials. *Journal of Computational and Applied Mathematics*, 90(1):57–94, 1998.

Yves Le Jan and Zhongmin Qian. Stratonovich's signatures of Brownian motion determine Brownian sample paths. *Probability Theory and Related Fields*, 157(1–2):209–223, 2013.

Daniel Levin, Terry Lyons, and Hao Ni. Learning from the past, predicting the statistics for the future, learning an evolving system. *arXiv preprint arXiv:1309.0260*, 2013.

Shujian Liao, Hao Ni, Marc Sabate-Vidales, Lukasz Szpruch, Magnus Wiese, and Baoren Xiao. Sig-Wasserstein GANs for conditional time series generation. *Mathematical Finance*, 34(2): 622–670, 2024.

Yaron Lipman, Ricky T. Q. Chen, Heli Ben-Hamu, Maximilian Nickel, and Matthew Le. Flow matching for generative modeling. In *ICLR*, 2023.

Chung I Lu and Julian Sester. Generative modelling of financial time series with structured noise and MMD-based signature learning. *Statistics & Risk Modeling*, 42(3-4):91–122, 2025.

Terry J. Lyons and Weijun Xu. Inverting the signature of a path. *Journal of the European Mathematical Society*, 20(7):1655–1687, 2018.

Terry J. Lyons, Michael Caruana, and Thierry Lévy. *Differential Equations Driven by Rough Paths*, volume 1908 of *Lecture Notes in Mathematics*. Springer, 2007.

James Morrill, Adeline Fermanian, Patrick Kidger, and Terry Lyons. A generalised signature method for multivariate time series feature extraction. *arXiv preprint arXiv:2006.00873*, 2020.

Hao Ni, Lukasz Szpruch, Marc Sabate-Vidales, Baoren Xiao, Magnus Wiese, and Shujian Liao. Sig-Wasserstein GANs for time series generation. In *ICAIF*, 2021.

William Peebles and Saining Xie. Scalable diffusion models with transformers. In *ICCV*, 2023.

Kashif Rasul, Calvin Seward, Ingmar Schuster, and Roland Vollgraf. Autoregressive denoising diffusion models for multivariate probabilistic time series forecasting. In *ICML*, 2021.

Yuki Tanaka, Ryuji Hashimoto, Takehiro Takayanagi, Zhe Piao, Yuri Murayama, and Kiyoshi Izumi. CoFinDiff: Controllable financial diffusion model for time series generation. In *IJCAI*, 2025.

Yusuke Tashiro, Jiaming Song, Yang Song, and Stefano Ermon. CSDI: Conditional score-based diffusion models for probabilistic time series imputation. In *NeurIPS*, 2021.

Jinsung Yoon, Daniel Jarrett, and Mihaela van der Schaar. Time-series generative adversarial networks. In *NeurIPS*, 2019.

Xinyu Yuan and Yan Qiao. Diffusion-TS: Interpretable diffusion for general time series generation. In *ICLR*, 2024.

A Proofs and auxiliary derivations

A.1 Proof of Proposition 8

Proposition 8 (Within–between variance decomposition)**.** For any statistic Φ,

$$\mathcal{E}_\Theta(\Phi) = 2\mathbb{E}_{S\sim\Phi_\#\bar{P}_\Theta}\left[\int_0^T \mathrm{Var}_{Q_\Theta}(X_t \mid S)\,dt\right] \tag{9}$$

$$= 2\mathbb{E}_S\left[\int_0^T \mathbb{E}_{\theta|S}\big[\mathrm{Var}_\theta(X_t \mid S)\big]\,dt + \int_0^T \mathrm{Var}_{\theta|S}\big(\mathbb{E}_\theta[X_t \mid S]\big)\,dt\right]. \tag{10}$$

Proof. Let $S := \Phi(X)$. By definition, conditional on S, X and $\tilde{X}$ are independent draws from $Q_\Theta(\cdot \mid S)$. Thus, for each $t \in [0,T]$,

$$\mathbb{E}\big[(X_t - \tilde{X}_t)^2 \mid S\big] = \mathrm{Var}(X_t \mid S) + \mathrm{Var}(\tilde{X}_t \mid S) + \big(\mathbb{E}[X_t \mid S] - \mathbb{E}[\tilde{X}_t \mid S]\big)^2 \tag{27}$$

$$= 2\mathrm{Var}(X_t \mid S). \tag{28}$$

Integrating over t and taking expectation over S gives eq. (9).

For the second identity, write the mixture conditional distribution through the posterior over parameters. The law of total variance gives

$$\mathrm{Var}_{Q_\Theta}(X_t \mid S) = \mathbb{E}_{\theta|S}\big[\mathrm{Var}_\theta(X_t \mid S)\big] + \mathrm{Var}_{\theta|S}\big(\mathbb{E}_\theta[X_t \mid S]\big). \tag{29}$$

Substituting this identity into eq. (9) yields eq. (10). □

A.2 Kernel formulas and numerical inputs

Table 4 lists the process-specific numerical inputs used in the kernel-based computation.

Table 4: **Kernel formulas and deterministic moments.** The table lists process-specific quantities for the kernel-based analysis of the posterior-sampling reconstruction error: the covariance kernel K_θ and the deterministic moments $\ell_j(m_\theta)$ for $(j \geq 1)$. Here $L^{(r)}(m_\theta)$ is determined by $\ell_0(m_\theta) = m_\theta(T)$ together with these values.

Process	m_θ(t)	G_t^θ	$K_\theta(t,s)$	$\ell_j(m_\theta)$
log-GBM	νt	σB_t	$\sigma^2 \min(s,t)$	$\nu\frac{T^{j+1}}{j+1}$
fBM	νt	σB_t^H	$\frac{\sigma^2}{2}(s^{2H} + t^{2H} - \lvert t-s\rvert^{2H})$	$\nu\frac{T^{j+1}}{j+1}$
OU	$\mu(1-e^{-\kappa t})$	$\sigma\int_0^t e^{-\kappa(t-u)}dB_u$	$\frac{\sigma^2}{2\kappa}\big(e^{-\kappa\lvert t-s\rvert} - e^{-\kappa(t+s)}\big)$	$\mu\left(\frac{T^j}{j} - \int_0^T t^{j-1}e^{-\kappa t}\,dt\right)$

The following lemma explains how these inputs determine $k^{(r)}(t)$ and $\Sigma^{(r)}$ in Proposition 11.

Lemma 16 (Linear-statistic covariance formulas)**.** The quantities $k^{(r)}(t)$ and $\Sigma^{(r)}$ in Proposition 11 are determined by the covariance kernel $K(\cdot,\cdot)$. More precisely,

$$\mathrm{Cov}(G_t, \ell_0(G)) = K(t,T), \tag{30}$$

$$\mathrm{Cov}(G_t, \ell_j(G)) = \int_0^T u^{j-1}K(t,u)\,du, \qquad j \geq 1. \tag{31}$$

Similarly,

$$\mathrm{Cov}(\ell_0(G), \ell_0(G)) = K(T,T), \tag{32}$$

$$\mathrm{Cov}(\ell_0(G), \ell_j(G)) = \int_0^T u^{j-1}K(T,u)\,du, \qquad j \geq 1, \tag{33}$$

$$\mathrm{Cov}(\ell_i(G), \ell_j(G)) = \int_0^T\int_0^T u^{i-1}v^{j-1}K(u,v)\,du\,dv, \qquad i,j \geq 1. \tag{34}$$

Proof. By Fubini,

$$\mathrm{Cov}(G_t, \ell_0(G)) = \mathrm{Cov}(G_t, G_T) = K(t,T), \tag{35}$$

$$\mathrm{Cov}(G_t, \ell_j(G)) = \mathrm{Cov}\left(G_t, \int_0^T u^{j-1} G_u \, du\right) = \int_0^T u^{j-1} K(t,u)\, du. \tag{36}$$

Similarly,

$$\mathrm{Cov}(\ell_0(G), \ell_0(G)) = K(T,T), \tag{37}$$

$$\mathrm{Cov}(\ell_0(G), \ell_j(G)) = \mathrm{Cov}(G_T, \ell_j(G)) = \int_0^T u^{j-1} K(T,u)\, du, \tag{38}$$

$$\mathrm{Cov}(\ell_i(G), \ell_j(G)) = \mathrm{Cov}\left(\int_0^T u^{i-1} G_u \, du, \int_0^T v^{j-1} G_v \, dv\right) \tag{39}$$

$$= \int_0^T \int_0^T u^{i-1} v^{j-1} K(u,v)\, du\, dv. \tag{40}$$

□

A.3 Proof of Lemma 12

Lemma 12 (Gaussian Hilbert-space projection)**.** Let B be a standard Brownian motion and define the Wiener integral $I(f) := \int_0^T f(s)\, dB_s$ for $f \in L^2([0,T])$. Then, I is linear and satisfies the Itô isometry, $\mathbb{E}[I(f)I(g)] = \langle f, g\rangle_{L^2([0,T])}$. For a closed subspace $M \subset L^2([0,T])$ and $f \in L^2([0,T])$,

$$\mathbb{E}\big[I(f) \mid \mathcal{F}(I(g) : g \in M)\big] = I(\Pi_M f), \tag{19}$$

$$\mathrm{Var}\big(I(f) \mid \mathcal{F}(I(g) : g \in M)\big) = \|f - \Pi_M f\|^2_{L^2([0,T])}, \tag{20}$$

where Π_M denotes the orthogonal projection onto M.

Proof. Since I is linear, $I(f) - I(\Pi_M f) = I(f - \Pi_M f)$. For any $g \in M$, the Itô isometry gives

$$\mathbb{E}\left[I(f - \Pi_M f)\, I(g)\right] = \langle f - \Pi_M f,\, g\rangle_{L^2([0,T])} = 0. \tag{41}$$

Hence $I(f - \Pi_M f)$ is uncorrelated with every $I(g)$, $g \in M$. Since these are jointly Gaussian, this implies independence, so $I(\Pi_M f)$ is the conditional expectation. The second identity follows from the Itô isometry. □

A.4 Proof of Proposition 13

We first state an auxiliary residual formula used below.

Lemma 17 (Legendre residual formula)**.** Define

$$a_r := \int_0^1 \|\mathbf{1}_{[0,u]} - \Pi_{\mathcal{H}_r} \mathbf{1}_{[0,u]}\|^2_{L^2([0,1])}\, du, \tag{42}$$

where $\mathcal{H}_r := \mathrm{span}\{1, t, \ldots, t^{r-1}\}$. Then,

$$a_r = \frac{r}{2(2r-1)(2r+1)}. \tag{43}$$

Proof. Let $L_n(x) = P_n(2x-1)$ be the shifted Legendre polynomials and $e_n(x) := \sqrt{2n+1}\, L_n(x)$ the $L^2([0,1])$-orthonormal basis. From the Legendre identity,

$$\frac{d}{dx}\big(L_{n+1}(x) - L_{n-1}(x)\big) = 2(2n+1)\, L_n(x). \tag{44}$$

Integrating from 0 to u,

$$c_n(u) := \int_0^u e_n(x)\, dx = \frac{L_{n+1}(u) - L_{n-1}(u)}{2\sqrt{2n+1}}. \tag{45}$$

Since $\mathbf{1}_{[0,u]} = \sum_{n=0}^{\infty} c_n(u)\, e_n$, the projection residual has squared norm $\sum_{n=r}^{\infty} c_n(u)^2$. Computing

$$\int_0^1 c_n(u)^2\, du = \frac{1}{4(2n+1)} \left(\frac{1}{2n+3} + \frac{1}{2n-1} \right) = \frac{1}{8} \left(\frac{1}{2n-1} - \frac{1}{2n+3} \right), \tag{46}$$

the sum telescopes:

$$a_r = \sum_{n=r}^{\infty} \frac{1}{8} \left(\frac{1}{2n-1} - \frac{1}{2n+3} \right) = \frac{1}{8} \left(\frac{1}{2r-1} + \frac{1}{2r+1} \right) = \frac{r}{2(2r-1)(2r+1)}. \tag{47}$$

□

Proposition 13 (Closed-form oracle posterior-sampling reconstruction error for log-GBM)**.** For log-GBM and $r \geq 1$, there exists a deterministic function $c_r : [0,T] \to \mathbb{R}^r$ such that

$$\mathbb{E}_\theta\big[X_t \mid L^{(r)}(X)\big] = c_r(t)^\top L^{(r)}(X), \qquad t \in [0,T]. \tag{21}$$

In particular, the conditional mean is determined entirely by the observed statistic $L^{(r)}(X)$ and does not depend on $\theta = (\mu, \sigma)$. Thus, the between-parameter conditional-mean variance term in Proposition 8 vanishes. Moreover, the integrated conditional variance is

$$\int_0^T \mathrm{Var}_\theta\big(X_t \mid L^{(r)}(X)\big)\, dt = a_r \sigma^2 T^2, \qquad a_r = \frac{r}{2(2r-1)(2r+1)}. \tag{22}$$

Substituting these identities into Proposition 8 yields the oracle error

$$\mathcal{E}_\Theta(L^{(r)}) = 2\, a_r\, T^2\, \mathbb{E}[\sigma^2]. \tag{23}$$

Proof. By Fubini, each linear statistic component satisfies

$$\ell_j(G) = \int_0^T t^{j-1} \sigma B_t\, dt = \sigma I(f_j), \qquad f_j(t) := \frac{T^j - t^j}{j}, \quad j \geq 1, \tag{48}$$

and $\ell_0(G) = \sigma B_T = \sigma I(f_0)$ with $f_0 \equiv 1$. Hence

$$S := L^{(r)}(X) = \nu\beta + \sigma(I(f_0), \ldots, I(f_{r-1}))^\top, \qquad \beta := (T, \ldots, T^r/r)^\top. \tag{49}$$

Therefore, for a fixed θ and $\mathcal{H}_r := \mathrm{span}\{f_0, \ldots, f_{r-1}\} = \mathrm{span}\{1, t, \ldots, t^{r-1}\}$,

$$\mathcal{F}(S) = \mathcal{F}(I(f) : f \in \mathcal{H}_r). \tag{50}$$

From $X_t = \nu t + \sigma I(\mathbf{1}_{[0,t]})$ and Lemma 12,

$$\begin{aligned}
\mathbb{E}_\theta[X_t \mid S] &= \nu t + \sigma I(\Pi_{\mathcal{H}_r} \mathbf{1}_{[0,t]}) && (51)\\
&= \nu t + \sigma I(c_r(t)^\top (f_0, \ldots, f_{r-1})) && (52)\\
&= \nu t + c_r(t)^\top (S - \nu\beta) && (53)\\
&= c_r(t)^\top S + \nu(t - c_r(t)^\top \beta) && (54)\\
&= c_r(t)^\top S. && (55)
\end{aligned}$$

The identity $c_r(t)^\top \beta = t$ follows from $1 \equiv f_0 \in \mathcal{H}_r$:

$$\begin{aligned}
t &= \langle \mathbf{1}_{[0,t]}, \mathbf{1} \rangle && (56)\\
&= \langle \Pi_{\mathcal{H}_r} \mathbf{1}_{[0,t]}, \mathbf{1} \rangle && (57)\\
&= \langle c_r(t)^\top (f_0, \ldots, f_{r-1}), \mathbf{1} \rangle && (58)\\
&= c_r(t)^\top \big(\langle f_0, \mathbf{1} \rangle, \ldots, \langle f_{r-1}, \mathbf{1} \rangle \big) && (59)\\
&= c_r(t)^\top \beta. && (60)
\end{aligned}$$

The resulting conditional mean is independent of $\theta = (\mu, \sigma)$, so the between-parameter conditional-mean variance term in Proposition 8 vanishes. For the variance, Lemma 12 and Lemma 17 prove

$$\begin{aligned}
\int_0^T \mathrm{Var}(X_t \mid S)\, dt &= \sigma^2 \int_0^T \|\mathbf{1}_{[0,t]} - \Pi_{\mathcal{H}_r} \mathbf{1}_{[0,t]}\|^2\, dt && (61)\\
&= \sigma^2 T^2 \int_0^1 \|\mathbf{1}_{[0,u]} - \Pi_{\mathcal{H}_r} \mathbf{1}_{[0,u]}\|^2\, du && (62)\\
&= a_r \sigma^2 T^2. && (63)
\end{aligned}$$

The second equality uses the change of variables $t = Tu$ and the corresponding identification of polynomial spans. The tower property yields $\mathcal{E}_\Theta(L^{(r)}) = 2 a_r T^2 \mathbb{E}_\pi[\sigma^2]$. □

A.5 Proof of Corollary 15

Corollary 15 (Multivariate log-GBM oracle)**.** Let B be a standard d-dimensional Brownian motion and consider $X_t = \nu t + \Sigma^{1/2} B_t$, where $\nu \in \mathbb{R}^d$ and Σ is a positive semidefinite diffusion covariance matrix. Let $L_d^{(r)}$ apply the linear statistic in Definition 9 to each coordinate, and extend Definition 7 using the squared Euclidean path loss $\int_0^T \|X_t - \tilde{X}_t\|_2^2 dt$, which sums the coordinate-wise squared errors. For $r \geq 1$ and a parameter prior π_Θ with finite second moments, the oracle posterior-sampling reconstruction error is

$$\mathcal{E}_\Theta\big(L_d^{(r)}\big) = \frac{rT^2}{(2r-1)(2r+1)} \mathbb{E}_{\pi_\Theta}\big[\mathrm{tr}(\Sigma)\big]. \tag{64}$$

Proof. The squared Euclidean loss is additive across coordinates, so

$$\mathcal{E}_\Theta\big(L_d^{(r)}\big) = \sum_{j=1}^{d} \mathbb{E}\left[\int_0^T (X_t^{(j)} - \tilde{X}_t^{(j)})^2 dt\right]. \tag{65}$$

Since the same temporal projection is used in every coordinate, each scalar residual is independent of the full LS observation at fixed θ. Thus, the scalar argument in Proposition 13 applies to each term with $\sigma^2 = \Sigma_{jj}$, and the conditional mean remains independent of θ. Summing the resulting scalar errors gives

$$\mathcal{E}_\Theta\big(L_d^{(r)}\big) = 2a_r T^2 \sum_{j=1}^{d} \mathbb{E}_{\pi_\Theta}[\Sigma_{jj}] = 2a_r T^2 \mathbb{E}_{\pi_\Theta}[\mathrm{tr}(\Sigma)]. \tag{66}$$

□

A.6 Piecewise-linear correction to the LS oracle for log-GBM

Proposition 18 (Piecewise-linear LS oracle for log-GBM)**.** Let X be the log-GBM process in eq. (16), with parameter prior π_Θ satisfying $\mathbb{E}[\nu^2 + \sigma^2] < \infty$. For an integer $n \geq 1$, let X^{lin} denote the piecewise-linear interpolation of the observations $(X_{t_i})_{i=0}^n$ at $t_i = iT/n$, as in Definition 3. For $1 \leq r \leq n$, there exists a deterministic function $c_{r,n} : [0,T] \to \mathbb{R}^r$ such that

$$\mathbb{E}_\theta\big[X_t^{\mathrm{lin}} \mid L^{(r)}(X^{\mathrm{lin}})\big] = c_{r,n}(t)^\top L^{(r)}(X^{\mathrm{lin}}), \qquad t \in [0,T]. \tag{67}$$

In particular, the conditional mean is determined entirely by the observed statistic $L^{(r)}(X^{\mathrm{lin}})$ and does not depend on $\theta = (\mu, \sigma)$. Thus, the between-parameter conditional-mean variance term in Proposition 8 vanishes. Moreover, the integrated conditional variance is

$$\int_0^T \mathrm{Var}_\theta\big(X_t^{\mathrm{lin}} \mid L^{(r)}(X^{\mathrm{lin}})\big) dt = a_{r,n}^{\mathrm{PL}} \sigma^2 T^2, \qquad a_{r,n}^{\mathrm{PL}} := a_r \left(1 + \frac{r^2-1}{3n^2}\right) - \frac{1}{6n}, \tag{68}$$

for $a_r = \frac{r}{2(2r-1)(2r+1)}$. Let $\mathcal{E}_{\Theta,n}^{\mathrm{PL}}(L^{(r)})$ denote the oracle posterior-sampling reconstruction error in Definition 7, with both the reconstruction target and the conditioning statistic computed from X^{lin}. Substituting these identities into Proposition 8 yields the oracle error

$$\mathcal{E}_{\Theta,n}^{\mathrm{PL}}\big(L^{(r)}\big) = \mathcal{E}_\Theta\big(L^{(r)}\big) \left(1 + \frac{r^2-1}{3n^2}\right) - \frac{T^2 \mathbb{E}[\sigma^2]}{3n}. \tag{69}$$

For fixed r, the piecewise-linear oracle differs from the continuous-time oracle by $O(n^{-1})$ as $n \to \infty$.

Proof. Fix θ and write $S = L^{(r)}(X^{\mathrm{lin}})$. Let P_n be the $L^2([0,T])$-orthogonal projection onto functions that are constant on each grid cell, and use I, $f_0, \ldots, f_{r-1}$, and β from Appendix A.4. Set $g_p = P_n f_p$ and $V_{r,n} = \mathrm{span}\{g_0, \ldots, g_{r-1}\}$. Piecewise-linear interpolation preserves the deterministic linear drift, so

$$X_t^{\mathrm{lin}} = \nu t + \sigma I\big(P_n \mathbf{1}_{[0,t]}\big), \qquad S = \nu\beta + \sigma\big(I(g_0), \ldots, I(g_{r-1})\big)^\top. \tag{70}$$

Write $\Pi_{V_{r,n}} P_n \mathbf{1}_{[0,t]} = c_{r,n}(t)^\top (g_0, \ldots, g_{r-1})^\top$. The functions g_p and hence the projection coefficients depend only on r, n, T. Since $g_0 = 1 \in V_{r,n}$ and $\beta_p = \langle g_p, 1\rangle$, orthogonal projection preserves

$$c_{r,n}(t)^\top \beta = \big\langle \Pi_{V_{r,n}} P_n \mathbf{1}_{[0,t]}, 1\big\rangle = \big\langle P_n \mathbf{1}_{[0,t]}, 1\big\rangle = t. \tag{71}$$

For $\sigma > 0$, Lemma 12 therefore gives

$$\mathbb{E}_\theta[X_t^{\mathrm{lin}} \mid S] = \nu t + c_{r,n}(t)^\top (S - \nu\beta) = c_{r,n}(t)^\top S. \tag{72}$$

The same identity holds when $\sigma = 0$, since then $X_t^{\mathrm{lin}} = \nu t$ and $S = \nu\beta$. This proves eq. (67) and makes the between-parameter conditional-mean variance zero.

The conditional covariance is unaffected by the deterministic drift. Using Lemma 12 and rescaling to $[0, 1]$ gives

$$\int_0^T \mathrm{Var}_\theta(X_t^{\mathrm{lin}} \mid S)dt = \sigma^2 T^2 \int_0^1 \left\| P_n \mathbf{1}_{[0,u]} - \Pi_{V_{r,n}} P_n \mathbf{1}_{[0,u]} \right\|^2_{L^2([0,1])} du, \tag{73}$$

where P_n and $V_{r,n}$ now refer to the rescaled grid.

To evaluate the projection residual, let Q_k be the degree-k Hahn polynomial with parameters $(0, 0, n-1)$, normalized by $Q_k(0) = 1$. The cellwise-constant functions $s \mapsto Q_k(\lfloor ns \rfloor)$ for $0 \le k < r$ form an orthogonal basis of $V_{r,n}$, since cell averaging preserves polynomial degree in the cell index. These polynomials are orthogonal on $\{0, \dots, n-1\}$ with uniform weights, and their standard norm formula [DLMF, Table 18.19.1] is

$$H_k := \sum_{i=0}^{n-1} Q_k(i)^2 = \frac{(n+k)!(n-k-1)!}{(2k+1)((n-1)!)^2}, \qquad 0 \le k \le n-1. \tag{74}$$

For $0 \le k < n$, these functions form a basis of the cellwise-constant functions and have squared $L^2([0, 1])$ norms H_k/n. Writing $F_k(u) := \int_0^u Q_k(\lfloor ns \rfloor)ds = \langle P_n \mathbf{1}_{[0,u]}, Q_k(\lfloor n\cdot \rfloor)\rangle$, orthogonality gives

$$\int_0^1 \left\| P_n \mathbf{1}_{[0,u]} - \Pi_{V_{r,n}} P_n \mathbf{1}_{[0,u]} \right\|^2 du = \sum_{k=r}^{n-1} E_k, \qquad E_k := \frac{n}{H_k} \int_0^1 F_k(u)^2 du. \tag{75}$$

Thus, only the components of degree $k \ge r$ contribute to the residual.

For $1 \le k \le n-1$, the antidifference identity of Koepf and Schmersau [1998, Corollary 1, p. 65], expressed in this normalization, gives

$$nF_k(i/n) = \sum_{j=0}^{i-1} Q_k(j) = \frac{n+k}{2(2k+1)} Q_{k-1}(i) - \frac{1}{2} Q_k(i) - \frac{n-1-k}{2(2k+1)} Q_{k+1}(i), \tag{76}$$

for $i = 0, \dots, n-1$, with the last term omitted when $k = n-1$. Also, $F_k(0) = F_k(1) = 0$ because Q_k is orthogonal to the constant polynomial.

Since F_k is piecewise linear, integration of its square on each cell gives

$$\begin{aligned} E_k &= \frac{1}{H_k} \sum_{i=0}^{n-1} F_k(i/n)^2 - \frac{1}{6H_k} \sum_{i=0}^{n-1} \big(F_k((i+1)/n) - F_k(i/n)\big)^2 \\ &= \frac{1}{n^2} \left(\frac{n^2 - k^2}{4(2k-1)(2k+1)} + \frac{1}{12} + \frac{n^2 - (k+1)^2}{4(2k+1)(2k+3)} \right) \\ &= \frac{3n^2 - k(k+1)}{6n^2(2k-1)(2k+3)}. \end{aligned} \tag{77}$$

The first equality uses the cellwise identity $\int_0^h (a + (b-a)t/h)^2 dt = h(a^2 + ab + b^2)/3$ and the zero endpoints. The second follows from eqs. (74) and (76), orthogonality, and $F_k((i+1)/n) - F_k(i/n) = Q_k(i)/n$.

Using eq. (68), direct substitution in eq. (77) gives $E_k = a_{k,n}^{\mathrm{PL}} - a_{k+1,n}^{\mathrm{PL}}$ and $a_{n,n}^{\mathrm{PL}} = 0$, so

$$\sum_{k=r}^{n-1} E_k = a_{r,n}^{\mathrm{PL}} = a_r \left(1 + \frac{r^2 - 1}{3n^2}\right) - \frac{1}{6n}. \tag{78}$$

This identity also holds for $r = n$, when the residual is zero and the sum is empty. Multiplying by $\sigma^2 T^2$ in eq. (73) proves eq. (68). Finally, Proposition 8 and the parameter-independent conditional mean give

$$\mathcal{E}_{\Theta,n}^{\mathrm{PL}}\big(L^{(r)}\big) = 2\mathbb{E}_S \mathbb{E}_{\theta|S}\big[a_{r,n}^{\mathrm{PL}} \sigma^2 T^2\big] = 2a_{r,n}^{\mathrm{PL}} T^2 \mathbb{E}[\sigma^2]. \tag{79}$$

Substituting eqs. (23) and (68) yields eq. (69). □

B Implementation details

B.1 Training

Datasets. Table 5 summarizes the training datasets used for the Gaussian-process and financial-data experiments. Both S&P 500 and FX paths are 252-trading-day cumulative log-return windows, multiplied by 10 during training. FX uses the paired GBP/USD and JPY/USD series from the Exchange Rate dataset spanning 1990–2016, obtained from the SigDiffusions repository [Barancikova et al., 2025].

Table 5: **Training dataset configurations.** All paths use $T = 1$. For Gaussian processes, parameters are sampled independently from Θ for each path. For 2D log-GBM, $j = 1, 2$ indexes coordinates and ρ denotes Brownian correlation. Length counts path points, including the initial point.

Type	Dataset	Configuration	Length	# Paths
Gaussian processes	Log-GBM	$\mu \in [1.5, 2.5],\ \sigma \in [1.5, 2.5]$	1001	100k
	fBM	$\mu \in [1.5, 2.5],\ \sigma \in [1.5, 2.5],\ H \in [0.25, 0.75]$	1001	100k
	OU	$\mu = 3,\ \sigma \in [1.5, 2.5],\ \kappa \in [0.5, 5.0]$	1001	100k
	2D log-GBM	$\mu_j, \sigma_j \in [1.5, 2.5], \rho \sim \mathrm{Uniform}(-0.7, 0.7)$	1001	100k
Financial data	S&P 500	2009-05 $\sim$ 2022-12 (Stride 1)	253	3190
	FX (2D)	1990–2016 (first 80%; stride 1)	253	5818

Model architecture. We adapt DiT [Peebles and Xie, 2023] to one-dimensional paths and use it as the flow matching velocity network. Paths are split into patches of size 8, projected to 128-dimensional tokens, processed by 7 Transformer blocks with 8 attention heads and FFN hidden dimension 512, and unpatchified back to the original path resolution. Sequences whose lengths are not divisible by 8 are zero-padded before patchification and cropped back after unpatchification. The time and log-signature embeddings are combined and injected through AdaLN-Zero modulation. For regression ablations, we use the same backbone and append one convolutional output layer.

Optimization. We train with AdamW using the PyTorch default hyperparameters, learning rate 2×10^{-4}, batch size 64, and no learning-rate scheduler. All learned models are trained for 100 epochs on a single A5000 GPU, taking approximately 3 hours per run.

Signature preprocessing. Log-signatures are computed with `iisignature`[1]. LS conditioning uses the TA log-signature coordinates corresponding to the linear statistics in Definition 9. For TLL, the lead-lag construction introduces repeated coordinate paths, which produce duplicate or linearly dependent log-signature coordinates. For example, by the TLL construction in Definition 6, iterated-integral coordinates associated with different words, such as $(1, 2)$ and $(1, 3)$, can evaluate identically or become linearly dependent. We therefore precompute a reduced TLL coordinate set on a fixed reference batch of 1000 randomly generated log-GBM paths by scanning the TLL log-signature coordinates one at a time and greedily retaining only those that add a new independent direction. Replacing this reference batch with drifted fBM or OU paths yields the same retained set. This preprocessing step is not tuned using downstream reconstruction performance; the reference batch is used only to detect algebraic or numerical redundancy in the TLL log-signature coordinates. The selected TLL indices are then fixed and reused for training and inversion. Finally, we apply the level-wise scaling map $x \mapsto \mathrm{sign}(x) \log(1 + k! \, |x|)$ to each retained level-k conditioning coordinate. The factor $k!$ compensates for the factorial decay of level-k signature terms [Morrill et al., 2020], while the signed logarithm is nearly linear for small magnitudes and suppresses only very large coordinates. Table 6 reports the resulting conditioning-vector dimensions.

[1] `https://github.com/bottler/iisignature`

Table 6: **Conditioning-vector dimensions by depth.** Each entry gives the number of conditioning coordinates at depth r. TLL full gives the raw TLL log-signature dimension, and TLL reduced gives the dimension after greedy removal of duplicate or linearly dependent coordinates.

Conditioning	$r=1$	$r=2$	$r=3$	$r=4$	$r=5$	$r=6$
LS	1	2	3	4	5	6
TA	2	3	5	8	14	23
TLL full	3	6	14	32	80	196
TLL reduced	2	4	9	19	43	93

B.2 Evaluation

Evaluation datasets. Table 7 summarizes the evaluation datasets used for the quantitative metrics and qualitative diagnostics defined in Appendix B.3. The S&P 500 and FX in-sample splits are subsampled training-period diagnostics rather than held-out generalization sets. The S&P 500 splits are evaluated on the original cumulative log-return scale.

Table 7: **Evaluation dataset configurations.** Length counts path points, including the initial point. The full FX dataset spans 1990–2016.

Type	Dataset	Configuration	Evaluation	Length	# Paths
Gaussian processes	Log-GBM	$\mu \in [1.5, 2.5],\ \sigma \in [1.5, 2.5]$	Quant.	1001	1000
		$\mu = 2.0,\ \sigma = 1.6$	Qual.		100
		$\mu = 2.0,\ \sigma = 2.0$	Qual.		100
		$\mu = 2.0,\ \sigma = 2.4$	Qual.		100
	fBM	$\mu \in [1.5, 2.5],\ \sigma \in [1.5, 2.5],\ H \in [0.25, 0.75]$	Quant.	1001	1000
		$\mu = 2.0,\ \sigma = 2.0,\ H = 0.30$	Qual.		100
		$\mu = 2.0,\ \sigma = 2.0,\ H = 0.50$	Qual.		100
		$\mu = 2.0,\ \sigma = 2.0,\ H = 0.70$	Qual.		100
	OU	$\mu = 3,\ \sigma \in [1.5, 2.5],\ \kappa \in [0.5, 5.0]$	Quant.	1001	1000
		$\mu = 3,\ \sigma = 2.0,\ \kappa = 0.5$	Qual.		100
		$\mu = 3,\ \sigma = 2.0,\ \kappa = 2.0$	Qual.		100
		$\mu = 3,\ \sigma = 2.0,\ \kappa = 5.0$	Qual.		100
	2D log-GBM	$\mu_j, \sigma_j \in [1.5, 2.5], \rho \sim \mathrm{Uniform}(-0.7, 0.7)$	Quant.	1001	1000
Financial data	S&P 500	In-sample: 2009-05 $\sim$ 2022-12 (Stride 63)	Quant./Qual.	253	51
		Out-of-sample: 2023-01 $\sim$ 2025-12 (Stride 5)	Quant./Qual.		100
	FX (2D)	In-sample: first 80% (stride 63)	Quant.	253	93
		Out-of-sample: last 20% (stride 5)	Quant.		254

Sampling protocol. Sampling uses 1000 explicit Euler steps. For quantitative evaluation, we generate 30 samples for each of 1000 reference paths, which takes less than 1 hour.

Deterministic baselines. Deterministic baselines return one reconstructed path for each target path. Regression ablations and `Insertion` are used only for univariate experiments; `Fourier` and `Legendre` are used for both univariate and multivariate experiments.

- Regression ablations use the same conditioning vector and backbone, but produce a single reconstruction through the supervised ℓ_2 regression objective. We use `Reg-TAd6` and `Reg-TLLd6` for univariate Gaussian-process experiments, and `Reg-TAd4` and `Reg-TLLd4` for S&P 500.

- For univariate experiments, `Fourier` (order $n = 6$) uses the signature of the augmented path $(t, \sin(t), \cos(t) - 1, x(t))$ up to depth 8; an order-n expansion requires depth $n + 2$. For multivariate experiments, we use order $n = 4$. Our implementation follows the public SigDiffusions codebase[2] [Barancikova et al., 2025].
- For univariate experiments, `Legendre` (order $n = 10$) uses the signature of the time-augmented path $(t, x(t))$ up to depth 12; an order-n expansion requires depth $n + 2$. For multivariate experiments, we use order $n = 8$. We implement this baseline directly from Barancikova et al. [2025].
- `Insertion` ($n = 12$ piecewise-linear pieces) uses the signature of the time-augmented path $(t, x(t))$ up to depth 12; a path with n piecewise-linear pieces requires depth n. Our implementation adapts the public Signatory codebase[3] [Kidger and Lyons, 2021].

B.3 Evaluation metrics

Let X_i denote the i-th observed test path and $\tilde{X}_i$ an inverted path associated with X_i; without the sample index, $\tilde{X}_i$ denotes one sampled inversion for probabilistic methods and the single reconstruction for deterministic methods. When all probabilistic samples are used, $\tilde{X}_{i,j}$ denotes the j-th sample, $j = 1, \dots, M$, with $M = 30$. Let Φ_r denote the conditioning statistic used by a method at depth r, and let $\|\cdot\|_2^2$ denote the discrete path norm used to approximate the integrated squared error in eq. (8).

Posterior-sampling reconstruction and conditioning diagnostics. Motivated by the posterior-sampling reconstruction-error analysis in Section 5, these diagnostics report empirical reconstruction error of the learned sampler against the oracle posterior-sampling reconstruction error, measure inter-sample variability, and verify consistency with the conditioning statistic.

- **Empirical reconstruction error.** We report the empirical reconstruction error

$$\frac{1}{NM} \sum_{i=1}^{N} \sum_{j=1}^{M} \|X_i - \tilde{X}_{i,j}\|_2^2.$$

  For LS conditioning, this is compared with the oracle posterior-sampling reconstruction error $\mathcal{E}_\Theta(L^{(r)})$.

- **Inter-sample spread ratio.** We define the inter-sample spread as

$$\frac{1}{N\binom{M}{2}} \sum_{i=1}^{N} \sum_{1 \le j < k \le M} \|\tilde{X}_{i,j} - \tilde{X}_{i,k}\|_2^2.$$

  The reported spread ratio divides this quantity by the corresponding empirical reconstruction error. A ratio near 1 provides evidence that the sampled inversions are not strongly contracted relative to the empirical reconstruction-error scale.

- **Conditioning consistency.** For each $\tilde{X}_{i,j}$ generated conditional on $\Phi_r(X_i)$, we report the median relative MSE of the conditioning statistic,

$$\underset{i,j}{\text{median}} \frac{\|\Phi_r(\tilde{X}_{i,j}) - \Phi_r(X_i)\|_2^2}{\|\Phi_r(X_i)\|_2^2} \times 100\%.$$

Process-structure diagnostics. These metrics evaluate whether the generated paths preserve process-level structure for Gaussian process families.

- **Parameter estimation error.** We apply process-specific estimators to the inverted paths and compare them with the ground-truth parameters of the conditioning paths. The σ estimators for log-GBM and drifted fBM with known H are Gaussian MLEs; for OU, σ is estimated by a Gaussian transition-residual MLE with κ and μ fixed to their known values. The

[2] https://github.com/BarbOra/SigDiffusions
[3] https://github.com/patrick-kidger/signatory

H estimator for drifted fBM with known σ instead uses a log-least-squares fit based on second-difference variance scaling. We report the mean absolute error (MAE)

$$\frac{1}{N}\sum_{i=1}^{N}\left|\widehat{\theta}_q(\tilde{X}_i)-\theta_{i,q}\right|.$$

- **Signature fidelity.** We measure whether the inverted path matches the reference path at evaluation depth r, using time-augmented signatures; we use $r=6$ for Gaussian-process diagnostics and $r=4$ for financial-data diagnostics.

$$\underset{i}{\mathrm{median}}\,\frac{\|S_{\mathrm{TA}}^{\leq r}(\tilde{X}_i)-S_{\mathrm{TA}}^{\leq r}(X_i)\|_2^2}{\|S_{\mathrm{TA}}^{\leq r}(X_i)\|_2^2}\times 100\%.$$

Stylized-fact, autocorrelation, and cross-series correlation errors. These metrics evaluate whether generated financial time series preserve the stylized facts, autocorrelation, and cross-series correlation of the reference paths. Let ΔX_i denote the daily log-return sequence of path window X_i, with returns computed separately for each coordinate in multivariate data. We report median absolute errors between generated and empirical windows, together with 10th–90th percentile ranges.

- **Realized volatility.** We compute realized volatility as $\|\Delta X_i\|_2$.
- **Skewness.** We compute the adjusted Fisher–Pearson skewness of ΔX_i.
- **Kurtosis.** We compute the bias-corrected Fisher excess kurtosis of ΔX_i.
- **Autocorrelation.** We compute the autocorrelation of ΔX_i and $|\Delta X_i|$ and average the absolute differences from the corresponding reference autocorrelations over lags 1–30.
- **Cross-series correlation.** For multivariate paths, we compute pairwise correlations between coordinates of ΔX_i and between coordinates of $|\Delta X_i|$.

Qualitative diagnostics. These diagnostics visualize distributional behavior that is difficult to summarize with a single scalar metric.

- **Path overlays.** We overlay inverted paths with the reference path to inspect sample diversity and path-level fidelity.
- **ACF plots.** For Gaussian processes, we plot the ACF of ΔX_i for log-GBM and drifted fBM, and the ACF of X_i for OU. For financial data, we plot the ACF of $|\Delta X_i|$ to visualize volatility clustering.
- **QQ plots.** We compare empirical quantiles of generated and reference samples to assess marginal distributional fit.

C Additional Gaussian process results

C.1 Quantitative results for univariate Gaussian processes

Theoretical-alignment diagnostics. Table 8 expands Table 1 by reporting full-sample empirical reconstruction-error estimates with normal-approximation 95% bootstrap half-widths, together with the inter-sample spread ratio (sample-to-sample to sample-to-reference MSE). Table 9 complements these reconstruction-error and spread diagnostics with a direct distributional comparison between generated samples and the numerical LS posterior using squared energy distance.

Overall, the LS empirical means stay close to the LS Oracle within small bootstrap uncertainty, while TA and TLL usually reduce path MSE under richer conditioning. The 95% bootstrap confidence intervals for the squared energy distance gaps contain zero in 8 of 9 process–depth settings, supporting agreement with the numerical LS posterior across most evaluated settings. The spread ratios fall within $[0.96, 1.01]$ across all settings, indicating that sample-to-sample MSE is close to sample-to-reference MSE.

Table 8: **Posterior-sampling reconstruction errors and spread ratios.** LS Oracle is analytic for log-GBM; for drifted fBM and OU, it reports the mean with 95% confidence intervals based on the t-distribution over 10 numerical seeds. LS, TA, and TLL report empirical reconstruction-error means with normal-approximation 95% bootstrap confidence intervals. Spread ratio is the ratio of sample-to-sample MSE to sample-to-reference MSE.

		Posterior-sampling reconstruction error				Spread ratio		
Process	r	LS Oracle	LS	TA	TLL	LS	TA	TLL
log-GBM	1	1.364	1.317±0.040	1.317±0.040	1.317±0.040	0.98	0.98	0.98
	2	0.545	0.547±0.014	0.547±0.014	0.553±0.016	0.98	0.98	0.99
	3	0.351	0.348±0.007	0.321±0.010	0.318±0.011	1.00	1.01	1.01
	4	0.260	0.261±0.005	0.197±0.006	0.197±0.006	0.99	1.00	1.01
	5	0.207	0.208±0.004	0.150±0.004	0.155±0.004	0.99	1.01	1.01
	6	0.172	0.173±0.003	0.135±0.003	0.144±0.004	1.01	1.00	1.01
fBM	1	1.486±0.001	1.488±0.050	1.488±0.050	1.488±0.050	0.99	0.99	0.99
	2	0.656±0.002	0.664±0.022	0.664±0.022	0.674±0.031	0.96	0.96	0.99
	3	0.450±0.002	0.456±0.017	0.454±0.025	0.442±0.026	0.96	0.98	0.99
	4	0.352±0.001	0.365±0.015	0.315±0.019	0.314±0.020	0.98	0.98	0.99
	5	0.291±0.002	0.307±0.013	0.252±0.016	0.255±0.016	0.98	1.00	1.00
	6	0.253±0.002	0.265±0.012	0.225±0.014	0.248±0.016	0.97	1.00	1.00
OU	1	1.163±0.006	1.143±0.036	1.143±0.036	1.143±0.036	0.98	0.98	0.98
	2	0.502±0.001	0.506±0.012	0.506±0.012	0.509±0.014	0.99	0.99	0.99
	3	0.334±0.001	0.333±0.007	0.305±0.011	0.301±0.011	1.00	1.00	1.00
	4	0.252±0.001	0.254±0.005	0.194±0.006	0.197±0.006	0.99	0.99	1.01
	5	0.202±0.001	0.204±0.004	0.150±0.004	0.155±0.004	0.99	1.00	1.01
	6	0.169±0.000	0.171±0.003	0.135±0.003	0.145±0.004	1.01	0.99	1.00

Table 9: **Squared energy distance gaps under LS conditioning.** We compare the squared energy distance between samples from our model and the numerical LS posterior against that between two independent sample sets from the same posterior. Entries report 95% bootstrap confidence intervals for the first distance minus the second, using 1,000 inputs and 30 samples per distribution.

r	log-GBM	fBM	OU
2	$[-9.2\text{e}{-4}, 1.2\text{e}{-3}]$	$[-1.1\text{e}{-3}, 9.4\text{e}{-4}]$	$[-6.6\text{e}{-4}, 1.4\text{e}{-3}]$
4	$[-4.9\text{e}{-4}, 5.8\text{e}{-4}]$	$[-2.0\text{e}{-4}, 8.8\text{e}{-4}]$	$[-5.0\text{e}{-4}, 5.7\text{e}{-4}]$
6	$[-4.1\text{e}{-4}, 3.4\text{e}{-4}]$	$[2.1\text{e}{-4}, 1.0\text{e}{-3}]$	$[-2.9\text{e}{-4}, 4.7\text{e}{-4}]$

Conditioning consistency. Table 10 shows that conditioning errors are generally small, although they tend to increase with depth under TLL conditioning.

Table 10: **Conditioning consistency.** Entries report the median relative MSE between the input conditioning statistic and the statistic recomputed from each generated path, expressed as percentages.

	log-GBM			fBM			OU		
r	LS	TA	TLL	LS	TA	TLL	LS	TA	TLL
1	0.03%	0.03%	0.03%	0.04%	0.04%	0.04%	0.01%	0.01%	0.01%
2	0.05%	0.05%	0.04%	0.06%	0.06%	0.05%	0.01%	0.01%	0.03%
3	0.04%	0.04%	0.24%	0.05%	0.05%	0.42%	0.01%	0.02%	0.12%
4	0.04%	0.13%	0.44%	0.07%	0.16%	0.72%	0.01%	0.09%	0.41%
5	0.04%	0.33%	1.75%	0.05%	0.34%	3.12%	0.01%	0.21%	1.40%
6	0.04%	0.72%	2.76%	0.05%	0.77%	4.85%	0.01%	0.32%	2.28%

Signature fidelity. Table 11 shows that the probabilistic methods achieve low signature errors, indicating that their generated samples remain close to the conditioning path in signature space.

Table 11: **Signature fidelity on Gaussian processes.** Entries report relative MSE of truncated time-augmented signatures up to depth 6, expressed as percentages.

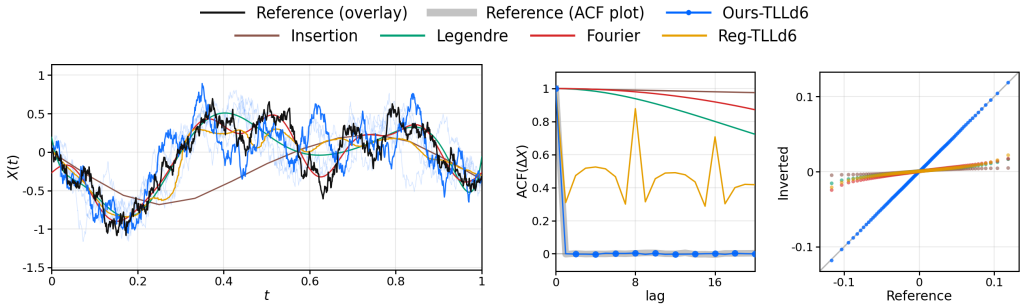
Method	log-GBM	fBM	OU
Deterministic			
`Fourier`	60.08%	57.85%	82.65%
`Legendre`	5.42%	6.27%	3.65%
`Insertion`	0.99%	1.01%	1.13%
`Reg-TAd6`	0.11%	0.15%	0.05%
`Reg-TLLd6`	0.18%	0.25%	0.08%
Probabilistic			
`Ours-TAd6`	0.03%	0.03%	0.01%
`Ours-TLLd6`	0.05%	0.06%	0.02%

The nonzero signature errors of deterministic baselines in Table 11 can be understood from their reconstruction mechanisms.

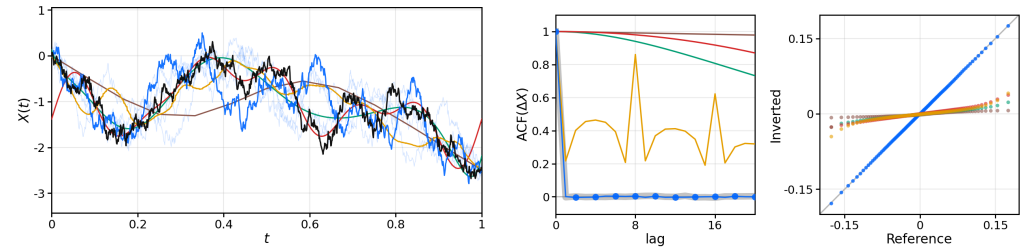
- **Fourier and Legendre** reconstruct finite-dimensional basis projections. For example, Fourier guarantees $R_F^N\left(S_F^{\leq N+2}(X)\right) = P_F^N(X)$, where R_F^N is the inversion map and P_F^N is the Fourier projection. Since $P_F^N(X) \neq X$ in general, their truncated signatures need not match. Empirically, this is consistent with the visibly different endpoint increment $X_T - X_0$ of the Fourier reconstruction in Figure 1.
- **Insertion** estimates local slopes using only the two highest signature levels without explicitly enforcing agreement at the lower levels.

C.2 Qualitative results across univariate Gaussian processes

This subsection reports path overlays, ACF plots, and QQ plots for the two most extreme settings of each Gaussian process under TLLd6 and TAd6 conditioning. Probabilistic inversions better preserve the characteristic roughness and fluctuation scales of the reference paths, whereas deterministic reconstructions often appear oversmoothed. The ACF and QQ plots generally support this pattern, with probabilistic inversions more closely matching the reference autocorrelation and marginal distributions. The periodic spikes in the regression baselines' ACF plots occur at multiples of the DiT patch size (8) and reflect patch-boundary discontinuities in the reconstructed paths.

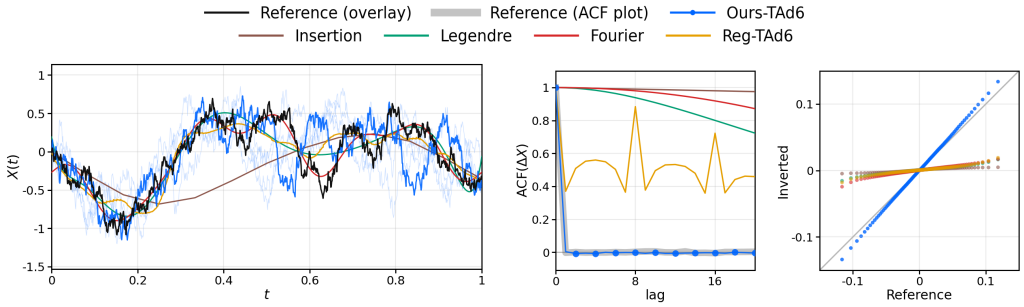


(a) log-GBM ($\sigma = 1.6, \mu = 2.0$)

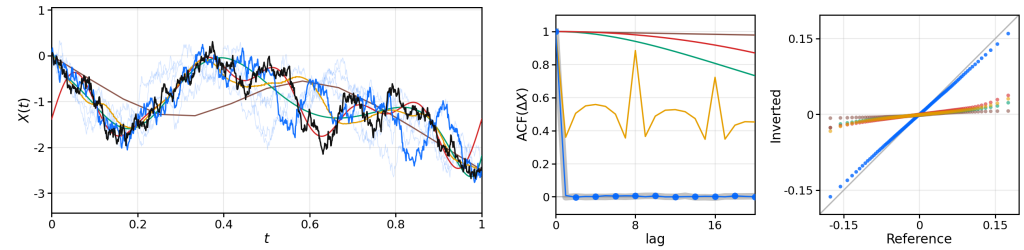


(b) log-GBM ($\sigma = 2.4, \mu = 2.0$)

Figure 4: **Path overlays, ACF plots, and QQ plots for log-GBM with TLL6 conditioning.** ACFs are computed from increments ΔX.

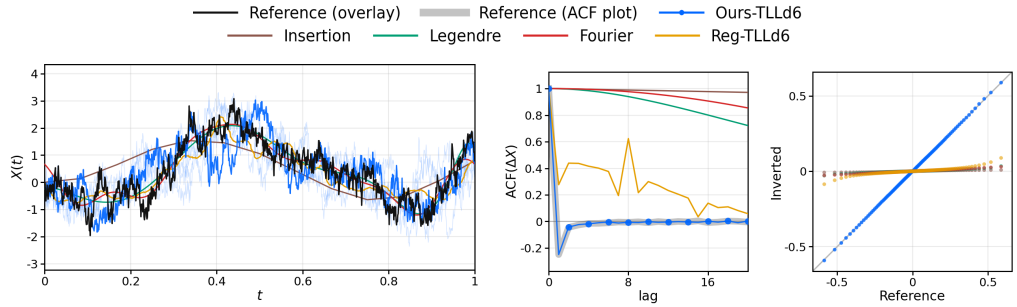


(a) log-GBM ($\sigma = 1.6, \mu = 2.0$)

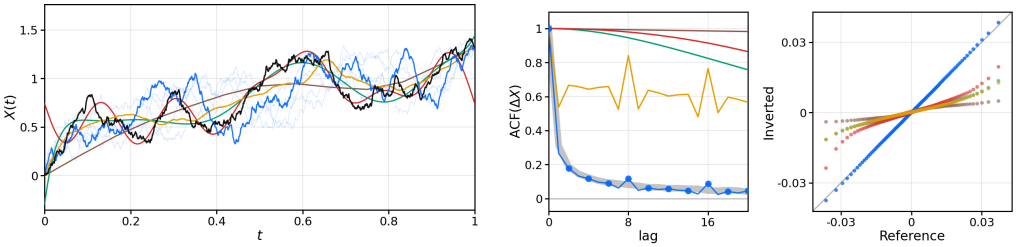


(b) log-GBM ($\sigma = 2.4, \mu = 2.0$)

Figure 5: **Path overlays, ACF plots, and QQ plots for log-GBM with TA6 conditioning.** ACFs are computed from increments ΔX.

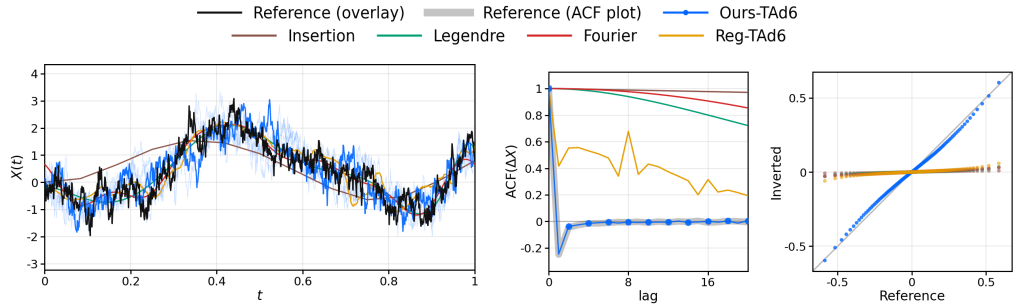


(a) fBM ($H = 0.3, \mu = 2.0, \sigma = 2.0$)

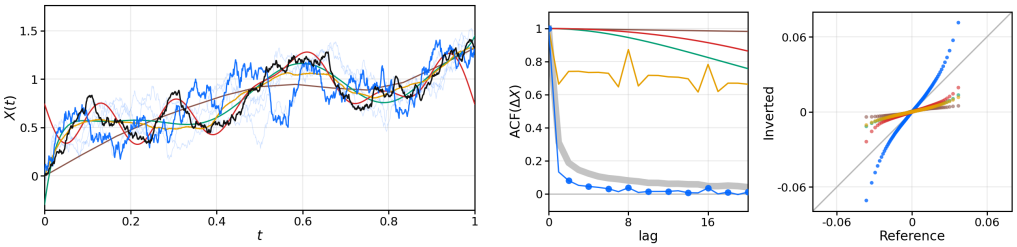


(b) fBM ($H = 0.7, \mu = 2.0, \sigma = 2.0$)

Figure 6: **Path overlays, ACF plots, and QQ plots for drifted fBM with TLL6 conditioning.** ACFs are computed from increments ΔX.

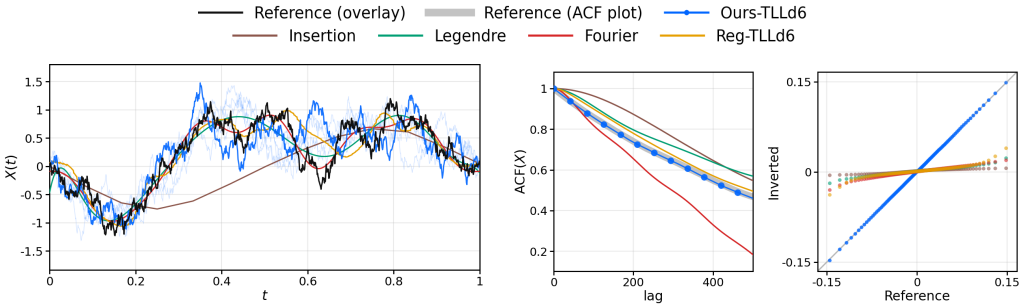


(a) fBM ($H = 0.3, \mu = 2.0, \sigma = 2.0$)

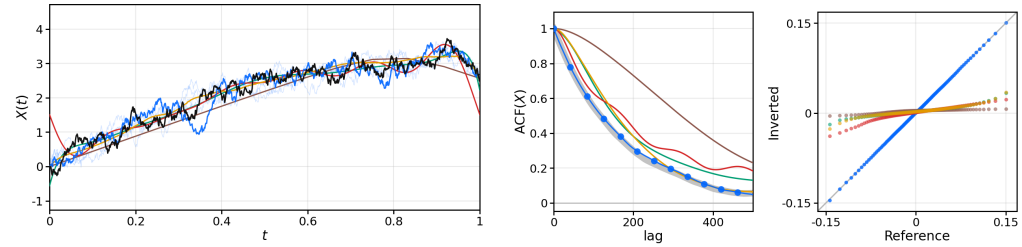


(b) fBM ($H = 0.7, \mu = 2.0, \sigma = 2.0$)

Figure 7: **Path overlays, ACF plots, and QQ plots for drifted fBM with TA6 conditioning.** ACFs are computed from increments ΔX.

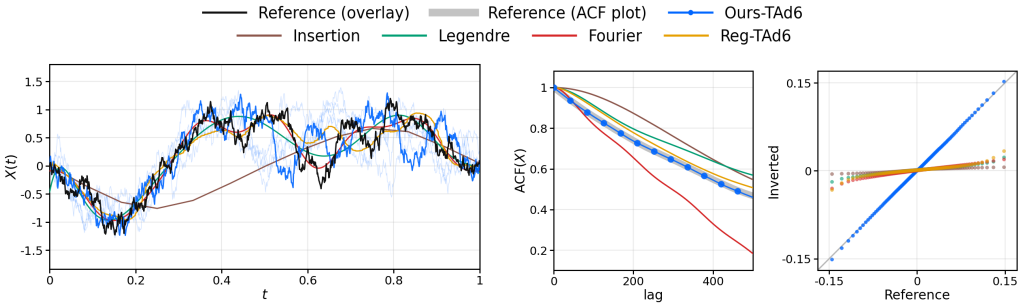


(a) OU ($\kappa = 0.5, \mu = 3.0, \sigma = 2.0$)

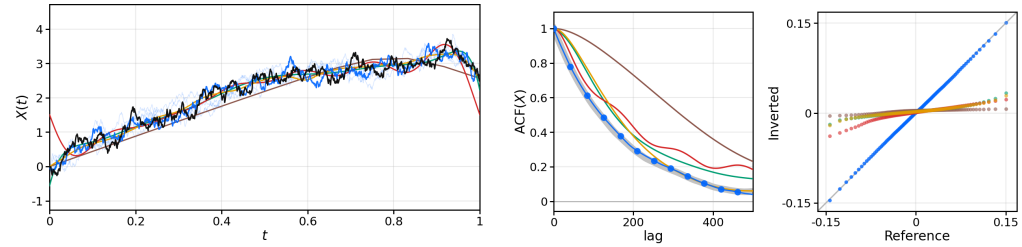


(b) OU ($\kappa = 5.0, \mu = 3.0, \sigma = 2.0$)

Figure 8: **Path overlays, ACF plots, and QQ plots for OU with TLL6 conditioning.** ACFs are computed from process values X.



(a) OU ($\kappa = 0.5, \mu = 3.0, \sigma = 2.0$)



(b) OU ($\kappa = 5.0, \mu = 3.0, \sigma = 2.0$)

Figure 9: **Path overlays, ACF plots, and QQ plots for OU with TA6 conditioning.** ACFs are computed from process values X.

C.3 Quantitative results for two-dimensional log-GBM

We compare LS and TA conditioning at depths $r = 2, 4, 6$, and evaluate `Ours-TAd4` against `Fourier` (order $n = 4$) and `Legendre` (order $n = 8$).

Across depths, the LS empirical reconstruction error closely matches the analytic oracle, while TA yields lower reconstruction errors (Table 12).

Table 12: **Posterior-sampling reconstruction errors for two-dimensional correlated log-GBM.** LS Oracle denotes the componentwise LS oracle in Corollary 15; LS Emp. and TA Emp. denote empirical errors under LS and TA conditioning.

r	LS Oracle	LS Emp.	TA Emp.
2	1.09	1.08	1.02
4	0.52	0.52	0.38
6	0.34	0.34	0.33

D Additional financial-data results

D.1 Quantitative results for S&P 500

Stylized-fact and autocorrelation errors. Tables 13 and 14 report stylized-fact and autocorrelation errors on the S&P 500 splits. Probabilistic inversions generally achieve lower median errors than the deterministic baselines across these metrics, indicating better recovery of the distributional and temporal characteristics of the reference paths. For reference, Table 15 reports the corresponding S&P 500 stylized facts.

Table 13: **In-sample S&P 500 stylized-fact and autocorrelation errors.** Entries report median absolute errors with 10th–90th percentile ranges in brackets. Bold entries indicate the lowest unrounded median error.

	Stylized-fact			Autocorrelation	
Method	Volatility	Skewness	Kurtosis	ACF(r)	ACF($\lvert r\rvert$)
Deterministic					
`Fourier`	0.11 [0.06, 0.19]	0.39 [0.07, 1.39]	1.89 [0.42, 4.74]	0.41 [0.34, 0.47]	0.23 [0.20, 0.29]
`Legendre`	0.12 [0.08, 0.20]	1.37 [0.19, 3.68]	5.24 [1.01, 20.59]	0.31 [0.19, 0.47]	0.18 [0.11, 0.38]
`Insertion`	0.13 [0.10, 0.21]	0.50 [0.17, 1.04]	3.19 [1.13, 7.31]	0.80 [0.66, 0.87]	0.50 [0.32, 0.75]
`Reg-TAd4`	0.10 [0.08, 0.16]	1.43 [0.22, 2.61]	11.33 [2.69, 32.69]	0.15 [0.11, 0.22]	0.11 [0.08, 0.15]
`Reg-TLLd4`	0.10 [0.07, 0.16]	1.01 [0.15, 2.64]	12.04 [3.74, 22.67]	0.14 [0.11, 0.23]	0.11 [0.08, 0.15]
Probabilistic					
`Ours-TAd4`	0.02 [0.00, 0.05]	0.30 [0.06, 0.78]	1.26 [0.19, 5.29]	0.08 [0.06, 0.10]	0.08 [0.06, 0.10]
`Ours-TLLd4`	**0.02 [0.00, 0.05]**	**0.29 [0.08, 0.75]**	**1.19 [0.29, 3.87]**	**0.08 [0.06, 0.10]**	**0.08 [0.06, 0.10]**

Table 14: **Out-of-sample S&P 500 stylized-fact and autocorrelation errors.** Entries report median absolute errors with 10th–90th percentile ranges in brackets. Bold entries indicate the lowest unrounded median error.

	Stylized-fact			Autocorrelation	
Method	Volatility	Skewness	Kurtosis	ACF(r)	ACF($\lvert r\rvert$)
Deterministic					
`Fourier`	0.08 [0.06, 0.16]	1.35 [0.95, 1.51]	2.37 [1.09, 17.36]	0.36 [0.34, 0.43]	0.24 [0.20, 0.27]
`Legendre`	0.10 [0.08, 0.16]	1.25 [0.26, 3.26]	9.61 [0.85, 20.02]	0.27 [0.18, 0.48]	0.19 [0.14, 0.32]
`Insertion`	0.11 [0.10, 0.17]	0.94 [0.19, 1.28]	2.77 [0.45, 18.47]	0.80 [0.61, 0.87]	0.57 [0.39, 0.78]
`Reg-TAd4`	0.09 [0.07, 0.13]	1.47 [0.32, 3.75]	11.58 [2.56, 29.35]	0.12 [0.09, 0.15]	0.11 [0.08, 0.15]
`Reg-TLLd4`	0.09 [0.08, 0.12]	1.11 [0.17, 2.83]	11.29 [3.15, 19.91]	0.10 [0.08, 0.14]	0.10 [0.08, 0.13]
Probabilistic					
`Ours-TAd4`	0.02 [0.01, 0.05]	**0.45 [0.06, 1.19]**	**2.35 [0.41, 14.25]**	**0.08 [0.06, 0.09]**	**0.08 [0.06, 0.10]**
`Ours-TLLd4`	**0.02 [0.00, 0.05]**	0.52 [0.11, 1.44]	2.42 [0.49, 13.10]	0.08 [0.06, 0.09]	0.08 [0.06, 0.10]

Table 15: **In-sample and out-of-sample S&P 500 reference stylized facts.** Statistics are computed over evaluation windows. Entries report medians with 10th–90th percentile ranges in brackets.

Split	Volatility	Skewness	Kurtosis
In-sample (2009–2022)	0.15 [0.11, 0.23]	-0.35 [-0.92, -0.13]	2.31 [0.81, 6.68]
Out-of-sample (2023–2025)	0.13 [0.12, 0.19]	-0.08 [-0.55, 0.69]	1.60 [-0.08, 17.09]

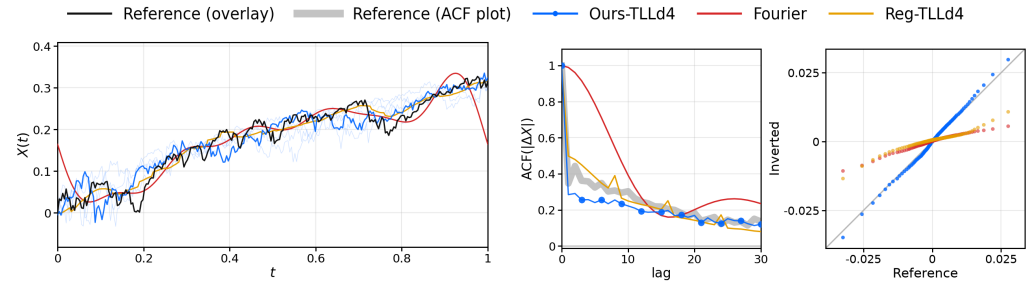
Signature fidelity. Table 16 shows that the learned methods achieve low signature errors in both in-sample and out-of-sample periods, indicating that their outputs remain close to the conditioning path in signature space on financial data.

Table 16: **Signature fidelity on in-sample and out-of-sample S&P 500.** Entries report relative MSE of truncated time-augmented signatures evaluated up to depth 4, expressed as percentages.

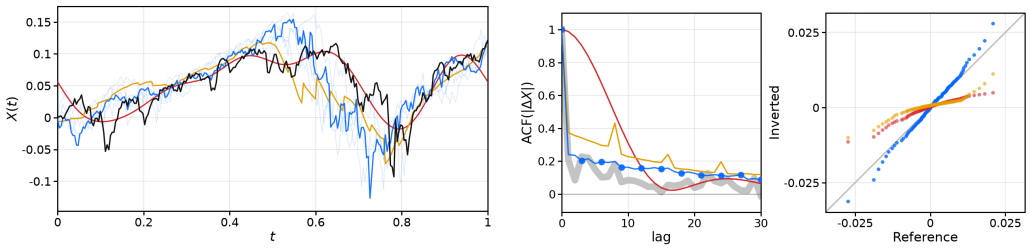
Method	In-sample (2009–2022)	Out-of-sample (2023–2025)
Deterministic		
`Fourier`	1.71%	3.71%
`Legendre`	0.08%	0.10%
`Insertion`	0.03%	0.03%
`Reg-TAd4`	<0.01%	<0.01%
`Reg-TLLd4`	<0.01%	0.03%
Probabilistic		
`Ours-TAd4`	<0.01%	0.01%
`Ours-TLLd4`	0.01%	0.02%

D.2 Qualitative results on S&P 500

This subsection reports path overlays, absolute-return ACF plots, and QQ plots for S&P 500 in-sample and out-of-sample examples under TLLd4 and TAd4 conditioning. In these examples, probabilistic inversions better reproduce local fluctuations, whereas deterministic reconstructions tend to be smoother. The absolute-return ACF and QQ plots generally support this pattern, with probabilistic inversions more closely matching the reference volatility clustering and marginal behavior.

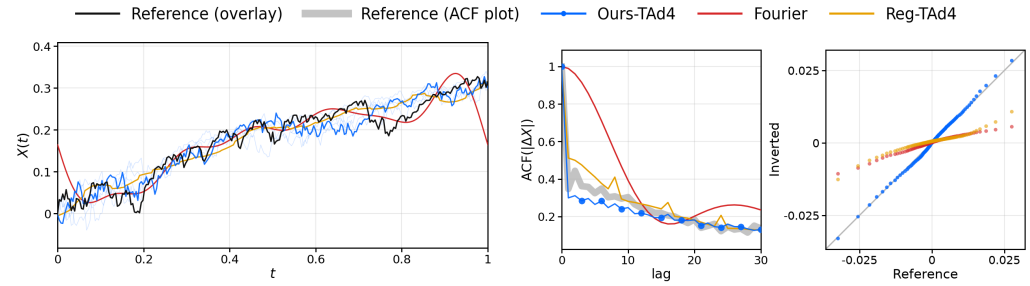


(a) S&P 500 (in-sample)

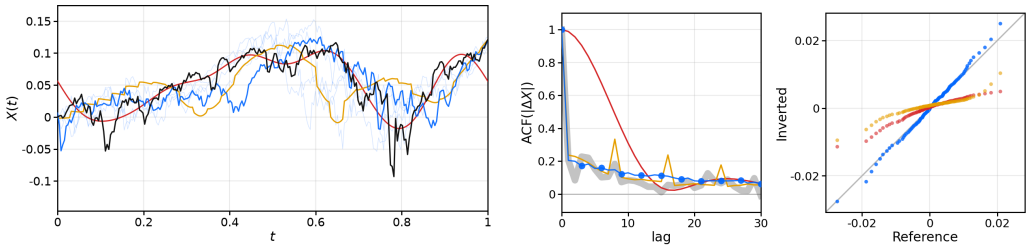


(b) S&P 500 (out-of-sample)

Figure 10: **Path overlays, ACF plots, and QQ plots for in-sample and out-of-sample S&P 500 with TLL4 conditioning.** ACFs are computed from absolute returns.



(a) S&P 500 (in-sample)



(b) S&P 500 (out-of-sample)

Figure 11: **Path overlays, ACF plots, and QQ plots for in-sample and out-of-sample S&P 500 with TA4 conditioning.** ACFs are computed from absolute returns.

D.3 Quantitative results for two-dimensional foreign-exchange data

We compare `Ours-TAd4` and `Ours-TLLd4` with `Fourier` (order $n = 4$) and `Legendre` (order $n = 8$) on paired GBP/USD and JPY/USD paths. Both conditional samplers achieve lower median errors than the deterministic baselines on all seven in-sample diagnostics (Table 17) and six of the seven out-of-sample diagnostics (Table 18). The exception is out-of-sample absolute-return cross-series correlation, for which `Legendre`has the lowest median error.

Table 17: **In-sample foreign-exchange stylized-fact, autocorrelation, and cross-series correlation errors.** Entries report median absolute errors with 10th–90th percentile ranges in brackets. $\rho(r)$ and $\rho(|r|)$ denote cross-series correlations of returns and absolute returns, respectively. Bold entries indicate the lowest unrounded median error.

	Stylized-fact			Autocorrelation		Cross-series correlation	
Method	Volatility	Skewness	Kurtosis	ACF(r)	ACF($\lvert r\rvert$)	$\rho(r)$	$\rho(\lvert r\rvert)$
`Fourier`	0.07 [0.06, 0.10]	0.50 [0.22, 0.97]	2.81 [1.66, 21.02]	0.48 [0.44, 0.56]	0.35 [0.30, 0.41]	0.25 [0.07, 0.59]	0.27 [0.05, 0.51]
`Legendre`	0.08 [0.06, 0.10]	1.38 [0.56, 2.20]	4.83 [1.42, 16.55]	0.38 [0.30, 0.53]	0.29 [0.23, 0.39]	0.23 [0.05, 0.49]	0.22 [0.07, 0.45]
`Ours-TAd4`	**0.01 [0.00, 0.02]**	0.43 [0.12, 0.96]	2.26 [0.60, 14.69]	0.07 [0.06, 0.08]	0.07 [0.06, 0.10]	0.13 [0.01, 0.27]	0.10 [0.01, 0.27]
`Ours-TLLd4`	0.01 [0.00, 0.02]	**0.42 [0.14, 0.98]**	**1.90 [0.76, 13.14]**	**0.07 [0.06, 0.08]**	**0.07 [0.06, 0.10]**	**0.10 [0.02, 0.24]**	**0.09 [0.02, 0.29]**

Table 18: **Out-of-sample foreign-exchange stylized-fact, autocorrelation, and cross-series correlation errors.** Entries report median absolute errors with 10th–90th percentile ranges in brackets. $\rho(r)$ and $\rho(|r|)$ denote cross-series correlations of returns and absolute returns, respectively.

	Stylized-fact			Autocorrelation		Cross-series correlation	
Method	Volatility	Skewness	Kurtosis	ACF(r)	ACF($\lvert r\rvert$)	$\rho(r)$	$\rho(\lvert r\rvert)$
`Fourier`	0.06 [0.03, 0.09]	0.95 [0.46, 3.00]	3.45 [2.38, 19.59]	0.46 [0.42, 0.51]	0.35 [0.30, 0.39]	0.56 [0.23, 0.93]	0.30 [0.06, 0.56]
`Legendre`	0.06 [0.04, 0.10]	1.59 [0.66, 3.06]	5.23 [1.58, 18.56]	0.39 [0.29, 0.54]	0.33 [0.24, 0.45]	0.39 [0.10, 0.81]	**0.19 [0.04, 0.38]**
`Ours-TAd4`	0.02 [0.01, 0.04]	0.48 [0.18, 2.15]	2.92 [1.65, 18.65]	0.07 [0.06, 0.08]	0.10 [0.08, 0.13]	0.17 [0.04, 0.30]	0.27 [0.12, 0.40]
`Ours-TLLd4`	**0.01 [0.00, 0.03]**	**0.45 [0.16, 2.12]**	**2.91 [1.72, 19.14]**	**0.07 [0.06, 0.08]**	**0.10 [0.08, 0.13]**	**0.14 [0.04, 0.32]**	0.30 [0.16, 0.42]

E Ablations and alternative modeling choices

E.1 Effect of log-signature preprocessing

Experimental setup. Higher-order log-signature coordinates can contain extreme values, so our main setup applies the level-wise scaling map described in Appendix B.1, including both factorial scaling and the signed logarithm. Table 19 compares level-wise scaling against no scaling on log-GBM, with the latter omitting this entire transformation.

Results. Table 19 shows that, with level-wise scaling, LS reconstruction errors remain close to the oracle even at greater depths, whereas without scaling they increasingly deviate from the oracle as depth increases. These results support the use of level-wise scaling to maintain stable reconstruction performance as conditioning depth increases.

Table 19: **Posterior-sampling reconstruction errors with and without level-wise scaling of LS conditioning statistics on log-GBM.** Entries report path MSE. Bold entries indicate the LS reconstruction error closest to the oracle at each depth.

Setup	$r = 4$	$r = 5$	$r = 6$
LS Oracle	0.260	0.207	0.172
w/o level-wise scaling	**0.260**	0.210	0.206
w/ level-wise scaling	0.261	**0.208**	**0.173**

E.2 Effect of the conditional distribution estimator

Experimental setup. We compare conditional diffusion with flow matching to assess how the choice of conditional distribution estimator affects reconstruction and parameter recovery. We use the same architecture, conditioning, and experimental protocol as the flow-matching model. The diffusion baseline uses the DDPM objective [Ho et al., 2020] with a cosine noise schedule and DDIM sampling. We evaluate DDIM sampling with and without predicted-x_0 clipping to $[-10, 10]$.

Results. Table 20 shows that, with clipping, conditional diffusion achieves reconstruction errors close to the LS oracle. Table 21 shows that it also achieves low parameter-estimation errors under TLL conditioning. However, unclipped DDIM is unstable in these experiments, while clipping imposes a fixed pointwise range not implied by the conditioning signature. Flow matching achieves comparable oracle alignment and equal or lower reported parameter-estimation errors without clipping.

Table 20: **Posterior-sampling reconstruction errors of conditional diffusion and flow matching under LS conditioning.** LS Oracle denotes the oracle posterior-sampling reconstruction error. DDPM (no clip) and DDPM (clip) use DDIM sampling without and with clipping of the predicted x_0 to $[-10, 10]$, respectively. FM denotes flow matching without clipping.

Process	r	LS Oracle	DDPM (no clip)	DDPM (clip)	FM
log-GBM	2	0.54	14.44	0.53	0.55
	4	0.26	11.34	0.25	0.26
	6	0.17	25.55	0.18	0.17
fBM	2	0.66	13.28	0.67	0.66
	4	0.35	14.68	0.37	0.36
	6	0.25	15.20	0.27	0.27

Table 21: **Parameter-estimation errors of conditional diffusion and flow matching under TLL conditioning.** DDPM (clip) denotes conditional diffusion with DDIM sampling and predicted-x_0 clipping to $[-10, 10]$; FM denotes flow matching without clipping. Bold entries indicate the lower reported error in each DDPM–FM pair, including ties.

	log-GBM σ		fBM σ		fBM H	
r	DDPM (clip)	FM	DDPM (clip)	FM	DDPM (clip)	FM
2	0.06	**0.04**	0.07	**0.05**	0.007	**0.006**
4	0.05	**0.04**	0.08	**0.05**	0.007	**0.006**
6	0.06	**0.05**	0.08	**0.07**	**0.007**	**0.007**

NeurIPS Paper Checklist

1. **Claims**

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

   Answer: [Yes]

   Justification: The abstract and introduction state the paper's main contributions and scope: probabilistic truncated-signature inversion, quantification of inversion ambiguity through posterior-sampling reconstruction error and derivation of theoretical upper bounds for Gaussian process families, and empirical validation on synthetic and real data (Sections 4 to 6).

   Guidelines:

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

2. **Limitations**

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

   Answer: [Yes]

   Justification: The paper discusses limitations in Section 7, including the theory's Gaussian-process and linear-statistics scope, the need to evaluate scalability to higher-dimensional paths, and extensions beyond Gaussian processes.

   Guidelines:

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

3. **Theory assumptions and proofs**

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

   Answer: [Yes]

   Justification: The theoretical setup and assumptions are stated in Sections 3 and 5. The paper provides complete proofs for its original theoretical results in Section 5 and Appendix A, while the classical tree-like equivalence result is cited to prior work.

   Guidelines:

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

4. **Experimental result reproducibility**

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

   Answer: [Yes]

   Justification: The paper provides the dataset configurations, training and sampling setup, baseline specifications, preprocessing details, model architecture, and evaluation metrics in Section 6.1 and Appendix B.

   Guidelines:

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

(d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

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

   Answer: [Yes]

   Justification: Code and reproduction instructions are publicly available at `https://github.com/Junoh-Kang/ProbSigInv`. The repository documents environment setup, synthetic-data generation, financial-data acquisition, training, inversion, and evaluation.

   Guidelines:

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

6. **Experimental setting/details**

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

   Answer: [Yes]

   Justification: The paper specifies the data splits, optimizer, learning rate, batch size, training epochs, conditioning depths, and sampling settings in Section 6.1 and Appendix B.

   Guidelines:

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

7. **Experiment statistical significance**

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

   Answer: [Yes]

   Justification: The paper reports normal-approximation 95% bootstrap confidence intervals for empirical posterior-sampling reconstruction errors and t-distribution-based 95% confidence intervals over 10 numerical seeds for the drifted-fBM and OU LS oracle estimates in

Table 8. It also reports 95% bootstrap confidence intervals for squared energy distance gaps in Table 9.

Guidelines:

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

8. **Experiments compute resources**

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

   Answer: [Yes]

   Justification: The paper reports in Appendix B that all learned models are trained for 100 epochs on a single A5000 GPU, taking approximately three hours per run. It also states that inversion uses 1000 sampling steps to generate 30 samples per reference path and completes in less than one hour.

   Guidelines:

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

9. **Code of ethics**

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

   Answer: [Yes]

   Justification: The work is methodological and uses synthetic stochastic processes together with S&P 500 and foreign-exchange log-return windows. It does not involve human subjects, personally identifiable data, sensitive private data, or high-risk deployed systems, and we do not identify any conflict with the NeurIPS Code of Ethics.

   Guidelines:

   - The answer [N/A] means that the authors have not reviewed the NeurIPS Code of Ethics.

- If the authors answer [No], they should explain the special circumstances that require a deviation from the Code of Ethics.
- The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader impacts**

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

    Answer: [No]

    Justification: The work proposes a framework for probabilistic inversion of time-series signatures using synthetic processes, S&P 500 log-return windows, and foreign-exchange series; we do not identify a direct deployment path or application-specific societal harms beyond those of generic time-series generative modeling.

    Guidelines:

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

11. **Safeguards**

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

    Answer: [N/A]

    Justification: The paper does not release data or models with a high risk for misuse, such as pretrained language models, image generators, or scraped Internet datasets.

    Guidelines:

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

12. **Licenses for existing assets**

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

    Answer: [Yes]

    Justification: The paper credits the existing assets used in the experiments, including `iisignature` for log-signature computation, prior work for deterministic baseline constructions, and S&P 500 log-return windows for real-data evaluation. We do not redistribute these external assets as part of the paper.

    Guidelines:

    - The answer [N/A] means that the paper does not use existing assets.
    - The authors should cite the original paper that produced the code package or dataset.
    - The authors should state which version of the asset is used and, if possible, include a URL.
    - The name of the license (e.g., CC-BY 4.0) should be included for each asset.
    - For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
    - If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, `paperswithcode.com/datasets` has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
    - For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.
    - If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New assets**

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

    Answer: [Yes]

    Justification: The released implementation is documented in the public repository at `https://github.com/Junoh-Kang/ProbSigInv`, including installation, data preparation, training, inversion, evaluation, and licensing information.

    Guidelines:

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

14. **Crowdsourcing and research with human subjects**

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

    Answer: [N/A]

    Justification: The paper does not involve crowdsourcing experiments or research with human subjects.

    Guidelines:

    - The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.

- Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
- According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

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

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

    Answer: [N/A]

    Justification: The paper does not involve crowdsourcing or research with human subjects, so IRB approval or equivalent human-subjects review is not applicable.

    Guidelines:

    - The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
    - Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
    - We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
    - For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.

16. **Declaration of LLM usage**

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

    Answer: [N/A]

    Justification: The core framework, theory, experiments, and evaluation do not involve LLMs as important, original, or non-standard components.

    Guidelines:

    - The answer [N/A] means that the core method development in this research does not involve LLMs as any important, original, or non-standard components.
    - Please refer to our LLM policy in the NeurIPS handbook for what should or should not be described.