When Statistics Meets AI: Estimating Fractional Ornstein–Uhlenbeck Processes with Competing Methods

Laszlo Markus and Daniel Boros

The fractional Ornstein–Uhlenbeck (fOU) process is a continuous-time mean-reverting stochastic process, satisfying the Langevin equation driven by fractional Brownian motion (fBM). It is widely used, e.g., in mathematical finance for modeling rough volatility, in physics for anomalous diffusion, and in hydrology and climate science for modeling long-range dependence. The process is governed by three primary parameters: the mean-reversion speed $\kappa > 0$, the Hurst exponent $H \in (0, 1)$ controlling memory structure, and the volatility $\sigma > 0$.

Estimating κ from discrete observations is a classical statistical problem with a rich literature. For the truly fractional case ($H \neq 1/2$), Kleptsina and LeBreton (2002) were the first to obtain the formula of the drift parameter's maximum likelihood estimation (MLE), assuming a continuous record of observations. To achieve the proof of consistency and asymptotic normality of the estimator in the full $(0, 1)$ for H lasted until Tanaka (2020). The case of discrete sampling is particularly challenging, as for $H \in (0, 1/2)$ the integral in the MLE is neither classical non Itô. Hu and Nualart(2010), Hu, Nualart, and Zhou (2019) introduced and studied the discretization of what they called the ergodic estimator. Through extensive simulation Xiao, Wang, and Yu (2018) analyze in detail the discretized estimators.

In the regime of large mean-reversion speed ($\kappa \gg 1$), estimation becomes particularly challenging. When $\kappa\Delta t \gg 1$, successive observations are nearly independent draws from the stationary marginal distribution, and the drift information is concentrated in the unobserved transient dynamics between samples. This creates a fundamental identifiability barrier that affects all methods, classical and learning-based alike.

We propose a neural amortized inference framework that addresses these challenges. Our contributions are:

1. A **multi-expert hybrid architecture** that routes predictions through specialized subnetworks based on the estimated θ regime, addressing the heterogeneous difficulty across the parameter space.
2. A **Gaussian mixture posterior refiner** that outputs calibrated predictive distributions, enabling principled uncertainty quantification.
3. A **universal feature representation** combining spectral, autocorrelation, signature, and analytic proxy features that generalizes across varying $(T, N, \Delta t)$ without architectural changes.
4. Comprehensive experiments demonstrating state-of-the-art accuracy on $\theta \in [0.1, 20]$ with a single forward pass.

Our model jointly estimates (κ, H) over the full prior $\kappa \in [0.1, 20]$.