PINNs for Rough Volatility

Zelin Wu

19^{th} March, 2026 @ TU Darmstadt

From BS to Rough Volatility Models

To model the price of the underlying assets:

$$dS_t = \mu S_t dt + \sigma S(t) dW(t), S(0) = s_0$$

From BS to Rough Volatility Models

To model the price of the underlying assets:

$$dS_t = \mu S_t dt + \sigma S(t) dW(t), S(0) = s_0$$

- Black-Scholes: σ is constant

From BS to Rough Volatility Models

To model the price of the underlying assets:

$$dS_t = \mu S_t dt + \sigma S(t) dW(t), S(0) = s_0$$

- Black-Scholes: σ is constant
- Local Vol: σ is a function e.g. CEV:

$$\sigma(t, S_t) = \bar{\sigma} S_t^{\gamma - 1}$$

From BS to Rough Volatility Models

To model the price of the underlying assets:

$$dS_t = \mu S_t dt + \sigma S(t) dW(t), S(0) = s_0$$

- Black-Scholes: σ is constant
- Local Vol: σ is a function e.g. CEV:

$$\sigma(t, S_t) = \bar{\sigma} S_t^{\gamma - 1}$$

- Stochastic Vol: σ follows another SDE. e.g. Heston:

$$d\sigma^2(t) = \kappa(\theta - \sigma^2(t))dt + \bar{\sigma}\sigma(t)d\bar{W}(t)$$

From BS to Rough Volatility Models

To model the price of the underlying assets:

$$dS_t = \mu S_t dt + \sigma S(t) dW(t), S(0) = s_0$$

- Black-Scholes: σ is constant
- Local Vol: σ is a function e.g. CEV:

$$\sigma(t, S_t) = \bar{\sigma} S_t^{\gamma - 1}$$

- Stochastic Vol: σ follows another SDE. e.g. Heston:

$$d\sigma^2(t) = \kappa(\theta - \sigma^2(t))dt + \bar{\sigma}\sigma(t)d\bar{W}(t)$$

- Rough Vol: σ follows an SDE driven by fractional B.M. $W^H(t)$

Why Roughness?

Rough Volatility models, first proposed in 2014[$GJR18$], has a dominant performance in addressing these two long-lasting open problems.

1. Aligns well with the stylized fact about the volatility surface while classical models fail. [$GJR18$]

Why Roughness?

Rough Volatility models, first proposed in 2014[*GJR*18], has a dominant performance in addressing these two long-lasting open problems.

1. Aligns well with the stylized fact about the volatility surface while classical models fail. [*GJR*18]
2. Solve the VIX & SPX Joint Calibration puzzle. [*GJR*20]

Fractional Brownian Motion

Definition 1 (Fractional Brownian Motion(FBM))

A FBM $\{W^H(t)\}_{t\geq 0}$is a continuous-time Gaussian Process with zero expectation and $W^H(0) = 0$ together with the following covariance function:

$$E[W^H(t)W^H(s)] = \frac{1}{2}(t^{2H} + s^{2H} - |t-s|^{2H}),$$

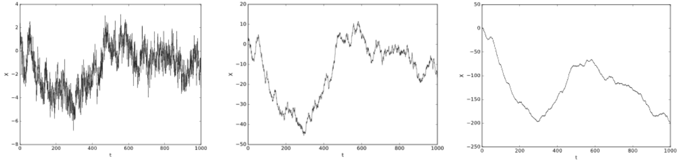
where $H \in [0, 1]$ is called the Hurst Parameter.

Definition 1 (Fractional Brownian Motion(FBM))

A FBM $\{W^H(t)\}_{t\geq 0}$ is a continuous-time Gaussian Process with zero expectation and $W^H(0) = 0$ together with the following covariance function:

$$E[W^H(t)W^H(s)] = \frac{1}{2}(t^{2H} + s^{2H} - |t-s|^{2H}),$$

where $H \in [0, 1]$ is called the Hurst Parameter.



H=0.15 H=0.55 H=0.95

Figure: Samples of FBM for different H

FBM Representations

There is no unique formula for a fBM. Among them, MVN(1968) is widely used:

$$W_t^H = \frac{1}{\Gamma(H+\frac{1}{2})}\left(\int_{-\infty}^{\infty}(t-s)_+^{H-\frac{1}{2}}dW_s - \int_{-\infty}^{\infty}(-s)_+^{H-\frac{1}{2}}dW_s\right)$$

FBM Representations

There is no unique formula for a fBM. Among them, MVN(1968) is widely used:

$$W_t^H = \frac{1}{\Gamma(H+\frac{1}{2})}(\int_{-\infty}^{\infty}(t-s)_+^{H-\frac{1}{2}}dW_s - \int_{-\infty}^{\infty}(-s)_+^{H-\frac{1}{2}}dW_s)$$

Sometimes we will use the "type-2" fBM, Riemann-Liouville process:

$$X_t^H = \frac{1}{\Gamma(H-\frac{1}{2})}\int_0^t (t-s)^{H-\frac{1}{2}}dW_s$$

FBM Representations

There is no unique formula for a fBM. Among them, MVN(1968) is widely used:

$$W_t^H = \frac{1}{\Gamma(H+\frac{1}{2})}(\int_{-\infty}^{\infty}(t-s)_+^{H-\frac{1}{2}}dW_s - \int_{-\infty}^{\infty}(-s)_+^{H-\frac{1}{2}}dW_s)$$

Sometimes we will use the "type-2" fBM, Riemann-Liouville process:

$$X_t^H = \frac{1}{\Gamma(H-\frac{1}{2})}\int_0^t (t-s)^{H-\frac{1}{2}}dW_s$$

Though X_t^H does not satisfy the covariance structure anymore, they both have α-Hölder continuity for $\alpha < H$ a.s., and their difference can be bounded $[P11]$:

$$\left\| \sup_{0\leq t\leq T} |(X_{S+t}^H - X_S^H) - (W_{S+t}^H - W_S^H)| \right\|_p \leq C_p S^{H-1} T$$

Rough Bergomi

Rough Bergomi model assumes

$$dS(t) = S(t)\sqrt{V(t)}dB(t)$$

$$dlog(V_t) = 2\nu dW_t^H$$

Rough Bergomi

Rough Bergomi model assumes

$$dS(t) = S(t)\sqrt{V(t)}dB(t)$$

$$dlog(V_t) = 2\nu dW_t^H$$

[BFG16]Rough Bergomi Model

$$V(t) = \xi_0(t)exp(\eta\sqrt{2H}\int_0^t (t-s)^{H-\frac{1}{2}}dW(s) - \frac{1}{2}\eta^2 t^{2H}),$$

where $\xi_0(t) = \mathbb{E}[V_t|\mathcal{F}_0]$ is the forward variance curve, usually set as constant during applications.

Rough Heston

Introduce the same rough kernel $(t-s)^{H-\frac{1}{2}}$ into Heston leads to

[ER19]Rough Heston Model

$$V_t = V_0 + \frac{1}{\Gamma(H+\frac{1}{2})}\Big(\int_0^t (t-s)^{H-\frac{1}{2}}\lambda(\theta - V_s)ds$$

$$+ \int_0^t (t-s)^{H-\frac{1}{2}}\lambda\nu\sqrt{V_s}dW(s)\Big).$$

Difficulties in Computation

Rough models are **non-Markovian** and do not have semi-martingale property.

- [*BLP*17] Proposed the **Hybrid Scheme**, which has been a benchmark.
- [*ZLCL*20] Used the Markovian approximation to simulate the rough Bergomi model by **Monte Carlo** simulation.
- [*HMT*21]Trained a **supervised NN** for the rough Bergomi model and its calibration.

Markovian Approximation

By Bernstein theorem on monotone functions, the rough kernel $t^{H-\frac{1}{2}}$ can be written as a Laplace transform of a measure.

$$t^{H-\frac{1}{2}} = \int_0^{\infty} e^{-xt} \mu_H(dx), \quad \mu_H(dx) = \frac{x^{-H-\frac{1}{2}}}{\Gamma(\frac{1}{2} - H)} dx$$

Markovian Approximation

By Bernstein theorem on monotone functions, the rough kernel $t^{H-\frac{1}{2}}$ can be written as a Laplace transform of a measure.

$$t^{H-\frac{1}{2}} = \int_0^\infty e^{-xt}\mu_H(dx), \quad \mu_H(dx) = \frac{x^{-H-\frac{1}{2}}}{\Gamma(\frac{1}{2}-H)}dx$$

And by Fubini's theorem,

$$\begin{aligned}\int_0^t (t-s)^{H-\frac{1}{2}}dW(s) &= \int_0^t \int_0^\infty e^{-x(t-s)}\mu_H(dx)dW(s) \\ &= \int_0^\infty \int_0^t e^{-x(t-s)}dW(s)\mu_H(dx)\end{aligned}$$

Markovian Approximation

By Bernstein theorem on monotone functions, the rough kernel $t^{H-\frac{1}{2}}$ can be written as a Laplace transform of a measure.

$$t^{H-\frac{1}{2}} = \int_0^\infty e^{-xt}\mu_H(dx), \quad \mu_H(dx) = \frac{x^{-H-\frac{1}{2}}}{\Gamma(\frac{1}{2}-H)}dx$$

And by Fubini's theorem,

$$\begin{aligned}\int_0^t (t-s)^{H-\frac{1}{2}}dW(s) &= \int_0^t \int_0^\infty e^{-x(t-s)}\mu_H(dx)dW(s) \\ &= \int_0^\infty \int_0^t e^{-x(t-s)}dW(s)\mu_H(dx)\end{aligned}$$

Note that $V_x(t) = \int_0^t e^{-x(t-s)}dW(s)$ is a solution to the OU process:

$$dV_x(t) = -xV_x(t)dt + dW(t), \quad V_x(0) = 0$$

Markovian Approximation

One natural idea is to discretize the $\int_0^\infty V_x(t)\mu_H(dx)$ into finite terms, i.e.

$$K(t) = t^{H-\frac{1}{2}} = \int_0^\infty e^{-xt}\mu_H(dx) \approx \sum_{i=1}^{n} \alpha_i e^{-\kappa_i t} = K^n(t)$$

Since we only care about the approximation performance on $[0, T]$, we may get these parameters α_i and κ_i from fitting.

Markovian Approximation

One natural idea is to discretize the $\int_0^\infty V_x(t)\mu_H(dx)$ into finite terms, i.e.

$$K(t) = t^{H-\frac{1}{2}} = \int_0^\infty e^{-xt}\mu_H(dx) \approx \sum_{i=1}^{n} \alpha_i e^{-\kappa_i t} = K^n(t)$$

Since we only care about the approximation performance on $[0, T]$, we may get these parameters α_i and κ_i from fitting.

Theorem 2 (Convergence Theorem)

[BB23] *For some 'nice' payoff function h, $\exists C$ s.t.*

$$|\mathbb{E}[h(S_T)] - \mathbb{E}[h(S_T^n)]| \leq C \int_0^T |K(t) - K^n(t)| dt$$

Markovian Approximation

Let $X(t) = log(\frac{V(t)}{\xi_0})$ and $K^n(t) = \sum_{i=1}^n \alpha_i e^{-\kappa_i t}$,
The Markovian Approximation of the Rough Bergomi Model is

$$\begin{cases} dS(t) = \sqrt{\xi_0} S(t) exp(\frac{1}{2} X(t)) dB(t), \quad S(0) = S_0 \\ dX(t) = \sum_{i=1}^{n} \alpha_i dV_i(t), \quad X(0) = 0 \\ dV_i(t) = -\kappa_i V_i(t) dt + \eta \sqrt{2H} dW(t), \quad V_i(0) = 0, \forall i \\ \langle dW(t), dB(t) \rangle = \rho dt \end{cases}$$

Feynman-Kac PDE

Denote $C(t, S, V_1, ..., V_n)$ the option pricing for a vanilla European Call option whose payoff function is $h(S_T) = (S_T - K)^+$.

By Feynman-Kac formula, we have

$$C_t - \sum_{i=1}^{n} \kappa_i V_i C_{v_i} + \frac{1}{2}\xi S^2 exp(\sum_{i=1}^{n} \alpha_i V_i) C_{ss}$$
$$+ \sqrt{2\xi H}\rho\eta S exp(\frac{1}{2}\sum_{i=1}^{n} \alpha_i V_i) \sum_{i=1}^{n} C_{sv_i} + 2H\eta^2 \sum_{i=1}^{n}\sum_{j=1}^{n} C_{v_i v_j} = 0,$$

together with $C(T, S, V_1, ..., V_n) = max((S - K), 0)\ \forall\ S, V_1, .., V_n$.

PINNs

Physics-Informed Neural Networks(PINNs)[RPK19] has shown promising results in solving high-dimensional problems.

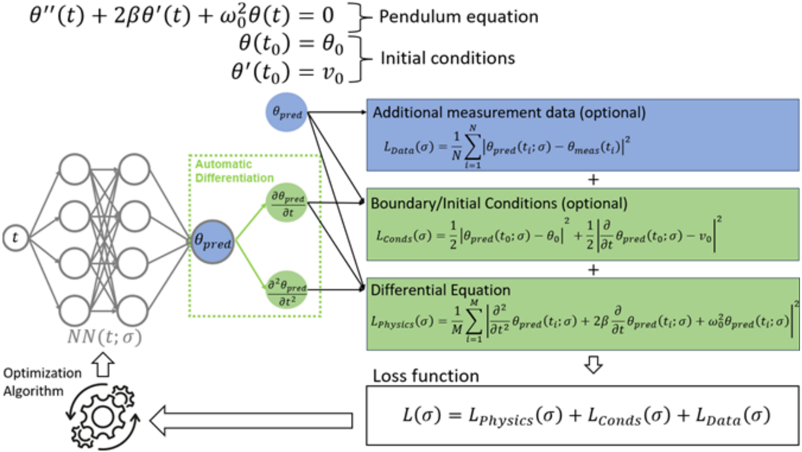


Figure: Illustration from Matlab

Volatility Skew

Definition 3 (At-the-money Volatility Skew)

$$\psi(\tau) = |\frac{\partial}{\partial k}\sigma_{BS}(k,\tau)|_{k=0}$$

where σ_{BS} is the implied volatility. $k = log(K/S_0)$ is the log moneyness. τ is the time-to-maturity.

.

Volatility Skew

Definition 3 (At-the-money Volatility Skew)

$$\psi(\tau) = |\frac{\partial}{\partial k}\sigma_{BS}(k,\tau)|_{k=0}$$

where σ_{BS} is the implied volatility. $k = log(K/S_0)$ is the log moneyness. τ is the time-to-maturity.

.

Empirically $\psi(\tau) \propto \tau^b$,and $b \approx -0.4$
As in $[BFG20]$, the skew for rough Bergomi is

$$\psi(\tau) = a\tau^{H-\frac{1}{2}} + b\tau^{2H},$$

When τ is small, roughly of order $H - \frac{1}{2}$ for $H < \frac{1}{2}$.

Numerical Verification

Set $H = 0.1$, observe $\psi(\tau) \propto \tau^{-0.4}$ as expected.

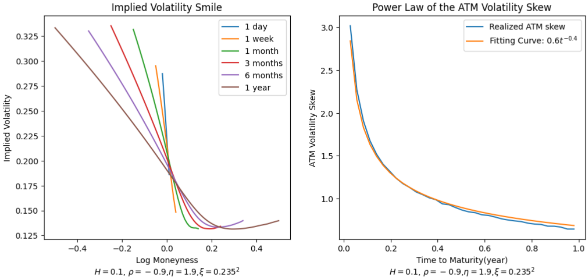


Figure: Markovian Approximation with 20 terms.

Compared to Hybrid Scheme

Method	Computation Cost	MAE
MC (10^8*samples*)	9h42m	-
Hybrid Scheme	12m02s	6.20E-3
Rough PINNs ($n = 5$)	4m22s	1.07E-3
Rough PINNs ($n = 10$)	6m41s	6.05E-4
Rough PINNs ($n = 20$)	11m47s	4.08E-4

Table: MAE Error Comparison w.r.t Monte Carlo Simulation