Rough differential equations for volatility

Signatures and Rough Paths: From Stochastics, Geometry and Algebra to ML

Edinburgh, 21 May 2025

Emilio Ferrucci

University of Oxford

DataSig

joint work with **Ofelia Bonesini**, **Ioannis Gasteratos**, and **Antoine Jacquier**

https://arxiv.org/abs/2412.21192

Rough paths to model stochastic dynamics

- Suppose we are interested in giving meaning to stochastic differential equations

$$\mathrm{d}Y_t = F(Y_t)\mathrm{d}X_t, \quad Y_0 = y_0. \tag{1}$$

  with X a given multidimensional stochastic process.
- We must only give meaning to the terms

$$\boldsymbol{X}^{\alpha_1 \dots \alpha_n}_{s,t} \text{“}=\text{”} \int_{s<u_1<\dots<u_n<t} \mathrm{d}X^{\alpha_1}_{u_1} \cdots \mathrm{d}X^{\alpha_n}_{u_n}, \quad |\alpha_1| + \dots + |\alpha_n| \leq 1$$

  with X^α $|\alpha|$-Hölder, e.g. through some kind of notion of stochastic $\int$.
- (1) will then automatically inherit meaning as

$$Y_t \approx Y_s + F_\alpha(Y_s)X^\alpha_{s,t} + \dots + F_{\alpha_1} \circ \cdots \circ F_{\alpha_n}(Y_s)\boldsymbol{X}^{\alpha_1 \dots \alpha_n}_{s,t}$$

- Is consistent with Itô and Stratonovich calculus and works for many other multidimensional stochastic processes beyond semimartingales, including $1/4 < H$-fBm[1].

[1] Coutin and Qian (2002).

DataSig

Rough volatility

- Option pricing: find dynamics for the (discounted) stock S

$$\mathrm{d}S_t = \sigma_t S_t \mathrm{d}W_t^{\mathbb{Q}}$$

such that the values of $\mathbb{E}^{\mathbb{Q}}\Phi(S_t)$ agree with those on the market, for liquid instruments Φ, e.g. call options $C(K,T)$.
  - Black-Scholes: $\sigma_t \equiv \sigma$, too simple, volatility surface;
  - Local volatility: $\sigma_t = \sigma(S_t, t)$;
  - Stochastic volatility: σ_t a Markovian diffusion process, correlated with but not determined by S (Bergomi, Heston, SABR,...).
- Rough volatility[2]: a stoch vol model in which σ_t has behaviour similar to $0.1 \approx H$-fBm on short timescales:
  - Good fit to market data (short dated smiles, term structure of the ATM vol skew) with relatively few parameters;
  - Statistical evidence from historical volatility.

[2] Gatheral, Jaisson, and Rosenbaum (2018).

DataSig

Rough volatility models

- In most rough vol models, σ is taken to be an explicit function of type-II fBm, e.g. rough Bergomi:

$$\sigma_t = \exp\big(\nu B^H_t - \tfrac{1}{2} c_H^2 \nu^2 t^{2H}\big)$$

where $\quad B^H_t = c_H \int_0^t (t-s)^{H-1/2} \mathrm{d}B, \qquad B = \rho W + \sqrt{1-\rho^2} W^\perp.$

- More complex models: take σ to satisfy a Volterra equation

$$\sigma_t = \sigma_0 + \int_0^t (t-s)^{H-1/2} \big(f(\sigma_s, s)\mathrm{d}s + g(\sigma_s, s)\mathrm{d}B_s\big).$$

DataSig

Challenges and limitations

- Challenges with rough vol:
  - $[\sigma, W] = \infty \implies$ no Stratonovich formulation or classical Wong-Zakai;
  - $H \leq 1/4, W \not\perp B^H \implies (W, B^H)$ not a classical Gaussian rp.
- Use regularity structures[3] to precisely subtract the divergent Itô-Stratonovich correction originating from $\mathbb{E}B_t^{H,\varepsilon}\dot{W}_t^{\varepsilon} \sim \rho\varepsilon^{H-1/2}$.
- However this comes at the cost of heavy tools, which become intractable for Volterra equations with low H.
- Moreover, Volterra equations driven by B with kernel $K(t,s)$ are different to ODEs driven by $\int_0^t K(t,s)\mathrm{d}B_s$ (even if B is C^∞), which are more basic and widespread.
- Since σ is one-dimensional, we could model it as an SDE with drift driven by B^H.
- However, currently no framework to treat general *joint* dynamics of price and vol, in the ODE sense, when the vol is rough.

[3] Bayer, Friz, Gassiat, Martin, and Stemper (2020).

DataSig

The Itô lift of an adapted process

- W d-dimensional brownian motion, $\boldsymbol{X}$ an e-dimensional stochastic rough path. Goal: jointly lift $\boldsymbol{X}$ and W to a geometric rough path $\overline{\boldsymbol{X}}$.
- Use Greek letters $\alpha, \beta, \gamma \in [d]$ for W, Latin letters $i, j, k \in [e]$ for $\boldsymbol{X}$. Define some terms with Stratonovich or Itô calculus. For $\alpha, \beta \in [d]$, $v \in [e]^\bullet$:

$$\begin{aligned} \overline{\boldsymbol{X}}^v &:= \boldsymbol{X}^v, \quad \overline{\boldsymbol{X}}^{\alpha\beta}_{s,t} := \int_s^t W^\alpha_{s,u} \circ \mathrm{d}W^\beta_u, \\ \overline{\boldsymbol{X}}^{v\alpha}_{s,t} &:= \int_s^t \boldsymbol{X}^v_{s,u} \mathrm{d}W^\alpha_u. \end{aligned} \tag{2}$$

- $\boldsymbol{X}$ of Hölder regularity $\in (1/3, 1/2]$ and deterministic[4]:

$$\alpha i = i \sqcup\!\sqcup \alpha - i\alpha \;\rightsquigarrow\; \overline{\boldsymbol{X}}^{\alpha i} := X^i X^\alpha - \overline{\boldsymbol{X}}^{i\alpha}$$

- When the regularity is $\leq 1/3$ there are more terms needed to define $\overline{\boldsymbol{X}}$. Impose some conditions on $\boldsymbol{X}$:

[4] Diehl, Oberhauser, and Riedel (2015).

The Itô lift of an adapted process II

Definition (Adapted H-integrable rough path)

An $\mathcal{F}_\bullet$*-adapted* H*-integrable (geometric) rough path* is an $\mathcal{F}_\bullet$-adapted, $\mathcal{G}^{\lfloor 1/H \rfloor}(\mathbb{R}^e)$-valued stochastic process $\boldsymbol{X}$ s.t. $\boldsymbol{X}_{s,t} \coloneqq \boldsymbol{X}_s^{-1} \otimes \boldsymbol{X}_t$ satisfies the Chen identity and

$$\sup_{0 \le s \le t \le T} \|\boldsymbol{X}^v_{s,t}\|^p_{L^p} \lesssim_{p,T} (t-s)^{pH \cdot \# v}, \qquad p \in [1, +\infty).$$

Theorem (Itô lift)

There exists a unique geometric rough path extending (2) on words $\alpha\beta$, v *and* $v\alpha$ *for* $v \in [e]^\bullet$. *It is* $\mathcal{F}_\bullet$*-adapted and*

$$|\overline{\boldsymbol{X}}^w_{st}(\omega)| \lesssim_{p,T,\omega} (t-s)^\lambda, \qquad \lambda < |w|,$$

with the constant of proportionality $\in L^p$ *for all* p.

DataSig

The Itô lift of an adapted process III

- Main idea: only words left are of the form $u\alpha v$ with $u, v \in [e]^\bullet$. Recursively express $u\alpha v$ as a ш-polynomial in words $u'\alpha, v'$ ($\rightsquigarrow \overline{\boldsymbol{X}}$ a polynomial in rough path terms already defined).

$$u\alpha j_1 \ldots j_n = u\alpha ш j_1 \ldots j_n - \sum_{k=1}^{n} (u ш j_1 \ldots j_k)\alpha j_{k+1} \ldots j_n.$$

- E.g. $$\begin{aligned} &i\alpha jk \\ &= i\alpha ш jk - (i ш j)\alpha k - (i ш jk)\alpha \\ &= i\alpha ш jk - ij\alpha k - ji\alpha k - ijk\alpha - jik\alpha - jki\alpha \\ &= i\alpha ш jk - [ij\alpha ш k - (ij ш k)\alpha] - [ji\alpha ш k - (ji ш k)\alpha] - ijk\alpha - jik\alpha - jki\alpha \\ &= i\alpha ш jk - ij\alpha ш k + ikj\alpha + kij\alpha - ji\alpha ш k + kji\alpha. \end{aligned}$$

- BDG + Kolmogorov + above statement $\rightsquigarrow$ Hölder regularity.
- Chen identity is inherited, shuffle property holds by construction.

A rough path for rough volatility

- Take $\boldsymbol{X}$ to be the canonical rough path above $X = B^H$ ($0.1 \approx H$-fBm correlated with X).

Proposition

$$\overline{\boldsymbol{X}}^{0^m \alpha 0^n}_{s,t} \text{``="} \int_s^t \frac{(X_u - X_s)^m}{m!} \frac{(X_t - X_u)^n}{n!} \mathrm{d}W^\alpha_u$$
$$:= \sum_{k=0}^{n} \frac{X_t^{n-k}}{m!k!(n-k)!} \int_s^t X^m_{su} (-X_u)^k W^\alpha_u.$$

- Consider general dynamics of the form

$$\begin{cases} \mathrm{d}S_t = \sigma_\alpha(S_t, V_t, t)\mathrm{d}\boldsymbol{W}^\alpha_t + g(S_t, V_t, t)\mathrm{d}t, \\ \mathrm{d}V_t = \tau(S_t, V_t, t)\mathrm{d}\boldsymbol{X}_t + \varsigma_\alpha(S_t, V_t, t)\mathrm{d}\boldsymbol{W}^\alpha_t + h(S_t, V_t, t)\mathrm{d}t. \end{cases} \tag{3}$$

DataSig

A rough path for rough volatility II

Proposition

S is given by Itô and Riemann integration as

$$S_t = S_0 + \int_0^t \sigma(S_u, V_u, u)\mathrm{d}W_u + \int_0^t \Big[\frac{1}{2}\sum_\alpha (\partial_S \sigma_\alpha \sigma_\alpha + \partial_V \sigma_\alpha \varsigma_\alpha) + g\Big](S_u, V_u, u)\mathrm{d}u.$$

- Proof: compare Davie expansions. Interesting to consider the Davie expansion of V ($1/3 < H \le 1/2$):

$$V_{uv} \approx \tau X_{uv} + \varsigma_\gamma W^\gamma_{uv} + h \cdot (v-u) + (\partial_V \varsigma_\beta \varsigma_\alpha + \partial_S \varsigma_\beta \sigma_\alpha) \int_u^v W^\alpha_{ur} \circ \mathrm{d}W^\beta_r + \partial_V \tau_j \tau_i \boldsymbol{X}^{ij}_{uv}$$
$$+ [\left(\begin{smallmatrix} 0 \\ \tau_k \end{smallmatrix}\right), \left(\begin{smallmatrix} \sigma_\gamma \\ \varsigma_\gamma \end{smallmatrix}\right)]^V \int_u^v X^k_{ur} \mathrm{d}W^\gamma_r + (\partial_S \tau_k \sigma_\gamma + \partial_V \tau_k \varsigma_\gamma) X^k_{uv} W^\gamma_{uv},$$

and if $[X, W] < \infty$ the second line is equal to

$$+\tau_k \partial_V \varsigma_\gamma \int_u^v X^k_{ur} \circ \mathrm{d}W^\gamma_r + (\partial_V \varsigma_\beta \varsigma_\alpha + \partial_S \varsigma_\beta \sigma_\alpha) \int_u^v W^\gamma_{ur} \circ \mathrm{d}X^k_r - \frac{1}{2}[\left(\begin{smallmatrix} 0 \\ \tau_k \end{smallmatrix}\right), \left(\begin{smallmatrix} \sigma_\gamma \\ \varsigma_\gamma \end{smallmatrix}\right)]^V [X^k, W^\gamma]_{uv}.$$

A rough path for rough volatility III

$$\begin{cases} \mathrm{d}S_t = \sigma_\alpha(S_t, V_t, t)\mathrm{d}\boldsymbol{W}_t^\alpha + g(S_t, V_t, t)\mathrm{d}t, \\ \mathrm{d}V_t = \tau(S_t, V_t, t)\mathrm{d}\boldsymbol{X}_t + \varsigma_\alpha(S_t, V_t, t)\mathrm{d}\boldsymbol{W}_t^\alpha + h(S_t, V_t, t)\mathrm{d}t. \end{cases}$$

- For some dynamics only the original rough path terms are needed: this occurs when $\varsigma = 0$ and τ does not depend on S. This has already been used for "simple" models ($V_t = f(X_t)$).[5]
- Correlation between V (rough process that feeds into the diffusion coeff. for S) and S can even be introduced if $X \perp\!\!\!\perp W$, thanks to ς. In this case, $\overline{\boldsymbol{X}}$ can be viewed as a Gaussian rough path with non i.d. components (the condition is $H_i + H_j > 1/2$, always verified if $H_j = 1/2$).
- Includes many models already in the literature, possible to consider multi-asset extensions.

[5] Fukasawa and Takano (2024).

Lead-lag approximations

- How to simulate RDEs? Euler scheme requires prior simulation of rough path terms and derivatives of order-10.
- Every geometric rough path $\boldsymbol{X}$ is the limit (in rough path topology) of the Stieltjes lifts of a sequence of smooth paths:

$$\boldsymbol{X}^{\alpha_1\ldots\alpha_n}_{s,t} = \lim_{\varepsilon\to 0}\int_{s<u_1<\ldots<u_n<t} \dot{X}^{\varepsilon,\alpha_1}_{u_1}\cdots\dot{X}^{\varepsilon,\alpha_n}_{u_n}\mathrm{d}u_1\cdots\mathrm{d}u_n$$
$$\overset{\text{ULT}}{\Longrightarrow} (\mathrm{d}Y^\varepsilon = F(Y^\varepsilon)\mathrm{d}X^\varepsilon) \xrightarrow{\varepsilon\to 0} (\mathrm{d}Y = F(Y)\mathrm{d}X)$$

- $\boldsymbol{X}$ Stratonovich lift: X^ε pwl interpolation, $X^\varepsilon = \varphi_\varepsilon \star X$, etc.
- Itô integrals of one-forms can also be approximated smoothly:[6]

$$\int f(W_t)\mathrm{d}W_t = \lim_{\varepsilon\to 0}\int f(W^{\varepsilon,-}_t)\dot{W}^\varepsilon_t\mathrm{d}t.$$

- Define $(X = B^H)$ for $t_{i+1} - t_i = \varepsilon$, $t \in [t_i, t_{i+1}]$ piecewise *lead-lag* approximations

$$W^\varepsilon_t := W_{t_i} + \frac{t - t_i}{\varepsilon}W_{t_i,t_{i+1}}, \quad X^{\varepsilon,-}_t := X^\varepsilon_{t-\varepsilon}.$$

[6] Flint, Hambly, and Lyons (2016).

Lead-lag approximations II

Theorem (Strong convergence of lead-lag)

$$\| |||\overline{\boldsymbol{X}} - \overline{\boldsymbol{X}}^{\varepsilon}|||_{\mathscr{C}} \|_{L^p} \lesssim \varepsilon^H$$

for all $p \geq 1$ and moreover the convergence is a.s. if the sequence of partitions is regular.

- Similar results using hybrid scheme approximations

$$\int_0^{t_k} (t_k - s)^{H-1/2} \mathrm{d}B_s$$
$$\approx \sum_{i=1}^{k-1} \int_{t_{i-1}}^{t_i} (t_k - s)^{H-1/2} \mathrm{d}s \cdot W_{t_{i-1},t_i} + \int_{t_{k-1}}^{t_k} (t_k - s)^{H-1/2} \mathrm{d}W_s$$

and for mollifier approximations.

Simulation

- Simulate $X^{\varepsilon,-}, W^{\varepsilon}$ on a grid and solve CDEs driven by them on a grid several times finer, using `Diffrax`[7].

$\rightsquigarrow$ notebook

[7] Kidger (2023).

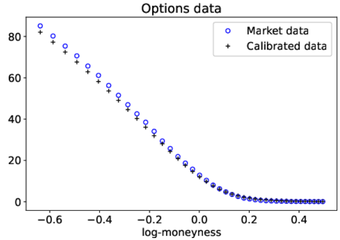
Calibration

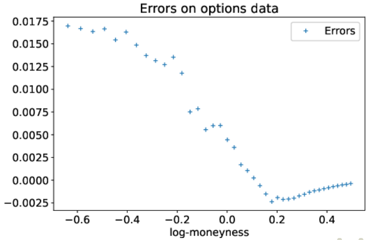
- Model:

$$\begin{cases} \dfrac{\mathrm{d}S_t}{S_t} = \sqrt{a(Z_t-b)^2+c}\circ \mathrm{d}W_t - \dfrac{1}{2}\left(a(Z_t-b)^2+c\right)\mathrm{d}t, \quad S_0 = s_0 > 0, \\ \mathrm{d}Z_t = \lambda(\theta - Z_t)\mathrm{d}t + \lambda\eta\sqrt{a(Z_t-b)^2+c}\circ \mathrm{d}B_t^H, \quad Z_0 = z_0 > 0, \end{cases}$$

- Minimise loss function:

$$\mathcal{L}(a,b,c,\lambda,\theta,\eta) := \sum_{j=1}^{L_i}\left(\mathrm{C}_{T,K_j}(a,b,c,\lambda,\theta,\eta) - \mathrm{C}^{\mathrm{obs}}_{T,K_j}\right)^2,$$





Thank you

Christian Bayer, Peter K. Friz, Paul Gassiat, Jorg Martin, and Benjamin Stemper (2020). **"A regularity structure for rough volatility".** In: *Mathematical Finance* 30.3, pp. 782–832. ISSN: 0960-1627. DOI: 10.1111/mafi.12233. URL: https://doi.org/10.1111/mafi.12233.

Laure Coutin and Zhongmin Qian (2002). **"Stochastic analysis, rough path analysis and fractional Brownian motions".** In: *Probability Theory and Related Fields* 122.1, pp. 108–140.

Joscha Diehl, Harald Oberhauser, and Sebastian Riedel (2015). **"A Lévy area between Brownian motion and rough paths with applications to robust nonlinear filtering and rough partial differential equations".** In: *Stochastic Processes and their Applications* 125.1, pp. 161–181. ISSN: 0304-4149. DOI: https://doi.org/10.1016/j.spa.2014.08.005. URL: https://www.sciencedirect.com/science/article/pii/S0304414914002002.

Guy Flint, Ben Hambly, and Terry J. Lyons (2016). **"Discretely sampled signals and the rough Hoff process".** In: *Stochastic Processes and Their Applications* 126.9, pp. 2593–2614. ISSN: 0304-4149,1879-209X. DOI: 10.1016/j.spa.2016.02.011. URL: https://doi.org/10.1016/j.spa.2016.02.011.

Masaaki Fukasawa and Ryoji Takano (2024). **"A partial rough path space for rough volatility".** In: *Electronic Journal of Probability* 29, pp. 1–28.

Jim Gatheral, Thibault Jaisson, and Mathieu Rosenbaum (2018). **"Volatility is rough".** In: *Quantitative Finance* 18.6, pp. 933–949.

Patrick Kidger (2023). ***Diffrax: Numerical Differential Equation Solvers for JAX.*** https://docs.kidger.site/diffrax/.