Asymptotic expansions for fractional stochastic volatility models

Blanka Horvath

Department of Mathematics, Imperial College London

3rd London-Paris Bachelier Workshop on Mathematical Finance

Based on joint works with
with Philipp Harms and Antoine Jacquier and with
Christian Bayer, Peter Friz, Archil Gulisashvili and Benjamin Stemper

September 29, 2016

Implied volatility

- Asset price process: $(S_t = \mathrm{e}^{X_t})_{t\geq 0}$, with $X_0 = 0$.
- No dividend, no interest rate.
- Black-Scholes-Merton (BSM) framework:

$$C_{\mathrm{BS}}(\tau, k, \sigma) := \mathbb{E}_0 \left(\mathrm{e}^{X_\tau} - \mathrm{e}^k\right)_+ = \mathcal{N}(d_+) - \mathrm{e}^k \mathcal{N}(d_-),$$

$$d_\pm := -\frac{k}{\sigma\sqrt{\tau}} \pm \frac{1}{2}\sigma\sqrt{\tau}.$$

- Spot implied volatility $\sigma_\tau(k)$: the unique (non-negative) solution to

$$C_{\mathrm{observed}}(\tau, k) = C_{\mathrm{BS}}(\tau, k, \sigma_\tau(k)).$$

- Implied volatility: unit-free measure of option prices.

Implied volatility

- Asset price process: $(S_t = \mathrm{e}^{X_t})_{t\geq 0}$, with $X_0 = 0$.
- No dividend, no interest rate.
- Black-Scholes-Merton (BSM) framework:

$$C_{\mathrm{BS}}(\tau, k, \sigma) := \mathbb{E}_0 \left(\mathrm{e}^{X_\tau} - \mathrm{e}^k\right)_+ = \mathcal{N}(d_+) - \mathrm{e}^k \mathcal{N}(d_-),$$

$$d_\pm := -\frac{k}{\sigma\sqrt{\tau}} \pm \frac{1}{2}\sigma\sqrt{\tau}.$$

- Spot implied volatility $\sigma_\tau(k)$: the unique (non-negative) solution to

$$C_{\mathrm{observed}}(\tau, k) = C_{\mathrm{BS}}(\tau, k, \sigma_\tau(k)).$$

- Implied volatility: unit-free measure of option prices.

However the implied volatility is not available in closed form for most models. Its asymptotic behaviour is available via (small/large k, τ) approximations.

Implied volatility ($\sigma_\tau(k)$) asymptotics as $|k| \uparrow \infty$, $\tau \downarrow 0$ or $\tau \uparrow \infty$:

- Hagan-Kumar-Lesniewski-Woodward (2003/2015): small-maturity for the SABR model
- Berestycki-Busca-Florent (2004): small-τ using PDE methods for diffusions.
- Henry-Labordère (2009): small-τ asymptotics using differential geometry.
- Forde et al.(2012), Jacquier et al.(2012): small/ large τ using large deviations.
- Lee (2003), Benaim-Friz (2009), Gulisashvili (2010-2012), Caravenna-Corbetta (2016), De Marco-Jacquier-Hillairet (2013): $|k| \uparrow \infty$.
- Laurence-Gatheral-Hsu-Ouyang-Wang (2012): small-τ in local volatility models.
- Fouque et al.(2000-2011): perturbation techniques for slow and fast mean-reverting stochastic volatility models.
- Mijatović-Tankov (2012): small-τ for jump models.
- Bompis-Gobet (2015): asymptotic expansions in the presence of both local and stochastic volatility using Malliavin calculus.

Related works:

- Deuschel-Friz-Jacquier-Violante (CPAM 2014), De Marco-Friz (2014): small-noise expansions using Laplace method on Wiener space (Ben Arous-Bismut approach).
- Baudoin-Ouyang (2010): small-noise expansions in a fractional setting
- Forde-Zhang (2015): large deviations in a fractional stochastic volatility setting
- Guennon-Jacquier-Roome (2015): large deviations in a fractional Heston model

Motivation

- Classical stochastic volatility models generate a constant short-maturity ATM skew and a large-maturity one proportional to τ^{-1};

- However, short-term data suggests a time decay of the ATM skew proportional to $\tau^{-\alpha}$, $\alpha \in (0, 1/2)$.

- One solution: adding volatility factors (risk of over-parameterisation). Gatheral's Double Mean-Reverting, Bergomi-Guyon, each factor acting on a specific time horizon.

- In the Lévy case (Tankov, 2010), the situation is different, as $\tau \downarrow 0$:
  - in the pure jump case with $\int_{(-1,1)} |x|\nu(\mathrm{d}x) < \infty$, then $\sigma_\tau^2(0) \sim c\tau$;
  - in the (α) stable case, $\sigma_\tau^2(0) \sim c\tau^{1-2/\alpha}$ for $\alpha \in (1, 2)$;
  - for out-of-the-money options, $\sigma_\tau^2(k) \sim \dfrac{k^2}{2\tau|\log(\tau)|}$.

Rough volatility models

- Gatheral-Jaisson-Rosenbaum and Bayer-Gatheral-Friz (2014,1015) proposed a fractional volatility model:

$$\begin{aligned} \mathrm{d}S_t &= S_t(\sigma_t \mathrm{d}Z_t + \mu_t dt), \\ \sigma_t &= \exp\left(X_t\right), \end{aligned} \tag{1}$$

  where

$$X_t = \mu W_t^H - \alpha(X_t - m)dt,$$

  for $\mu, \alpha > 0$, $m \in \mathbb{R}$ for a Bm Z and a fBm motion W^H with Hurst parameter H.
- Time series of the Oxford-Man SPX realised variance as well as implied volatility smiles of the SPX suggest that $H \in (0, 1/2)$: **short-memory** volatility.
- Is not statistically rejected by Ait-Sahalia-Jacod's test (2009) for Itô diffusions.
- Main drawback: loss of Markovianity ($H \neq 1/2$) rules out PDE techniques, and Monte Carlo is computationally intensive.

Today's menu

General setting:

$$\begin{aligned} dX_t^\epsilon &= b_1(\epsilon^{\kappa_1}, X_t^\epsilon, Y_t^\epsilon)dt + \epsilon^\beta \left(\sigma_{11}(X_t^\epsilon, Y_t^\epsilon)dW_t^{H_1} + \sigma_{12}(X_t^\epsilon, Y_t^\epsilon)dW_t^{H_2}\right) \\ dY_t^\epsilon &= b_2(\epsilon^{\kappa_2}, X_t^\epsilon, Y_t^\epsilon)dt + \epsilon^\beta \left(\sigma_{21}(X_t^\epsilon, Y_t^\epsilon)dW_t^{H_1} + \sigma_{22}(X_t^\epsilon, Y_t^\epsilon)dW_t^{H_2}\right). \end{aligned} \tag{2}$$

General setting:

$$dX_t^\epsilon = b_1(\epsilon^{\kappa_1}, X_t^\epsilon, Y_t^\epsilon)dt + \epsilon^\beta \left(\sigma_{11}(X_t^\epsilon, Y_t^\epsilon)dW_t^{H_1} + \sigma_{12}(X_t^\epsilon, Y_t^\epsilon)dW_t^{H_2}\right) \tag{2}$$
$$dY_t^\epsilon = b_2(\epsilon^{\kappa_2}, X_t^\epsilon, Y_t^\epsilon)dt + \epsilon^\beta \left(\sigma_{21}(X_t^\epsilon, Y_t^\epsilon)dW_t^{H_1} + \sigma_{22}(X_t^\epsilon, Y_t^\epsilon)dW_t^{H_2}\right).$$

Motivation: "Classical" ($H_i = \frac{1}{2}$) case: Deuschel-Friz-Jacquier-Violante (2011)

$$d\mathrm{X}_t^\epsilon = b(\epsilon, \mathrm{X}_t^\epsilon)dt + \epsilon \sum_{i=1}^m \sigma_i(\mathrm{X}_t^\epsilon)dW_t^i, \quad \mathrm{X}_0^\epsilon = x_0^\epsilon \in \mathbb{R}^d$$

General setting:

$$dX_t^\epsilon = b_1(\epsilon^{\kappa_1}, X_t^\epsilon, Y_t^\epsilon)dt + \epsilon^\beta\left(\sigma_{11}(X_t^\epsilon, Y_t^\epsilon)dW_t^{H_1} + \sigma_{12}(X_t^\epsilon, Y_t^\epsilon)dW_t^{H_2}\right)$$
$$dY_t^\epsilon = b_2(\epsilon^{\kappa_2}, X_t^\epsilon, Y_t^\epsilon)dt + \epsilon^\beta\left(\sigma_{21}(X_t^\epsilon, Y_t^\epsilon)dW_t^{H_1} + \sigma_{22}(X_t^\epsilon, Y_t^\epsilon)dW_t^{H_2}\right). \quad (2)$$

Motivation: "Classical" ($H_i = \frac{1}{2}$) case: Deuschel-Friz-Jacquier-Violante (2011)

$$d\mathrm{X}_t^\epsilon = b(\epsilon, \mathrm{X}_t^\epsilon)dt + \epsilon\sum_{i=1}^m \sigma_i(\mathrm{X}_t^\epsilon)dW_t^i, \quad \mathrm{X}_0^\epsilon = x_0^\epsilon \in \mathbb{R}^d$$

Fractional ($H_i = H \in (\frac{1}{4}, 1)$) case: Baudoin-Ouyang (2015)

$$d\mathrm{X}_t^\epsilon = b(\epsilon, \mathrm{X}_t^\epsilon)dt + \epsilon\sum_{i=1}^m \sigma_i(\mathrm{X}_t^\epsilon)d(W^H)_t^i, \quad \mathrm{X}_0^\epsilon = x_0^\epsilon \in \mathbb{R}^d$$

General setting:

$$dX_t^\epsilon = b_1(\epsilon^{\kappa_1}, X_t^\epsilon, Y_t^\epsilon)dt + \epsilon^\beta \left(\sigma_{11}(X_t^\epsilon, Y_t^\epsilon)dW_t^{H_1} + \sigma_{12}(X_t^\epsilon, Y_t^\epsilon)dW_t^{H_2}\right)$$
$$dY_t^\epsilon = b_2(\epsilon^{\kappa_2}, X_t^\epsilon, Y_t^\epsilon)dt + \epsilon^\beta \left(\sigma_{21}(X_t^\epsilon, Y_t^\epsilon)dW_t^{H_1} + \sigma_{22}(X_t^\epsilon, Y_t^\epsilon)dW_t^{H_2}\right). \quad (2)$$

Motivation: "Classical" ($H_i = \frac{1}{2}$) case: Deuschel-Friz-Jacquier-Violante (2011)

$$d\mathrm{X}_t^\epsilon = b(\epsilon, \mathrm{X}_t^\epsilon)dt + \epsilon \sum_{i=1}^m \sigma_i(\mathrm{X}_t^\epsilon)dW_t^i, \quad \mathrm{X}_0^\epsilon = x_0^\epsilon \in \mathbb{R}^d$$

Fractional ($H_i = H \in (\frac{1}{4}, 1)$) case: Baudoin-Ouyang (2015)

$$d\mathrm{X}_t^\epsilon = b(\epsilon, \mathrm{X}_t^\epsilon)dt + \epsilon \sum_{i=1}^m \sigma_i(\mathrm{X}_t^\epsilon)d(W^H)_t^i, \quad \mathrm{X}_0^\epsilon = x_0^\epsilon \in \mathbb{R}^d$$

Assumptions made: $b(\varepsilon, \cdot) \to \sigma_0(\cdot)$ and $x_0^\varepsilon \to x_0$ as $\varepsilon \to 0$, the weak Hörmander condition for $\{\sigma_0, \sigma_1, \ldots, \sigma_d\}$ at x_0, and vector fields are C^∞-bounded (can be relaxed).

General setting:

$$dX_t^\epsilon = b_1(\epsilon^{\kappa_1}, X_t^\epsilon, Y_t^\epsilon)dt + \epsilon^\beta \left(\sigma_{11}(X_t^\epsilon, Y_t^\epsilon)dW_t^{H_1} + \sigma_{12}(X_t^\epsilon, Y_t^\epsilon)dW_t^{H_2}\right)$$
$$dY_t^\epsilon = b_2(\epsilon^{\kappa_2}, X_t^\epsilon, Y_t^\epsilon)dt + \epsilon^\beta \left(\sigma_{21}(X_t^\epsilon, Y_t^\epsilon)dW_t^{H_1} + \sigma_{22}(X_t^\epsilon, Y_t^\epsilon)dW_t^{H_2}\right). \quad (2)$$

Motivation: "Classical" ($H_i = \frac{1}{2}$) case: Deuschel-Friz-Jacquier-Violante (2011)

$$d\mathrm{X}_t^\epsilon = b(\epsilon, \mathrm{X}_t^\epsilon)dt + \epsilon \sum_{i=1}^m \sigma_i(\mathrm{X}_t^\epsilon)dW_t^i, \quad \mathrm{X}_0^\epsilon = x_0^\epsilon \in \mathbb{R}^d$$

Fractional ($H_i = H \in (\frac{1}{4}, 1)$) case: Baudoin-Ouyang (2015)

$$d\mathrm{X}_t^\epsilon = b(\epsilon, \mathrm{X}_t^\epsilon)dt + \epsilon \sum_{i=1}^m \sigma_i(\mathrm{X}_t^\epsilon)d(W^H)_t^i, \quad \mathrm{X}_0^\epsilon = x_0^\epsilon \in \mathbb{R}^d$$

Assumptions made: $b(\varepsilon, \cdot) \to \sigma_0(\cdot)$ and $x_0^\varepsilon \to x_0$ as $\varepsilon \to 0$, the weak Hörmander condition for $\{\sigma_0, \sigma_1, \ldots, \sigma_d\}$ at x_0, and vector fields are C^∞-bounded (can be relaxed).

Our main interest: $H_1 = \frac{1}{2}, H_2 \neq \frac{1}{2}$.

Our setting:

$$
\begin{aligned}
dX_t^\epsilon &= b_1(\epsilon^{\kappa_1}, Y_t^\epsilon)dt + \epsilon^\beta \sigma_1(Y_t^\epsilon)dW_t \\
dY_t^\epsilon &= b_2(\epsilon^{\kappa_2}, Y_t^\epsilon)dt + \epsilon^\beta \sigma_2 dW_t^H.
\end{aligned}
\tag{2}
$$

Our setting:

$$\begin{aligned} dX_t^\epsilon &= b_1(\epsilon^{\kappa_1}, Y_t^\epsilon)dt + \epsilon^\beta \sigma_1(Y_t^\epsilon)dW_t \\ dY_t^\epsilon &= b_2(\epsilon^{\kappa_2}, Y_t^\epsilon)dt + \epsilon^\beta \sigma_2 dW_t^H. \end{aligned} \qquad (2)$$

σ_1 α-Hölder continuous, $\alpha \in (0,1]$, $\sigma_2 > 0$, (but can be extended to $\sigma_2(\cdot)$ bounded and elliptic) conditions on b_1, b_2 dictated by scaling and existence, and $H \neq 1/2$, particular interest in $H < 1/2$.

Our setting:

$$\begin{aligned} dX_t^\epsilon &= b_1(\epsilon^{\kappa_1}, Y_t^\epsilon)dt + \epsilon^\beta \sigma_1(Y_t^\epsilon)dW_t \\ dY_t^\epsilon &= b_2(\epsilon^{\kappa_2}, Y_t^\epsilon)dt + \epsilon^\beta \sigma_2 dW_t^H. \end{aligned} \quad (2)$$

σ_1 α-Hölder continuous, $\alpha \in (0,1]$, $\sigma_2 > 0$, (but can be extended to $\sigma_2(\cdot)$ bounded and elliptic) conditions on b_1, b_2 dictated by scaling and existence, and $H \neq 1/2$, particular interest in $H < 1/2$.
To introduce correlation we consider $\tilde{B}$ and B independent, and set $W = \bar{\rho}\tilde{B} + \rho B$ and

$$W^H = \int_0^t K(t,s)B_s$$

where K the Volterra kernel of the (standard) fBm W^H.

Our setting:

$$
\begin{aligned}
dX_t^{\epsilon} &= b_1(\epsilon^{\kappa_1}, Y_t^{\epsilon})dt + \epsilon^{\beta}\sigma_1(Y_t^{\epsilon})dW_t \\
dY_t^{\epsilon} &= b_2(\epsilon^{\kappa_2}, Y_t^{\epsilon})dt + \epsilon^{\beta}\sigma_2 dW_t^H.
\end{aligned} \tag{2}
$$

σ_1 α-Hölder continuous, $\alpha \in (0,1]$, $\sigma_2 > 0$, (but can be extended to $\sigma_2(\cdot)$ bounded and elliptic) conditions on b_1, b_2 dictated by scaling and existence, and $H \neq 1/2$, particular interest in $H < 1/2$.
To introduce correlation we consider $\tilde{B}$ and B independent, and set $W = \bar{\rho}\tilde{B} + \rho B$ and

$$
W^H = \int_0^t K(t,s)B_s
$$

where K the Volterra kernel of the (standard) fBm W^H.

Remark: The setting of Forde-Zhang '16 is included. With a leap of faith Gatheral-Jaisson-Rosenbaum '14 and Bayer-Friz-Gatheral '15 as well.

Our setting:

$$
\begin{aligned}
dX_t^\epsilon &= b_1(\epsilon^{\kappa_1}, Y_t^\epsilon)dt + \epsilon^\beta \sigma_1(Y_t^\epsilon)dW_t, \quad X_0^\epsilon = x_0^\epsilon \\
dY_t^\epsilon &= b_2(\epsilon^{\kappa_2}, Y_t^\epsilon)dt + \epsilon^\beta \sigma_2 dW_t^H, \qquad Y_0^\epsilon = y_0^\epsilon.
\end{aligned} \tag{2}
$$

Our setting:

$$
\begin{aligned}
dX_t^\epsilon &= b_1(\epsilon^{\kappa_1}, Y_t^\epsilon)dt + \epsilon^\beta \sigma_1(Y_t^\epsilon)dW_t, \quad X_0^\epsilon = x_0^\epsilon \\
dY_t^\epsilon &= b_2(\epsilon^{\kappa_2}, Y_t^\epsilon)dt + \epsilon^\beta \sigma_2 dW_t^H, \qquad Y_0^\epsilon = y_0^\epsilon.
\end{aligned} \tag{2}
$$

Our setting:

$$
\begin{aligned}
dX_t^\epsilon &= b_1(\epsilon^{\kappa_1}, Y_t^\epsilon)dt + \epsilon^\beta \sigma_1(Y_t^\epsilon)dW_t, \quad X_0^\epsilon = x_0^\epsilon \\
dY_t^\epsilon &= b_2(\epsilon^{\kappa_2}, Y_t^\epsilon)dt + \epsilon^\beta \sigma_2 dW_t^H, \qquad Y_0^\epsilon = y_0^\epsilon.
\end{aligned} \tag{2}
$$

Examples: Consider $(X_0, Y_0) = (0, y_0)$ and

$$
dX_t = -\frac{Y_t^2}{2}dt + Y_t dW_t, \qquad dY_t = (a + bY_t)dt + cdW_t^H. \tag{3}
$$

Our setting:

$$\begin{aligned} dX_t^\epsilon &= b_1(\epsilon^{\kappa_1}, Y_t^\epsilon)dt + \epsilon^\beta \sigma_1(Y_t^\epsilon)dW_t, \quad X_0^\epsilon = x_0^\epsilon \\ dY_t^\epsilon &= b_2(\epsilon^{\kappa_2}, Y_t^\epsilon)dt + \epsilon^\beta \sigma_2 dW_t^H, \quad Y_0^\epsilon = y_0^\epsilon. \end{aligned} \tag{2}$$

Examples: Consider $(X_0, Y_0) = (0, y_0)$ and

$$dX_t = -\frac{Y_t^2}{2}dt + Y_t dW_t, \qquad dY_t = (a + bY_t)dt + cdW_t^H. \tag{3}$$

- Define $(X_t^\epsilon, Y_t^\epsilon) := (\epsilon^{2H-1}X_{\epsilon^2 t}, Y_{\epsilon^2 t}) \Rightarrow$ (2) with $\kappa_1 = 2H+1$, $\kappa_2 = 2$ and $\beta = 2H$, $(x_0^\epsilon, y_0^\epsilon) = (0, y_0)$:

$$dX_t^\epsilon = -\epsilon^{2H+1}\frac{(Y_t^\epsilon)^2}{2}dt + \epsilon^{2H}Y_t^\epsilon dW_t, \qquad dY_t^\epsilon = \epsilon^2(a + bY_t^\epsilon)dt + \epsilon^{2H}cdW_t^H. \tag{4}$$

Our setting:

$$\begin{aligned} dX_t^\epsilon &= b_1(\epsilon^{\kappa_1}, Y_t^\epsilon)dt + \epsilon^\beta \sigma_1(Y_t^\epsilon)dW_t, \quad X_0^\epsilon = x_0^\epsilon \\ dY_t^\epsilon &= b_2(\epsilon^{\kappa_2}, Y_t^\epsilon)dt + \epsilon^\beta \sigma_2 dW_t^H, \quad Y_0^\epsilon = y_0^\epsilon. \end{aligned} \tag{2}$$

Examples: Consider $(X_0, Y_0) = (0, y_0)$ and

$$dX_t = -\frac{Y_t^2}{2}dt + Y_t dW_t, \qquad dY_t = (a + bY_t)dt + c dW_t^H. \tag{3}$$

- Define $(X_t^\epsilon, Y_t^\epsilon) := (\epsilon^{2H-1}X_{\epsilon^2 t}, Y_{\epsilon^2 t}) \Rightarrow$ (2) with $\kappa_1 = 2H+1$, $\kappa_2 = 2$ and $\beta = 2H$, $(x_0^\epsilon, y_0^\epsilon) = (0, y_0)$:

$$dX_t^\epsilon = -\epsilon^{2H+1}\frac{(Y_t^\epsilon)^2}{2}dt + \epsilon^{2H} Y_t^\epsilon dW_t, \qquad dY_t^\epsilon = \epsilon^2(a + bY_t^\epsilon)dt + \epsilon^{2H} c dW_t^H. \tag{4}$$

- Define $(X_t^\epsilon, Y_t^\epsilon) := (\epsilon^{2H}X_t, \epsilon^H Y_t) \Rightarrow$ (2) with $\kappa_1 = 0$, $\kappa_2 = \beta = H$, $(x_0^\epsilon, y_0^\epsilon) = (0, \epsilon^H y_0)$:

$$dX_t^\epsilon = -\frac{(Y_t^\epsilon)^2}{2}dt + \epsilon^H Y_t^\epsilon dW_t, \qquad dY_t^\epsilon = (a\epsilon^H + bY_t^\epsilon)dt + \epsilon^H c dW_t^H. \tag{5}$$

Our setting:

$$dX_t^\epsilon = b_1(\epsilon^{\kappa_1}, Y_t^\epsilon)dt + \epsilon^\beta \sigma_1(Y_t^\epsilon)dW_t, \quad X_0^\epsilon = x_0^\epsilon$$
$$dY_t^\epsilon = b_2(\epsilon^{\kappa_2}, Y_t^\epsilon)dt + \epsilon^\beta \sigma_2 dW_t^H, \quad Y_0^\epsilon = y_0^\epsilon. \tag{2}$$

Examples: Consider $(X_0, Y_0) = (0, y_0)$ and

$$dX_t = -\frac{Y_t^2}{2}dt + Y_t dW_t, \qquad dY_t = (a + bY_t)dt + c dW_t^H. \tag{3}$$

- Define $(X_t^\epsilon, Y_t^\epsilon) := (\epsilon^{2H-1}X_{\epsilon^2 t}, Y_{\epsilon^2 t}) \Rightarrow$ (2) with $\kappa_1 = 2H+1$, $\kappa_2 = 2$ and $\beta = 2H$, $(x_0^\epsilon, y_0^\epsilon) = (0, y_0)$:

$$dX_t^\epsilon = -\epsilon^{2H+1}\frac{(Y_t^\epsilon)^2}{2}dt + \epsilon^{2H} Y_t^\epsilon dW_t, \qquad dY_t^\epsilon = \epsilon^2(a + bY_t^\epsilon)dt + \epsilon^{2H} c dW_t^H. \tag{4}$$

- Define $(X_t^\epsilon, Y_t^\epsilon) := (\epsilon^{2H}X_t, \epsilon^H Y_t) \Rightarrow$ (2) with $\kappa_1 = 0$, $\kappa_2 = \beta = H$, $(x_0^\epsilon, y_0^\epsilon) = (0, \epsilon^H y_0)$:

$$dX_t^\epsilon = -\frac{(Y_t^\epsilon)^2}{2}dt + \epsilon^H Y_t^\epsilon dW_t, \qquad dY_t^\epsilon = (a\epsilon^H + bY_t^\epsilon)dt + \epsilon^H c dW_t^H. \tag{5}$$

Remark: Scaling (4) $\Rightarrow$ short-time (cf. Forde-Zhang (2015)), scaling (5) $\Rightarrow$ tails.

Theorem (Harms-H-Jacquier)

Consider an SDE of the form (2). Then the density of X_T^ε admits an expansion

$$f_\varepsilon(T,x) = \exp\left(-\frac{\Lambda(x)}{\varepsilon^{2\beta}} + \frac{\Lambda'(x)\widehat{X}_T}{\varepsilon^{\beta}}\right)\varepsilon^{-\min(\kappa_1,\beta)}\left(c_0 + \mathcal{O}(\varepsilon^{\delta(\kappa_1,\beta)})\right), \quad \text{as } \varepsilon \to 0,$$

where

$$\Lambda(x) = \inf\left\{\frac{1}{2}\|\mathrm{k}\|^2_{\mathcal{H}_H}, \mathrm{k} \in \mathcal{K}^x_{x_0,y_0}\right\} = \frac{1}{2}\|\mathrm{k}_0\|^2_{\mathcal{H}_H},$$

and

$$\mathrm{d}\widehat{X}_t = \left[\partial_x b\left(0,\phi_t^{\mathrm{k}_0}\right) + \partial_x\sigma\left(\phi_t^{\mathrm{h}_0}\right)\cdot\dot{\mathrm{k}}_0(t)\right]\widehat{X}_t\mathrm{d}t + \partial_{\varepsilon^\beta} b\left(0,\phi_t^{\mathrm{k}_0}\right)\mathrm{d}t, \quad \widehat{X}_0 = \partial_{\varepsilon^\beta}x_0^\varepsilon|_{\varepsilon=0},$$

where ϕ^{k_0} denotes the ODE solution of the same SDE (2) replacing $\varepsilon^\beta \mathrm{d}W$ by $\dot{\mathrm{k}}_0$ and x_0^ε by x_0.

Corollary: Varadhan-type asymptotics

Corollary (short-time asymptotics in Stein-Stein) $dY_t = (a + bY_t)dt + cdW_t^H$

Consider the Stein-Stein model (X_t, Y_t) as in (3) with $X_0 = 0$, $Y_0 = y_0 > 0$. Then in a neighbourhood of (x_0, y_0) the density of X_t satisfies the following asymptotic expansion as $t \to 0$

$$f_X(t,x) = \exp\left(\frac{\Lambda(x)}{t^{2H}}\right) t^{-H} \left(\frac{1}{2\pi} + \mathcal{O}(t^{\delta(H,H+1/2)})\right)$$

where

$$\Lambda(x) = \inf\left\{\frac{1}{2}\|\mathrm{k}\|^2_{\mathcal{H}_H}, \mathrm{k} \in \mathcal{K}^x_{x_0,y_0}\right\}.$$

Corollary: Varadhan-type asymptotics

Corollary (short-time asymptotics in Stein-Stein) $dY_t = (a + bY_t)dt + cdW_t^H$

Consider the Stein-Stein model (X_t, Y_t) as in (3) with $X_0 = 0$, $Y_0 = y_0 > 0$. Then in a neighbourhood of (x_0, y_0) the density of X_t satisfies the following asymptotic expansion as $t \to 0$

$$f_X(t,x) = \exp\left(\frac{\Lambda(x)}{t^{2H}}\right) t^{-H}\left(\frac{1}{2\pi} + \mathcal{O}(t^{\delta(H,H+1/2)})\right)$$

where

$$\Lambda(x) = \inf\left\{\frac{1}{2}\|\mathrm{k}\|^2_{\mathcal{H}_H}, \mathrm{k} \in \mathcal{K}^x_{x_0,y_0}\right\}.$$

Proof: Take $T = 1$, $\varepsilon^2 = t$ and consider $(X_t^\epsilon, Y_t^\epsilon) := (\epsilon^{2H-1}X_{\epsilon^2 t}, Y_{\epsilon^2 t})$ with $X_0^\epsilon = 0$, $Y_0^\epsilon = y_0 > 0$. $\Rightarrow$ Short-time scaling:

$$dX_t^\epsilon = -\epsilon^{2H+1}\frac{(Y_t^\epsilon)^2}{2}dt + \epsilon^{2H}Y_t^\epsilon dW_t, \qquad dY_t^\epsilon = \epsilon^2(a + bY_t^\epsilon)dt + \epsilon^{2H}cdW_t^H, \quad (4)$$

Note that the drift vanishes in the limit $\epsilon \to 0$ and $x_0^\varepsilon = x_0 = 0$.
$\Rightarrow (\widehat{X}_t, \widehat{Y}_t) \equiv 0$, so that there is no $1/\epsilon^\beta = 1/t^{\beta/2}$ term in the exponential.

Corollary: tail asymptotics

Corollary (tail expansion in Stein-Stein) $dY_t = (a + bY_t)dt + cdW_t^H$

Consider the Stein-Stein model (3) with $X_0 = 0$, $Y_0 = y_0 > 0$. Then as $x \to \infty$,

$$f_X(t,x) = \exp\left(-c_1 x + c_2 x^{1/2}\right) \frac{1}{x^{1/2}} \left(c_0 + \mathcal{O}\left(x^{1/2}\right)\right)$$

where $c_1 := \Lambda(1)$, $c_2 := \widehat{X}_t \Lambda'(1)$.

Note that the expression on the RHS is independent of the Hurst-parameter!

Corollary: tail asymptotics

Corollary (tail expansion in Stein-Stein) $dY_t = (a + bY_t)dt + cdW_t^H$

Consider the Stein-Stein model (3) with $X_0 = 0$, $Y_0 = y_0 > 0$. Then as $x \to \infty$,

$$f_X(t, x) = \exp\left(-c_1 x + c_2 x^{1/2}\right) \frac{1}{x^{1/2}} \left(c_0 + \mathcal{O}\left(x^{1/2}\right)\right)$$

where $c_1 := \Lambda(1)$, $c_2 := \widehat{X}_t \Lambda'(1)$.

Note that the expression on the RHS is independent of the Hurst-parameter!

Proof: Consider $(X_T^\epsilon, Y_T^\epsilon) := (\epsilon^{2H} X_T, \epsilon^H Y_t)$ with $X_0^\epsilon = \epsilon^{2H} X_0$ and $Y_0^\epsilon = \epsilon^H Y_0$.

$$dX_t^\epsilon = -\frac{(Y_t^\epsilon)^2}{2} dt + \epsilon^H Y_t^\epsilon dW_t, \qquad dY_t^\epsilon = (a\epsilon^H + bY_t^\epsilon)dt + \epsilon^H c dW_t^H, \qquad (5)$$

Note that $X_t^\varepsilon \overset{\Delta}{=} \varepsilon^{2H} X_t$. $\Rightarrow \mathbb{P}(X_t^\varepsilon \geq y) = \mathbb{P}(X_t \geq y/\varepsilon^{2H})$, $\Rightarrow$
$f_X(t, y/\varepsilon^{2H}) = \varepsilon^{2H} f_\varepsilon(t, y)$. Take $y = 1$, that is $x := \varepsilon^{-2H}$. By the theorem,
$f_\varepsilon(t, 1) \approx \exp\left(-\frac{\Lambda(1)}{\varepsilon^2 H} + \dots\right) \frac{1}{\varepsilon^H}$, hence
$f_X(t, x) \approx \exp\left(-\frac{\Lambda(1)}{\varepsilon^{2H}} + \dots\right) \varepsilon^H = \exp\left(-\Lambda(1)x + \dots\right) \frac{1}{x^{1/2}}$.

From density to implied volatility: small-time

Recall the Black-Scholes density expansion:

$$f_{\mathrm{BS}}(t,x) \sim t^{-1/2} \exp\left(-\frac{1}{2t}\left(\frac{x-x_0}{\sigma}\right)^2\right), \quad \text{as } t \to 0, \text{ for any } x \in \mathbb{R}.$$

(We normalise the spot here, so that $x_0 = 0$).

Our theorem (corollary) says that in the Stein-Stein model (3), we have

$$f_{\mathrm{X}}(t,x) \sim \mathrm{cst}\ t^{-H} \exp\left(-\frac{d^2(x_0,y_0;x)}{2t^{2H}}\right), \quad \text{as } t \to 0.$$

From density to implied volatility: small-time

Recall the Black-Scholes density expansion:

$$f_{\mathrm{BS}}(t,x) \sim t^{-1/2} \exp\left(-\frac{1}{2t}\left(\frac{x-x_0}{\sigma}\right)^2\right), \quad \text{as } t \to 0, \text{ for any } x \in \mathbb{R}.$$

(We normalise the spot here, so that $x_0 = 0$).

Our theorem (corollary) says that in the Stein-Stein model (3), we have

$$f_{\mathrm{X}}(t,x) \sim \text{cst } t^{-H} \exp\left(-\frac{d^2(x_0,y_0;x)}{2t^{2H}}\right), \quad \text{as } t \to 0.$$

Matching the leading-orders gives

$$\sigma_{\mathrm{BS}}(t,x) \sim \frac{|x|}{d(x_0,y_0;x)} t^{H-1/2} \quad \text{as } t \to 0.$$

From density to implied volatility: tails

Recall the Black-Scholes density expansion:

$$f_{\mathrm{BS}}(t,x) \sim \exp\left(-\frac{x^2}{2\sigma^2 t} - \frac{x}{4}\right) \quad \text{as } x \to \infty, \text{ for any } t > 0.$$

Our theorem (corollary) says that in the Stein-Stein model (3), we have

$$f_{\mathrm{X}}(t,x) \sim \frac{\text{cst}}{x^{1/2}} \exp\left(-c_1 x + c_2\sqrt{x}\right), \quad \text{as } x \to \infty.$$

From density to implied volatility: tails

Recall the Black-Scholes density expansion:

$$f_{\mathrm{BS}}(t,x) \sim \exp\left(-\frac{x^2}{2\sigma^2 t} - \frac{x}{4}\right) \quad \text{as } x \to \infty, \text{ for any } t > 0.$$

Our theorem (corollary) says that in the Stein-Stein model (3), we have

$$f_{\mathrm{X}}(t,x) \sim \frac{\mathrm{cst}}{x^{1/2}} \exp\left(-c_1 x + c_2\sqrt{x}\right), \quad \text{as } x \to \infty.$$

Matching the leading-orders gives

$$-c_1 x + c_2\sqrt{x} \sim -\frac{x^2}{2\sigma^2 t} - \frac{x}{4},$$

and we recover Roger Lee's formula independently of the Hurst exponent in (3).

Outlook: Moderate regimes

- Moderate Regimes (in the sense of Friz-Gerhold-Pinter '16) interpolate between out-of-the-money calls with fixed strike $\left(\log \frac{K}{S_0}\right) = k > 0$ and at-the-money $k = 0$ calls: Now $k_t = ct^\theta \Rightarrow$ MOTM (for $0 < \theta < \frac{1}{2}$) and AATM (for larger θ)
- Reflects market data: options closer expiry $\Rightarrow$ strikes closer to the money first observed by Mijatović-Tankov on FX markets
- The moderate regime (MOTM) permits explicit computations for the rate function $\Lambda(k)$ in terms of the model parameters
  Moderate deviations$\Rightarrow$ Advantage over OTM (large deviations) case where the $\Lambda(k)$ often related to geodesic distance problems and not explicitly available.
- MOTM expansions naturally involve quantities very familiar to practitioners, notably spot (implied) volatility, implied volatility skew ...
- In some cases (fractional volatility models) the scaling θ permits a fine-tuning to understand the behavior and derivatives of the energy function.

Moderate regimes for rough volatility

Rescalings $\Longrightarrow$ We tacitly agreed to consider $\mathbb{P}\left(X_t \approx t^{1/2-H}x\right)$. Now it is only a small step to consider instead (for some suitable small $\theta > 0$)

$$\mathbb{P}\left(X_t \approx t^{1/2-H+\theta}x\right).$$

Moderate regimes for rough volatility

Rescalings $\implies$ We tacitly agreed to consider $\mathbb{P}\left(X_t \approx t^{1/2-H}x\right)$. Now it is only a small step to consider instead (for some suitable small $\theta > 0$)

$$\mathbb{P}\left(X_t \approx t^{1/2-H+\theta}x\right).$$

Theorem (Bayer-Friz-Gulisashvili-H-Stemper)

Consider a moderately out-of-the-money call $k_t = x^{1/2-H+\theta}$; $\theta \in (0, H)$ resp. $\theta \in (0, \frac{2H}{3})$. Then as $t \to 0$, the following (non-Markovian extention of Osajima-energy-expansion) holds

$$\log c(k_t, t) \approx \frac{1}{2}\Lambda''(0)\frac{x^2}{t^{2H-2\theta}} + \frac{1}{6}\Lambda'''(0)\frac{x^3}{t^{2H-3\theta}},$$

where we have explicit expressions: $\Lambda''(0) = \frac{1}{\sigma_0})$ and $\Lambda'''(0) = -\rho\frac{6\sigma_0'}{\sigma_0^4}\langle K, 1\rangle$.

Here K denotes the Volterra kernel and $\langle K, 1\rangle := \int_0^1\int_0^t K(t,s)dsdt$.

Notations

- $\mathcal{H}$: absolutely continuous paths $[0,T] \to \mathbb{R}^2$ starting at 0 such that $\left\|\dot{\mathrm{h}}\right\|_{\mathcal{H}}^2 < \infty$.
- $\mathcal{H}_H := K_H \mathcal{H}$ and $\mathrm{k} := K_H \mathrm{h}$, where K_H denotes the Volterra kernel.
- For fixed $(x_0, y_0) \in \mathbb{R}^2$, ϕ^{k} is the (unique) ODE solution to

$$\dot{\phi}_t^{\mathrm{k}} = \sigma_0\left(\phi_t^{\mathrm{k}}\right)\mathrm{d}t + \sum_{i=1}^{m} \sigma_i\left(\phi_t^{\mathrm{h}}\right)\mathrm{dk}_t^i, \quad \phi_0^{\mathrm{k}} = (x_0, y_0).$$

- Denote $\psi^{\mathrm{k}} := \Pi_1 \phi^{\mathrm{k}}$ its projection on to the first coordinate X.
- $\mathcal{K}_{\mathrm{a}} := \left\{\mathrm{k} \in \mathcal{H}_H : \psi_T^{\mathrm{k}} = a \in \mathbb{R}\right\} \neq \emptyset$ ("by Hörmander condition").
- $\Lambda(\mathrm{a}) := \inf\left\{\frac{1}{2}\|\mathrm{k}\|_H^2 : \mathrm{k} \in \mathcal{K}_{\mathrm{a}}\right\}$.

Proof of the theorem 1

$$dX_t = -\epsilon^{2H+1}\frac{1}{2}Y_t^2 dt + \epsilon^{2H} Y_t dW_t, \qquad dY_t = \epsilon^{2H} dW_t^H,$$

with the same initial condition $X_0 = Y_0 = 0$.

$$\text{Density: } f_\varepsilon(T,x) = \exp\left[-\frac{\Lambda(x)}{\varepsilon^{4H}} + \frac{\Lambda'(x)\widehat{X}_T}{\varepsilon^{2H}}\right]\varepsilon^{-2H}\left(c_0 + \mathcal{O}(\varepsilon^{2H})\right).$$

Proof of the theorem 1

$$dX_t = -\epsilon^{2H+1}\frac{1}{2}Y_t^2 dt + \epsilon^{2H} Y_t dW_t, \qquad dY_t = \epsilon^{2H} dW_t^H,$$

with the same initial condition $X_0 = Y_0 = 0$.

$$\text{Density: } f_\varepsilon(T,x) = \exp\left[-\frac{\Lambda(x)}{\varepsilon^{4H}} + \frac{\Lambda'(x)\widehat{X}_T}{\varepsilon^{2H}}\right]\varepsilon^{-2H}\left(c_0 + \mathcal{O}(\varepsilon^{2H})\right).$$

Proof: Take $x \in \mathbb{R}$ and a C^∞-bounded function F such that $F(x) = 0$.

$$f_\varepsilon(T,x)\mathrm{e}^{-F(x)/\varepsilon^{4H}} = \frac{1}{2\pi\varepsilon^{2H}}\int_{\mathbb{R}}\mathbb{E}\left\{\exp\left[\mathrm{i}(\zeta,0)\cdot\left(\frac{X_T^\varepsilon - (x,0)}{\varepsilon^{2H}}\right) - \frac{F(X_T^\varepsilon)}{\varepsilon^{4H}}\right]\right\}\mathrm{d}\zeta.$$

Choose F such that $F(\cdot) + \Lambda_{x_0}(\cdot)$ has a non-degenerate minimum at z. This implies that $\mathrm{k} \mapsto F(\phi_T^{\mathrm{k}}(x_0,y_0)) + \frac{1}{2}\|\mathrm{k}\|^2_{\mathcal{H}_H}$ has a non-degenerate minimum at $\mathrm{k}_0 \in H$.

(For instance $F(z) = \lambda|z-x|^2 - [\Lambda_{x_0,y_0}(z) - \Lambda_{x_0,y_0}(x)]$ with $\lambda > 0$).

Proof of the theorem 2

Replace $\varepsilon^{2H}\mathrm{d}B$ ($B := (W, W^H)$) in the SDE by $\varepsilon^{2H}\mathrm{d}W + \dot{\mathrm{k}}_0$.
Call the corresponding Girsanov-transformed process $\widetilde{Z}_t^\varepsilon = (\widetilde{X}_t^\varepsilon, \widetilde{Y}_t^\varepsilon)$:

$$d\widetilde{X}^\varepsilon = -\varepsilon^{2H+1}\frac{1}{2}\widetilde{Y}^2 dt + \widetilde{Y}^\varepsilon(\varepsilon^{2H}dW_t + (\dot{\mathrm{k}}_0)_1), \qquad d\widetilde{Y} = \epsilon^{2H}dW_t^H + (\dot{\mathrm{k}}_0)_2.$$

Girsanov factor

$$\mathcal{G} = \exp\left(-\frac{1}{\epsilon^{2H}}\int_0^T \psi(\mathrm{k}_0)_t dB_t - \frac{1}{2\epsilon^{4H}}\|\mathrm{k}_0\|^2_{\mathcal{H}_H}\right).$$

Therefore

$$\begin{aligned} f(x,T)e^{-F(x)/4\varepsilon^{4H}} &= \frac{1}{2\pi\epsilon^{2H}}\int_{\mathbb{R}}\mathbb{E}\left[e^{\epsilon^{2H}i\zeta(\widetilde{X}_T - x) - \varepsilon^{-4H}F(\widetilde{X}_T)}\mathcal{G}\right]d\zeta \\ &= \frac{1}{2\pi\varepsilon^{2H}}\int_{\mathbb{R}}\mathbb{E}\left[e^{(*)}\right]d\zeta \end{aligned}$$

where

$$(*) = \varepsilon^{2H}i\zeta(\widetilde{X}_T - x) - \epsilon^{-4H}F(\widetilde{X}_T) - \epsilon^{-2H}\int_0^T \psi(\gamma)_t dB_t - \epsilon^{-4H}\frac{1}{2}\|\gamma\|^2_{1/2,H}.$$

Proof of the theorem 2

Replace $\varepsilon^{2H}\mathrm{d}B$ ($B := (W, W^H)$) in the SDE by $\varepsilon^{2H}\mathrm{d}W + \dot{\mathrm{k}}_0$.
Call the corresponding Girsanov-transformed process $\widetilde{Z}_t^\varepsilon = (\widetilde{X}_t^\varepsilon, \widetilde{Y}_t^\varepsilon)$:

$$d\widetilde{X}^\varepsilon = -\varepsilon^{2H+1}\frac{1}{2}\widetilde{Y}^2 dt + \widetilde{Y}^\varepsilon(\varepsilon^{2H} dW_t + (\dot{\mathrm{k}}_0)_1), \qquad d\widetilde{Y} = \epsilon^{2H} dW_t^H + (\dot{\mathrm{k}}_0)_2.$$

Girsanov factor

$$\mathcal{G} = \exp\left(-\frac{1}{\epsilon^{2H}}\int_0^T \psi(\mathrm{k}_0)_t dB_t - \frac{1}{2\epsilon^{4H}}\|\mathrm{k}_0\|^2_{\mathcal{H}_H}\right).$$

By a stochastic Taylor expansion of $\widetilde{Z}_t^\varepsilon = (\widetilde{X}_t^\varepsilon, \widetilde{Y}_t^\varepsilon)$ for $\varepsilon^{2H} \to 0$,

$$\exp\left(\frac{-F\left(\widetilde{X}_t^\varepsilon\right)}{\varepsilon^{4H}}\right) = \exp\left[\frac{-1}{\varepsilon^{4H}}\left(F(x) - \varepsilon^{2H}\int_0^T \psi(\mathrm{k}_0)_t \mathrm{d}B_t - \varepsilon^{2H}\widehat{X}_T \cdot \Lambda'_{x_0}(x) + \mathcal{O}(\varepsilon^{4H})\right)\right]$$

Proof of the theorem 2

Replace $\varepsilon^{2H}\mathrm{d}B$ ($B := (W, W^H)$) in the SDE by $\varepsilon^{2H}\mathrm{d}W + \dot{\mathrm{k}}_0$.
Call the corresponding Girsanov-transformed process $\widetilde{Z}_t^\varepsilon = (\widetilde{X}_t^\varepsilon, \widetilde{Y}_t^\varepsilon)$:

$$d\widetilde{X}^\varepsilon = -\varepsilon^{2H+1}\frac{1}{2}\widetilde{Y}^2 dt + \widetilde{Y}^\varepsilon(\varepsilon^{2H}dW_t + (\dot{\mathrm{k}}_0)_1), \qquad d\widetilde{Y} = \epsilon^{2H}dW_t^H + (\dot{\mathrm{k}}_0)_2.$$

Girsanov factor

$$\mathcal{G} = \exp\left(-\frac{1}{\epsilon^{2H}}\int_0^T \psi(\mathrm{k}_0)_t dB_t - \frac{1}{2\epsilon^{4H}}\|\mathrm{k}_0\|^2_{\mathcal{H}_H}\right).$$

By a stochastic Taylor expansion of $\widetilde{Z}_t^\varepsilon = (\widetilde{X}_t^\varepsilon, \widetilde{Y}_t^\varepsilon)$ for $\varepsilon^{2H} \to 0$,

$$\exp\left(\frac{-F\left(\widetilde{X}_t^\varepsilon\right)}{\varepsilon^{4H}}\right) = \exp\left[\left(\frac{1}{\varepsilon^{2H}}\int_0^T \psi(\mathrm{k}_0)_t dB_t + \frac{1}{\varepsilon^{2H}}\widehat{X}_T \cdot \Lambda'_{x_0}(x) + \mathcal{O}(1)\right)\right].$$

The rest of the proof follows Ben Arous' proof for X_T^ε.

Thank you!