QuantMinds International

Intercontinential O2 London, November 18, 2024

Rough volatility workshop

Lecture 3: Affine forward variance models and their microstructural foundation

Jim Gatheral
Department of Mathematics

Baruch COLLEGE
The City University of New York

Outline of Lecture 3

- The microstructural foundation of affine forward variance models
- Characteristic function methods
  - Option pricing
  - The ATM skew
  - The skew-stickiness ratio
- Diamonds and the forest expansion
- Moment computations

A microstructural foundation for affine stochastic volatility models

- [Jaisson and Rosenbaum][10] first showed that affine stochastic volatility models could arise as limits of Hawkes process-based models of order flow.
- In the following, we both generalize and attempt to give intuition for their argument.

Hawkes processes

- Dating from the 1970's, Hawkes processes are jump processes where the jump arrival rate is self-exciting.

- One of the first applications was to the modeling of earthquakes.

The Hawkes process-based microstructure model of Jaisson and Rosenbaum

[Jaisson and Rosenbaum][10] consider the following simple model of price formation:

- Order arrivals are modeled as a counting process
  - Buy order arrivals cause the price to increase
  - Sell order arrivals cause the price to decrease
  - All orders are unit size
- The order arrival process is self-exciting
  - The price process is a bivariate Hawkes process.

The stock price process

Specifically, with $X_t = \log S_t$,

$$dX_t = m_X dt + dN_t^+ - dN_t^-$$

where $N^\pm$ are counting processes with arrival rates $\lambda_t^\pm$, and m_X is determined by the martingale condition on $S = e^X$.

The order arrival rate process

$$\lambda_t = \mu + \int_0^t \varphi(t-s)\, d\mathbf{N}_s.$$

where $\lambda = \{\lambda^+, \lambda^-\}$ and $\mathbf{N} = \{N^+, N^-\}$. The kernel φ is a 2×2 matrix.

- The order arrival process is self-exciting.
  - As orders arrive, the order arrival rate increases.
  - In the absence of new orders, the order arrival rate decays according to some Hawkes kernel φ.
- Jaisson and Rosenbaum show that that in a suitable scaling limit, and with a suitable choice of the kernel φ, this model tends to the rough Heston model.

Affine forward intensity (AFI) models

- In analogy to stochastic volatility models in forward variance form, [Gatheral and Keller-Ressel][9] define the forward intensity model

$$dX_t = -\lambda_t m_X dt + dJ_t^+ - dJ_t^-, \tag{1}$$
$$d\xi_t(T) = \kappa(T-t)\left(\gamma^+ d\widetilde{J}_t^+ + \gamma^- d\widetilde{J}_t^-\right). \tag{2}$$

where κ is an integrable, decreasing non-zero kernel.

- $\gamma^\pm$ are positive constants
- jumps can have various sizes; the jump size measures are $\zeta_\pm$
- m_X is determined by the martingale condition on $S = e^X$

- The $\widetilde{J}_t^\pm$ denote the *compensated* order flow processes, i.e.

$$\widetilde{J}_t^\pm := J_t^\pm - m_\pm \int_0^t \lambda_s ds,$$

where

$$m_\pm = \int_{\mathbb{R}_{\geqslant 0}} x\, \zeta_\pm(dx).$$

Variance and jump intensity

Denote the variance per unit time of the process X_t by V_t. Then

$$V_t\, dt = \mathrm{var}[dJ_t^+ - dJ_t^-] = \lambda_t \left\{v^+ + v^-\right\} dt =: \lambda_t\, v_J\, dt,$$

where

$$v^\pm = \int_{\mathbb{R}_{\geqslant 0}} x^2\, \zeta_\pm(dx) - m_\pm^2$$

are the variance of positive and negative jump sizes respectively.

Continuing the analogy with stochastic volatility, $\xi_t(u)$ is linked to V_t by

$$\xi_t(u) = \mathbb{E}_t\left[V_u\right]. \tag{3}$$

Setting

$$J_t^X = J_t^+ - J_t^-, \qquad \widetilde{J}_t^v = \gamma^+ \widetilde{J}_t^+ + \gamma^- \widetilde{J}_t^-, \tag{4}$$

the affine forward intensity (AFI) model may be rewritten as

$$\begin{aligned} dX_t &= -\lambda_t m_X dt + dJ_t^X, \\ d\xi_t(T) &= \kappa(T-t) d\widetilde{J}_t^v. \end{aligned} \tag{5}$$

High-frequency limit of the AFI model

Consider new processes J^ϵ such that

$$\lambda^\epsilon = \frac{1}{\epsilon}\,\lambda; \quad \zeta^\epsilon(dx) = \zeta\left(\frac{dx}{\sqrt{\epsilon}}\right).$$

Thus in the limit $\epsilon \to 0$,

- jump sizes are very small and jumps are very frequent.
- the martingale component of dX_t may be approximated by $\sqrt{V_t}\,dZ_t$
- $d\widetilde{J}^v_t$ may be approximated by dY_t for some diffusion process Y.

High frequency limit of the AFI model

In the limit, we obtain

$$\begin{aligned} dX_t &= -\frac{1}{2}\,V_t\,dt + \sqrt{V_t}\,dZ_t, \\ d\xi_t(T) &= \kappa(T-t)\,dY_t, \end{aligned}$$

where

$$\begin{aligned} \mathrm{var}[dY_t] = \mathrm{var}[d\tilde{J}^V_t] &= \lambda_t\,\left[{\gamma^+}^2\,v^+ + {\gamma^-}^2\,v^-\right]\,dt \\ &= V_t\,\left[\frac{{\gamma^+}^2\,v^+ + {\gamma^-}^2\,v^-}{v^+ + v^-}\right]\,dt. \end{aligned}$$

Then

$$d\xi_t(T) = \eta\,\kappa(T-t)\,\sqrt{V_t}\,dW_t$$

where

$$\eta^2 = \frac{{\gamma^+}^2\,v^+ + {\gamma^-}^2\,v^-}{v^+ + v^-}.$$

As for the correlation between dZ_t and dW_t, we first compute

$$\mathbb{E}\left[dJ^+_t\,d\tilde{J}^+_t\right] = \lambda_t\,v^+\,dt; \quad \mathbb{E}\left[dJ^-_t\,d\tilde{J}^-_t\right] = \lambda_t\,v^-\,dt$$

so

$$\begin{aligned} \mathbb{E}\left[dX_t\,d\tilde{J}^v_t\right] &= \lambda_t\,(\gamma^+\,v^+ - \gamma^-\,v^-)\,dt \\ &= \mathbb{E}\left[\sqrt{V_t}\,dZ_t\,\eta\,\sqrt{V_t}\,dW_t\right] =: \rho\,\eta\,V_t\,dt, \end{aligned}$$

where

$$\rho = \frac{1}{\sqrt{v^+ + v^-}} \, \frac{\gamma^+ v^+ - \gamma^- v^-}{\sqrt{\gamma^{+2} v^+ + \gamma^{-2} v^-}}.$$

Example: The bivariate Hawkes process of of Jaisson and Rosenbaum

Consider the case of a bivariate Hawkes process (J^+, J^-) with unit jump size (i.e., $\zeta_\pm(dx) = \delta_1(dx)$). Then in the above limit, as $\epsilon \to 0$, the process converges to

$$\begin{aligned} dX_t &= -\frac{1}{2} V_t \, dt + \sqrt{V_t} \, dZ_t, \\ d\xi_t(T) &= \eta \sqrt{V_t} \, \kappa(T-t) \, dW_t, \end{aligned}$$

where $d\langle Z, W \rangle_t = \rho \, dt$ and

$$\eta^2 = \frac{1}{2}\left[\gamma^{+2} + \gamma^{-2}\right]; \quad \rho = \frac{\gamma^+ - \gamma^-}{\sqrt{2\left(\gamma^{+2} + \gamma^{-2}\right)}}.$$

Near instability of Hawkes kernel in the limit

- So far, we have shown how AFV models arise naturally as limits of AFI models.
- Now we show that in order to get stochastic (as opposed to constant) volatility, the AFI model Hawkes process needs to be nearly unstable.

Consider the (generalized) Hawkes process

$$\begin{aligned} \lambda_t &= \mu + \int_0^t \varphi(t-s) \, dJ_s^v \\ &= \mu + \hat{\gamma} \int_0^t \varphi(t-s) \, \lambda_s \, ds + \int_0^t \varphi(t-s) \, d\tilde{J}_s^v \end{aligned}$$

where $\hat{\gamma} = \gamma^+ m_+ + \gamma^- m_-$.

Following [Bacry et al.][2], we rewrite this last equation symbolically as

$$\lambda = \mu + \hat{\gamma} \, (\varphi \star \lambda) + \varphi \star d\tilde{J}^v.$$

Rearranging gives

$$(1 - \hat{\gamma} \, \varphi\star)\lambda = \mu + \varphi \star d\tilde{J}^v$$

and applying the Laplace transform gives

$$(1-\hat{\gamma}\,\hat{\varphi})\hat{\lambda} = \hat{\mu} + \hat{\varphi}\,\widehat{d\tilde{J}^v}.$$

which may be rearranged as

$$\hat{\lambda} = \hat{\mu} + \hat{\psi}\,\hat{\mu} + \frac{1}{\hat{\gamma}}\,\hat{\psi}\,\widehat{\tilde{J}^v}$$

where

$$\hat{\psi} = \frac{\hat{\gamma}\hat{\varphi}}{1-\hat{\gamma}\,\hat{\varphi}}.$$

Then

$$v_J\,\hat{\lambda} = v_J\,\hat{\mu} + \hat{\gamma}\,\hat{\kappa}\,\hat{\mu} + \hat{\kappa}\,\widehat{\tilde{J}^v}$$

where

$$\hat{\kappa} = \frac{v_J}{\hat{\gamma}}\,\hat{\psi} = \frac{v_J\hat{\varphi}}{1-\hat{\gamma}\,\hat{\varphi}}.$$

Inverting the Laplace transform, and recalling that $V_t = v_J\,\lambda_t$, we obtain

$$V_u = v_J\,\mu + \hat{\gamma}\,\mu\int_0^u \kappa(u-s)\,ds + \int_0^u \kappa(u-s)\,d\tilde{J}^v_s.$$

Computing the conditional expectation wrt $\mathcal{F}_t$,

$$\begin{aligned}\xi_t(u) &= \mathbb{E}_t\left[V_u\right]\\ &= v_J\,\mu + \hat{\gamma}\,\mu\int_0^u \kappa(u-s)\,ds + \eta\int_0^t \kappa(u-s)\,\sqrt{V_s}\,dW_s\end{aligned}$$

and so $d\xi_t(u) = \kappa(u-t)\,d\tilde{J}^v_t$, the dynamics of an AFI model.

Now

$$\hat{\kappa} = \frac{v_J\hat{\varphi}}{1-\hat{\gamma}\,\hat{\varphi}} \implies \hat{\varphi} = \frac{\hat{\kappa}}{v_J + \hat{\gamma}\,\hat{\kappa}}.$$

Recall that the kernel of our generalized Hawkes process is $\hat{\gamma}\,\hat{\varphi}$. The stability condition is then

$$\hat{\gamma}\int_{\mathbb{R}_{\geqslant 0}} \varphi(\tau)\,d\tau = \hat{\gamma}\,\hat{\varphi}(0) = \frac{\hat{\gamma}\,\hat{\kappa}}{v_J + \hat{\gamma}\,\hat{\kappa}} \to 1 \text{ as } \epsilon \to 0$$

since in that limit, $v_J \sim \epsilon$ and $\hat{\gamma} \sim \sqrt{\epsilon}$.

Conversely, $\hat{\gamma}\,\hat{\varphi}(0) \to a < 1$ as $\epsilon \to 0$ only if $\kappa \sim \sqrt{\epsilon}$. Then in the limit, $\kappa \to 0$ and volatility is deterministic.

> **Near instability**
>
> The high frequency limit of the AFI model is a non-trivial AFV model if and only if the Hawkes process is nearly unstable.

Affine processes

The following explanation is due to Martin Keller-Ressel:

> An *affine process* can be described as a Markov process whose log-characteristic function is an affine function of its initial state vector.

And here's a definition of the word *affine* from Wikipedia:

> In geometry, an affine transformation or affine map or an affinity (from the Latin, *affinis*, "connected with") between two vector spaces (strictly speaking, two affine spaces) consists of a linear transformation followed by a translation:
>
> $$x \mapsto A\,x + b$$

Affine CGF

Let $X_t = \log S_t$. According to Definition 2.2 of [Gatheral and Keller-Ressel][9], a forward variance model has an *affine cumulant generating function* determined by $g(t;u)$, if its conditional cumulant generating function is of the form

$$\log \mathbb{E}_t\left[e^{u(X_T - X_t)}\right] = \int_t^T g(T-s;u)\,\xi_t(s)ds.$$

for all $u \in [0,1]$, $0 \le t \le T$ and $g(.\,;u)$ is $\mathbb{R}_{\leqslant 0}$-valued and continuous on $[0,T]$ for all $T > 0$ and $u \in [0,1]$.

- The restriction $u \in [0,1]$ is for mathematical convenience. We will later allow complex u.

When is a forward variance model affine?

Theorem 2.4 of [Gatheral and Keller-Ressel][10] states that a forward variance model has an affine CGF if and only if it takes the form

$$\frac{dS_t}{S_t} = \sqrt{V_t}\, dZ_t$$
$$d\xi_t(u) = \sqrt{V_t}\, \kappa(u-t)\, dW_t$$

for some deterministic, non-negative decreasing kernel κ, which satisfies $\int_0^T \kappa(r)dr < \infty$ for all $T > 0$.

Moreover, $g(.\,; u) : \mathbb{R}_{\geqslant 0} \to \mathbb{R}_{\leqslant 0}$ in the definition (1) of the CGF is the unique global continuous solution of the convolution Riccati equation

$$g = \frac{1}{2}(u^2 - u) + \rho\, u\, (\kappa \star g) + \frac{1}{2}\, (\kappa \star g)^2.$$

where we have dropped the arguments for clarity.

- That is, $g = g(\tau; u)$ and $(\kappa \star g) = (\kappa \star g)(\tau; u)$

Derivation of the Riccati equation

From the definition of the CGF,

$$M_t = \mathbb{E}_t\left[e^{u\, X_T}\right] = \exp\left\{u\, X_t + \int_t^T \xi_t(s)\, g(T-s; u)\, ds\right\} =: \exp\{u\, X_t + G_t\}$$

is a conditional expectation and thus a martingale in t.

Applying Itô's Lemma to M gives

$$\frac{dM_t}{M_t} = u\, dX_t + dG_t + \frac{u^2}{2}\, d\langle X\rangle_t + \frac{1}{2}\, d\langle G\rangle_t + u\, d\langle X, G\rangle_t.$$

Now

$$\begin{aligned} dX_t &= -\frac{1}{2}\, V_t\, dt + \sqrt{V_t}\, dZ_t \\ dG_t &= -\xi_t(t)\, g(T-t; u)\, dt + \int_t^T d\xi_t(s)\, g(T-s; u)\, ds \\ &= -V_t\, g(T-t; u)\, dt + \int_t^T \kappa(s-t)\, \sqrt{V_t}\, dW_t\, g(T-s; u)\, ds. \end{aligned}$$

We compute

$$
\begin{aligned}
d\langle X\rangle_t &= V_t\,dt \\
d\langle G\rangle_t &= V_t\,dt\left(\int_t^T \kappa(s-t)\,g(T-s;u)\,ds\right)^2 \\
d\langle X,G\rangle_t &= \rho\,V_t\,dt\int_t^T \kappa(s-t)\,g(T-s;u)\,ds.
\end{aligned}
$$

Imposing $\mathbb{E}\left[dM_t\right]=0$ and letting $\tau=T-t$ gives

$$
0 = V_t\,dt\left\{-\frac{1}{2}\,u+\frac{1}{2}\,u^2-g(\tau;u)+\rho\,u\,(\kappa\star g)(\tau;u)+\frac{1}{2}\,(\kappa\star g)(\tau;u)^2\right\}
$$

where the convolution integral is given by

$$
(\kappa\star g)(\tau;u)=\int_0^\tau \kappa(\tau-s)\,g(s;u)\,ds.
$$

- It is almost obvious why the CGF is affine if an only if the forward variance process is of the form $d\xi_t(u)=\sqrt{V_t}\,\kappa(u-t)\,dW_t$.

The convolution Riccati equation

Rearranging gives

$$
g(\tau;u)=\frac{1}{2}\,u(u-1)+\rho\,u\,(\kappa\star g)(\tau;u)+\frac{1}{2}\,(\kappa\star g)(\tau;u)^2,
$$

as required.

Example: The rough Heston model (with $\lambda=0$)

In this case, with $\alpha=H+\frac{1}{2}$, $\kappa(\tau)=\frac{\eta}{\Gamma(\alpha)}\,\tau^{\alpha-1}$ and

$$
\begin{aligned}
\eta\,h(\tau;u):=(\kappa\star g)(\tau;u)&=\frac{\eta}{\Gamma(\alpha)}\int_0^\tau(\tau-s)^{\alpha-1}\,g(s;u)\,ds \\
&=\eta\,I^\alpha g(\tau;u).
\end{aligned}
$$

Inverting this gives $g(\tau;u)=D^\alpha h(\tau;u)$.

The convolution integral Riccati equation then reads

$$
D^\alpha h(\tau;u)=\frac{1}{2}\,u\,(u-1)+\rho\,\eta\,u\,h(\tau;u)+\frac{1}{2}\,\eta^2\,h(\tau;u)^2,
$$

consistent with [El Euch and Rosenbaum][4].

An aside: Fractional calculus

Define the fractional integral and differential operators:

$$I^\alpha f(t) = \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} f(s)\, ds; \quad D^\alpha f(t) = \frac{d}{dt} I^{1-\alpha} f(t).$$

The fractional integral is a natural generalization of the ordinary integral using the Cauchy formula for repeated integration:

$$\begin{aligned} I^n f(t) &:= \int_0^t dt_1 \int_0^{t_1} \dots dt_{n-1} \int_0^{t_{n-1}} f(t_n)\, dt_n \\ &= \frac{1}{(n-1)!} \int_0^t (t-s)^{n-1} f(s)\, ds. \end{aligned}$$

The induction step

$$\begin{aligned} I^{n+1} f(t) &:= \int_0^t I^n f(s)\, ds \\ &= \frac{1}{(n-1)!} \int_0^t ds \int_0^s (s-r)^{n-1} f(r)\, dr \\ &= \frac{1}{(n-1)!} \int_0^t f(r)\, dr \int_r^t (s-r)^{n-1}\, ds \\ &= \frac{1}{n!} \int_0^t (t-r)^n f(r)\, dr. \end{aligned}$$

Characteristic function methods

- The primary reason that affine models are popular, is that they have easy to compute characteristic functions.
- Given the characteristic function, many computations become much easer.

Computing option prices from the characteristic function

[Lewis][11] originally derived the following formula, which is a special (but most useful) case of [Carr and Madan][3].

> **Formula (2.10) of Lewis**
>
> $$C_t(S,K,T) = S - \sqrt{SK}\frac{1}{\pi}\int_0^\infty \frac{du}{u^2+\frac{1}{4}} \operatorname{Re}\left[e^{-iuk}\varphi_t\left(T; u - i/2\right)\right]$$

with $k = \log\left(\frac{K}{S}\right)$.

Another way to compute implied volatility

The Lewis formula allows us to derive an elegant implicit expression for the Black-Scholes implied volatility of an option in any model for which the characteristic function is known.

Substituting the characteristic function for the Black-Scholes process into the Lewis formula gives

$$C_{BS}(S,K,\tau) = S - \sqrt{SK}\frac{1}{\pi}\int_0^\infty \frac{du}{u^2+\frac{1}{4}}\,\mathrm{Re}\left[e^{-iuk}e^{-\frac{1}{2}\left(u^2+\frac{1}{4}\right)\sigma_{BS}^2\tau}\right]$$

with $\tau = T - t$.

Then, from the definition of implied volatility, wlog setting $t = 0$, we must have

(1)

$$\int_0^\infty \frac{du}{u^2+\frac{1}{4}}\,\mathrm{Re}\left[e^{-iuk}\left(\varphi\left(T;u-i/2\right) - e^{-\frac{1}{2}\left(u^2+\frac{1}{4}\right)\Sigma(k,T)}\right)\right] = 0,$$

where we have defined the implied total variance $\Sigma(k,T) = \sigma_{\mathrm{BS}}(k,T)^2\,T$.

- Equation (1) gives us a simple but implicit relationship between the implied volatility surface and the characteristic function of the underlying stock process.
  - We may efficiently compute the structure of at-the-money implied volatility and the at-the-money volatility skew in terms of the characteristic function (at least numerically) without having to explicitly compute option prices.
- In practice, computing the option price and numerically inverting the Black-Scholes formula to get the implied volatility is faster.

Computing the at-the-money volatility skew

Following [The Volatility Surface][8], differentiating (1) with respect to k and evaluating at $k = 0$ gives

$$\int_0^\infty du\left\{\frac{u\,\mathrm{Im}\left[\varphi(T;u-i/2)\right]}{u^2+\frac{1}{4}} + \frac{1}{2}\left.\frac{\partial\Sigma(k,T)}{\partial k}\right|_{k=0} e^{-\frac{1}{2}\left(u^2+\frac{1}{4}\right)\Sigma(0,T)}\right\} = 0$$

Integrating the second term explicitly we get

$$\mathcal{S}(T) := \left.\frac{\partial \sigma_{BS}}{\partial k}\right|_{k=0} = -e^{\frac{1}{8}\sigma_{BS}^2 T}\sqrt{\frac{2}{\pi}}\frac{1}{\sqrt{T}}\int_0^\infty du\,\frac{u\,\mathrm{Im}\left[\varphi(T;u-i/2)\right]}{u^2+\frac{1}{4}}$$

Example: Black-Scholes

In the Black-Scholes case,

$$\mathrm{Im}\left[\varphi(T;u-i/2)\right] = \mathrm{Im}\left[e^{-\frac{1}{2}\left(u^2+1/4\right)\sigma^2 T}\right] = 0$$

Then

$$\mathcal{S}(T) = \left.\frac{\partial \sigma_{BS}(k,T)}{\partial k}\right|_{k=0} = 0 \;\; \forall T>0.$$

The skew-stickiness ratio (SSR)

- Denote the Black-Scholes implied volatility of an option by $\sigma_{\mathrm{BS,t}}(k,T)$
- Market makers, when updating option prices using the Black-Scholes formula, typically consider two effects.
  - First, the explicit spot effect
  $$\frac{\partial C_{\mathrm{BS,t}}}{\partial S_t}\,\delta S_t,$$
  - Second, the change in implied volatility conditional on a change in the spot
  $$\frac{\partial C_{\mathrm{BS,t}}}{\partial \sigma_{\mathrm{BS}}}\,\mathbb{E}\left[\delta\sigma_{\mathrm{BS,t}}|\delta S_t\right].$$
- Market makers can estimate this using a simple regression:
$$\delta\sigma_t(T) = \beta_t(T)\,\frac{\delta S_t}{S_t} + \text{noise} =: \beta_t(T)\,\delta X_t + \text{noise}.$$
- As before, we define the ATM volatility skew
$$\mathcal{S}_t(T) = \left.\frac{\partial}{\partial k}\sigma_{\mathrm{BS,t}}(k,T)\right|_{k=0}.$$
- Lorenzo Bergomi calls

$$\mathcal{R}_t(T) = \frac{\beta_t(T)}{\mathcal{S}_t(T)}. \tag{6}$$

the *skew-stickiness ratio* or *SSR*.

The SSR in terms of the characteristic function

In [Friz and Gatheral][6], we show that the SSR $\mathcal{R}_t(T)$ has the following representation in terms of the characteristic function.

Let $\psi = \log \varphi$. Then

$$\mathcal{R}_t(T) = -\frac{\int_{\mathbb{R}^+} \frac{da}{a^2+\frac{1}{4}}\, \Re\left[\rho\, D_t^\xi \psi_t(T; a - \mathrm{i}/2)\, \exp\{\psi_t(T; a - \mathrm{i}/2)\}\right]}{\int_{\mathbb{R}^+} \frac{a\, da}{a^2+1/4}\, \Im\left[\exp\{\psi_t(T; a - \mathrm{i}/2)\}\right]}. \tag{7}$$

Here, denoting the Fréchet derivative by δ,

$$D_t^\xi := \frac{1}{\sqrt{V_t}} \int_t^T du\, f_t(\xi)\, \kappa(u-t)\, \frac{\delta}{\delta \xi_t(u)}, \tag{8}$$

Sensitivity of the rough Heston SSR to model dynamics

- The rough Heston kernel takes the form

  $$\kappa(\tau) = \nu\, \tau^{\alpha-1}\, E_{\alpha,\alpha}\left(-\lambda\, \tau^\alpha\right),$$

  where $E_{\alpha,\alpha}(\cdot)$ is the Mittag-Leffler function and $\alpha = H + \frac{1}{2}$.

- We choose three values of the parameter $\lambda\, (= 0, 1, 2)$.

  - Then find values of H and ν such that the resulting parameter sets Π_0, Π_1 and Π_2 generate the almost identical 1-month, 3-month, 6-month, and 12-month smiles.

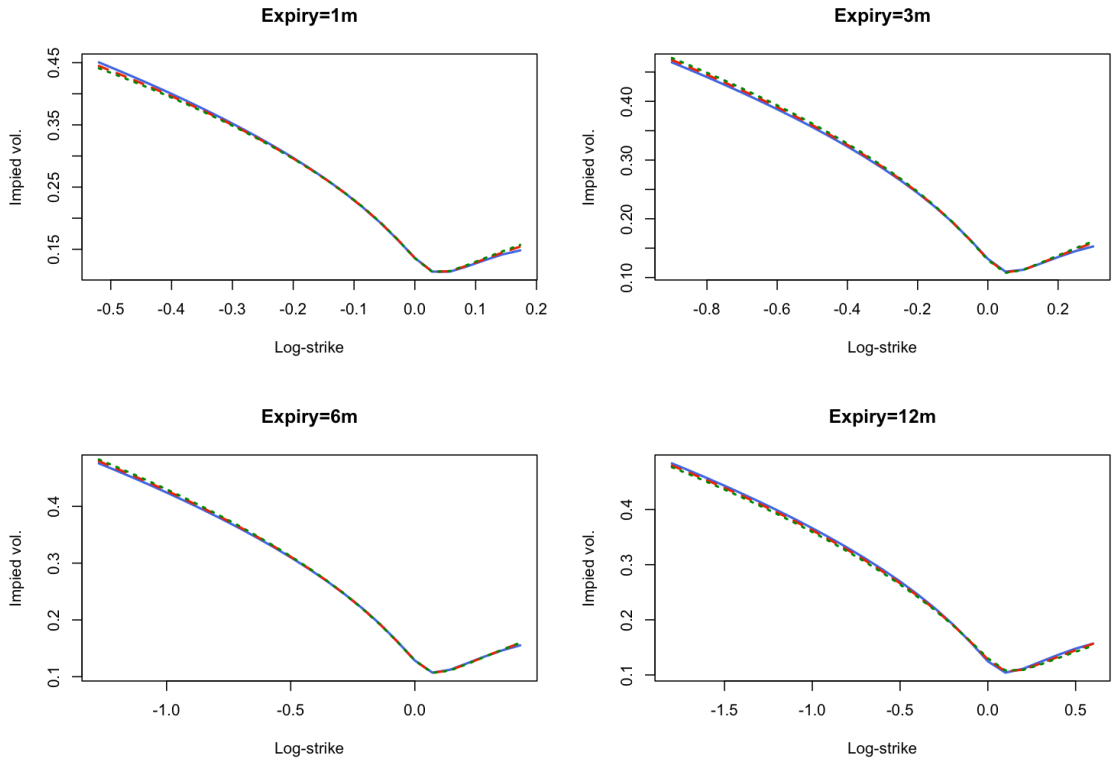

Figure 1. Almost identical rough Heston smiles generated by the parameter sets Π_0, Π_1, and Π_2, in blue, red and green respectively.

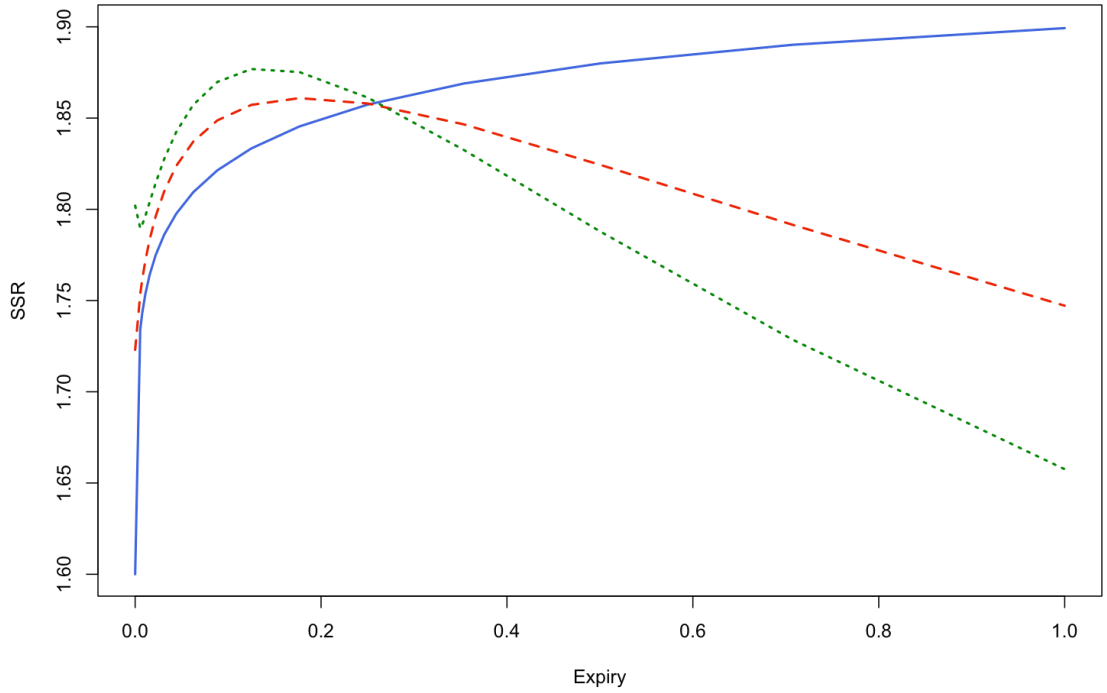

Figure 2. The blue, red, and green SSR plots correspond to parameter sets Π_0, Π_1, and Π_2 respectively.

Diamond products and the forest expansion

- The interest of the forest expansion is that the characteristic function may be expressed as a sum of forests of diamond trees.
  - For amy model, in complete generality.

The diamond product

Let A_t and B_t be continuous semimartingales. Then

$$(A \diamond B)_t(T) = \mathbb{E}_t \left[\int_t^T d\langle A, B \rangle_s \right].$$

- Note that $(A \diamond B)_t(T)$ is not in general a martingale in t (for fixed T).

- Thus for example, with $X = \log S$,

$$\begin{aligned}(X \diamond X)_t(T) &= \mathbb{E}_t \left[\int_t^T d\langle X, X \rangle_u \right] \\ &= \int_t^T \xi_t(u)\, du \\ &= M_t(T).\end{aligned}$$

Properties of the diamond product

- Commutative: $A \diamond B = B \diamond A$.
- Non-associative: $(A \diamond B) \diamond C \neq A \diamond (B \diamond C)$.
- $A \diamond B$ depends only on the respective martingale parts of A and B.
- $A \diamond B$ is in general not a martingale.

Variance swap dynamics

- As before

$$(X \diamond X)_t(T) = M_t(T) = \int_t^T \xi_t(u)\, du,$$

  where the $\xi_t(u)$ are forward variances.

- Thus

$$\begin{aligned} d(X \diamond X)_t(T) &= dM_t(T) \\ &= \int_t^T d\xi_t(u)\, du - \xi_t(t)\, dt \\ &= \int_t^T d\xi_t(u)\, du - V_t\, dt, \end{aligned}$$

a specific example where a diamond product is not a martingale.

- Financially, the variance contract gets seasoned, picking up the $V_t\, dt$ term.

Products of products

- We can easily make diamond products of diamond products:

$$\begin{aligned} (X \diamond (X \diamond X))_t(T) &= (X \diamond M)_t(T) \\ &= \mathbb{E}_t \left[\int_t^T d\langle X, M \rangle_s \right] \\ ((X \diamond X) \diamond (X \diamond X))_t(T) &= (M \diamond M)_t(T) \\ &= \mathbb{E}_t \left[\int_t^T d\langle M, M \rangle_s \right] \\ (S \diamond (X \diamond X))_t(T) &= (S \diamond M)_t(T) \\ &= \mathbb{E}_t \left[\int_t^T d\langle S, M \rangle_s \right], \end{aligned}$$

and so on.

- Note that drift terms (or bounded variation terms) do not contribute to diamond products.
  - This makes computation much easier.
- We will often drop the explicit dependence on T or even on t. So for example, it is to be understood that

$$(X \diamond M) = \mathbb{E}_t \left[\int_t^T d\langle X, M \rangle_s \right].$$

Diamond products as covariances

- Diamond (or autocovariance) functionals are intimately related to conventional covariances.

> **Lemma 1**

> Let A and B be martingales in the same filtered probability space. Then
>
> $$(A \diamond B)_t(T) = \mathbb{E}_t\left[A_T B_T\right] - A_t\, B_t = \mathrm{cov}_t\left[A_T, B_T\right].$$

- By finding the appropriate martingales, it is thus always possible to re-express diamond autocovariance functionals in terms of covariances of terminal quantities. For example, it is easy to show that

$$((X \diamond X) \diamond (X \diamond X))_t(T) = (M \diamond M)_t(T) = \mathrm{var}_t\left[\langle X\rangle_T\right].$$

  - A very neat general result!

Autocovariance functionals vs covariances

- Covariances are typically easy to compute using simulation.
- Diamond products are expressible directly in terms of the formulation of a model in forward variance form.

The leverage swap again

- An easy corollary of Lemma 1 is:

$$\mathrm{cov}\left[M, S \middle| \mathcal{F}_t\right] = (S \diamond M)_t(T).$$

- It follows that the fair value of the leverage swap is given by

$$\mathfrak{L}_t(T) = \frac{1}{S_t}(S \diamond M)_t(T).$$

The $\mathbb{G}$-forest expansion

> **Theorem 1.1 of [Friz et al.][5]**
>
> Let Y_T be a real-valued, $\mathcal{F}_T$-measurable random variable with associated martingale $Y_t = \mathbb{E}_t\left[Y_T\right]$. Under natural integrability conditions, with a, b small enough, there is a.s. convergence of
>
> $$\log \mathbb{E}\left[e^{aY_T + b\langle Y\rangle_T} \middle| \mathcal{F}_t\right] = aY_t + b\langle Y\rangle_t + \sum_{k\geq 2} \mathbb{G}_t^k(T), \qquad (9)$$
>
> where

$$\mathbb{G}^2 = \left(\frac{1}{2}a^2 + b\right)(Y \diamond Y)_t(T),$$
$$\mathbb{G}^k = \frac{1}{2}\sum_{j=2}^{k-2} \mathbb{G}^{k-j} \diamond \mathbb{G}^j + (a\,Y \diamond \mathbb{G}^{k-1}) \text{ for } k > 2. \tag{10}$$

Idea of the proof

- For a generic (continuous) semimartingale Z, sufficiently integrable, let
$$\Lambda_t^T = \log \mathbb{E}_t\left[e^{Z_{t,T}}\right].$$
- Then, noting that $\Lambda_T^T = 0$,
$$\mathbb{E}_t\left[e^{Z_T}\right] = \mathbb{E}_t\left[e^{Z_T + \Lambda_T^T}\right] = e^{Z_t + \Lambda_t^T}.$$
- The stochastic logarithm $\mathcal{L}\left(\mathbb{E}_\bullet(Z_T)\right) = Z + \Lambda^T + \frac{1}{2}\left\langle Z + \Lambda^T\right\rangle$ is a martingale.
- Thus,
$$\begin{aligned}\Lambda_t^T &= \mathbb{E}_t\left[Z_{t,T} + \tfrac{1}{2}\left\langle Z + \Lambda^T\right\rangle_{t,T}\right] \\ &= \mathbb{E}_t\left[Z_{t,T}\right] + \tfrac{1}{2}((Z + \Lambda^T) \diamond (Z + \Lambda^T))_t(T).\end{aligned}$$
- Now with $Z = \epsilon a Y + \epsilon^2\, b\,\langle Y\rangle$ we get
$$\Lambda_t^T(\epsilon) = \epsilon a\,\mathbb{E}_t\left[Y_{t,T}\right] + \epsilon^2\, b\,(Y \diamond Y)_t(T) + \tfrac{1}{2}\left(\epsilon a Y + \Lambda_t^T(\epsilon)\right)_t^{\diamond 2}(T)\,.$$
- Put $\Lambda_t^T(\epsilon) = \epsilon^2\mathbb{G}_t^2 + \epsilon^3\mathbb{G}_t^3 + \ldots$, and match coefficients of ϵ^n.
  - $[\epsilon^2]$: $\mathbb{G}_t^2 = b\,(Y \diamond Y)_t(T) + \frac{1}{2}a^2\,(Y \diamond Y)_t(T)$.
  - $[\epsilon^3]$: $\mathbb{G}_t^3 = (a\,Y \diamond \mathbb{G}^2)_t(T)$.
  - $[\epsilon^4]$: $\mathbb{G}_t^4 = (a\,Y \diamond \mathbb{G}^3)_t(T) + \frac{1}{2}(\mathbb{G}^2 \diamond \mathbb{G}^2)_t(T)$.
- We see the recursion formula emerge!

The $\widetilde{\mathbb{F}}$ expansion

A corollary is:

> **Corollary 3.1 of [Alòs et al.][1]**
>
> The cumulant generating function (CGF) is given by

$$\psi_t(T;a) = \log \mathbb{E}_t\left[e^{\mathrm{i}\, a\, X_T}\right] = \mathrm{i}\, a\, X_t - \frac{1}{2} a\,(a+\mathrm{i})\, M_t(T) + \sum_{\ell=1}^{\infty} \tilde{\mathbb{F}}_\ell(a).$$

where the $\tilde{\mathbb{F}}_\ell$ satisfy the recursion $\tilde{\mathbb{F}}_0 = -\frac{1}{2} a(a+\mathrm{i})\, M_t$ and for $k > 0$,

$$\tilde{\mathbb{F}}_\ell = \frac{1}{2} \sum_{j=0}^{\ell-2} \left(\tilde{\mathbb{F}}_{\ell-2-j} \diamond \tilde{\mathbb{F}}_j\right) + \mathrm{i}a\, \left(X \diamond \tilde{\mathbb{F}}_{\ell-1}\right).$$

- A formal expression for the characteristic function for any stochastic volatility model written in forward variance form.
  - Useful in practice if the diamond products are expressible in closed form (AFV models for example).

Variance and gamma swaps

The variance swap is given by the fair value of the log-strip:

$$\mathbb{E}_t\left[X_T\right] = (-\mathrm{i})\, \psi_t{}'(T;0) = X_t - \frac{1}{2}\, w_t(T)$$

and the gamma swap (wlog set $X_t = 0$) by

$$\mathbb{E}_t\left[X_T\, e^{X_T}\right] = (-\mathrm{i})\, \frac{d}{da}\, \psi_t(T;a)\Big|_{a=-\mathrm{i}}.$$

- The point is that we can in principle compute such moments for any stochastic volatility model written in forward variance form, whether or not there exists a closed-form expression for the characteristic function.

The gamma swap

We can compute the gamma swap as

$$\mathbb{E}\left[X_T\, e^{X_T}\Big|\, \mathcal{F}_t\right] = (-\mathrm{i})\, \frac{d}{da}\, \psi_t(T;a)\Big|_{a=-\mathrm{i}}.$$

It is easy to see that only trees containing a single M leaf will survive in the sum after differentiation when $a = -\mathrm{i}$ so that

$$\sum_{k=1}^{\infty} \tilde{\mathbb{F}}_k{}'(-\mathrm{i}) = \frac{1}{2} \sum_{k=1}^{\infty} (X\diamond)^k M$$

where $(X\diamond)^k M$ is defined recursively for $k > 0$ as $(X\diamond)^k M = X \diamond (X\diamond)^{k-1} M$.

- For example, $(X\diamond)^3 M = (X \diamond (X \diamond (X \diamond M)))$.

Then the fair value of a gamma swap is given by

$$\mathcal{G}_t(T) = 2\,\mathbb{E}_t\left[X_T\, e^{X_T}\right] = w_t(T) + \sum_{k=1}^{\infty} (X\diamond)^k M.$$

- This expression allows for explicit computation of the gamma swap for any model written in forward variance form.

The leverage swap

We deduce that the fair value of a leverage swap is given by

(2)

$$\mathfrak{L}_t(T) = \mathcal{G}_t(T) - M_t(T) = \sum_{k=1}^{\infty} (X\diamond)^k M.$$

- The leverage swap is expressed explicitly in terms of covariance functionals of the spot and vol. processes.
  - If spot and vol. processes are uncorrelated, the fair value of the leverage swap is zero.

- The leverage swap may be easily estimated from the volatility smile along the lines of [Fukasawa][7] or alternatively by integration if we have fitted some curve to the smile.

- We will now use (2) to compute an explicit expression for the value of a leverage swap in the rough Heston model with $\lambda = 0$.

The rough Heston model (with $\lambda = 0$) in forward variance form

Recall that in forward variance form, the rough Heaston model with $\lambda = 0$ reads

$$\begin{aligned}\frac{dS_t}{S_t} &= \sqrt{V_t}\, dZ_t \\ d\xi_t(u) &= \frac{\nu}{\Gamma(\alpha)}\,(u-t)^{\alpha-1}\,\sqrt{V_t}\, dW_t.\end{aligned}$$

- The rough Heston model (with $\lambda = 0$ turns out to be even more tractable than the classical Heston model!

Computation of autocovariance functionals

Apart from $\mathcal{F}_t$ measurable terms (abbreviated as `drift'), we have

$$
\begin{aligned}
dX_t &= \sqrt{V_t}\, dZ_t + \text{drift} \\
dM_t &= \int_t^T d\xi_t(u)\, du + \text{drift} \\
&= \frac{\nu}{\Gamma(\alpha)} \sqrt{V_t} \left(\int_t^T \frac{du}{(u-t)^\gamma} \right) dW_t + \text{drift} \\
&= \frac{\nu\,(T-t)^\alpha}{\Gamma(1+\alpha)} \sqrt{V_t}\, dW_t + \text{drift}.
\end{aligned}
$$

The first order forest

There is only one tree in the forest $\mathbb{F}_1$.

$$
\begin{aligned}
\mathbb{F}_1 = (X \diamond M)_t(T) &= \mathbb{E}_t \left[\int_t^T d\langle X, M \rangle_s \right] \\
&= \frac{\rho\,\nu}{\Gamma(1+\alpha)}\, \mathbb{E}_t \left[\int_t^T V_s\, (T-s)^\alpha\, ds \right] \\
&= \frac{\rho\,\nu}{\Gamma(1+\alpha)} \int_t^T \xi_t(s)\, (T-s)^\alpha\, ds.
\end{aligned}
$$

The second order forest

There are two trees in $\mathbb{F}_2$. The first tree is

$$
\begin{aligned}
(M \diamond M)_t(T) &= \mathbb{E}_t \left[\int_t^T d\langle M, M \rangle_s \right] \\
&= \frac{\nu^2}{\Gamma(1+\alpha)^2} \int_t^T \xi_t(s)\, (T-s)^{2\,\alpha}\, ds.
\end{aligned}
$$

The second tree $\left(X \diamond (X \diamond M)\right)_t(T)$ is more complicated.

Define for $j \geq 0$

$$
I_t^{(j)}(T) := \int_t^T ds\, \xi_t(s)\, (T-s)^{j\,\alpha}.
$$

In terms of $I_t^{(j)}(T)$

We may then rewrite the above expressions as

$$(X \diamond M)_t(T) = \frac{\rho\,\nu}{\Gamma(1+\alpha)}\, I_t^{(1)}(T)$$
$$(M \diamond M)_t(T) = \frac{\nu^2}{\Gamma(1+\alpha)^2}\, I_t^{(2)}(T).$$

A little more computation gives

$$\begin{aligned}
(X \diamond (X \diamond M))_t(T) &= \mathbb{E}_t\left[\int_t^T d\Big\langle X, I^{(1)}\Big\rangle_s\right] \\
&= \frac{\rho^2\,\nu^2}{\Gamma(1+\alpha)}\,\frac{\Gamma(1+\alpha)}{\Gamma(1+2\,\alpha)}\int_t^T ds\,\mathbb{E}_t\left[\int_s^T V_s\,(T-s)^{2\,\alpha}\,ds\right] \\
&= \frac{\rho^2\,\nu^2}{\Gamma(1+\alpha)\,\Gamma(\alpha)}\int_t^T ds\,\mathbb{E}_t\left[\int_s^T V_s\,\frac{(T-u)^{\alpha}}{(u-s)^{\gamma}}\,du\right] \\
&= \frac{\rho^2\,\nu^2}{\Gamma(1+2\,\alpha)}\int_t^T ds\;\xi_t(s)\,(T-s)^{2\,\alpha} \\
&= \frac{\rho^2\,\nu^2}{\Gamma(1+2\,\alpha)}\, I_t^{(2)}(T).
\end{aligned}$$

One can be easily convinced that each tree in the level-k forest $\mathbb{F}_k$ is $I^{(k)}$ multiplied by a simple prefactor.

The third order forest

For example, continuing to the forest $\mathbb{F}_3$, we have the following.

$$(M \diamond (X \diamond M))_t(T) = \frac{\rho\,\nu^3\,\Gamma(1+2\,\alpha)}{\Gamma(1+\alpha)^2\,\Gamma(1+3\,\alpha)}\, I_t^{(3)}(T)$$
$$(X \diamond (X \diamond (X \diamond M)))_t(T) = \frac{\rho^3\,\nu^3}{\Gamma(1+3\,\alpha)}\, I_t^{(3)}(T)$$
$$(X \diamond (M \diamond M))_t(T) = \frac{\rho\,\nu^3\,\Gamma(1+2\,\alpha)}{\Gamma(1+\alpha)^2\,\Gamma(1+3\,\alpha)}\, I_t^{(3)}(T).$$

In particular, we easily identify the pattern

$$(X\diamond)^k M_t = \frac{(\rho\,\nu)^k}{\Gamma(1+k\,\alpha)}\, I_t^{(k)}(T).$$

The leverage swap under rough Heston

Using (2), we have

$$\begin{aligned}\mathfrak{L}_t(T) &= \sum_{k=1}^{\infty} (X\diamond)^k M \\ &= \sum_{k=1}^{\infty} \frac{(\rho\,\nu)^k}{\Gamma(1+k\,\alpha)} \int_t^T du\,\xi_t(u)\,(T-u)^{k\,\alpha} \\ &= \int_t^T du\,\xi_t(u)\,\{E_\alpha(\rho\,\nu\,(T-u)^\alpha) - 1\}\end{aligned}$$

where $E_\alpha(\cdot)$ denotes the Mittag-Leffler function.

- A closed-form formula for the leverage swap!

The normalized leverage contract

Given the form of the expression for the leverage contract, it is natural to normalize by the variance contract. We therefore define

$$L_t(T) = \frac{\mathfrak{L}_t(T)}{M_t(T)}.$$

In the special case of the rough Heston model with a flat forward variance curve,

$$L_t(T) = E_{\alpha,2}(\rho\,\nu\,\tau^\alpha) - 1,$$

where $E_{\alpha,2}(\cdot)$ is a generalized Mittag-Leffler function, independent of the reversion level θ. We further define an nth order approximation to $L_t(T)$ as

$$L_t^{(n)}(T) = \sum_{k=1}^{n} \frac{(\rho\,\nu\,\tau^\alpha)^k}{\Gamma(2+k\,\alpha)}.$$

Implement the approximate formulae

In [1]:
```
ltT.raw <- function(H,eta,rho,n)function(tau){
    k <- (1:n)
    alpha <- H+1/2
    x <- rho * eta * tau^alpha
    vec <- x^k/gamma(2+k*alpha)
    return(sum(vec))
}

ltT <- function(H,eta,rho,n){Vectorize(ltT.raw(H,eta,rho,n))}
```

Some R-code

In [2]:
```
setwd("./QRV")
```

In [3]:
```
source("BlackScholes.R")
source("BlackFormula.R")
source("fwdVarCurve.R")
source("FukasawaRobust.R")
source("GammaKernel.R")
source("HQE.R")
source("Lewis.R")
source("plotIvols.R")
source("roughHestonPadeLambda.R")
```

In [4]:
```
library(repr)
library(colorspace)
library(MittagLeffleR)
library(stinepack)
```

Set up nice colors

In [5]:
```
my.col <- sequential_hcl(5, palette="Batlow")
bl <- "royalblue"
rd <- "red2"
pk <- "hotpink1"
gr <- "green4"
br <-"brown"
pu <- "purple"
or <- "orange"
```

A numerical example

We now perform a numerical computation of the value of the leverage swap using the forest expansion in the rough Heston model with the following parameters, calibrated to the SPX options market as of May 19, 2017:

$$H = 0.0474; \quad \nu = 0.2910; \quad \rho = -0.6710.$$

In [6]:
```
H.20170519 <- 0.0236
nu.20170519 <- 0.3266
rho.20170519 <- -0.6510
params.rHeston <- list(H=H.20170519,nu=nu.20170519,rho=rho.20170519)
```

Plot of successive approximations

In [7]:
```
library(repr)
options(repr.plot.width=10,repr.plot.height=7)
```

In [8]:
```
curve(ltT(H.20170519,nu.20170519,rho.20170519,1)(x),from=0,to=3,col=br,lwd=1
     xlab=expression(paste("Time to expiry ",tau)),ylab=expression(L[t](T)))
curve(ltT(H.20170519,nu.20170519,rho.20170519,2)(x),from=0,to=3,col=bl,lwd=1
curve(ltT(H.20170519,nu.20170519,rho.20170519,3)(x),from=0,to=3,col=gr,lwd=1
curve(ltT(H.20170519,nu.20170519,rho.20170519,20)(x),from=0,to=3,col=rd,lwd=
```

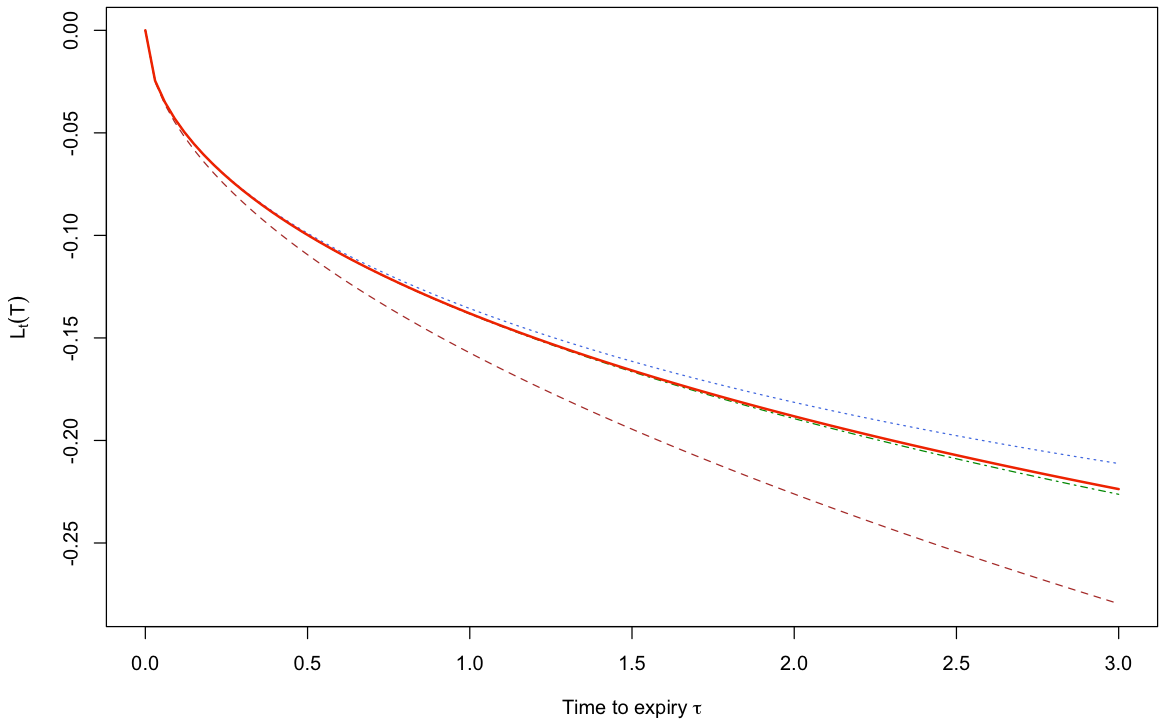

Figure 3: Successive approximations to the (absolute value of) the normalized rough Heston leverage swap. The solid red line is the exact expression $L_t(T)$; $L_t^{(1)}(T)$, $L_t^{(2)}(T)$, and $L_t^{(3)}(T)$ are brown dashed, blue dotted and dark green dash-dotted lines respectively.

The leverage swap in a general AFV model

From the definition of affine forward variance model,

$$\psi_t(T;a) = \log \mathbb{E}_t\left[e^{\mathrm{i}a\,(X_{t,T})}\right] = \int_t^T \xi_t(u)\, g(T-s;a)\, du = (\xi \star g)_t(T)$$

where g satisfies the convolution Riccati equation

$$g(\tau;a) = -\tfrac{1}{2}a\,(a+\mathrm{i}) + \rho\,\mathrm{i}a\,(\kappa\star g)(\tau;a) + \tfrac{1}{2}\left((\kappa\star g)(\tau;a)\right)^2, \quad \tau \geq 0. \quad (11)$$

Now recall that

$$\mathfrak{L}_t(T) = -2\,\mathrm{i}\,\left\{\psi_t'(T;-\mathrm{i}) + \psi_t'(T;0)\right\}.$$

Let $h(t;a) = (\kappa\star g)(t;a)$. Then, differentiating the convolution Riccati equation wrt a,

$$g'(\tau;a) = -a - \tfrac{\mathrm{i}}{2} + i\,\rho\,h(\tau;a) + i\,\rho\,a\,h'(\tau;a) + h(\tau;a)\,h'(\tau;a). \quad (12)$$

Since $g(\tau;0) = g(\tau;-\mathrm{i}) = 0$, it follows that $h(\tau;0) = h(\tau;-\mathrm{i}) = 0$, and thus

$$g'(\tau;0) = -\tfrac{i}{2}; \quad g'(\tau;-\mathrm{i}) = \tfrac{\mathrm{i}}{2} + \rho\, h'(\tau;-\mathrm{i}).$$

The leverage swap is then given by

$$\mathfrak{L}_t(T) = -2\,\mathrm{i}\,\rho \int_t^T \xi_t(u)\, h'(T-u;-\mathrm{i})\, du.$$

Denote $g'(\tau;-\mathrm{i})$ by g'. Then, from the convolution Riccati equation with $a = -\mathrm{i}$,

$$g'(\tau) = \tfrac{\mathrm{i}}{2} + \rho\,(\kappa \star g')(\tau), \tag{13}$$

a *linear* convolution equation for g'.

- Can be solved (in principle) using Laplace transforms.

Fukasawa's formula for the gamma contract

- [Fukasawa][7] derives an expression for the value of a generalized European payoff in terms of implied volatilities.
- He derives the formula for the variance contract we saw before

$$M_t(T) = -2\,\mathbb{E}_t\left[\log\frac{S_T}{S_t}\right] = \int_{-\infty}^{\infty} dz\, N'(z)\,\Sigma\left(g_-(z)\right).$$

- As another application, he also derives the following expression for the value of a gamma contract.

$$\mathcal{G}_t(T) = 2\,\mathbb{E}_t\left[\frac{S_T}{S_t}\log\frac{S_T}{S_t}\right] = \int_{-\infty}^{\infty} dz\, N'(z)\,\Sigma\left(g_+(z)\right).$$

  (note g_+ instead of g_- in the variance swap case).

  - In particular, if we have a parameterization of the volatility smile (such as SVI or Vola Dynamics), computing the fair value of the covariance swap is straightforward.

Model calibration using the leverage swap

- The leverage swap is easily computed in affine forward variance models.

- It may also be estimated from European option prices using the Fukasawa formulae.
  - This depends in practice on the quality of variance and gamma swap estimates.
- The parameters of the model may then be estimated. In the case of rough Heston, 4 parameters to estimate using 50 or so expirations.

Load volatility smiles from 15-Feb-2023

In [9]:
```
load("spxIvols20230215.rData")
ivolData <- spxIvols20230215
head(ivolData)
```

A data.frame: 6 × 7

	Expiry	Texp	Strike	Bid	Ask	Fwd	CallMid
	<int>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
1	20230216	0.002737851	1000	NA	7.793085	4146.742	NA
2	20230216	0.002737851	1200	NA	6.813266	4146.742	NA
3	20230216	0.002737851	1400	NA	5.987566	4146.742	NA
4	20230216	0.002737851	1600	NA	5.273554	4146.742	NA
5	20230216	0.002737851	1800	NA	4.644049	4146.742	NA
6	20230216	0.002737851	2000	NA	4.080578	4146.742	NA

Estimated variance and gamma curves

In [10]:
```
expiries <- sort(unique(ivolData$Texp))
vs <- varSwap.Robust(ivolData)$vs.mid
gs <- gammaSwap.Robust(ivolData)$gs.mid

head(vs)
head(gs)
```

0.036529328507355 · 0.0317776298748159 · 0.019801436839558 · 0.0216205797598485 · 0.0239817142815479 · 0.0260070933624724

0.0363327515229832 · 0.0315435433905739 · 0.0196544532312524 · 0.0213902911973152 · 0.0236727054980354 · 0.0255246287563552

Plot the curves

In [11]:
```
plot(expiries,vs,type="b",pch=20,col=bl,ylab="Swap values",xlab="Maturity (y
lines(expiries,gs,lty=1,type="b",pch=20,col=rd)
legend("topleft",inset=0.02,c("Variance swap","Gamma swap"),col=c(bl,rd),lty
```

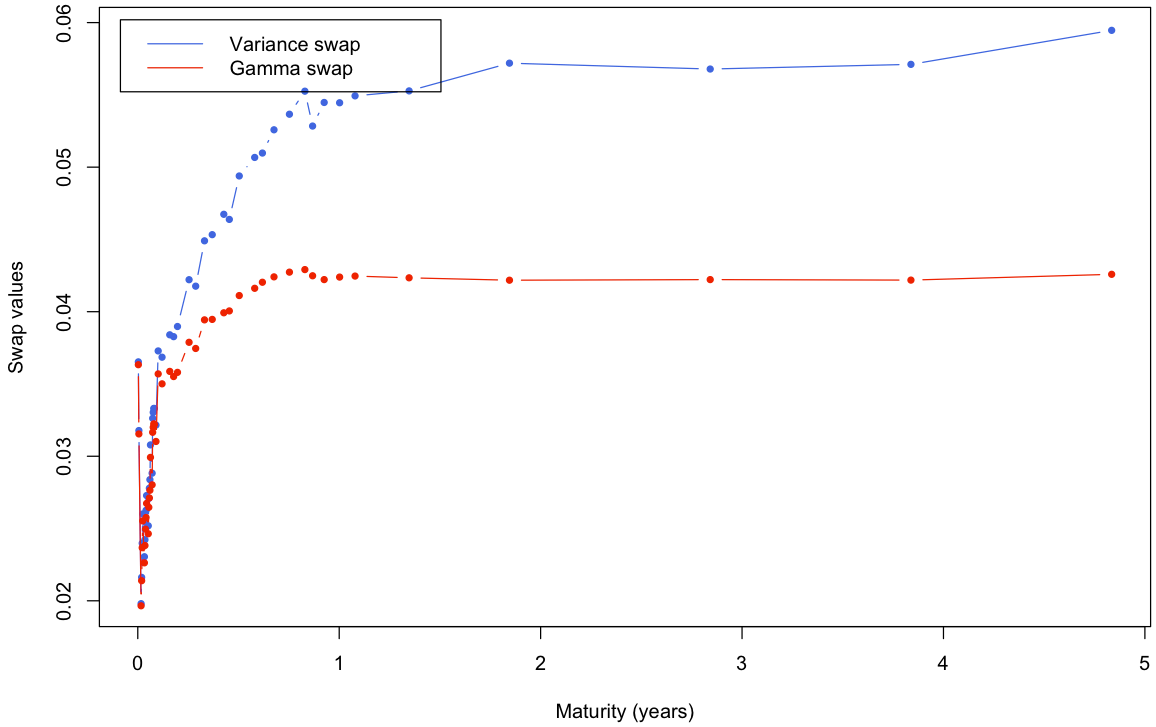

Figure 4: SPX variance curve (blue) and gamma swap curve (red) as of February 15, 2023 estimated using Fukasawa's robust methodology.

Leverage swap term structure

In [12]:
```
plot(expiries,gs-vs,type="b",pch=20,col=gr,lwd=2,
     ylab="Leverage swap",xlab="Maturity (years)")
```

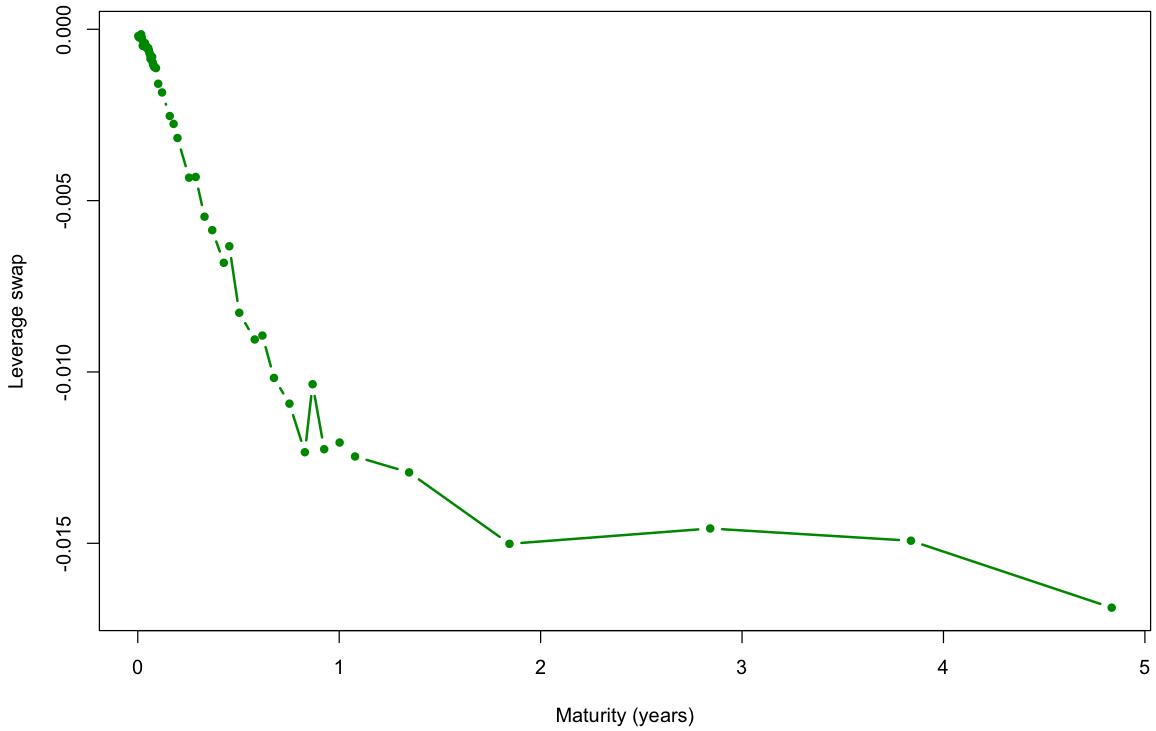

Figure 5: SPX leverage swap curve as of February 15, 2023 estimated using Fukasawa's robust methodology.

Calibration of rough Heston using the leverage contract

- We now use the above leverage swap curve to estimate the parameters of the rough Heston model.
- The shape of the curve indicates that we will need the rough Heston model with $\lambda > 0$ to get a decent fit.
- It can be shown that the value of the leverage contract under rough Heston is gven by

$$\mathfrak{L}_t(T) = \frac{\rho\,\nu}{\lambda'} \int_t^T \xi_t(u)\,\left[1 - E_\alpha(-\lambda'\,(T-u)^\alpha)\right]\,du, \tag{14}$$

  where $\lambda' = \lambda - \rho\,\eta$.

  - A (natural) generalization of our earlier result for the case $\lambda > 0$.

The normalized leverage contract

- Again, we define

$$L_t(T) = \frac{\mathfrak{L}_t(T)}{M_t(T)}.$$

- In the special case of the rough Heston model with a flat forward variance curve,

$$L_t(T) = \frac{\rho\,\nu}{\lambda'}\left(1 - E_{\alpha,2}(-\lambda'\,\tau^{\alpha})\right)$$

where $E_{\alpha,2}(\cdot)$ is a generalized Mittag-Leffler function.

Implement the normalized leverage contract formula

In [13]:
```
lev.norm <- function(params,xi)function(T){

    H <- params$H
    nu <- params$nu
    rho <- params$rho
    lambda <- params$lambda
    al <- H+1/2
    lamp <- lambda-rho*nu

    res <- (1-mlf(-lamp*T^al,al,2))*rho*nu/lamp
    return(res)

}
```

Here is the empirical estimated leverage:

In [14]:
```
lev.est <- gs/vs-1
```

Rough Heston parameter optimization

In [15]:
```
obj <- function(paramvec){

    H <- paramvec[1]
    nu <- paramvec[2]
    rho <- paramvec[3]
    lambda <- paramvec[4]
    params <- list(H=H,nu=nu,rho=rho,lambda=lambda)

    l.model <- lev.norm(params)(expiries)
    res <- sum((l.model - lev.est)^2/(expiries)^(0.9))*1e6
    return(res)
}
```

In [16]:
```
(res.optim <- optim(c(.05,.25,-.64,.3),obj,method =  "L-BFGS-B",
 lower=c(.0001,.01,-.999,0),upper=c(1,10,0,10)))
```

$par	0.511599077350975 · 1.04560609788258 · -0.971373372481705 · 2.23552496279593
$value	11204.3325174843
$counts	**function:** 27 **gradient:** 27
$convergence	0
$message	'CONVERGENCE: REL_REDUCTION_OF_F <= FACTR*EPSMCH'

Notice how fast the calibration is!

The optimized parameters are:

In [17]:
```
fit.rHeston <- as.list(res.optim$par)
names(fit.rHeston) <- c("H","nu","rho","lambda")
fit.rHeston
save(fit.rHeston,file="fit_rHeston.rData")
```

$H	0.511599077350975
$nu	1.04560609788258
$rho	-0.971373372481705
$lambda	2.23552496279593

Plot the rough Heston fit

In [18]:
```
plot(expiries,lev.est,type="p",col=gr,cex=1,
     pch=20,xlab="Time to expiry",ylab="Leverage swap fair value")
lines(expiries,lev.norm(fit.rHeston)(expiries),col=rd,lwd=2)
```

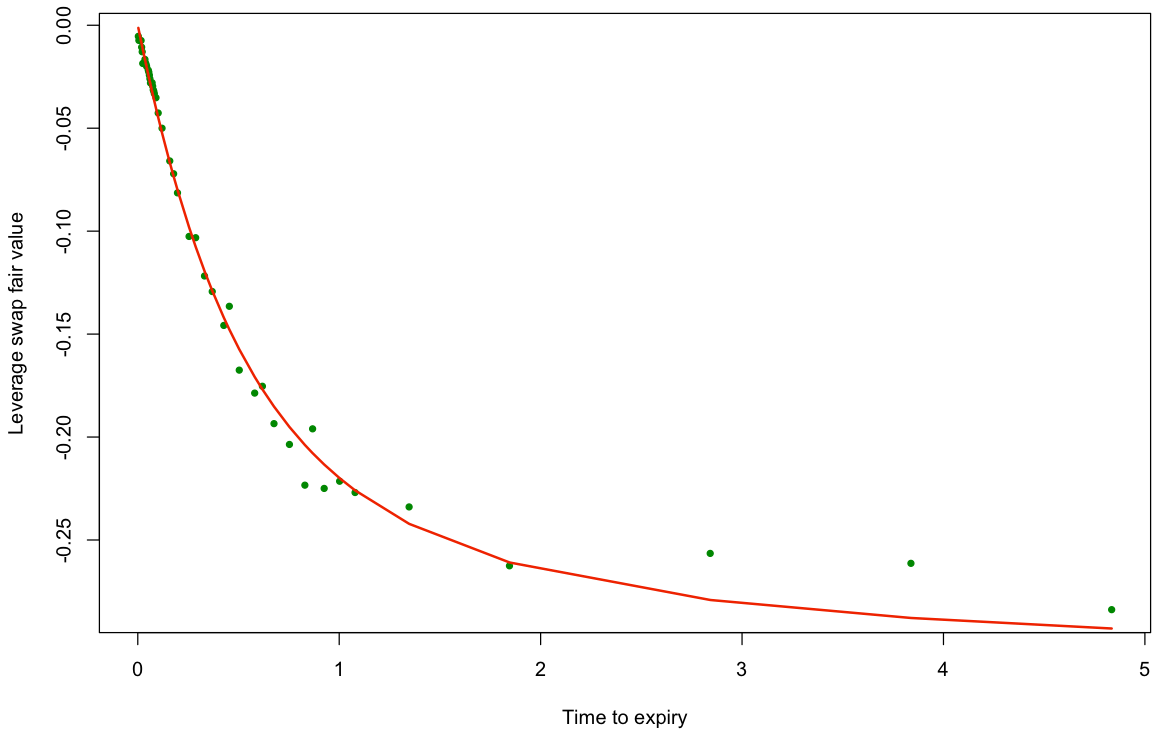

Figure 6: The green points are robust Fukasawa leverage contract estimates; the red curve is the rough Heston fit.

Comments on the calibrated parameters

- I don't believe these parameters.
  - The best fit is almost classical Heston!
  - One issue is that the Fukasawa robust methodology extrapolates smiles at constant volatility in the wings.
- However, we will see in Lecture 4 that these calibrated parameters do indeed give the best fit!
- We can only conclude that rough Heston is an unreasonable model.
  - We already know it has unreasonable dynamics!

Summary of lecture 3

- There is a one-to-one correspondence between Hawkes-process based AFI models and AFV models.
  - Jaisson and Rosenbaum's rough Heston model is one example.
- To get a non-trivial stochastic volatility model as a limit, we need near-instability of the Hawkes kernel.

- Affine models have easy to compute characteristic functions.
    - Computations of key quantities become much easier.
- Diamonds and the forest expansion allow easy computation of model quantities that can be compared with market values
    - Potentially easy calibration.
    - As many matching conditions as market option expirations.

References

1. ^ Elisa Alòs, Jim Gatheral, and Radoš Radoičić, Exponentiation of conditional expectations under stochastic volatility, *Quantitative Finance* **20**(1) 13–27.
2. ^ Emmanuel Bacry, Iacopo Mastromatteo, and Jean-François Muzy, Hawkes processes in finance, *Market Microstructure and Liquidity* **1**(01) 1550005 (2015).
3. ^ Peter Carr and Dilip Madan, Option valuation using the Fast Fourier Transform, *Journal of Computational Finance* **2**(4), 61–73 (1999).
4. ^ Omar El Euch and Mathieu Rosenbaum, The characteristic function of rough Heston models, *Mathematical Finance*, **29**(1), 3–38 (2019).
5. ^ Peter Friz, Jim Gatheral, and Radoš Radoičić, Forests, cumulants, martingales *The Annals of Probability* **50**(4) 1418-1445 (2022).
6. ^ Peter K. Friz and Jim Gatheral, Computing the SSR, *SSRN:4872776* , (2024).
7. ^ Masaaki Fukasawa, The normalizing transformation of the implied volatility smile, *Mathematical Finance* **22**(4) 753-762 (2012).
8. ^Jim Gatheral, *The Volatility Surface: A Practitioner's Guide* , John Wiley and Sons, Hoboken, NJ (2006).
9. ^ Jim Gatheral and Martin Keller-Ressel, Affine forward variance models, *Finance and Stochastics*, **23**(3) 501-533 (2019).
10. ^ Thibault Jaisson and Mathieu Rosenbaum, Rough fractional diffusions as scaling limits of nearly unstable heavy tailed Hawkes processes, *The Annals of Applied Probability* **26**(5) 2860-2882 (2016).
11. ^ Alan L. Lewis, *Option Valuation under Stochastic Volatility with Mathematica Code* , Finance Press: Newport Beach, CA (2000).