From quadratic Hawkes processes to super-Heston rough volatility models with Zumbach effect

ADITI DANDAPANI, PAUL JUSSELIN and MATHIEU ROSENBAUM*
École Polytechnique, Palaiseau, France

(Received 13 July 2019; accepted 2 October 2020; published online 14 May 2021)

Building log-normal-like rough volatility models with proper Zumbach effect using a microstructural approach

1. Introduction

Since the paper (Gatheral et al. 2018), it has been well accepted that volatility is rough. This means that log-volatility essentially behaves as fractional Brownian motion with Hurst parameter of order 0.1, see also for example Bennedsen et al. (2016), Da Fonseca and Zhang (2019), Glasserman and He (2018), Livieri et al. (2018). There are microstructural foundations for rough volatility that use Hawkes processes to create a microscopic model for asset prices. In this vein, El Euch et al. (2018) consider four stylized facts concerning market microstructure: the high degree of endogeneity of markets, the no-arbitrage property, buying/selling asymmetry and the long memory of the market order flow generated by metaorders. They show that when only the three first stylized facts are taken into account, one obtains the Heston model for the scaling limit of the price process. When the long memory property of the flow is added, the limit is the rough Heston model introduced and developed in El Euch and Rosenbaum (2018, 2019). In the rough Heston model, the spot variance $V_t$ can be written as follows:

$$V_t = V_0 + \frac{\lambda}{\Gamma(1 - \alpha)} \int_0^t (t - s)^{\alpha - 1} (\theta_0(s) - V_s) \, ds + \nu \sqrt{V_t} \, dB_s,$$

where $\lambda$ and $\nu$ are some positive constants, $\theta_0$ is a deterministic function, $\alpha \in (1/2, 1)$ and $B$ is a Brownian motion. The rough behavior is due to the singular kernel $(t - s)^{\alpha - 1}$ which is the same as that appearing in the Mandelbrot–van Ness representation of a fractional Brownian motion with Hurst parameter $\alpha - 1/2$. More recently, assuming only that the order flow is driven by a linear Hawkes process and that there is no statistical arbitrage on the market, it is shown in Jusselin and Rosenbaum (2018) that the price necessarily follows a rough Heston model. In fact, as far as we know, all the works on microstructural foundations of rough volatility have hitherto produced a rough Heston model.

However, in the context of rough models, there are other aspects of volatility that one could wish to understand from a microstructural perspective. A first point is to go beyond the square root associated with the dynamics of the volatility in the rough Heston model (1). A particularly interesting case is when an additional additive or multiplicative term appears, enhancing the square root and leading to fatter volatility tails, see Abi Jaber and El Euch (2019), Blanc et al. (2017).

Another important stylized fact of financial time series is the feedback of price returns on volatility. This phenomenon is introduced in Zumbach (2010) where he measures the impact of price trends on future volatility, see also Lynch and Zumbach (2003), Zumbach (2009). It is demonstrated that price trends induce an increase of volatility. We refer to this property as the Zumbach effect. In the literature, see notably...
Chicheportiche and Bouchaud (2012), a way to reinterpret the Zumbach effect is to consider that the predictive power of past squared returns on future volatility is stronger than that of past volatility on future squared returns. To check this on data, one typically shows that the covariance between past squared price returns and future realized volatility (over a given duration) is larger than that between past realized volatility and future squared price returns, see Blanc et al. (2017), Chicheportiche and Bouchaud (2012), El Euch et al. (2020) for more details. We refer to this version of Zumbach effect as the weak Zumbach effect.

It has been proved in El Euch et al. (2020) that the rough Heston model reproduces the weak form of Zumbach effect. However, this is not obtained through a feedback effect, which is the motivating phenomenon in the original paper by Zumbach (2010). It is only due to the dependence between price and volatility created by the correlation of the Brownian motions driving their dynamics. In particular in the rough Heston model, the conditional law of the volatility depends on the past dynamic of the price only through the past volatility, see El Euch and Rosenbaum (2018). From now on, we speak about the strong Zumbach effect when the conditional law of future volatility depends not only on the past volatility trajectory but also on past returns.

Inspired by the methodology of El Euch et al. (2018), our goal in this paper is to propose microstructural foundations for the strong Zumbach effect. We also wish to consider models in which the instantaneous volatility of variance is equal to the classical square root term of Heston-like models multiplied by a non-trivial process in order to enhance volatility tails. Any model satisfying the latter property will be called a super-Heston model.

A convenient way to build a microscopic model, encoding the Zumbach effect and leading naturally to super-Heston rough volatility, is to use a quadratic Hawkes based price process, in the spirit of Blanc et al. (2017). More precisely, we consider the following microstructural model for the price \( (P_t)_{t \geq 0} \): it is piecewise constant with sizes of price jumps independent and identically distributed taking values \( 1 \) or \( -1 \) with probability 1/2. The jump times are of a counting process \( N \). We assume that \( N \) is a quadratic Hawkes process as introduced in Blanc et al. (2017) and Ogata (1981). This means the intensity \( (\lambda_s)_{s \geq 0} \) of \( N \) is given by

\[
\lambda_s = \mu + \int_0^s \phi(t-s) \, dN_t + Z_s^2,
\]

where \( \phi \) and \( k \) are two non-negative measurable functions with support on \( \mathbb{R}_+ \) and \( \mu > 0 \). In the definition of the intensity, the linear term with kernel \( \phi \) enables us to model the self-exciting nature of the order flow. The component \( Z_s \) is a moving average of past returns. It can be thought of as a proxy for price return over a given time horizon. If the price has been essentially trending in the past, \( Z_s \) is large, leading to high intensity. On the contrary if it has been oscillating, \( Z_s \) is close to zero and there is low feedback from the returns to the intensity. Hence \( Z_s \) can obviously be understood as a (strong) Zumbach term. Note that of course one can think that positive and negative price trends have different impact on the volatility. However for simplicity in this paper we neglect this asymmetry. Finally recall that the stability condition for Model (2) is \( \| \phi \|_1 + \| k \|_2^2 \) being strictly smaller than one, see Blanc et al. (2017).

Note that if we forget the quadratic term \( Z_s^2 \) in the intensity, we are left with a linear Hawkes process as in Jaisson and Rosenbaum (2016). In this case, at the scaling limit, if the kernel \( \phi \) is heavy tailed and if we are near instability, meaning \( \| \phi \|_1 \) tends to one with the time parameter driving the asymptotics, the rescaled intensity process converges in law to a rough dynamic similar to (1), see El Euch et al. (2018) and Jaisson and Rosenbaum (2016). When the kernel norm \( \| \phi \|_1 \) is fixed and strictly smaller than one, a deterministic limiting model is obtained. Thus we see that being in the near instability regime is crucial so that roughness can arise from the kernel \( \phi \). Recall that this regime corresponds to a high degree of endogeneity in the market, see Filimonov and Sornette (2012), Hardiman et al. (2013), Jaisson and Rosenbaum (2015) and Jaisson and Rosenbaum (2016).

In Blanc et al. (2017), the authors study the long-term behavior of the intensity of quadratic Hawkes processes. That is, on the time horizon \([0, T]\) letting \( T \) tend to infinity, they are interested in the limiting dynamics of \( (\lambda_{tT})_{t \in [0,1]} \), which can be viewed as the macroscopic (squared) volatility. They work in a setting where \( \| \phi \|_1 + \| k \|_2^2 = 2\gamma < 1 \) is fixed, not depending on \( T \). Based on PDE techniques, they obtain a diffusion process power-law marginal distributions and a strong Zumbach effect for the asymptotic volatility. More precisely, their limiting model \( (\hat{P}_t, V_t) \) for price and volatility is as follows: \( d\hat{P}_t = \sqrt{V_t} dB_t \) with

\[
V_t = \mu + (Z_t)^2 + \int_0^t \gamma \beta e^{-\beta(t-s)} V_s \, ds
\]

\[
Z_t = \int_0^t \sqrt{\gamma} e^{-(t-s)/\alpha} / \alpha^2 \, d\hat{P}_s,
\]

with \( B \) a Brownian motion and \( \alpha, \beta \) some positive parameters defining the functions \( \phi \) and \( k \), taken as exponential in Blanc et al. (2017).

In this paper, we wish to go beyond the case treated in Blanc et al. (2017) from which we draw inspiration. We describe further relevant limiting price dynamics that can be generated from quadratic Hawkes processes. We focus on finding a microscopic basis for super-Heston rough volatility processes with strong Zumbach effect. Our goal is to establish connections between the micro-parameters of the quadratic Hawkes dynamics and macro-phenomena such as the roughness of the volatility and the strong Zumbach effect.

We first focus in section 2 on the purely quadratic case, when \( \phi \) is equal to zero. Choosing appropriate scaling parameters, we obtain the following limiting model: \( d\hat{P}_t = \sqrt{V_t} dB_t \) with

\[
V_t = \mu + Z_t^2
\]

\[
Z_t = \sqrt{\gamma} \int_0^t k(t-s) \, d\hat{P}_s,
\]

where \( \gamma \in (0, 1) \) is related to the scaling of the kernel \( k \). In contrast to the purely linear case, we do not need any sort
of near instability here so that a stochastic volatility model arises at the scaling limit. In (3), the strong Zumbach effect is naturally encoded since the volatility is a functional of past price returns through $Z$. We also have that the quadratic feedback of price returns on volatility implies that $V_t$ is of super-Heston type (essentially log-normal here). This can be seen for instance when $\mu = 0$ where we get

$$Z_t = \sqrt{T} \int_0^t k(t-s) |Z_s| dB_s.$$  

Moreover taking for example $k = t^{H+1/2,\lambda}$ for $H \in (0,1/2)$ and $\lambda > 0$ with $t^{\alpha,\beta}$ the Mittag–Leffler function, \footnote{see El Euch and Rosenbaum (2019) for a reminder and connections with the Mandelbrot–van Ness representation of fractional Brownian motion.} we get that the volatility has Hölder regularity $H - \varepsilon$ for any $\varepsilon > 0$. Thus, from a natural microscopic dynamic, we are able to obtain a super-Heston rough volatility model with strong Zumbach effect at the macroscopic limit.

We then investigate the limiting models arising from quadratic Hawkes processes with non-vanishing linear part. Knowing that roughness can be obtained from the linear part quadratic Hawkes processes with non-vanishing linear part. Interestingly, we also show that when $h(t) = 0$ for any $t$, the $V_t$ behavior as a super-Heston rough volatility model.

2. Asymptotic behavior of purely quadratic Hawkes models

In this section, we investigate the possible scaling limits of purely quadratic Hawkes based price processes. This corresponds to (2) with $\phi = 0$. We devote a specific section to this case since it enables us to convey some of our main ideas in a simplified setting. More precisely, we consider $(NT^3)_{T \geq 0}$ with intensity given by

$$\lambda^T_t = \mu + (Z^T_t)^2, \quad \text{with } Z^T_t = \int_0^t k_T(t-s) dP^T_t. \quad (5)$$

For any $T$, the existence of the process $(NT^3, P^T_t)$ can be obtained from Jacod (1975). We are interested in the long time behavior of the price $P^T_t$ and of its intensity $\lambda^T_t$. Before stating the main result of this section, we first discuss (in a non-rigorous manner) our scaling procedure.

2.1. Scaling procedure

The scaling procedure consists in finding appropriate factors $\omega_T$ so that the sequence $\omega_T \lambda^T_t$ converges towards a non-degenerate limit. Assume $\omega_T \lambda^T_t$ converges towards some process $V_t$. Since $[P^T_t] = NT^3$, we have

$$(P^T)_T = \int_0^T T \lambda^T_t \, ds.$$ 

Thus we expect the martingale $P^T_t = \sqrt{T} P^T_t$ to converge since its bracket does. Let $P$ be its limit. Since we wish to get $P$ continuous, we need $\omega_T / T$ to go to zero. From the convergence of $(\sqrt{\omega_T} Z^T_t)^2$, we expect that of

$$\sqrt{\omega_T} Z^T_t = \int_0^t k_T(T(t-s)) \sqrt{T} \, dP^T_t$$

too, which requires $k_T(T \cdot) \sqrt{T}$ to converge. This leads us to consider, as in Blanc \emph{et al.} (2017), a sequence of kernels $k_T$ of the form

$$k_T = k(\cdot / T) \sqrt{T}$$

for some $\gamma > 0$ and $\omega_T = 1$ (since we observe that $\omega_T$ plays eventually no role). Finally passing to the limit in (5) we obtain the following candidate for our limiting process:

$$V_t = \mu + Z^*_t; \quad \text{with } Z^*_t = \int_0^t k(t-s) \, dP_t.$$

2.2. Assumptions and results in the purely quadratic case

We now give our exact assumptions, the second of them being purely technical.

**Assumption 2.1** (i) The sequence of kernels $(k_T)_{T \geq 0}$ is given by

$$k_T = \sqrt{T} k \left( \frac{\cdot}{T} \right),$$

with $\gamma \in (0,1)$ and $k$ a non-negative measurable function such that $\|k\|_2 = 1$. Furthermore $\mu_T = \mu > 0$. 

\footnote{We use the notation $L^p$ without reference to the underlying domain when no confusion is possible.}
The function $k$ belongs to $L^{2+r}$ for some $\varepsilon > 0$ and for any $0 < t < t' \leq 1$,
\[ \int_0^{t'} |k(t' - s) - k(t - s)|^2 \, ds < C|t' - t|^r, \]
for some $r > 0$ and $C > 0$ and
\[ \frac{1}{\eta} \int_0^1 |k(t)|^2 \tau^{2\alpha} \, dt + \int_0^1 \int_0^1 \frac{|k(t) - k(s)|^2}{|t - s|^{1+2\eta}} \, ds \, dt < +\infty \]
for some $\eta \in (0, 1)$.

Note that for $\alpha \in (1/2, 1)$ and $\lambda > 0$, the Mittag–Leffler function $f^{\alpha, \lambda}$ satisfies Assumption 2.1 ii) for any $\varepsilon \in (0, (2\alpha - 1)/(1 - \alpha))$, $\eta \in (0, \alpha - 1/2)$ and $r = 2\alpha - 1$.

Under Assumption 2.1, for any $T$, we have $\|k_T\|_2 = \gamma < 1$. So the stability condition is not violated at the limit. We now state the main result of this section. Consider the rescaled processes
\[ X^T_t = \frac{N^T_t}{T} \quad \text{and} \quad P^T_t = \frac{1}{\sqrt{T}} P^T_t. \]
We have the following theorem.

**Theorem 2.2** Under Assumption 2.1, as $T$ goes to infinity, the sequence $(X^T, P^T)_{T \geq 0}$ converges in law for the Skorohod topology on $[0, 1]$ towards some processes $(X, P)$ satisfying the following properties:

- $X$ is almost surely continuously differentiable.
- There exists a Brownian motion $B$ such that $P_t = \int_0^t \sqrt{V_s} dB_s$,

where $V$ is the derivative of $X$ and the unique continuous solution of
\[ V_t = \mu + Z_t^2, \]
\[ Z_t = \int_0^t \sqrt{\gamma} k(t - s) \sqrt{V_s} dB_s, \quad \text{on} \ [0, 1]. \quad (6) \]

- For any $\varepsilon > 0$, if $k = f^{H+1/2, \lambda}$ with $H \in (0, 1/2)$ and $\lambda > 0$, $V$ has almost surely $H - \varepsilon$ Hölder regularity.

Theorem 2.2 will be generalized in section 3 and its proof is given in section 5.1.

### 2.3. Discussion of theorem 2.2

- Quadratic Hawkes models share many similarities with GARCH and QARCH models, see Engle (1982), Engle and Bollerslev (1986), and Sentana (1995). However, from Theorem 2.2, we see that we do not need to be in the near instability regime $\|k_T\|_2 + \|\phi_T\|_1 \to 1$ in order to obtain a stochastic model at the scaling limit, while it is required in the GARCH setting, see Nelson (1990).
- In the limiting model (6), volatility and price are driven by the same Brownian motion $B$. This is in contrast to the GARCH case or to that of nearly unstable Hawkes processes where a new Brownian motion appears in the volatility dynamic, see El Euch et al. (2018). Compared to the GARCH situation, the difference essentially lies in the very constrained law of the returns here.
- The Zumbach effect is obviously present in the limiting model: the volatility is purely driven by the returns via the term $Z_t$.
- The use of Mittag–Leffler type kernels as in the last point of Theorem 2.2 is very standard in the rough volatility literature, see, e.g. Jaisson and Rosenbaum (2016). It enables us to obtain at the limit a rough behavior for the sample paths of the volatility process.
- When $k(t) = \sqrt{2c} e^{-\varsigma t}$, Model (6) is that of Blanc et al. (2017) with $\phi = 0$. Therefore Theorem 2.2 extends the results of Blanc et al. (2017) to any kernel $k$ with suitable integrability conditions. In the next section, we provide an even more general extension that encompasses the case $\phi \neq 0$ and clearly shows the super-Heston nature of the dynamic (6).

### 3. General quadratic Hawkes models: the stable case

We now study the asymptotic behavior of a sequence of general quadratic Hawkes models for which the stability condition is not violated at the limit. We consider $(N^T)_{T \geq 0}$ with intensity given by (2) (with parameters depending on $T$) where $\|\phi_T\|_1 + \|k_T\|_2$ is a fixed constant strictly smaller than one. As in the previous section, we first give intuitions about our scaling procedure.

#### 3.1. Suitable scaling in the general case

Using a scaling factor $\omega_T$, the rescaled intensity becomes
\[ \omega_T \lambda_T^T = \mu_T \omega_T + \int_0^t \phi_T(T(t - s)) T_T \lambda_T^T \, ds \]
\[ + \int_0^t \phi_T(T(t - s)) \sqrt{\omega_T} T \, d \left( \frac{\omega_T^{1/2}}{T^{1/2}} M_T^T \right) \]
\[ + T^{1/2} Z_t^2, \quad (7) \]
where
\[ M_T^T = N_T^T - \int_0^t \lambda_T^T \, ds. \]
Assume that $(\omega_T \lambda_T^T)_{T \geq 0}$ converges and consider the processes $M_T^{\ast T} = M_T^T \sqrt{T}$ and $P_T^T = P_T^T \sqrt{T}$. We have
\[ \langle P_T^T \rangle_t = \langle M_T^{\ast T} \rangle_t = \int_0^t \omega_T \lambda_T^T \, ds \quad \text{and} \quad \langle P_T^T, M_T^{\ast T} \rangle_t = 0. \]

Thus we expect $P_T^T$ and $M_T^{\ast T}$ to converge towards two martingales $M$ and $P$ such that $\langle M, P \rangle = 0$. As in the previous section, to obtain continuous martingales $M$ and $P$, we pick $\omega_T$ such that $\omega_T/T$ tends to zero.

One of our goals being to preserve Zumbach effect in the limit of (7), we need a non-degenerate behavior for the
feedback term $\sqrt{\omega T}Z^T_{sT}$. We have
\[
\sqrt{\omega T}Z^T_{sT} = \int_0^{t^*} k_T(T(t-s))\sqrt{T}dP^h_{sT},
\]
which leads us again to the specification
\[
k_T = k(\cdot T) \sqrt{T},
\]
for some positive $\gamma$. Now, if $\sqrt{\omega T}Z^T_{sT}$ converges, according to (7) we should also obtain convergence of
\[
\mu_T\omega_T + \int_0^{t^*} \phi_T(T(t-s))\omega_T\lambda^T_{sT}ds + \int_0^{t^*} \phi_T(T(t-s))\sqrt{\omega T}dM^*_{sT}.
\]
Thus, since both $\omega_T\lambda^T_{sT}$ and $M^*_{sT}$ are expected to converge, we set $\mu_T = \mu/\omega_T$ and must ensure the convergence of both $\phi_T(T(t-s))T\omega_T\lambda^T_{sT}$ and $\phi_T(T(t-s))\sqrt{\omega T}$. Because $\omega_T/T$ tends to zero, the first integral dominates the second one. Consequently we only need to take care of the first integral and again we can take $\omega_T = 1$. A logical specification is therefore
\[
\phi_T = \phi(\cdot T)(\beta/T)
\]
for some positive $\beta$. Passing to the limit in equation (7) we expect the following limiting model:
\[
V_t = \mu + \int_0^t \beta \phi(t-s)V_s ds + Z^2_t,
\]
with $Z_t = \int_0^t \sqrt{k}(t-s) dP_s$.

### 3.2. Assumptions and results in the presence of a linear component in the stable case

We now give our exact assumptions which are quite similar to those in the previous section.

**Assumption 3.1** (i) The sequence of kernels is given by
\[
k_T(t) = \sqrt{T} k \left( \frac{t}{T} \right), \quad \phi_T(t) = \frac{\beta}{T} \phi \left( \frac{t}{T} \right),
\]
with $0 < \gamma + \beta < 1$ and $k$ and $\phi$ non-negative measurable such that $\|k\|_2^2 = \|\phi\|_1 = 1$. Furthermore $\mu_T = \mu > 0$.

(ii) Assumption 2.1 (ii) holds.

Assumption 3.1 implies that the stability condition is not violated at the limit. Nevertheless, from a rescaling perspective, the choice of kernels $\phi_T$ and $k_T$ does not seem really natural at first sight. It would be probably more satisfactory to consider kernel sequences of the form $\alpha_T k$ and $\phi_T$ (with $k$ and $\phi$ not depending on $T$) and then investigate the limit of $\omega_T\lambda^T_{sT}$ as in El Euch et al. (2018), Jaisson and Rosenbaum (2015), and Jaisson and Rosenbaum (2016). This would imply here $\phi_T(T(t-s))T = \alpha_T \phi(T(t-s))T$. According to Tauberian theorems, see, e.g., Bingham et al. (1989), $\alpha_T \phi(T(t-s))T$ can only converge in that case towards a power-law function of the form $t^{-\delta}$ for some positive $\delta$ and $\phi$ has to be such that $\phi(t) \sim_{+.\infty} t^{-\delta}$ up to a slowly varying function. But recall that $\phi$ must be integrable and so we need $\delta \geq 1$. However, such choice would lead to difficulties for defining properly the limit of the integral
\[
\int_0^{t^*} \phi(\cdot T)(T(t-s))\omega_T\lambda^T_{sT}ds.
\]
To be able to consider such types of natural but technically more intricate rescaling procedures, we will drop the stability assumption in section 4 where we work in the nearly unstable case.

Let us define the rescaled process $X^T_t = N_{Tt}/T$. We have the following theorem whose proof is given in section 5.2.

**Theorem 3.2** Under assumption 3.1, the sequence $(X^T, p^{T}_{Tt})_{T \geq 0}$ is C-tight for the Skorohod topology on $[0,1]$ as $T$ goes to infinity, with the following properties for any limit point $(X, P)$:

- $X$ is almost surely continuously differentiable.
- There exists a Brownian motion $B$ such that
\[
P_t = \int_0^t \sqrt{V_s} dB_s,
\]
where $V$ is the derivative of $X$ and the unique continuous solution of
\[
V_t = \mu + H_t + Z^2_t, \quad \text{with } H_t = \int_0^t 2\phi(t-s)V_s ds
\]
and
\[
Z_t = \int_0^t \sqrt{k}(t-s)\sqrt{V_s} dB_s, \quad \text{on } [0,1].
\]

- For any $\varepsilon > 0$, if $k = \varepsilon^{H+1/2\lambda}$ with $H \in (0,1/2)$ and $\lambda > 0$, $V$ has almost surely $H-\varepsilon$ Hölder regularity.

### 3.3. Discussion of theorem 3.2

- Compared to theorem 2.2, only one new term appears in the volatility equation (8). It comes from the self-exciting part in the Hawkes dynamic. Thus the elements in the discussion of the purely quadratic case in section 2.3 remain valid here.
- Let us consider the case where $k$ is a continuously differentiable kernel with $0 < \kappa(0) < +\infty$. Using integration by parts and Fabini’s theorem, we can write
\[
Z_t = \int_0^t \sqrt{k}(0)\sqrt{V_s} dB_s + \int_0^t \int_u^t \sqrt{k'(s-u)} \sqrt{V_s} dB_s ds
\]
and
\[
Z^2_t = \int_0^t \int_u^t \sqrt{k'(s-u)} \sqrt{V_s} dB_s ds.
\]
Therefore $Z$ is a semi-martingale and up to a finite variation term we have
\[
Z^2_t = \int_0^t 2\sqrt{k}(0)Z_s \sqrt{V_s} dB_s.
We see that the quadratic feedback term in the Hawkes dynamic induces a super-Heston type volatility because of the multiplicative term $Z_t$ in front of the $\sqrt{V_t}$ in the equation above.

- Let us take the kernel $k$ as the Mittag--Leffler function $f^{a,\lambda}$ with $a \in (1/2, 1)$ and $\lambda > 0$ and $\phi(t) = \kappa e^{-\kappa t}$ for some $\kappa > 0$. Adapting Theorem 2.1 in El Euch and Rosenbaum (2018) we get for any $h$ and $t_0$ positive

$$Z_{t_0+h} = \xi_{t_0}(h) + \tilde{Z}_h,$$

with $\tilde{Z}_h = \int_0^h \sqrt{f^{a,\lambda}}(h-s) \, dP_{s+h_0}$

and

$$\xi_{t_0}(h) = Z_{t_0} + \int_0^h f^{a,\lambda}(h-s) \theta_{t_0}(s) \, ds$$

where

$$\theta_{t_0}(h) = -Z_{t_0} + \frac{\alpha}{\lambda \Gamma(1 - \alpha)} \int_0^{t_0} (t_0 - s + h)^{-1-a} (Z_s - Z_{t_0}) \, ds - (h + t_0)^{1-a} \frac{Z_{t_0}}{\lambda \Gamma(1 - \alpha)}.$$ 

Then we can write the forward volatility as

$$V_{t+h} = H_t e^{-\mu t} + \left(\xi_{t_0}(h)\right)^2 + 2\xi_{t_0}(h)\tilde{Z}_h + \mu + \tilde{H}_h + (\tilde{Z}_h)^2$$

with

$$\tilde{H}_h = \int_0^h \phi(h-s) V_{s+h_0} \, ds.$$ 

The function $\xi_{t_0}$ only depends on $(Z_s)_{0 \leq s \leq t_0}$ and cannot be expressed as a function of $(V_s)_{0 \leq s \leq t_0}$. So we get from (9) that conditional on the history of the market from time 0 to $t_0$, the law of $(V_{t+h})_{h \geq 0}$ does depend on past returns and not only through past volatility. It means models (6) and (8) can reproduce the strong Zumbach effect. So when $k$ is a Mittag--Leffler function, model (8) is a super-Heston type rough volatility model with strong Zumbach effect.

In the case of exponential kernels $k(t) = \sqrt{2} e^{-\sqrt{2} t}$ and $\phi(t) = \kappa e^{-\kappa t}$ using similar computations, we prove that

$$V_{t+h} = \mu + \tilde{Z}_h + \tilde{H}_t + e^{-2\sqrt{2}h} Z_{t_0}^2 + 2Z_{t_0} e^{-\sqrt{2}h} + e^{-\sqrt{2}h} H_{t_0}.$$ 

- Finally remark that we do not prove uniqueness in law of the limit points $(X, P)$ in general. However, we can show uniqueness in the special case $\phi = 0$. This particular case can be treated because $Z$ is the solution of a stochastic Volterra equation which admits a unique strong solution, see section 5.1 for details and Jaber et al. (2019) for more results about uniqueness of rough equations.

4. Nearly unstable quadratic Hawkes models

We now focus on the case where the instability condition becomes almost violated at the limit. Let us consider a sequence of quadratic Hawkes processes $(N_T^T)_{T \geq 0}$ such that

$$\|\phi_T\|_1 + \|k_T\|_2^{1/2} \to 1.$$ 

Contrary to the sections before, we wish to work here with a natural renormalization (at least for $\phi$, see comments below assumption 3.1) and therefore take $\phi_T$ of the form $\phi_T = \beta T \phi$ with $\beta_T \in (0, 1)$ and $\|\phi_T\|_1 = 1$. We also assume that $\phi$ is high-tailed ($\phi(x) \sim x^{-(1 + \nu)}$ with $\nu \in (0, 1)$) as $x$ tends to infinity since this type of kernels leads to rough volatility in the case of linear Hawkes processes, see El Euch et al. (2018) and Jaisson and Rosenbaum (2016). Again, we start with insights about the suitable scaling procedure.

4.1. An adapted scaling procedure in the nearly unstable case

Let $a_T = \|k_T\|_2 + \|\phi_T\|_1$. We have

$$\mathbb{E}[\lambda_T^2] = \mu_T + \int_0^t (k_T^2 + \phi_T(t-s)) \mathbb{E}[\lambda_T^2] \, ds$$

and therefore

$$\mathbb{E}[\lambda_T^2] \leq \frac{\mu_T}{1-a_T}.$$ 

So we naturally define the following renormalized processes:

$$\lambda_T^{sT} = \frac{1-a_T}{\mu_T} \lambda_T^{ST}, \quad \Lambda_T^{sT} = \frac{1}{T} \int_0^T \lambda_T^{sT} \, ds$$

and

$$X_T^{sT} = \frac{1-a_T}{T\mu_T} N_T^{sT}.$$ 

Let us assume that $\lambda_T^{sT}$ converges to some $V$. We can then expect that $\Lambda_T^{sT}$ and $X_T^{sT}$ converge to some $\Lambda$ and $X$. Consider the rescaled martingales

$$M_t^{sT} = \sqrt{\frac{1-a_T}{T\mu_T}} M_t^{ST}$$

and $P_t^{sT} = \sqrt{\frac{1-a_T}{T\mu_T}} P_t^{ST}$,

where $M_t^{ST} = N_t^{ST} - \int_0^t \lambda_T^{sT} \, ds$. Since $[M_t^{ST}] = [P_t^{ST}] = N_t^{ST}$, we have $[M_t^{sT}] = [P_t^{sT}] = X_t^{sT}$. Moreover, $\langle M_t^{sT}, P_t^{sT} \rangle = 0$ and so $M_t^{sT}$ and $P_t^{sT}$ are likely to converge towards some martingales $M$ and $P$ with same bracket $X$ and such that $\langle M, P \rangle = 0$.

Let

$$\psi_T = \sum_{i \geq 1} \phi_T^i.$$ 

Using proposition 2.1 in Jaisson and Rosenbaum (2015), we deduce from (2) that

$$\lambda_T^t = \mu_T + (Z_T^t)^2 + \int_0^t \psi_T(t-s) (\mu_T + (Z_T^s)^2) \, ds$$

$$+ \int_0^t \psi_T(t-s) \, dM_T^{sT}.$$
So we have
\[ \lambda_t = (1 - \alpha_T) + \frac{1 - \alpha_T}{\mu_T} (Z_{\alpha_T}^T)^2 + \int_0^t (1 - \alpha_T) T \psi_T(T(t - s)) (1 + \frac{1}{\mu_T} (Z_{\alpha_T}^T)^2) ds + \int_0^t T(1 - \alpha_T) \psi_T(T(t - s)) \frac{1}{\sqrt{T \mu_T (1 - \alpha_T)}} dM_s^T. \]

The function \( T \psi_T(T \cdot) \) has \( L^1 \) norm equal to \((1 - \beta_T)^{-1}\). Therefore \( T(1 - \alpha_T) \psi_T(T) \) is non-vanishing only provided \( 1 - \beta_T \) is of order \( 1 - \alpha_T \). Consequently we set \( \beta_T = 2 \alpha_T - 1 \) (so that \( \beta_T < \alpha_T \)). Since \( \| \psi_T \|_1 = \beta_T \rightarrow 1 \) then \( \| k_T \|^2 \rightarrow 0 \). However we will see that the sequence \( k_T \) still plays a role in the limit.

In (10) the first integral is
\[ \int_0^T T(1 - \alpha_T) \psi_T(T(t - s)) ds. \]

It already appears in the case of a purely linear Hawkes process. We know from Jaisson and Rosenbaum (2015) and Jusselin and Rosenbaum (2018) that this term is crucial in the limiting behavior of the intensity and that a necessary condition to obtain a non-trivial scaling limit for it is that \( T^\alpha (1 - \alpha_T) \) tends to a positive constant. Under this specification, we need to impose additionally that \( T \mu_T (1 - \alpha_T) \) converges in order to obtain a non-degenerate asymptotic limit for the last integral in (10).

We now study the terms containing the quadratic feedback:
\[ \frac{1 - \alpha_T}{\mu_T} (Z_{\alpha_T}^T)^2 \quad \text{and} \quad \frac{1 - \alpha_T}{\mu_T} \int_0^t T \psi_T(T(t - s)) (Z_{\alpha_T}^T)^2 ds. \]

Since \( \| T \psi_T(T \cdot) \|_1 = (1 - \beta_T)^{-1} \) which tends to infinity, the second term dominates the first one. To make the second term converge, we need a proper behavior of \( Z_t^T = Z_t^T / \sqrt{T \mu_T} \). We have
\[ Z_t^T = \sqrt{\frac{T}{1 - \alpha_T}} \int_0^t k_T(T(t - s)) dP_{s^T}. \]

Thus we wish \( \sqrt{\frac{T}{1 - \alpha_T}} k_T(Tt) \) to converge and are naturally lead to assume that \( k_T \) is of the form
\[ k_T = k(\cdot / T) \frac{1 - \alpha_T}{T}. \]

Furthermore for some \( K > 0 \) and \( \alpha \in (0, 1) \),
\[ \lim_{s \to +\infty} \alpha \mu^2 \int_0^s \phi(s) ds = K. \]

(ii) The sequence of kernels \( (k_T)_{T \geq 0} \) satisfies \( k_T = k(t / T) \sqrt{\frac{1 - \alpha_T}{T}} \) with \( k \) a non-negative continuously differentiable function such that \( \| k \|_2 = 1 \) (in particular \( k(0) < +\infty \)).

(iii) Let \( \delta = K \frac{T(1 - \alpha_T)}{\alpha} \). There are two positive constants \( \lambda \) and \( \mu^* \) such that
\[ \lim_{T \to +\infty} (1 - \alpha_T) T^\alpha = \lambda \delta \quad \text{and} \quad \lim_{T \to +\infty} T^{1 - \alpha} \mu_T = \mu^* \delta^{-1}. \]

The choice of \( \delta \) in Point (iii) is just for convenience of notation in the results and proofs.

Recall that from Lemma 4.3 in Jaisson and Rosenbaum (2016), under assumption 4.1, the function
\[ F^T(t) = \int_0^t T(1 - \alpha_T) \psi_T(Ts) ds \]
converges towards \( \frac{1}{2} F^{0, \lambda} \) where
\[ F^{0, \lambda}(t) = \int_0^t f^{0, \lambda}(s) ds. \]

So
\[ T(1 - \alpha_T) \psi_T(Ts) \sim \frac{1}{2} f^{0, \lambda}(s). \]

This provides us intuition for the form of the limit \((V, Z)\) of (10)–(11):
\[ V_t = \int_0^t \frac{1}{2} f^{0, \lambda}(t - s) (1 + Z^2_s) ds + \int_0^t \frac{1}{2} f^{0, \lambda}(t - s) \frac{1}{\sqrt{k(s) \lambda \mu}} dM_s, \]

with \( Z_t = \int_0^t k(t - s) dP_s \), and where \( M \) and \( P \) are martingales such that \((M, P) = 0 \) and \((M_t, P_t) = (P) = \int_0^t V_s ds \).

We eventually state the main result of this section whose proof is given in section 5.3.

**Theorem 4.2** Under assumption 4.1, the sequence \((X^T, M^T, P^T)_{T \geq 0}\) is C-tight for the Skorohod topology on \([0, 1]\) as \( T \) goes to infinity, with the following properties for any limit point \((X, M, P)\):

- We have \((M) = (P) = X\) and
\[ X_t = \int_0^t \frac{1}{2} F^{0, \lambda}(t - s) (1 + Z^2_s) ds + \int_0^t \frac{1}{2} f^{0, \lambda}(t - s) \frac{1}{\sqrt{k(s) \lambda \mu}} M_s ds \]
with \( Z_t = \int_0^t k(t - s) dP_s \).

- If \( \alpha \in (1/2, 1) \), the process \( X \) is almost surely continuously differentiable with derivative \( V \) and up
to an enlargement of the filtration there exists two Brownian motions $B^{(1)}$ and $B^{(2)}$ such that $V$ is solution of

$$V_t = \int_0^t \left( \frac{1}{2} f_{\alpha, \lambda}^\alpha (t - s) \right) ds + \frac{1}{\sqrt{\lambda \mu}} \sqrt{V_s} dB_s^{(1)}$$

with

$$Z_t = \int_0^t k(t - s) \sqrt{V_s} dB_s^{(2)}.$$

Moreover, for any $\varepsilon > 0$, $V$ has almost surely $\alpha - \frac{1}{2} - \varepsilon$ Hölder regularity.

### 4.3. Discussion of theorem 4.2

- The form of the feedback is not the same in model (12) as in model (8). In (8), it is instantaneous through the $Z_t^2$ term while in (12) it is digested via a convolution with a fractional kernel.
- In model (12), price and volatility are driven by two different Brownian motions. This additional Brownian motion comes from the rescaling of the linear part of the intensity, as already observed for example in Jaisson and Rosenbaum (2015).
- Rough volatility appears for very different reasons in model (8) and model (12). In (12), the origin of rough volatility is the fat tail of the kernel $\phi$ while in (8) it arises from the behavior of the kernel $k$ in zero. Moreover it is clear from the proof of the last point of theorem 4.2 that the regularity of $Z$ has no influence on that of $V$.
- As computed in the previous section, we can write

$$Z_t = \int_0^t k(0) \sqrt{V_s} dB_s^{(2)} + \int_0^t \int_0^s k'(s - u) \sqrt{V_u} dB_u^{(2)} ds.$$

Therefore $Z$ is a semi-martingale and up to a finite variation term, we have

$$dZ_t^2 = 2k(0)Z_t \sqrt{V_t} dB_t^{(2)}.$$

Furthermore using integration by parts, we get

$$\int_0^t f_{\alpha, \lambda}^\alpha (t - s) Z_s^2 ds = \int_0^t F_{\alpha, \lambda} (t - s) dZ_s^2.$$

So up to a finite variation term, we have in model (12)

$$V_t = \int_0^t f_{\alpha, \lambda}^\alpha (t - s) \frac{1}{\sqrt{\lambda \mu}} \sqrt{V_s} dB_s^{(1)} + \int_0^t F_{\alpha, \lambda} (t - s) k(0)Z_s \sqrt{V_s} dB_s^{(2)}.$$

Thus as in model (6) and (8), the quadratic feedback term in the volatility dynamic induces that model (12) is a super-Heston type rough volatility model. Note however that in that case, the super-Heston and rough components are not the same.

- Using lemma A.2 in El Euch and Rosenbaum (2019), when $\alpha \in (1/2, 1)$, we get that equation (12) is equivalent to

$$V_t = V_0 + \frac{1}{\Gamma(\alpha)} \int_0^t (t - s)^{\alpha - 1} \lambda Z_t^2 + \theta_0(s) - V_s) ds$$

$$+ \frac{1}{\Gamma(\alpha)} \int_0^t (t - s)^{\alpha - 1} \sqrt{V_s} \frac{1}{\sqrt{\lambda \mu}} dB_s^{(1)}$$

with $\theta_0$ a deterministic function. In the case $k(t) = \sqrt{2} e^{-t^2}$ with $\nu > 0$, direct adaptation of theorem 2.1 in El Euch and Rosenbaum (2018) gives that

$$V_{h+h} = V_h + \frac{1}{\Gamma(\alpha)} \int_0^h (h - s)^{\alpha - 1} \lambda (Z_s^2 + 2Z_s^2 \nu e^{-s^2}) ds$$

$$+ \frac{1}{\Gamma(\alpha)} \int_0^h (h - s)^{\alpha - 1} \sqrt{V_{h+s}} \frac{1}{\sqrt{\lambda \mu}} dB_s^{(1)},$$

where $\theta_0(h)$ is equal to

$$\theta_0(t_0 + h) = \frac{\alpha}{\Gamma(1 - \alpha)} \int_0^h (t_0 - v + h)^{1-\alpha} (V_v - V_{h}) dv$$

$$+ \frac{(t + h)^{1-\alpha}}{\lambda \Gamma(1 - \alpha)} (V_0 - V_{h}) + Z_h^2 e^{-2\nu h}$$

and

$$\tilde{Z}_h = \int_0^h k(h - s) dP_{h+s}.$$
5.2. Proof of theorem 3.2

We proceed in three steps. First we prove the sequence \((X^T, P^{sT})_{T \geq 0}\) is C-tight for the Skorohod topology. Then we show the results about the dynamics of the limit points. Finally we establish the regularity properties of the limit points.

5.2.1. Tightness of the sequence \((X^T, P^{sT})_{T \geq 0}\). We consider the processes

\[
\Lambda^T_t = \int_0^t \lambda^T_s \, ds \quad \text{and} \quad Z^T_t = \int_0^t (k(t-s) \, dP^s)
\]

defined for \(t \in [0, 1]\). Remark that \(\Lambda^T_t\) is the predictable compensator of \(X^T\). We have the following equality:

\[
\mathbb{E}[\lambda^T_t] = \mu_T + \int_0^t (\kappa_2 + \phi_T)(t-s) \, d\mathbb{E}[\lambda^T_s] \, ds.
\]

Thus

\[
\mathbb{E}[\lambda^T_t] \leq \frac{\mu}{1 - \gamma - \beta}.
\]

and consequently

\[
\mathbb{E}[X^T_T] = \mathbb{E}[\Lambda^T_T] \leq \frac{\mu}{1 - \gamma - \beta}.
\]

Since the processes \(X^T\) and \(\Lambda^T_t\) are increasing for any \(T\), using the last inequality, we deduce from theorem VI-3.21 together with proposition VI-3.35 in Jacod and Shiryaev (2013) that \((X^T)^{\Lambda^T}_{T \geq 0}\) and \((\Lambda^T)^{\Lambda^T}_{T \geq 0}\) are tight. Moreover since \(|\Delta X^T_t| + |\Delta \Lambda^T_t| \leq 1/2\) almost surely on \([0, 1]\), according to proposition VI-3.26 in Jacod and Shiryaev (2013), \((X^T)^{\Lambda^T}_{T \geq 0}\) and \((\Lambda^T)^{\Lambda^T}_{T \geq 0}\) are C-tight. The tightness of \((M^T)^{\Lambda^T}_{T \geq 0}\) and \((P^T)^{\Lambda^T}_{T \geq 0}\) follows from theorem VI-4.13 in Jacod and Shiryaev (2013) using that \(\langle M^T, \lambda^T \rangle = \langle P^T, \lambda^T \rangle = \Lambda^T_t\). We then get C-tightness because \(|\Delta M^T_t| + |\Delta P^T_t| \leq 2/2\). Finally \((X^T, \Lambda^T_t, M^T_t, P^T_t)_{T \geq 0}\) is C-tight for the Skorohod topology on \([0, 1]\).

We also show that the sequence \((Z^T, \Lambda^T_t)_{T \geq 0}\) is tight for the \(L^2([0,1])\) topology. For this, inspired by Abi Jaber et al. (2019), we consider the Sobolev–Slobodeckij norm defined for any measurable function \(f\) by

\[
\|f\|_{W^{1,2}([0,1])} = \left( \int_0^1 f(s)^2 \, ds + \int_0^1 \int_0^1 \left| \frac{f(t) - f(s)}{|t-s|^{1+2q}} \right|^2 \, ds \, dt \right)^{1/2}.
\]

We recall that the closed balls of \(\| \cdot \|_{W^{1,2}([0,1])}\) are relatively compact in \(L^2([0,1])\), see Flandoli and Gatarek (1995).

Therefore it is enough to show that \(\mathbb{E}[\|Z^T_t\|^2_{W^{1,2}([0,1])}]_{T \geq 0}\) is uniformly bounded to conclude the tightness of \((Z^T, \Lambda^T_t)_{T \geq 0}\) in \(L^2([0,1])\).

For any \(t \in [0, 1]\), we have using Ito’s formula

\[
\mathbb{E}[(Z^T_t)^2] = \int_0^t k^2(t-s) \mathbb{E}[X^T_t] \, ds \leq \frac{\mu}{1 - \gamma - \beta} \mathbb{E}[|k|_2^2]
\]

and for \(0 \leq s \leq t \leq 1\)

\[
Z^T_t - Z^T_s = \int_s^t (k(t-u) - k(s-u)) \, dP^T_u.
\]

Then we get

\[
\mathbb{E}[(Z^T_t - Z^T_s)^2] = \int_s^t k^2(t-u) \mathbb{E}[X^T_u] \, du + \int_0^s (k(t-u) - k(s-u))^2 \mathbb{E}[\lambda^T_u] \, du.
\]

Using that \(\mathbb{E}[\lambda^T_u] \leq \frac{\mu}{1 - \gamma - \beta}\) we obtain

\[
\mathbb{E}[(Z^T_t - Z^T_s)^2] \leq \frac{\mu}{1 - \gamma - \beta} \left( \int_s^t k^2(t-u) \, du + \int_0^s (k(t-u) - k(s-u))^2 \, du \right).
\]

According to Abi Jaber et al. (2019) we have

\[
\int_0^1 \int_0^1 \int_s^t k(s \vee t - u)^2 \, du \, ds \, dt \leq \frac{1}{\eta} \int_0^1 |k(t)|^2 t^{-2q} \, dt
\]

and

\[
\int_0^1 \int_0^1 \int_s^t \frac{|k(t-u) - k(s-u)|^2}{|t-s|^{1+2q}} \, du \, ds \, dt \leq \int_0^1 \int_0^1 \int_s^t \frac{|k(t) - k(s)|^2}{|t-s|^{1+2q}} \, du \, ds \, dt,
\]

which is bounded from Assumption 3.1 \((ii)\). Finally using Fubini’s theorem we deduce that \(\mathbb{E}[\|Z^T\|^2_{W^{1,2}([0,1])}]_{T \geq 0}\) is bounded. So \((Z^T, \Lambda^T_t)_{T \geq 0}\) is tight in \(L^2([0,1])\).

Before going to the next step we prove the following lemma.

**Lemma 5.1** The sequence of martingales \(X^T - \Lambda^T_t\) converges to 0 uniformly in probability on \([0, 1]\).

**Proof** Since \(N^T_t = \int_0^t \lambda^T_s \, ds\) is a true martingale, from Doob’s inequality we get

\[
\mathbb{E} \left[ \sup_{t \in [0,1]} \left| X^T_t - \Lambda^T_t \right|^2 \right] \leq \frac{1}{T^2} \mathbb{E}[N^T_T].
\]

Using that \(\mathbb{E}[N^T_T] = T \mathbb{E}[\Lambda^T_T]\) we deduce

\[
\mathbb{E} \left[ \sup_{t \in [0,1]} \left| X^T_t - \Lambda^T_t \right|^2 \right] \leq \frac{\mu}{T^2 (1 - \gamma - \beta)}
\]

which concludes the proof.
5.2.2. Dynamic of the limit points. We now consider $(X, X, M, P, Z)$ a limit point of $(X_t^T, \Lambda_t^T, M^T_t, P^T_t, Z^T_t)_{t \geq 0}$. Using Skorohod representation theorem and the fact that $(X, X, M, P)$ is continuous, we may consider that almost surely $(X_t^T, \Lambda_t^T, M^T_t, P^T_t)_{t \geq 0}$ converges uniformly on $[0, 1]$ towards $(X, X, M, P)$ and $(Z^T_t)_{t \geq 0}$ converges in $L^2([0, 1])$ towards $Z$:

$$\sup_{t \in [0, 1]} |X_t^T - X_t| \to 0, \quad \sup_{t \in [0, 1]} |M_t^T - M_t| \to 0,$$

$$\sup_{t \in [0, 1]} |P_t^T - P_t| \to 0$$

Towards (surely)

From corollary IX-1.19 in Jacod and Shiryaev (2013) we have that $M_t$ is adapted. Moreover using Ito’s formula together with Fubini’s theorem we have

$$\int_0^t \beta \phi(t-s) X_s \, ds = \int_0^t F_1(t-s) \, dX_s,$$

So we obtain the almost sure uniform convergence of $(\Lambda_t^T)_{t \geq 0}$ towards (15). Consequently, using lemma 5.1, we deduce

$$X_t = \int_0^t \mu + Z_t^2 \, ds + \int_0^t F_1(t-s) \, dX_s$$

and eventually

$$X_t = \int_0^t \mu + Z_t^2 + \int_0^t \beta \phi \psi(u) \, dX_u \, ds.$$

Thus $X$ is absolutely continuous with respect to the Lebesgue measure with derivative $V$ given by

$$V_t = \mu + Z_t^2 + \int_0^t \beta \phi \psi(u) \, dX_u.$$

Letting $\psi = \sum_{i \geq 1} \beta \phi \psi_i$ we have

$$V_t = \mu + Z_t^2 + \int_0^t \psi(u) Z_s^2 \, ds.$$

The boundedness of $(E[Z_s^2])_{t \in [0, 1]}$ gives that $(V_t)_{t \in [0, 1]}$ is uniformly bounded in $L^1$.

We now prove that

$$Z_t = \int_0^t k(t-s) \, dP_s.$$

Using Cauchy–Schwarz inequality, the convergence of $(Z^T_t)_{t \geq 0}$ implies that almost surely, uniformly in $t \in [0, 1]$,

$$\int_0^t Z_s^T \, ds \to \int_0^t Z_s \, ds.$$

From Ito’s formula we get

$$\int_0^t Z_s^T \, ds = \int_0^t (k(t-s) P_s^T \, ds$$

and using equation (14) we deduce that it converges almost surely uniformly towards

$$\int_0^t k(t-s) P_s \, ds.$$

Since $F_2(t) = \int_0^t \gamma k^2(s) \, ds < 1$ we have

$$\int_0^t \int_0^s (k(s-u))^2 \, dX_u \, ds = \int_0^t F_2(t-s) \, dX_s < \infty.$$

So we can use the stochastic Fubini theorem and show that

$$\int_0^t \int_0^s k(s-u) \, dP_s \, ds = \int_0^t k(t-s) P_s \, ds.$$
Thus almost surely, for any $t \in [0, 1],$
\[\int_0^t Z_s \, ds = \int_0^t \int_0^s k(s - u) \, dP_u \, ds\]
and
\[Z_t = Z_0 + \int_0^t \sqrt{\gamma} k(t - s) \, dP_s.\]
Moreover, from theorem V-3.9 in Revuz and Yor (2013), there exists a Brownian motion $B$ such that
\[P_t = \int_0^t \sqrt{V} \, dB_s\]
and finally we get
\[Z_t = \int_0^t \sqrt{\gamma} k(t - s) \sqrt{V} \, dB_s.\]

We recall that $(E[V], \mu)$ is bounded in $L^2$. So using assumption 3.1 ii) together with theorems 3.1 and 3.3 in Zhang (2010) we obtain that the process $Z$ is continuous. Therefore using (16) $V$ is also continuous. This concludes this part of the proof.

5.2.3. Regularity property. We now consider that $k$ is given by $f^{H+1/2,\lambda}$ for $H \in (0, 1/2)$ and $\lambda > 0$. We can write
\[
\int_0^t Z_s \, ds = \int_0^t \int_0^s k(s - u) \, dP_u \, ds.
\]
Since $P_t = \int_0^t \sqrt{V} \, dB_s$, $P$ has the same regularity as a Brownian motion. We can thus use the same arguments as in section 4.4 in Jajson and Rosenbaum (2016) to deduce that $Z$, and therefore $V$, are $H - \epsilon$ Hölder for any $\epsilon > 0$.

5.3. Proof of theorem 4.2
We proceed again in three steps. First we show that the sequence $(X^T, M^{sT}, P^{sT})_{T \geq 0}$ is $C$-tight for the Skorohod topology. Then we prove the results about the dynamics of the limit points and finally those on the regularity of the limit points.

5.3.1. Tightness of $(X^T, M^{sT}, P^{sT})_{T \geq 0}$. Recall the definition of the renormalized processes
\[\lambda^s_T = \frac{1 - \alpha_T}{\mu_T} \lambda_T, \quad \Lambda^s_T = \frac{1 - \alpha_T}{T/\mu_T} \int_0^T \lambda^s_t \, ds,
\]
\[X^T = \frac{1 - \alpha_T}{T/\mu_T} M^s_T, \quad Z^s_T = Z_T / \sqrt{T},
\]
\[M^{sT} = \sqrt{\frac{1 - \alpha_T}{T/\mu_T}} M^s_T \quad \text{and} \quad P^{sT} = \sqrt{\frac{1 - \alpha_T}{T/\mu_T}} P^s_T.
\]
We have
\[\mathbb{E}[\lambda^s_T] \leq \mu_T + \int_0^t (k^s_T(t - s) + \phi_T(t - s)) \mathbb{E}[\lambda^s_T] \, ds.
\]
Thus
\[\mathbb{E}[\lambda^s_T] \leq \frac{\mu_T}{1 - \|\phi_T\|_1 - \|k_T\|_2^2}
\]
and consequently
\[\mathbb{E}[\lambda^s_T] \leq \frac{1 - \alpha_T}{1 - \beta_T - \|k_T\|_2^2} = 1.
\]
So
\[\mathbb{E}[X^T_t] = \mathbb{E}[\Lambda^s_T] \leq 1,
\]
which gives the tightness of the sequences $(X^T_t)_{t \in [0, 1]}$ and $(\Lambda^s_T)_{t \in [0, 1]}$, both of them being increasing. Actually we can get $C$-tightness since $|\Delta X^T_t| + |\Delta \Lambda^s_T| \leq \frac{T}{\sqrt{T}}$ that goes to zero as $T$ goes to infinity. Remark that lemma 5.1 still holds under assumption 4.1.

The tightness of $(M^{sT})_{T \geq 0}$ and $(P^{sT})_{T \geq 0}$ follows from theorem VI-4.13 in Jacod and Shiryaev (2013) because $(M^{sT}) = (P^{sT}) = \Lambda^s_T$ and $(M^{sT}, P^{sT}) = 0$. We then obtain $C$-tightness since $|\Delta M^{sT}| + |\Delta P^{sT}| \leq 2/T$. Finally $(X^T, M^{sT}, P^{sT})_{T \geq 0}$ is $C$-tight for the Skorohod topology on $[0, 1]$.

5.3.2. Dynamics of the limit points. We now take $(X, M, P)$ a limit point of $(X^T, M^{sT}, P^{sT})_{T \geq 0}$. Since $(X, M, P)$ is continuous, according to the Skorohod representation theorem, we can consider that $(X^T, M^{sT}, P^{sT})_{T \geq 0}$ converges almost surely uniformly towards $(X, M, P, P)$:
\[
\sup_{r \in [0, 1]} |X^T_t - X_t| \to 0, \quad \sup_{r \in [0, 1]} |M^{sT}_t - M_t| \to 0, \quad \sup_{r \in [0, 1]} |P^{sT}_t - P_t| \to 0, \quad T \to +\infty.
\]
From corollary IX-1.19 in Jacod and Shiryaev (2013), we have that $M$ and $P$ are local martingales. Moreover since $(M^{sT}) = (P^{sT}) = X^T$, we have $|M_t| = |P_t| = X$ and $(M, P) = 0$ using corollary VI-6.29 in Jacod and Shiryaev (2013). Because $M$ and $P$ are continuous, we deduce
\[\langle M \rangle = |M| = |P| = |X|.
\]
Because $\mathbb{E}[X^T_t] \leq 1$ is uniformly bounded in $T$, we get that $X_t$ is in $L^1$ and so $M$ and $P$ are true martingales. In addition, the Dambis–Dubin–Schwarz theorem gives the existence of two independent Brownian motions $B^{(1)}$ and $B^{(2)}$ such that
\[M_t = B^{(1)}_t, \quad P_t = B^{(2)}_t.
\]
Recall that for $F^T(t) = \int_0^t T(1 - \alpha_T) \psi(Ts) \, ds$ we have
\[\Lambda^{sT}_t = t(1 - \alpha_T) + \int_0^t F^T(t - s) \, ds + \int_0^t F^T(t - s) \, dM^{sT}_t
\]
and
\[\int_0^t F^T(t - s)(Z^{sT}_t)^2 \, ds + \int_0^t (1 - \alpha_T)(Z^{sT}_t)^2 \, ds.
\]
According to lemma 4.3 in Jaisson and Rosenbaum (2016), we have the uniform convergence
\[ \int_0^t F^T(t-s)ds \to \int_0^t \frac{1}{2} F^{\alpha, \lambda}(t-s)ds. \]

Using integration by parts, we obtain
\[ Z_s^T = k(0)P_s^T + \int_0^t k(t-s)P_s^T ds. \]

Assumption 4.1 i) implies that \( k' \) is bounded on \([0, 1]\). As a consequence of (17) we have that almost surely, \( Z^T \) converges uniformly on \([0, 1]\) towards
\[ k(0)P_t + \int_0^t k(t-s) dP_s ds = \int_0^t k(t-s) dP_s \]
which is continuous. This convergence together with lemma 4.3 in Jaisson and Rosenbaum (2016) implies that almost surely, uniformly in \( t \) \( Z^T \) converges uniformly in \([0, 1]\),
\[ \int_0^t F^T(t-s)(Z_s^T)^2 ds \to \int_0^t F^{\alpha, \lambda}(t-s)Z_s^2 ds \]
and
\[ (1 - a_T) \int_0^t (Z_s^T)^2 ds \to 0. \]

We now prove that \( \int_0^t \frac{F^T(t-s)}{\sqrt{(1-a_T)\mu_T}} dM_s^T \) converges uniformly in probability towards
\[ \int_0^t \frac{f^{\alpha, \lambda}(t-s)}{2\sqrt{X_T^T}} M_s ds. \]

Using integration by parts we have
\[ \int_0^t \frac{F^T(t-s)}{\sqrt{(1-a_T)\mu_T}} dM_s^T = \int_0^t \frac{f^T(t-s)}{\sqrt{(1-a_T)\mu_T}} M_s^T ds. \]

Remark that
\[ \int_0^t f^T(t-s)M_s^T ds - \int_0^t f^{\alpha, \lambda}(t-s) \frac{1}{2} M_s ds \]
can be written
\[ \int_0^t \frac{1}{2} f^{\alpha, \lambda}(t-s)(M_s^T - M_s) ds + \int_0^t (f^T(t-s) - \frac{1}{2} f^{\alpha, \lambda}(t-s)) M_s^T ds. \]

The first term in (18) goes almost surely uniformly to zero using (17) and the fact that \( f^{\alpha, \lambda} \in L^1 \). Applying integration by parts again we obtain
\[ \int_0^t (f^T(t-s) - \frac{1}{2} f^{\alpha, \lambda}(t-s)) M_s^T ds \]
and using Burkholder–Davis–Gundy inequality we get (\( C \) denotes here a positive constant that varies from line to line)
\[ \mathbb{E} \left[ \sup_{t \in [0,1]} \left( \int_0^t \left( F^T(t-s) - \frac{1}{2} f^{\alpha, \lambda}(t-s) \right) dM_s^T \right)^2 \right] \]
\[ \leq C \int_0^t \left( F^T(t-s) - \frac{1}{2} f^{\alpha, \lambda}(t-s) \right)^2 \frac{1 - a_T}{\mu_T} ds \]
\[ \leq C \int_0^t \left( F^T(t-s) - \frac{1}{2} f^{\alpha, \lambda}(t-s) \right)^2 ds. \]

This converges to zero according to lemma 4.3 in Jaisson and Rosenbaum (2016). So we have proved that
\[ X_t = \int_0^t \frac{1}{2} f^{\alpha, \lambda}(t-s)(1 + Z_s^2) ds + \int_0^t f^{\alpha, \lambda}(t-s) \frac{1}{2\sqrt{\lambda T}} M_s ds \]
with
\[ Z_s^2 = \int_0^t k(t-s) dP_s. \]

5.3.3. Regularity property. We can write \( X \) as
\[ X_t = \int_0^t \frac{1}{2} f^{\alpha, \lambda}(t-s)(s + \int_0^s Z_u^2 du + M_s) ds. \]

Since \( Z \) is continuous, \( \int_0^t Z_u^2 du \) is continuously differentiable. So using the same arguments as in sections 4.3 and 4.4 in Jaisson and Rosenbaum (2016) replacing \( s \) by \( s + \int_0^s Z_u^2 du \), we obtain that almost surely, \( X \) is differentiable with derivative \( V \) satisfying
\[ V_t = \int_0^t \frac{1}{2} f^{\alpha, \lambda}(t-s)(1 + Z_s^2) ds + \int_0^t f^{\alpha, \lambda}(t-s) \frac{1}{2\sqrt{\lambda T}} M_s ds. \]

We get the stated result using theorem V-3.9 in Revuz and Yor (2013) which gives the existence of two independent Brownian motions \( B_1^{(1)} \) and \( B_2^{(2)} \) such that
\[ M_t = \int_0^t \sqrt{V_t} dB_t^{(1)} \text{ and } P_t = \int_0^t \sqrt{V_t} dB_t^{(2)}. \]

The regularity property of \( V \) is also deduced using the same arguments as in sections 4.3 and 4.4 in Jaisson and Rosenbaum (2016).
Acknowledgements

The authors are grateful to Jean-Philippe Bouchaud and Jim Gatheral for many interesting discussions. The editors wish to thank Jim Gatheral for providing the feature graph which used data from the Oxford-Man Institute of Quantitative Finance, Realized Library.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The authors gratefully acknowledge financial support of the ERC Grant 679836 Staqamof and of the Chair Analytics and Models for regulation.

References

Abi Jaber, E., Cuchiero, C., Larsson, M. and Pulido, S., Existence and stability for stochastic Volterra equations of convolution type with jumps. Preprint, 2019.

Abi Jaber, E. and El Euch, O., Multi-factor approximation of rough volatility models. SIAM J. Financ. Math., 2019, 10(2), 309–349.

El Euch, O. and Rosenbaum, M., The characteristic function of rough Heston models. Math. Finance, 2019, 29(1), 3–38.

Engle, R.F., Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. Econometrica, 1982, 50(4), 987–1007.

Engle, R.F. and Bollerslev, T., Modelling the persistence of conditional variances. Econom. Rev., 1986, 5(1), 1–50.

Filimonov, V. and Sornette, D., Quantifying reflexivity in financial markets: toward a prediction of flash crashes. Phys. Rev. E., 2012, 85(5), 056108.

Filandoli, F. and Gatarek, D., Martingale and stationary solutions for stochastic Navier–Stokes equations. Probab. Theory. Relat. Fields., 1995, 102(3), 367–391.

Gatheral, J., Jaissle, T. and Rosenbaum, M., Volatility is rough. Quant. Finance, 2018, 18(6), 933–949.

Glasserman, P. and He, P., Buy rough, sell smooth. Working paper, 2018.

Hardiman, S.J., Bercot, N. and Bouchaud, J.-P., Critical reflexivity in financial markets: A Hawkes process analysis. Eur. Phys. J. B, 2013, 86(10), 442.

Jaber, E.A., Larsson, M., Pulido, S., et al. Affine Volterra processes. Ann. Appl. Probab., 2019, 29(5), 3155–3200.

Jacod, J., Multivariate point processes: Predictable projection, Radon–Nikodym derivatives, representation of martingales. Z. Wahrscheinlichkeit Verwandte Gebiete, 1975, 31(3), 235–253.

Jacod, J. and Shiryaev, A., Limit Theorems for Stochastic Processes, Vol. 288, 2013 (Springer Science & Business Media).

Jaissle, T. and Rosenbaum, M., Limit theorems for nearly unstable Hawkes processes. Ann. Appl. Probab., 2015, 25(2), 600–631.

Jaissle, T. and Rosenbaum, M., Rough fractional diffusions as scaling limits of nearly unstable heavy tailed Hawkes processes. Ann. Appl. Probab., 2016, 26(5), 2860–2882.

Jusselin, P. and Rosenbaum, M., No-arbitrage implies power-law market impact and rough volatility. Math. Finance, 2018.

Livieri, G., Mouti, S., Pallavicini, A. and Rosenbaum, M., Rough volatility: Evidence from option prices. ISEE Trans., 2018, 50(9), 767–776.

Lynch, P.E. and Zumbach, G., Market heterogeneities and the causal structure of volatility. Quant. Finance, 2003, 3(4), 320–331.

Nelson, D.B., ARCH models as diffusion approximations. J. Econometrics, 1990, 45(1-2), 7–38.

Ogata, Y., On Lewis’ simulation method for point processes. IEEE Trans. Inform. Theory, 1981, 27(1), 23–31.

Revuz, D. and Yor, M., Continuous Martingales and Brownian Motion, Vol. 293, 2013 (Springer Science & Business Media).

Sentana, E., Quadratic ARCH models. Rev. Econ. Stud., 1995, 62(4), 639–661.

Zhang, X., Stochastic Volterra equations in banach spaces and stochastic partial differential equation. J. Funct. Anal., 2010, 258(4), 1361–1425.

Zumbach, G., Time reversal invariance in finance. Quant. Finance, 2009, 9(5), 505–515.

Zumbach, G., Volatility conditional on price trends. Quant. Finance, 2010, 10(4), 431–442.