CREATES Research Paper 2009-25

Stochastic volatility of volatility in continuous time

Ole E. Barndorff-Nielsen and Almut E. D. Veraart

School of Economics and Management
Aarhus University
Bartholins Allé 10, Building 1322, DK-8000 Aarhus C
Denmark

Stochastic volatility of volatility in continuous time

Ole E. Barndorff-Nielsen
The T.N. Thiele Centre for Mathematics in Natural Science,
Department of Mathematical Sciences,
& CREATES,
Aarhus University
oebn@imf.au.dk

Almut E. D. Veraart *
CREATES, School of Economics and Management,
Aarhus University
averaart@creates.au.dk

June 7, 2009

Abstract

This paper introduces the concept of stochastic volatility of volatility in continuous time and, hence, extends standard stochastic volatility (SV) models to allow for an additional source of randomness associated with greater variability in the data. We discuss how stochastic volatility of volatility can be defined both non–parametrically, where we link it to the quadratic variation of the stochastic variance process, and parametrically, where we propose two new SV models which allow for stochastic volatility of volatility. In addition, we show that volatility of volatility can be estimated by a novel estimator called *pre–estimated spot variance based realised variance*.

Keywords: Stochastic volatility; volatility of volatility; non-Gaussian Ornstein–Uhlenbeck process; superposition; leverage effect; Lévy processes.

JEL classification: C10, C13, C14, G10.

1 Introduction

Stochastic volatility (SV) models have been widely used in finance in the last decade since they are particularly suitable for coping with many stylised facts of asset returns such as time–varying volatility, volatility clusters and the leverage effect, i.e. the usually negative correlation between asset prices and volatility. This paper extends this line of investigation by introducing an additional source of randomness representing *stochastic volatility of volatility*. Here we view stochastic volatility of volatility as expressing the possibility or fact that there is greater variability – i.e. more volatility – in the data structure under study than might initially be surmised. In modelling terms this means that we consider the initial thinking as embodied in a (classical) SV model and want to describe the extra variability by a further source of randomness.

There are basically two ways of thinking about volatility of volatility: One way is motivated by a non–parametric point of view where we measure this extra source of randomness by the quadratic variation of the variance process (QVV, thereafter), which is analogous to measuring the variability of an asset by means of the quadratic variation (QV) of the logarithmic asset price. An alternative

*Financial support by the Center for Research in Econometric Analysis of Time Series, CREATES, funded by the Danish National Research Foundation is gratefully acknowledged.

way of thinking about stochastic volatility of volatility would be by means of parametric models. Both approaches will be studied in this paper.

The literature on stochastic volatility models comprises at least three branches of research where stochastic volatility of volatility is mentioned: First, there are the time series models for realised variance (RV), i.e. the sum of squared returns, which allow for heterogeneous variance, see e.g. Corsi (2004), Corsi et al. (2008), Bollerslev, Kretschmer, Pigorsch & Tauchen (2009). In that literature, the authors find that the variance of realised variance is time varying and possibly stochastic.They conclude that having a heterogeneous variance of RV (in addition to non–Gaussian errors in their time series model) results in a very good forecasting performance of their model for RV. Second, there is a new area of research where the existence of a time–varying variance risk premium is linked to the existence of stochastic volatility of volatility, see e.g. Drechsler & Yaron (2008) and Bollerslev, Tauchen & Zhou (2009). Third, there is recent research on general inference problems in stochastic volatility models, see e.g. Mykland & Zhang (2009), where stochastic volatility of volatility is mentioned. In particular, it is a key quantity when it comes to making inference on the *spot volatility*. Hence, by studying volatility of volatility thoroughly one can get a better understanding of financial market volatility itself.

To the best of our knowledge, the present paper is the first which contains a systematic treatment of stochastic volatility of volatility models in continuous time. We will proceed as follows. First of all, we start with a conceptual part, where we define stochastic volatility of volatility both in a parametric model and non–parametrically. In the parametric context, we study two new stochastic volatility models, which are based on the Barndorff-Nielsen & Shephard (2001) model, (BNS model, thereafter) and which account for stochastic volatility of volatility. Analytical properties of the resulting model classes are discussed and we also point out how stochastic volatility of volatility can be used for introducing leverage type effects in a novel way. Next, we sketch how such models can be extended to a multivariate set up. By generalising the two specific model classes, we then define stochastic volatility of volatility more formally in two generic parametric SV models. In particular, we show that an additional source of randomness can be introduced into SV models either by spatial or by temporal scaling of the driving process.

Next, we study stochastic volatility of volatility non–parametrically by means of the quadratic variation of the squared volatility process and we discuss its relationship to parametric models which allow for stochastic volatility of volatility.

Furthermore, we turn our attention to the question of how volatility of volatility can be estimated based on high frequency financial data. In particular, we develop a methodology for estimating the quadratic variation of the squared volatility consistently by using a novel two time scale approach.

Finally we give an extended outlook on further research questions in the context of the new concept of stochastic volatility of volatility.

2 A brief review on stochastic volatility models

There are basically two classes of asset price models in the literature, which allow for stochastic volatility. In these, stochastic volatility is introduced either by stochastic spatial scaling of a semimartingale $S = (S_t)_{t\geq 0}$ or by stochastic temporal scaling. The first approach results in models of the type $\int_0^t \sigma_{s-} dS_s$, where $\sigma = (\sigma_s)_{s\geq 0}$ is a stochastic volatility process, and the latter leads to models of the type S_{τ_t}, where $\tau = (\tau_t)_{t\geq 0}$ is a stochastic time change. Since the first type of models is now more widely used in the literature (see e.g. Barndorff-Nielsen & Shephard (2002, 2007), Jacod (2008), Aït-Sahalia & Jacod (2009)) and since it has a bigger potential for generalisation to multivariate models, we entirely focus on that class in order to introduce stochastic variance of variance. In particular, we will assume in the following that the logarithmic asset price $Y = (Y_t)_{t\geq 0}$ is given by an Itô semimartingale

$$dY_t = a_t dt + \sigma_{t-} dW_t + dJ_t, \tag{1}$$

which is defined on a probability space $(\Omega, \mathcal{A}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, where $a = (a_t)_{t\geq 0}$ is a predictable drift process, $\sigma = (\sigma_t)_{t\geq 0}$ is a predictable stochastic volatility process and $J = (J_t)_{t\geq 0}$ is the jump component of the Itô semimartingale. Note that an Itô semimartingale is defined as a semimartingale whose characteristics are absolutely continuous with respect to the Lebesgue measure (see e.g. Jacod (2008)). So, for the jump component, we assume that

$$J_t = \int_0^t \int_{\mathbb{R}} \kappa(\delta(s,x))(\widetilde{\mu}(ds,dx) - \widetilde{\nu}(ds,dx)) + \int_0^t \int_{\mathbb{R}} (\delta(s,x) - \kappa(\delta(s,x)))\widetilde{\mu}(ds,dx),$$

where $\nu(ds,dx) = dsF_s(dx)$ and δ is a predictable map from $\Omega \times \mathbb{R}_+ \times \mathbb{R}$ on $\mathbb{R}$ such that the predictable random measure $F_s(\omega,dx)$ is the restriction to $\mathbb{R} \setminus \{0\}$ of the image of the Lebesgue measure on $\mathbb{R}$ by the map $x \mapsto \delta(\omega,t,x)$, and $\widetilde{\mu}$ is a Poisson random measure with predictable compensator $\widetilde{\nu}$. Also, κ is a continuous truncation function, which is bounded with compact support and $\kappa(x) = x$ on a neighbourhood of x.

The variation of financial markets, which is often referred to as squared *volatility*, is usually measured by means of the quadratic variation of the logarithmic price process. In our modelling framework, the quadratic variation (QV) (denoted by $[\cdot]$) is given by

$$[Y]_t = \sigma_t^{2+} + \sum_{0 \leq s \leq t} (\Delta J_s)^2, \tag{2}$$

where $\sigma_t^{2+} = \int_0^t \sigma_s^2 ds$ is the integrated squared stochastic volatility process and where $\Delta J_s = J_s - J_{s-}$ denotes the jump of J at time s. Taking the square root of the quadratic variation $\sqrt{[Y]_t}$ leads to a measure of the *volatility* of the asset price.

3 Parametric stochastic volatility of volatility

First of all, we study stochastic volatility of volatility in a parametric model. In order to get a better understanding of what this new concept comprises, we study two concrete examples of SV models which allow for this additional source of randomness first before we turn to a more formal definition of the new concept of stochastic volatility of volatility in parametric SV models.

3.1 Volatility modulated non–Gaussian Ornstein Uhlenbeck processes

In order to define stochastic volatility of volatility in a concrete parametric SV model, we choose the Barndorff-Nielsen & Shephard (2001, 2002) model (BNS model), based on a non–Gaussian Ornstein–Uhlenbeck process, as a base model, since it has been widely used in the finance literature. Note that non–Gaussian Ornstein Uhlenbeck processes (as well as the often used CIR process (Cox et al. (1985), Heston (1993)) have an exponentially declining autocorrelation function which contradicts empirical findings of long memory in volatility. However, by studying a superposition of such processes as in Barndorff-Nielsen & Shephard (2002), one can easily overcome this problem (see also, Section A.3).

This paper proposes two new classes of stochastic volatility models which are given by *volatility modulated non–Gaussian Ornstein Uhlenbeck processes*.

Model 1: The stochastic volatility process σ satisfies $\sigma_t^2 = V_t$ for a stochastic process $V = (V_t)_{t \geq 0}$, where

$$dV_t = -\lambda V_t dt + \omega_{\lambda t-} dL_{\lambda t}; \tag{3}$$

$(L_t)_{t \geq 0}$ is a Lévy subordinator and $(\omega_t)_{t \geq 0}$ denotes a stationary non–negative stochastic volatility process, which is assumed to be independent of L.

Model 2: The stochastic volatility process σ is defined by $\sigma_t^2 = U_t$ for a process $U = (U_t)_{t \geq 0}$ with

$$dU_t = -\lambda U_t dt + dL_{\tau_{\lambda t}}; \tag{4}$$

$(L_t)_{t \geq 0}$ is a Lévy subordinator, $\tau = (\tau_t)_{t \geq 0}$ is a time change process with stationary increments, which is independent of L.

Remark The processes ω^2 and τ can be interpreted as the *stochastic variability of variance*, in Model 1 and Model 2, respectively. Clearly, when $\omega_t \equiv 1$ or $\tau_{\lambda t} = \lambda t$, we obtain the well-known BNS model. Both ω and τ can be driven by a Brownian motion or (and) a jump process. E.g. we can think of ω (or ω^2) being a CIR process.

Note that SV models satisfying (3), will generally not belong to the class of *affine models* (see e.g. Duffie et al. (2003), Kallsen (2006)). However, SV models satisfying (4) are affine as soon as the time change is affine. Such models are analytically tractable and, hence, of particular interest for various applications in financial mathematics.

In the following, we study some properties of the stochastic volatility processes V and U. In particular, we derive representation results for the integrated stochastic volatility process.

3.2 Properties of Model 1

First, we study the stochastic volatility process V defined in (3). From standard arguments, we deduce the following representation:

$$V_t = V_0 e^{-\lambda t} + \int_0^t e^{-\lambda(t-s)} \omega_{\lambda s-} dL_{\lambda s}.$$

We now define the process $H = (H_t)_{t \geq 0}$ by

$$H_t = \int_0^t \omega_{\lambda s-} dL_{\lambda s} = \int_0^{\lambda t} \omega_{s-} dL_s,$$

which is clearly a semimartingale, but in general not a Lévy process. In the following we will refer to H as the *background driving volatility modulated Lévy process* (BDVMLP) of V. Then, the stationary version of V can be written as

$$V_t = \int_{-\infty}^t e^{-\lambda(t-s)} dH_s,$$

where L is suitably extended to the negative half line (see Barndorff-Nielsen & Shephard (2001)).

Remark Clearly, L is a Markov process. However, the Markov property is not preserved under stochastic integration. In particular, V is no longer a Markov process. However, the bivariate process (V, ω) satisfies the Markov property if ω is itself a Markov process.

Note that we can easily derive an expression for the increment process of V. In particular, we have for any $T \geq 0$

$$V_{t+T} - V_t = \left(e^{-\lambda T} - 1\right) V_t + \int_t^{t+T} e^{-\lambda(t+T-s)} \omega_{\lambda s} dL_{\lambda s}.$$

3.2.1 Integrated process

In the finance literature, integrated volatility is regarded as the main object of interest since it essentially measures the accumulated variance over a certain period of time (usually a day). So, this section analyses main properties of this key quantity in our new modelling framework. In the following, we will use the notation $V^+ = (V_t^+)_{t \geq 0}$ for the integrated process

$$V_t^+ = \int_0^t V_s ds.$$

First of all, we derive a different representation of the integrated process.

Proposition 1 *The integrated process can be written as*

$$V_t^+ = \epsilon(t,\lambda)V_0 + \int_0^t \epsilon(t-s,\lambda)\omega_{\lambda s} dL_{\lambda s} = \epsilon(t,\lambda)V_0 + \int_0^t \epsilon(t-s,\lambda) dH_s,$$

where $\epsilon(t,\lambda) = \frac{1}{\lambda}\left(1-e^{-\lambda t}\right)$ *and, also, as*

$$V_t^+ = \frac{1}{\lambda}\left(\int_0^{\lambda t} \omega_s dL_s + V_0 - V_t\right) = \frac{1}{\lambda}\left(H_t + V_0 - V_t\right).$$

The proof of the above Proposition is straightforward and, therefore, not given here.

These different representations of V^+ are interesting, since they shed some light on the joint behaviour of V and V^+. Recall that $H_t = \int_0^t \omega_{\lambda s} dL_{\lambda s}$. So clearly, V_t and H_t have identical jumps (breaks), they co–break, i.e. $\Delta V_t = \Delta H_t = \omega_{\lambda t}\Delta L_{\lambda t}$; but V and H are not cointegrated (see Granger (1981)). However, V^+ and H are in fact cointegrated since

$$\lambda V_t^+ - H_t = V_0 - V_t.$$

So, roughly, for large λt, λV_t^+ will have the same distribution as the BDVMLP H_t, where the error in this approximation is a stationary process. Now we can clearly see what kind of influence the stochastic variance of variance has in the new modelling set up: While the long–run behaviour of integrated volatility in the classical BNS model is described by the background driving Lévy process, our new model allows for a greater flexibility in the sense that it can allow for processes which have stationary, but not necessarily independent increments in the long run behaviour of the integrated variance.

Finally, since L is a nonnegative process, the integrated process V^+ is bounded below by the quantity $\frac{1}{\lambda}\left(1-e^{-\lambda t}\right)V_0$.

Remark The new SV model (3) is in fact analytically tractable despite its greater generality compared to the standard BNS model. In particular, it is possible to derive the second order structure of both V and V^+ explicitly and we also get a (quasi) explicit expression for the corresponding characteristic functionals. These results are presented in Section A.1 in the appendix.

3.2.2 Leverage through stochastic volatility of volatility

Next, we focus on the fact that stochastic volatility of volatility can be used for introducing the leverage effect into stochastic volatility models in a novel way. The (usually negative) correlation between asset returns and volatility has been found in many empirical studies, see e.g. Black (1976), Christie (1982) and Nelson (1991) among others and, more recently, by Harvey & Shephard (1996), Bouchaud et al. (2001), Tauchen (2004, 2005), Yu (2005) and Bollerslev et al. (2006).

So far, leverage type effects have usually been introduced by directly correlating the driving process of the volatility with the driving process of the asset prices (as e.g. in the Heston (1993) model). Introducing leverage in the BNS model is slightly more complicated since the volatility is driven by a subordinator and the price is driven by a Brownian motion which are inherently independent from each other (by the Lévy – Khinchine formula). Hence Barndorff-Nielsen & Shephard (2001) suggested to add a jump component to the asset price, which is given by the subordinator which drives the volatility multiplied by a (negative) constant. Hence, such a structure assumes linear dependence between asset price and volatility. However, having an additional random factor in the stochastic volatility model, i.e. the stochastic variance of variance makes it possible to introduce leverage type effects indirectly and independently of the fact whether we want to have a jump component in the model for the logarithmic asset price. In order to illustrate this, let us look at a small example.

Example For simplicity, we just focus on the Brownian semimartingale component $P = (P_t)_{t\geq 0}$, which we define by

$$dP_t = \sqrt{V}_{t-} dW_t,$$

and for the volatility process we work with the following model

$$dV_t = -\lambda V_t dt + \omega_{\lambda t} dL_{\lambda t},$$
$$d\omega_t^2 = \alpha(\beta - \omega_t^2)dt + \gamma\omega_t B_{\alpha t},$$

for parameters $\lambda, \alpha, \beta, \gamma > 0$ and a Brownian motion $B = (B_t)_{t\geq 0}$ with $d[B_{\alpha\lambda}, W]_t = \widetilde{\rho} dt$, for $\widetilde{\rho} \in [-1, 1]$ and all the other quantities defined as above. Clearly P has zero mean and an application of Itô's formula leads to the following result for higher moments of order $n \in \mathbb{R}, n \geq 2$ (provided they exist):

$$\mathbb{E}\left(P_t^n\right) = \frac{n(n-1)}{2} \int_0^t \mathbb{E}\left(P_s^{n-2} V_s\right) ds.$$

In particular, we have

$$Cov(P_t, V_t) = \mathbb{E}(P_t V_t) = -\lambda \int_0^t \mathbb{E}(P_s V_s) ds + \lambda \mathbb{E}(L_1) \int_0^t \mathbb{E}\left(P_s \omega_{\lambda s}\right) ds \not\equiv 0,$$

since

$$\mathbb{E}\left(P_u \omega_{\lambda u}\right) = \mathbb{E}\left(\int_0^u P_s d\omega_{\lambda s}\right) + \mathbb{E}\left(\int_0^u \sqrt{V_{s-}} d[W, \omega_\lambda]_s\right) \not\equiv 0.$$

So, we see that we can have a non–zero correlation between the asset price and the volatility, even if the volatility is jump driven and there are no jumps in the logarithmic asset price.

3.3 Properties of Model 2

Now, we turn our attention to the properties of the SV Model 2, where stochastic volatility of volatility is introduced by a stochastic time change. Clearly, the stochastic volatility process defined in (4) can be represented as

$$U_t = U_0 e^{-\lambda t} + \int_0^t e^{-\lambda(t-s)} dL_{\tau_{\lambda s}}.$$

Similarly to Model 1, we can define the BDVMLP $H = (H_t)_{t\geq 0}$ by $H_t = L_{\tau_{\lambda t}}$. Then the stationary version of U can be written as

$$U_t = \int_{-\infty}^t e^{-\lambda(t-s)} dH_s.$$

3.3.1 Integrated process

Since the integrated process, which is denoted by $U^+ = (U_t^+)_{t\geq 0}$, where

$$U_t^+ = \int_0^t U_s ds,$$

is the key object of interest, we also study its properties in the framework of Model 2. Similarly as before, we obtain the following representation result.

Proposition 2 *The integrated process can be written as*

$$U_t^+ = \epsilon(t, \lambda) U_0 + \int_0^t \epsilon(t-s, \lambda) dL_{\tau_{\lambda s}} = \epsilon(t, \lambda) U_0 + \int_0^t \epsilon(t-s, \lambda) dH_s,$$

where $\epsilon(t, \lambda) = \frac{1}{\lambda}\left(1 - e^{-\lambda t}\right)$ *and, also, as*

$$U_t^+ = \frac{1}{\lambda}\left(L_{\tau_{\lambda t}} - L_{\tau_0} + U_0 - U_t\right) = \frac{1}{\lambda}(H_t + U_0 - U_t).$$

The result above follows from a straightforward calculation, hence the proof is omitted.

As in Model 1, we get that U^+ and H are cointegrated and for large λt, λU_t^+ will nearly have the same distribution as the BDVMLP H_t. The error in this approximation is again given by a stationary process. So, also in Model 2 do we get that the stochastic variance of variance influences the long run behaviour of integrated variance U^+: The limiting process is given by the BDVMLP H and not just by the BDLV L as in the BNS model.

Remark Also Model 2 has nice analytical properties. In particular, we can derive its second order structure and characteristic functionals in (quasi) explicit form, see Section A.2.5.

3.3.2 Leverage through stochastic volatility of volatility

As we have seen in section 3.2.2, stochastic volatility of volatility can be used for introducing leverage type effects indirectly into asset price models by allowing for dependence between the driving process of the volatility of volatility and the driving process of the logarithmic asset price. Clearly, this methodology works in Model 2, too. E.g. if we consider a logarithmic asset price given by a Brownian semimartingale $P = (P_t)_{t\geq 0}$, where $P_t = \int_0^t \sqrt{U_{s-}} dW_s$, then the leverage appears in the quantity

$$Cov(P_t, U_t) = \mathbb{E}(P_t U_t).$$

From Itô's formula, we can easily derive that

$$d(P_t U_t) = P_{t-} dU_t + U_{t-} dP_t + d[U,P]_t = -\lambda P_{t-} U_{t-} dt + P_{t-} dL_{\tau_{\lambda t}} + U_{t-} dP_t + d[U,P]_t.$$

Hence, we obtain

$$\mathbb{E}(P_t U_t) = -\lambda \int_0^t \mathbb{E}\left(P_s U_s\right) ds + \mathbb{E}\left(\int_0^t P_{s-} dL_{\tau_{\lambda s}}\right).$$

Here

$$\mathbb{E}\left(\int_0^t P_{s-} dL_{\tau_{\lambda s}}\right) = \mathbb{E}\left(\mathbb{E}\left(\int_0^t P_{s-} dL_{\tau_{\lambda s}} \middle| \tau\right)\right) = \mathbb{E}(L_1) \int_0^t \mathbb{E}\left(P_s \tau_{\lambda s}\right) ds,$$

which is non–zero if we allow for dependence between (the driving process of) τ and the Brownian motion driving P. In particular, if we specify a parametric model for τ as the one below, we can compute the correlation between P and U by solving an integral equation.

Example A stochastic volatility model which accounts for stochastic variance of variance and leverage could be defined by

$$\begin{aligned} dY_t &= a_t dt + \sqrt{U_{t-}} dW_t + dJ_t, \\ dU_t &= -\lambda U_t dt + dL_{\tau_{\lambda t}}, \\ d\tau_t &= \xi_t dt, \\ d\xi_t^2 &= \alpha(\beta - \xi_t^2) dt + \gamma \xi_t dB_{\alpha t}, \end{aligned}$$

where all quantities are defined as in the previous example and in particular, where B and W are correlated. Note that this model belongs to the class of *affine* models.

3.4 Multivariate models

Next, we study multivariate extensions of stochastic volatility models which account for stochastic volatility of volatility. Here, we entirely restrict our attention to Ornstein–Uhlenbeck type processes, and, in particular, to multivariate extensions of (3).

3.4.1 Notation

Before we define the new model, we have to introduce some notation which we choose along the lines of Pigorsch & Stelzer (2008). Let $M_n(\mathbb{R})$ denote the set of real $n \times n$ dimensional matrices, where $n \in \mathbb{N}$. Then we denote by $\mathbb{S}_n$ the subset of symmetric and invertible $n \times n$ matrices. Furthermore, we define $\sigma(A)$ as the spectrum of a matrix $A \in M_n(\mathbb{R})$. Finally we introduce the notation for integration with respect to a matrix. Let $M_{m,n}(\mathbb{R})$ denote the set of real $m \times n$ dimensional matrices, where $m, n \in \mathbb{N}$. Let $L = (L_t)_{t \geq 0} \in M_{n,r}(\mathbb{R})$ denote a semimartingale and let $A = (A_t)_{t \geq 0} \in M_{m,n}(\mathbb{R})$ and $B = (B_t)_{t \geq 0} \in M_{r,s}(\mathbb{R})$ denote adapted processes which are integrable with respect to L. Then $C = (C_t)_{t \geq 0}$ with $C_t = \int_0^t A_u dL_u B_u$ is in $M_{m,s}(\mathbb{R})$ with elements

$$C_{ij,t} = \sum_{k=1}^{n} \sum_{l=1}^{r} \int_0^t A_{ik,u} B_{lj,u} dL_{kl,u}.$$

3.4.2 Stochastic integral model

In order to extend Model 1 to the multivariate case, we generalise the model proposed by Barndorff-Nielsen & Stelzer (2007) and Pigorsch & Stelzer (2008) and write

$$dV_t = \left(AV_{t-} + V_{t-}A^T\right) dt + \Xi_{t-}^{1/2} dL_t \Xi_{t-}^{1/2},$$

where A is a $d \times d$–matrix (for $d \in \mathbb{N}$) describing the mean reversion coefficient and L is the driving matrix subordinator (see Barndorff-Nielsen & Pérez-Abreu (2008)), while Ξ is a positive definite, stationary stochastic volatility of volatility matrix. We write $\Xi_t = \Xi_t^{1/2} \Xi_t^{1/2}$ for the corresponding unique square root decomposition for positive definite matrices $\Xi^{1/2}$, see Barndorff-Nielsen & Stelzer (2007) and the discussion therein. Also, we assume that the initial value V_0 is a positive semidefinite matrix.

Proposition 3 *Let L be a matrix subordinator with $\mathbb{E}(\max(\log(||L_1||), 0)) < \infty$ and $A \in M_d(\mathbb{R})$ such that $\sigma(A) \subset (-\infty, 0) + i\mathbb{R}$, and let Ξ denote a positive definite, stationary, matrix-valued stochastic process, whose components are independent of the components of L. Then the stochastic differential equation of generalised OU type*

$$dV_t = \left(AV_{t-} + V_{t-}A^T\right) dt + \Xi_t^{1/2} dL_t \Xi_t^{1/2},$$

has a unique stationary solution

$$V_t = \int_{-\infty}^t e^{A(t-s)} \Xi_t^{1/2} dL_s \Xi_t^{1/2} e^{A^T(t-s)}.$$

The integrated process can then easily be computed. Let $\mathcal{A} : \mathbb{S}_s \to \mathbb{S}_d$ denote an operator such that $X \mapsto AX + XA^T$. Then

$$dV_t = \mathcal{A} V_{t-} dt + \Xi_t^{1/2} dL_t \Xi_t^{1/2},$$

and

$$V_t - V_0 = \int_0^t dV_s = \int_0^t \mathcal{A} V_s ds + \int_0^t \Xi_s^{1/2} dL_s \Xi_s^{1/2}.$$

So, we have the following result: Under the same assumptions as in Proposition 3, we get the following representation for the integrated multivariate volatility process with $H_t = \int_0^t \Xi_s^{1/2} dL_s \Xi_s^{1/2}$:

$$V_t^+ = \int_0^t V_s ds = -\mathcal{A}^{-1} \left(H_t - V_t + V_0\right).$$

As in the univariate model, it would also be interesting to study superpositions of multivariate volatility modulated non–Gaussian OU process by extenting recent work by Barndorff-Nielsen & Stelzer (2009).

3.5 Formal definition of stochastic volatility of volatility in a parametric model

After, we have studied two particular SV models, which allow for stochastic volatility of volatility, we now turn to a more formal definition of stochastic volatility of volatility.

In order to obtain a precise definition of stochastic volatility of volatility in a (semi–) parametric set up, we have to specify a corresponding base model for the stochastic volatility process. Here we propose two generic stochastic volatility models. The first model is given by the following stochastic differential equation (SDE):

$$d\sigma_t^2 = b_t dt + \omega_{t-} f\left(\sigma_{t-}^2\right) dZ_t, \tag{5}$$

where $b = (b_t)_{t\geq 0}$ is a predictable drift process, possibly describing the mean reversion of the volatility process, $f : \mathbb{R}_+ \to \mathbb{R}_+$ is a known deterministic function (usually a power/root function), $Z = (Z)_{t\geq 0}$ is a Lévy process and $\omega = (\omega_t)_{t\geq 0}$ is a positive, predictable and stationary semimartingale, which is *independent* of Z. Furthermore, we always work under additional regularity conditions which ensure that $\sigma^2 > 0$ (e.g. when σ^2 is a generalised non–Gaussian Ornstein–Uhlenbeck process or a generalised CIR process satisfying the positivity condition).

Definition 4 *In a stochastic volatility model given by (5), we define the* stochastic variance of variance *process* $\Xi = (\Xi_t)_{t\geq 0}$ *by*

$$\Xi_t = \omega_t^2, \quad \textit{for all } t \geq 0.$$

Remark Generally, we have

$$\omega_{t-}^2 = \frac{\left(d\sigma_t^2\right)^2}{f\left(\sigma_{t-}^2\right)^2 (dZ_t)^2} \qquad \text{or, equivalently,} \qquad [\sigma^2]_t - [\sigma^2]_0 = \int_0^t \omega_{t-}^2 f\left(\sigma_{t-}^2\right)^2 d[Z]_t.$$

This implies that, only in special cases (namely when Z is a standard Brownian motion and $f \equiv 1$), do we have that the quadratic variation of the squared volatility process equals the integrated stochastic variance of variance up to a constant, i.e. $[\sigma^2]_t - [\sigma^2]_0 = \int_0^t \omega_t^2 dt$.

The second class of volatility models is defined by

$$d\sigma_t^2 = b_t dt + f\left(\sigma_{t-}^2\right) dZ_{\tau_t}, \tag{6}$$

where b, f, Z are defined as in (5) and where $\tau = (\tau_t)_{t\geq 0}$ is a time change process (i.e. an increasing, right–continuous semimartingale with $\tau_0 = 0$), which has stationary increments and is *independent* of Z.

Definition 5 *In a stochastic volatility model given by (6), we define the* stochastic variance of variance *process* $\Xi = (\Xi_t)_{t\geq 0}$ *by*

$$\Xi_t = \tau_t, \quad \textit{for all } t \geq 0.$$

Remark In general, we have in a model of type (6) that

$$\frac{\left(d\sigma_t^2\right)^2}{f\left(\sigma_{t-}^2\right)^2} = (dZ_{\tau_t})^2 \qquad \text{or, equivalently,} \qquad d[\sigma^2]_t = f\left(\sigma_{t-}^2\right)^2 (dZ_{\tau_t})^2.$$

In the special case, when Z is a Brownian motion and $f \equiv 1$, we get $[\sigma^2]_t - [\sigma^2]_0 = \tau_t$, i.e. the stochastic variance of variance is given by the quadratic variation.

Note that the quadratic variation of the time changed process corresponds to the time changed quadratic variation of the original process if the time change is absolutely continuous.

We have seen that stochastic volatility of volatility can be introduced either by stochastic spatial or by stochastic temporal scaling of the Lévy process which drives the volatility process. The first approach implies introducing the multiplicative component ω and the latter implies introducing the time change τ in the volatility model. While the first class of models as defined in (5) does generally not belong to the class of affine models, the second class of SV models, given by (6), can be placed into an affine form if the time change process is chosen from the affine class of processes. Clearly, one will have to specify concrete parametric models for ω and τ (as given in the examples above), if one wants to refer to a true *parametric* model for stochastic variance of variance.

4 Non–parametric stochastic volatility of volatility

In this section we study stochastic volatility of volatility from a non–parametric point of view. In order to define this new concept, we proceed as follows. Rather than working with a concrete SV model, we only assume that the logarithmic asset price is given by an Itô semimartingale.

Definition 6 *Let the logarithmic asset price be given by equation (1). The* quadratic variation of variance *(QVV) is defined as the quadratic variation of the squared stochastic volatility process, i.e.* $\left[\sigma^2\right]$.

The QVV can be seen as a non– or semi–parametric definition of variance of variance, which is in line with the approach of viewing the quadratic variation of the logarithmic asset price as a measure of the stochastic variance. If the QVV is in fact stochastic, we view it as non–parametric stochastic variance of variance.

Note that the parametric and the non–parametric definition of stochastic variance of variance are closely linked to each other, but they are generally not the same. In particular, we have seen that the quadratic variation of σ^2 is strongly related to the (integrated) parametric stochastic variance of variance, given by ω^2 and τ. However, QVV only equals (integrated) parametric variance of variance if the volatility defined by (5) or (6) is driven by a standard Brownian motion (rather than a jump process) and if $f \equiv 1$.

5 Estimating the quadratic variation of the variance process

After we have studied the concept of stochastic volatility of volatility in detail, we now turn to the question of how we can actually measure it. Here, we will solely focus on non–parametric (stochastic) variance of variance, i.e. on QVV.

Estimating stochastic volatility based on high frequency observations has been studied extensively in the last decade and, in that context, a quantity called *realised variance*, see e.g. Andersen et al. (2001) and Barndorff-Nielsen & Shephard (2002), and its extensions have been in the research focus. Following this stream of literature, let us assume that we observe the logarithmic asset price $Y = (Y_t)_{t\geq 0}$ given by (1) at times $i\delta_n$ for $i = 0, 1, \ldots, \lfloor T/\delta_n \rfloor$, for some $T > 0$ and where $\delta_n > 0$ and $\delta_n \to 0$ as $n \to \infty$. The realised variance (RV) is then defined as the sum of squared returns over a time interval $[0, t]$ for $0 \leq t \leq T$, i.e.

$$RV_t^n = RV_t^n(Y, \delta_n) = \sum_{i=1}^{\lfloor t/\delta_n \rfloor} \left(\delta_i^n Y\right)^2,$$

where $\delta_i^n Y = Y_{i\delta_n} - Y_{(i-1)\delta_n}$, denotes the ith return/increment of Y. From standard arguments, we get that

$$RV_t^n \xrightarrow{ucp} [Y]_t, \quad \text{as } n \to \infty,$$

where the convergence is uniform on compacts in probability (ucp) (see Protter (2004)) and $[Y]$ is the quadratic variation of Y given by (2). The concept of realised variance has sequentially been extended to realised multipower variation (Barndorff-Nielsen & Shephard (2006), Barndorff-Nielsen et al. (2006), Jacod (2008),Veraart (2009)) and truncated realised variance (Mancini (2001, 2006), Jacod (2008)) in order to construct estimators which are robust towards jumps and estimate the continuous part of the quadratic variation only. Further generalisations and related concepts have been introduced to make such estimators robust to market microstructure noise. However, this aspect is beyond the scope of this paper and we just refer to Bandi & Russell (2008), Zhang et al. (2005), Hansen & Lunde (2006), Jacod et al. (2009), Barndorff-Nielsen et al. (2008*a*,*b*) and the references therein for details on that issue.

The concept of realised variance and its generalisations can also be used for constructing estimators of the spot variance σ_t^2 rather than the accumulated, i.e. integrated, quadratic volatility σ_t^{2+}. If one computes a suitable realised multipower variation statistic over a local window and scales the resulting statistic appropriately, as described below, one can easily derive a spot volatility estimator. This property plays a key role in constructing an estimator for the QVV. In fact, the idea we introduce for estimating the quadratic variation of the squared volatility process is based on three steps: First, we will construct a time series of spot variance estimators, where the spot variance estimators are based on high frequency data, computed on a fine time scale of length $\delta_n > 0$ and where $\delta_n \to 0$ as $n \to \infty$. Next, we introduce a second, sparser time scale of length Δ_n, where $\Delta_n = O\left(\delta_n^C\right)$ for a constant $C > 0$. Finally, we compute the realised variance of the increments of the spot variance estimators on that sparser time scale. For an appropriate choice of C, this will result in a consistent estimator of QVV. In the following, we will call such estimators *pre–estimated spot–variance based realised variance* (PSRV). Clearly, rather than computing the RV on that second, sparser time scale, one could also compute other types of (truncated) realised (multi-) power variation, depending on what kind of underlying structure one assumes for σ^2.

5.1 Estimating the quadratic variation of the squared volatility when we observe the squared volatility

If we hypothetically assume that we observe the volatility process σ^2, which is a semimartingale, over a time interval $[0, T]$ at times $i\Delta_n$ for $i = 1, \ldots, \lfloor T/\Delta_n \rfloor$ (for some $\Delta_n > 0$ such that $\Delta_n \to 0$, as $n \to \infty$), then we can deduce from standard arguments that

$$RV_t^n(\sigma^2, \Delta_n) = \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} \left(\Delta_i^n \sigma^2\right)^2 \overset{ucp}{\longrightarrow} [\sigma^2]_t, \quad \text{as } n \to \infty,$$

where $\Delta_i^n \sigma^2 = \sigma_{i\Delta_n}^2 - \sigma_{(i-1)\Delta_n}^2$ for $i = 0, 1, \ldots, \lfloor t/\Delta_n \rfloor$.

However, since volatility is unobservable, we have to replace the squared volatility above by a consistent spot variance estimator, which will be described in the following section.

5.2 Estimating spot variance

We will estimate the quadratic variation of the stochastic variance by the realised variance $RV_t^n(\sigma^2, \Delta_n)$, where σ^2 is replaced by a spot variance estimator. However, this only works when we estimate spot volatility on a finer time scale than the one we use for computing the realised variance.

Recall, that we denote by $\delta_n > 0$ the mesh size of the fine time scale, at which we observe the logarithmic asset price Y. Here, $\delta_n \to 0$ as $n \to \infty$. Next, we construct a second time scale, which is more sparse and which has mesh size Δ_n. In particular, we define

$$\Delta_n = O\left(\delta_n^C\right), \quad \text{for } 0 < C < 1.$$

E.g. we could choose $\Delta_n = \lfloor \delta_n^{C-1} \rfloor \delta_n$, where $\lfloor \cdot \rfloor$ denotes the floor function.

In the literature, we find various consistent estimators of spot variance. Such estimators are constructed by choosing any of the well-known consistent estimators of integrated variance and then using a locally averaged version of them. In particular, all these estimators are computed over a local window of size $K_n\delta_n$, where we require $K_n \to \infty$, satisfying $\delta_n \to 0$ and $K_n\delta_n \to 0$ as $t \to \infty$. Therefore, we choose a sequence of integers

$$K_n = O_P\left(\delta_n^B\right), \quad \text{for} \quad -1 < B < 0.$$

Let $s \in [0, t]$. In the absence of jumps and noise in the price process, spot variance can be consistently estimated by *locally averaged realised variance*, i.e.

$$\widehat{\sigma}_{s,n}^{2,RV} = \frac{1}{2K_n\delta_n} \sum_{i=\lfloor s/\delta_n \rfloor - K_n}^{\lfloor s/\delta_n \rfloor + K_n} \left(\delta_i^n Y\right)^2.$$

In the absence of noise, but in the presence of jumps, *locally averaged realised bipower variation* (Lee & Mykland (2008), Veraart (2009)) and *locally averaged truncated realised variance* (Aït-Sahalia & Jacod (2009)) provide consistent estimates of spot variance. The first is defined by

$$\widehat{\sigma}_{s,n}^{2,RBV} = \frac{1}{2K_n\delta_n} \sum_{i=\lfloor s/\delta_n \rfloor - K_n}^{\lfloor s/\delta_n \rfloor + K_n} \left|\delta_i^n Y\right| \left|\delta_{i+1}^n Y\right|.$$

and the latter is for some $c > 0$ and $\epsilon \in \left(0, \frac{1}{2}\right)$ given by

$$\widehat{\sigma}_{s,n}^{2,TRV} = \frac{1}{2K_n\delta_n} \sum_{i=\lfloor s/\delta_n \rfloor - K_n}^{\lfloor s/\delta_n \rfloor + K_n} \left(\delta_i^n Y\right)^2 \mathbb{1}_{\{\left|\delta_i^n Y\right| < c\delta_n^\epsilon\}}. \tag{7}$$

If we allow for both jumps and noise in the price process, we can choose locally averaged modulated realised bipower variation as in Jacod et al. (2009), Podolskij & Vetter (2008). Under some regularity conditions, many spot variance estimators satisfy a central limit theorem which allows us to read off the rate of convergence easily. Details on the specific rate of convergence for a variety of spot variance estimators are given in Bandi & Renò (2008).

5.3 Main consistency result

The following proposition contains the main result of this section. It basically says that the realised variance evaluated over a sparse time grid based on estimated spot variances over a fine time grid is a consistent estimator of the quadratic variation of the squared volatility process.

We will prove the consistency of our estimator for two very important classes of stochastic volatility processes: for Brownian motion and for pure jump Lévy–driven stochastic volatility models. To be more precise, we formulate the following assumptions.

Assumption (V1) The stochastic variance σ^2 (or the stochastic volatility σ) is given by a Brownian semimartingale defined by

$$d\sigma_t^2 = b_t dt + \gamma_t dB_t,$$

where $b = (b_t)_{t\geq 0}$ and $\gamma = (\gamma_t)_{t\geq 0}$ are càdlàg and square–integrable and $B = (B_t)_{t\geq 0}$ is a standard Brownian motion (possibly correlated with W).

Assumption (V2) The stochastic variance σ^2 is given by

$$d\sigma_t^2 = b_t dt + \gamma_t dL_t,$$

where $b = (b_t)_{t\geq 0}$ is càdlàg and square–integrable and $L = (L_t)_{t\geq 0}$ is a Lévy subordinator.

Definition 7 *Let $\widehat{\sigma}^2(\delta_n)$ denote the one–sided locally average truncated realised variance estimator based on a time grid of length $\delta_n > 0$, defined by*

$$\widehat{\sigma}^2(\delta_n)_s = \frac{1}{K_n \delta_n} \sum_{j=\lfloor s/\delta_n \rfloor - K_n}^{\lfloor s/\delta_n \rfloor} (\delta_i^n Y)^2 \, \mathbb{1}_{\{|\delta_i^n Y| < c\delta_n^{\epsilon}\}},$$

for $s \in [0,T]$, $c > 0$, $0 < \epsilon < 1/2$. We choose a sequence $(K_n)_{n\in\mathbb{N}}$ such that $K_n = O\left(\delta_n^B\right)$ for some $-1 < B < 0$. That ensures that

$$K_n \to \infty \qquad \text{and} \qquad K_n \delta_n \to 0, \qquad \text{as } n \to \infty.$$

Proposition 8 *Let Y denote the logarithmic asset price defined in (1), which we observe over a time interval $[0,T]$ for some $T > 0$ at times $i\delta_n$ for $i = 0, 1, \ldots, \lfloor T/\delta_n \rfloor$. The corresponding high frequency returns of Y are denote by $\delta_i^n Y = Y_{i\delta_n} - Y_{(i-1)\delta_n}$. Let $\sigma = (\sigma_t)_{t\geq 0}$ be a stochastic volatility process which satisfies either assumption (V1) or assumption (V2). Let $\widehat{\sigma}^2(\delta_n)$ denote the spot variance estimator satisfying Definition 7. Assume that the Blumenthal–Getoor index of J is strictly smaller than $\min\{1, \frac{4\epsilon-1}{2\epsilon}\}$, where ϵ is chosen as in Definition 7. Let $\Delta_n = O\left(\delta_n^C\right)$ denote the mesh of a sparser time grid, where $0 < C < 1$. Furthermore, we have the following assumptions:*

- *If the volatility process satisfies assumption (V1), we assume that*

$$-\frac{1}{2} < B < 0, \qquad \text{and} \qquad 0 < C < -\frac{B}{2}.$$

- *If the volatility process satisfies assumption (V2), we assume that*

$$-\frac{2}{2+\alpha} < B < 0, \qquad \text{and} \qquad 0 < C < -\frac{B}{2},$$

  for some $\alpha > 1$, which is also strictly greater than the Blumenthal–Getoor index of the driving Lévy process of σ^2.

Then, the pre–estimated spot variance based realised variance satisfies:

$$RV_t(\widehat{\sigma}^2(\delta_n), \Delta_n) \xrightarrow{\mathbb{P}} [\sigma^2]_t, \quad \text{as } n \to \infty.$$

Proof The proof is given in the Appendix (Section C). □

So, we can conclude that the pre–estimated spot variance based realised variance is a consistent estimator of the quadratic variation of the stochastic variance.

6 Concluding remarks

6.1 Summary

This paper contains a systematic treatment of a new concept in continuous time financial econometrics: stochastic variance of variance. This concept is of relevance in various areas of finance as described above, such as volatility forecasting, modelling of a time varying variance risk premium, modelling of leverage type effects and inference on spot volatility. We have seen that stochastic variance of variance can be studied both from the perspective of a parametric SV model, where it is essentially given by an additional stochastic temporal or spatial scalar of the driving process of the SV, or from a non–parametric perspective, where it is defined as the quadratic variation of the stochastic variance process (QVV).

As concrete cases of parametric SV models which allow for stochastic volatility of volatility, we propose two new models given by volatility modulated non–Gaussian Ornstein–Uhlenbeck processes. A detailed study of such processes is provided and we also discuss extensions to multivariate models.

In order to estimate QVV consistently, we propose to use the so–called pre–estimated spot–variance based realised variance, which is based on a novel two time scale framework. This approach is potentially applicable to wider fields in finance, e.g. we could think of applying tests for jumps or estimating the activity of a driving jump component not just to asset price data, but to estimated spot variance time series and hence deduce further information on the volatility process.

6.2 Discussion

After having introduced the new concept of stochastic volatility of volatility, we would like to point out some topics for future research.

6.2.1 Remarks on the definition of parametric variance of variance

First of all, there is the issue whether we want to allow for *dependence* between the parametric variance of variance, given by the processes ω^2 and τ, respectively, and the driving process Z of the volatility process. In this paper, we have restricted our attention to the case of independence in order to make it easier to understand the concept of introducing an additional source of randomness to a stochastic volatility model. But extensions to allow for dependence are clearly possible.

Next, there is the question of allowing for *multi–factor* stochastic volatility models and also for *long memory* in the volatility. The generic stochastic volatility models defined in (5) and (6) are only one factor models and generally do not allow for long memory. An simple approach for extending such models is to study a *superposition* of such generic stochastic volatility models, as e.g. studied in the OU context in Barndorff-Nielsen & Shephard (2001, 2002), Barndorff-Nielsen & Stelzer (2009) and in Section A.3 in the Appendix.

6.2.2 Remarks on estimating the processes ω and τ

So far, we have seen how QVV can be estimated consistently by using the new estimator proposed in this paper. However, it might also be interesting to estimate the processes ω or τ directly, provided we believe that such a model assumption is realistic. If the squared stochastic volatility process σ^2 is driven by a Brownian motion, then the quadratic variation of σ^2 equals the integrated ω^2 or τ. Clearly, in such a framework, the estimation technique we have introduced in this paper works well (at least theoretically). Note that such models for σ^2 have been widely used in the literature. E.g. there are many recent papers where the volatility itself is modelled as a Brownian semimartingale (e.g. Barndorff-Nielsen et al. (2006), Kinnebrock & Podolskij (2008), Mykland & Zhang (2009), Jacod (2008)).

However, if the SV model is assumed to be driven by a jump process, then the quadratic variation of σ^2 contains both the jumps and the stochastic variance of variance process. In order to estimate stochastic variance of variance in such a set up, one has to extend the realised variance to a (scaled) truncated realised power variation in order to get rid of the jumps. In fact, this problem is related to making inference on a process $\widetilde{\sigma} = (\widetilde{\sigma}_t)_{t\geq 0}$ when we observe a process $\widetilde{Y} = (\widetilde{Y}_t)_{t\geq 0}$ with $\widetilde{Y}_t = \int_0^t \widetilde{\sigma}_{s-} d\widetilde{L}_s$ in discrete time, where $\widetilde{L} = (\widetilde{L}_t)_{t\geq 0}$ is a purely discontinuous Lévy process. Under the assumption that the driving Levy process is stable and that $\widetilde{\sigma}$ and $\widetilde{L}$ are independent, this has been studied by Woerner (2003). Recent work by Jacod & Protter (1998), Jacod (2004, 2007) touches upon more general cases. However, the problem of making inference on $\widetilde{\sigma}$ when $\widetilde{Y}$ is driven by an arbitrary purely discontinuous Lévy processes has – to the best of our knowledge – not been solved yet.

6.2.3 Remarks on model estimation

Next, we want to comment on model estimation. In particular, we are interested in the case when the stochastic volatility model is given by Model 1 or 2, i.e. when we have a non–Gaussian OU process which allows for stochastic variance of variance. It turns out that despite the additional source of randomness, Model 1 and Model 2 are still analytically tractable. It is, in particular, possible to compute the conditional and unconditional moments of the volatility process and also moments of the price process explicitly. Since these computations are fairly lengthy, we only present them in the Appendix (Section A). Based on these results, model parameters can be estimated by using a quasi–maximum likelihood estimation or the generalised method of moments. An implementation of these estimation techniques and empirical work will be left for future research.

6.3 Remarks on option pricing in the presence of stochastic variance of variance

Also, we would like to comment briefly on the use of stochastic variance of variance models for option pricing. Nicolato & Venardos (2003) studied in detail how option pricing can be done when the stochastic volatility process is, under the risk neutral probability measure, given by a non–Gaussian Ornstein Uhlenbeck process. Can we use similar methods when we account for an additional source of randomness in BNS type models? One widely used technique for option pricing is the use of transform–based methods (see e.g. Nicolato & Venardos (2003)), where the price of a contract is expressed in terms of an integral transform of the Laplace exponent of the underlying price process. We have provided formulas for the characteristic function in Model 1 for both V and V^+ (see Proposition 17 and Proposition 19 in the appendix), which are essential for deriving the characteristic function of the price process. However, due to the additional source of randomness in form of stochastic variance of variance, it is difficult to evaluate these characteristic functions explicitly. For practical applications, it might therefore be necessary to use numerical methods to evaluate option prices in a stochastic variance of variance set up.

6.3.1 Remarks on the relationship between stochastic variance of variance and leverage type effects

Finally, we comment on the relationship between stochastic variance of variance and leverage type effects. We have seen in Sections 3.2.2 and 3.3.2 that stochastic variance of variance can be used for introducing leverage type effects into stochastic volatility models in a novel way. In particular, it is possible to obtain a *stochastic* quadratic covariation between the asset price and the squared volatility process $[Y, \sigma^2]$, which can be expressed in terms of the stochastic variance of variance and which can be regarded as a stochastic leverage type effect. Alternatively, such a structure can also be obtained by randomising the correlation between the driving process of the asset price and the driving process of the stochastic volatility process, as discussed in Veraart & Veraart (2009). From a conceptual point of view, the concept of stochastic volatility of volatility is clearly different from the concept of stochastic correlation or stochastic leverage. However, how one can disentangle these two effects in practice is an interesting question for future research.

References

Aït-Sahalia, Y. & Jacod, J. (2009), 'Testing for jumps in a discretely observed process', *Annals of Statistics* **37**(1), 184–222.

Andersen, T. G., Bollerslev, T., Diebold, F. X. & Ebens, H. (2001), 'The distribution of realized stock return volatility', *Journal of Financial Economics* **61**, 43–76.

Bandi, F. M. & Renò, R. (2008), Nonparametric stochastic volatility. Working paper (June 2008). Available at SSRN: http://ssrn.com/abstract=1158438.

Bandi, F. M. & Russell, J. R. (2008), 'Microstructure noise, realised volatility, and optimal sampling', *Review of Economic Studies* **75**(2), 339 – 369.

Barndorff-Nielsen, O. E., Graversen, S. E., Jacod, J., Podolskij, M. & Shephard, N. (2006), A central limit theorem for realised power and bipower variations of continuous semimartingales, *in* Y. Kabanov & R. Lipster, eds, 'From Stochastic Analysis to Mathematical Finance, Festschrift for Albert Shiryaev', Springer.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A. & Shephard, N. (2008*a*), 'Designing realised kernels to measure the ex-post variation of equity prices in the presence of noise', *Econometrica* **76**, 1481–1536.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A. & Shephard, N. (2008*b*), 'Multivariate realised kernels: consistent positive semi-definite estimators of the covariation of equity prices with noise and non-synchronous trading'. Working paper.

Barndorff-Nielsen, O. E. & Pérez-Abreu, V. (2008), 'Matrix subordinators and related upsilon transformations', *Theory of Probability and Its Applications* **52**, 1–23.

Barndorff-Nielsen, O. E. & Shephard, N. (2001), 'Non–Gaussian Ornstein–Uhlenbeck–based models and some of their uses in financial economics (with discussion)', *Journal of the Royal Statistical Society Series B* **63**, 167–241.

Barndorff-Nielsen, O. E. & Shephard, N. (2002), 'Econometric analysis of realised volatility and its use in estimating stochastic volatility models', *Journal of the Royal Statistical Society* **B 64**, 253–280.

Barndorff-Nielsen, O. E. & Shephard, N. (2006), 'Econometrics of testing for jumps in financial economics using bipower variation.', *Journal of Financial Econometrics* **4**, 1–30.

Barndorff-Nielsen, O. E. & Shephard, N. (2007), Variation, jumps, market frictions and high frequency data in financial econometrics, *in* R. Blundell, P. Torsten & W. K. Newey, eds, 'Advances in Economics and Econometrics. Theory and Applications, Ninth World Congress', Econometric Society Monographs, Cambridge University Press.

Barndorff-Nielsen, O. E. & Stelzer, R. (2007), 'Positive–definite matrix processes of finite variation', *Probability and Mathematical Statistics* **27**(1), 3–43.

Barndorff-Nielsen, O. E. & Stelzer, R. (2009), Multivariate supOU processes. Preprint.

Black, F. (1976), 'Studies of stock market volatility changes', *Proceedings of the Business and Economic Statistics Section, American Statistical Association* pp. 177–181.

Bollerslev, T., Kretschmer, U., Pigorsch, C. & Tauchen, G. (2009), 'A discrete–time model for daily S&P 500 returns and realized variations: jumps and leverage effects', *Journal of Econometrics* **150**(2), 151–166.

Bollerslev, T., Litvinova, J. & Tauchen, G. (2006), 'Leverage and volatility feedback effects in high-frequency data', *Journal of Financial Econometrics* **4**(3), 353–384.

Bollerslev, T., Tauchen, G. & Zhou, H. (2009), 'Expected stock returns and variance risk premia', *Review of Financial Studies* . In press.

Bouchaud, J.-P., Matacz, A. & Potters, M. (2001), 'Leverage effect in financial markets: The retarded volatility model', *Physical Review Letters* **87**(228701).

Christie, A. A. (1982), 'The stochastic behavior of common stock variances', *Journal of Financial Economics* **10**, 407–432.

Corsi, F. (2004), A simple long memory model for realized volatility, Technical report, University of Southern Switzerland.

Corsi, F., Mittnik, S., Pigorsch, C. & Pigorsch, U. (2008), 'The volatility of realized volatility', *Econometrics Reviews* **27**(1–3), 46–78.

Cox, J. C., Ingersoll, J. E. J. & Ross, S. A. (1985), 'A theory of the term structure of interest rates', *Econometrica* **53**(2), 385–408.

Drechsler, I. & Yaron, A. (2008), What's vol got to do with it. AFA 2009 San Francisco Meetings Paper.

Duffie, D., Filipovic, D. & Schachermayer, W. (2003), 'Affine processes and applications in finance', *Annals of Applied Probability* **13**, 984–1053.

Granger, C. J. (1981), 'Some properties of time series data and their use in econometric model specification', *Journal of Econometrics* **16**, 121–130.

Hansen, P. R. & Lunde, A. (2006), 'Realized variance and market microstructure noise', *Journal of Business and Economic Statistics* **24**, 127–218.

Harvey, A. & Shephard, N. (1996), 'Estimation of an asymmetric stochastic volatility model for asset returns', *Journal of Business and Economic Statistics* **14**, 429–434.

Heston, S. L. (1993), 'A closed–form solution for options with stochastic volatility with applications to bond and currency options', *Review of Financial Studies* **6**, 327–343.

Jacod, J. (2004), 'The Euler scheme for Lévy driven stochastic differential equations: Limit theorems', *The Annals of Probability* **32**(3A), 1830 – 1872.

Jacod, J. (2007), 'Asymptotic properties of power variations of Lévy processes', *ESAIM: Probability and Statistics* **11**, 173–196.

Jacod, J. (2008), 'Asymptotic properties of realized power variations and related functionals of semimartingales', *Stochastic Processes and their Applications* **118**(4), 517–559.

Jacod, J., Li, Y., Mykland, P. A., Podolskij, M. & Vetter, M. (2009), 'Estimation of volatility functionals in the simultaneous presence of microstructure noise and jumps', *Stochastic Processes and their Applications* **119**(7), 2249–2276.

Jacod, J. & Protter, P. (1998), 'Asymptotic error distribution for the Euler method for stochastic differential equations', *The Annals of Probability* **26**(1), 267 – 307.

Kallsen, J. (2006), A didactic note on affine stochastic volatility models, *in* Y. Kabanov, R. Liptser & J. Stoyanov, eds, 'From Stochastic Calculus to Mathematical Finance', Springer, Berlin, pp. 343–368.

Kinnebrock, S. & Podolskij, M. (2008), 'A note on the central limit theorem for bipower variation of general functions', *Stochastic processes and their applications* **118**, 1056–1070.

Lee, S. S. & Mykland, P. A. (2008), 'Jumps in financial markets: A new nonparametric test and jump dynamics', *Review of Financial Studies* **21**(6), 2535–2563.

Mancini, C. (2001), 'Disentangling the jumps of the diffusion in a geometric jumping brownian motion', *Giornale dell Istituto Italiano degli Attuari* **LXIV**, 19–47.

Mancini, C. (2006), Non parametric threshold estimation for models with stochastic diffusion coefficient and jumps. Working paper. arXiv:math/0607378v1.

Mykland, P. & Zhang, L. (2009), 'Inference for continuous semimartingales observed at high frequency: A general approach', *Econometrica* . Forthcoming.

Nelson, D. B. (1991), 'Conditional heteroscedasticity in asset returns: a new approach', *Econometrica* **59**(2), 347–370.

Nicolato, E. & Venardos, E. (2003), 'Option pricing in stochastic volatility models of the Ornstein–Uhlenbeck type', *Mathematical Finance* **13**, 445–466.

Pigorsch, C. & Stelzer, R. (2008), 'On the definition, stationary distribution and second order structure of positive semi–definite Ornstein–Uhlenbeck type processes', *Bernoulli* . Forthcoming.

Podolskij, M. & Vetter, M. (2008), 'Estimation of volatility functionals in the simultaneous presence of microstructure noise and jumps', *Bernoulli* . Forthcoming.

Protter, P. E. (2004), *Stochastic Integration and Differential Equations*, second edn, Springer, London.

Pruitt, W. E. (1981), 'The growth of random walks and Lévy processes', *The Annals of Probability* **9**(6), 948–956.

Revuz, D. & Yor, M. (2001), *Continuous Martingales and Brownian Motion*, third edn, Springer, Berlin.

Tauchen, G. (2004), Remarks on recent developments in stochastic volatility: statistical modelling and general equilibrium. Working paper, Department of Economics, Duke University.

Tauchen, G. (2005), Stochastic volatility and general equilibrium. Working paper, Department of Economics, Duke University.

Veraart, A. E. D. (2009), 'Inference for the jump part of quadratic variation of Itô semimartingales', *Econometric Theory* . Forthcoming.

Veraart, A. E. D. & Veraart, L. A. M. (2009), Stochastic volatility and stochastic leverage. Aarhus University, CREATES working paper RP2009–20.

Woerner, J. H. C. (2003), Purely discontinuous Lévy processes and power variation: inference for integrated volatility and the scale parameter. 2003–MF–08 Working Papers Series in Mathematical Finance, University of Oxford.

Yu, J. (2005), 'On leverage in a stochastic volatility model', *Journal of Econometrics* **127**(2), 165–178.

Zhang, L., Mykland, P. A. & Aït-Sahalia, Y. (2005), 'A tale of two time scales: determining integrated volatility with noisy high–frequency data', *Journal of the American Statistical Association* **100**, 1394–1411.

APPENDIX

A Cumulants and moments

Throughout this section we only present the results of the various cumulants and characteristic functionals and omit the rather lengthy proofs since they consist of straightforward computations.

A.1 Conditional and unconditional moments of V and V^+

In order to get a better understanding of the new volatility of volatility model, we derive various conditional and unconditional moments of the processes V and U.

A.1.1 Notation

In order to simplify the exposition, we fix the following notation. We denote by $\kappa_i = \kappa_i(X_1)$ the ith cumulant of X_1 for $i \in \mathbb{N}$. In particular, we write $\kappa_i := \frac{1}{\lambda}\kappa_i(L_\lambda) = \kappa_i(L_1)$ for a Lévy process L.

Remark In order to ensure that our model (3) is uniquely identified, we have to set the variance (or the mean) of L_1 to a fixed value. Otherwise one could always multiply ω by a constant and scale L approximately, which results in an identification problem. For convenience, we will later set $\kappa_2(L_1) = Var(L_1) = 1$.

Furthermore, we will write $\gamma(h) = Cov(\omega_t, \omega_{t+h})$ for the covariance function of ω. We will carry out all computations for a general stationary volatility process ω, which is independent of the subordinator L. However, for some illustrating examples, we will sometimes impose the following condition:

Assumption (C) on the covariance function of ω: The covariance function of ω is given by

$$\gamma(h) = Cov\left(\omega_t, \omega_{t+h}\right) = C\exp(-\alpha h),$$

for constants $C, \alpha > 0$. In particular, this assumption is satisfied when ω is an Ornstein–Uhlenbeck (OU) process or a Constant Elasticity of Variance (CEV) process.

A.1.2 Moments of V conditional on V_0

First of all, we derive the mean, variance and autocovariance of V, conditional on the initial value V_0.

Proposition 9 *Let V be a solution to (3). Then the conditional mean, variance and covariance are given by*

$$\begin{aligned}
\mathbb{E}\left(V_t|\,V_0\right) &= e^{-\lambda t}V_0 + \kappa_1(L_1)\mathbb{E}(\omega_0)\left(1-e^{-\lambda t}\right),\\
Var\left(V_t|\,V_0\right) &= \frac{1}{2}\kappa_2(L_1)\mathbb{E}\left(\omega_0^2\right)\left(1-e^{-2\lambda t}\right) + \kappa_1^2(L_1)\int_0^{\lambda t}\gamma(y)\left(e^{-y}-e^{-2\lambda t+y}\right)dy,\\
Cov\left((V_t, V_{t+h})|\,V_0\right) &= \frac{1}{2}\kappa_2(L_2)\mathbb{E}\left(\omega_0^2\right)\left(1-e^{-2\lambda t}\right)e^{-\lambda h}\\
&\quad + \kappa_1^2(L_1)e^{-\lambda h}\int_0^{\lambda t}\left(e^{-y}-e^{-2\lambda t+y}\right)\gamma(y)dy\\
&\quad + \lambda^2\kappa_1^2(L_1)e^{-2\lambda t-\lambda h}\int_t^{t+h}\int_{u-t}^{u} e^{2\lambda u-\lambda x}\gamma(\lambda x)dxdu\\
Cor\left((V_t, V_{t+h})|\,V_0\right) &= e^{-\lambda h}\left(1+\frac{1}{Var(V_t|V_0)}\lambda^2\kappa_1^2(L_1)e^{-2\lambda t}\int_t^{t+h}\int_{u-t}^{u} e^{2\lambda u-\lambda x}\gamma(\lambda x)dxdu\right).
\end{aligned}$$

Corollary 10 *Under condition (C) the variance and the covariance simplify to*

$$\begin{aligned}
Var\left(V_t|\,V_0\right) &= \frac{1}{2}\kappa_2(L_1)\mathbb{E}\left(\omega_0^2\right)\left(1-e^{-2\lambda t}\right)\\
&\quad + \frac{\kappa_1^2(L_1)C}{(1-\alpha^2)}\left(1-\alpha+(1+\alpha)e^{-2\lambda t}-2e^{-\lambda(1+\alpha)t}\right),\\
Cov\left((V_t, V_{t+h})|\,V_0\right) &= e^{-\lambda h}\left(Var\left(V_t|\,V_0\right)\right.\\
&\quad \left. + \lambda^2\kappa_1^2(L_1)\frac{C}{\alpha^2-1}\left(e^{-\lambda t(1+\alpha)}-1-e^{-\lambda(\alpha(t+h)-h+t)}+e^{-\lambda h(\alpha-1)}\right)\right).
\end{aligned}$$

A.1.3 Moments of V conditional on ω

Next, we compute the moments of V, when we condition on ω. Clearly, if ω was deterministic, these results would also hold unconditionally.

Proposition 11 *The conditional mean, variance and covariance are given by*

$$\begin{aligned}
\mathbb{E}\left(V_t|\,\omega\right) &= \lambda\kappa_1(L_1)\int_{-\infty}^{t} e^{-\lambda(t-s)}\omega_{\lambda s}ds,\\
Var\left(V_t|\,\omega\right) &= \lambda\kappa_2(L_1)e^{-2\lambda t}\int_{-\infty}^{t} e^{2\lambda s}\omega_{\lambda s}^2 ds,\\
Cov\left(V_t, V_{t+h}|\,\omega\right) &= \lambda\kappa_2(L_1)e^{-\lambda h}e^{-2\lambda t}\int_{-\infty}^{t} e^{2\lambda u}\omega_{\lambda u}^2 du,\\
Cor\left(V_t, V_{t+h}|\,\omega\right) &= e^{-\lambda h}.
\end{aligned}$$

A.1.4 Unconditional moments

Finally, we compute the unconditional mean, variance and covariance of the generalised non-Gaussian Ornstein Uhlenbeck process V.

Proposition 12 *Let V be the stationary solution to (3). Then:*

$$\begin{aligned}
\mathbb{E}\left(V_t\right) &= \kappa_1(L_1)\mathbb{E}(\sigma_0^2), \\
Var\left(V_t\right) &= \frac{1}{2}\kappa_2(L_1)\mathbb{E}\left(\omega_0^2\right) + \kappa_1^2(L_1)\int_0^\infty e^{-y}\gamma(y)dy, \\
Cov\left(V_t, V_{t+h}\right) &= e^{-\lambda h}\left(\frac{1}{2}\kappa_2(L_1)\mathbb{E}\left(\omega_0^2\right)\right. \\
&\left. +\kappa_1^2(L_1)\left(\int_0^\infty e^{-y}\gamma(y)dy + \lambda^2\int_{-\infty}^0\int_0^h e^{\lambda(u+s)}\gamma(\lambda|s-u|)dsdu\right)\right).
\end{aligned}$$

Corollary 13 *Under assumption (C), the results above simplify to*

$$\begin{aligned}
\mathbb{E}\left(V_t\right) &= \kappa_1(L_1)\mathbb{E}(\omega_0), \\
Var\left(V_t\right) &= \frac{1}{2}\kappa_2(L_1)\mathbb{E}\left(\omega_0^2\right) + \frac{C\kappa_1^2(L_1)}{1+\alpha}, \\
Cov\left(V_t, V_{t+h}\right) &= \frac{1}{2}\kappa_2(L_1)\mathbb{E}\left(\omega_0^2\right)e^{-\lambda h} + \frac{\kappa_1^2(L_1)C}{1-\alpha^2}\left(e^{-\alpha\lambda h} - \alpha e^{-\lambda h}\right).
\end{aligned}$$

A.1.5 Cumulants of the integrated process

First, we study the mean and the variance, conditional on the initial value V_0.

Proposition 14 *The conditional mean and variance of the integrated process are given by:*

$$\begin{aligned}
\mathbb{E}\left(V_t^+\middle| V_0\right) &= \frac{1}{\lambda}\left(1-e^{-\lambda t}\right)\left(V_0 - \kappa_1(L_1)\mathbb{E}\left(\omega_0\right)\right) + \kappa_1(L_1)\mathbb{E}\left(\omega_0\right)t, \\
Var\left(V_t^+\middle| V_0\right) &= \kappa_2(L_1)\mathbb{E}\left(\omega_0^2\right)\frac{1}{\lambda^2}\left(-\frac{1}{2}e^{-2\lambda t} + 2e^{-\lambda t} + \lambda t - \frac{3}{2}\right) \\
&+ 2\kappa_1^2(L_1)\int_0^t\int_s^t e^{-\lambda(u-s)}\int_0^{\lambda s}\left(e^{-y} - e^{-2\lambda s+y}\right)\gamma(y)dyduds \\
&+ 2\lambda^2\kappa_1^2(L_1)\int_0^t\int_s^t e^{-\lambda(u+s)}\int_u^s\int_{x-u}^x e^{2\lambda x-\lambda y}\gamma(\lambda y)dydxduds.
\end{aligned}$$

Finally, we compute the mean, variance and covariance of the integrated process, which are given as follows.

Proposition 15 *Assume that Y satisfies (3). Let*

$$G(h) := G(h, \lambda, \kappa_1, \gamma(\cdot)) = Cov(V_t, V_{t+h}) - \frac{1}{2}\kappa_2(L_1)\mathbb{E}\left(\omega_0^2\right)e^{-\lambda h}.$$

Then:

$$\begin{aligned}
\mathbb{E}\left(V_t^+\right) &= \kappa_1(L_1)\mathbb{E}\left(\omega_0\right)t, \\
Var(V_t^+) &= \kappa_2(L_1)\mathbb{E}\left(\omega_0^2\right)\frac{1}{\lambda^2}\left(e^{-\lambda t} - 1 + \lambda t\right) + 2\int_0^t\int_s^t G(u-s)duds, \\
Cov\left(V_t^+, V_{t+h}^+\right) &= \frac{1}{2}\kappa_2(L_1)\mathbb{E}\left(\omega_0^2\right)\frac{1}{\lambda^2}\left(e^{-\lambda(t+h)} + e^{-\lambda t} - e^{-\lambda h} - 1 + 2\lambda t\right) \\
&+ \int_0^t\int_t^{t+h} G(u-s)duds.
\end{aligned}$$

A.2 Characteristic functions and functionals

In this section, we turn our attention to characteristic functionals and functions of V and V^+. Recall that we are assuming independence of L and ω throughout this section.

A.2.1 Characteristic function of V

We start by computing the characteristic function of V, conditional on the initial value V_0 and the volatility process ω.

Proposition 16 *The conditional characteristic function of V is given by:*

$$\mathbb{E}(\exp(i\theta V_t)|V_0, \omega_s, 0 \leq s \leq \lambda t) = \exp\left(i\theta e^{-\lambda t} V_0\right) \exp\left(\int_0^{\lambda t} \psi_Z\left(\theta e^{-\lambda t} e^s \omega_s\right) ds\right),$$

where $\psi(\cdot)$ is the characteristic exponent of L.

Clearly, the conditional cumulant function is then given by

$$\log\left(\mathbb{E}(\exp(i\theta V_t)|V_0, \omega_s, 0 \leq s \leq \lambda t)\right) = i\theta e^{-\lambda t} V_0 + \int_0^{\lambda t} \psi_Z\left(\theta e^{-\lambda t} e^s \omega_s\right) ds.$$

A.2.2 Characteristic functional

Next, we study more general characteristic functionals. Let f be an arbitrary function. We define $f \bullet V_t = \int_0^t f(s) dV_s$ and $f \bullet V = \int_0^\infty f(s) dV_s$.

Proposition 17 *The characteristic functional is given by*

$$\begin{aligned}&\mathbb{E}\left(\exp\left(if \bullet V_t\right)\middle| V_0, \omega_{\lambda s}, 0 \leq s \leq t\right) \\ &= \exp\left(-i\lambda V_0 \int_0^t f(x) e^{-\lambda x} dx\right) \exp\left(\int_0^t \psi_L\left(\left(f(s) - \lambda \int_0^{x-s} f(x+s) e^{-\lambda x} dx\right) \omega_{\lambda s}\right) ds\right),\end{aligned}$$

and when we take $t \to \infty$ in the formula above, we immediately get that

$$\begin{aligned}&\mathbb{E}\left(\exp\left(if \bullet V\right)\middle| V_0, \omega_{\lambda s}, s \geq 0\right) \\ &= \exp\left(-i\lambda V_0 \int_0^\infty f(x) e^{-\lambda x} dx\right) \exp\left(\int_0^\infty \psi_L\left(\left(f(s) - \lambda \int_0^\infty f(x+s) e^{-\lambda x} dx\right) \omega_{\lambda s}\right) ds\right).\end{aligned}$$

So, in particular, for $f(x) = \theta$, we can easily derive the characteristic function of Y.

Corollary 18 *From integrating out ω, we get*

$$\begin{aligned}\mathbb{E}\left(\exp\left(if \bullet V_t\right)\middle| V_0\right) &= \exp\left(-i\lambda V_0 \int_0^t f(x) e^{-\lambda x} dx\right) \\ &\quad \mathbb{E}\left(\exp\left(\int_0^t \psi_L\left(\left(f(s) - \lambda \int_0^{t-s} f(x+s) e^{-\lambda x} dx\right) \omega_{\lambda s}\right) ds\right)\right).\end{aligned}$$

A.2.3 Characteristic function of V^+

Proposition 19 *The characteristic function of the integrated process is given by*

$$\begin{aligned}&\mathbb{E}\left(\exp\left(i\theta \int_0^t V_u du\right)\middle| V_0\right) \\ &= \phi\left(\exp\left(\frac{\theta}{\lambda}\left(1 - \exp(-\lambda t)\right)\right)\right) \mathbb{E}\left(\exp\left(\int_0^{\lambda t} \psi_L\left(\frac{\theta}{\lambda}\left(1 - e^{-(\lambda t - s)}\right) \omega_s ds\right)\right)\right),\end{aligned}$$

where ϕ is the characteristic function of V_0.

A.2.4 Cumulant functional

Let f denote an *arbitrary* function.

Proposition 20 *The conditional cumulant functional is given by*

$$\log\left(\mathbb{E}\left(\exp\left(if\bullet\int_0^t V_s ds\right)\middle|\,(\omega_s)_{s\geq 0}, V_0\right)\right)$$
$$= iV_0\int_0^\infty f(t)e^{-\lambda t}dt + \int_0^\infty \psi_{L_\lambda}\left(\sigma^2_{\lambda s}\int_0^\infty f(t+s)e^{-\lambda t}dt\right)ds.$$

Unconditionally on ω, we have

$$\mathbb{E}\left(\exp\left(if\bullet\int_0^t V_s ds\right)\middle|\,V_0\right)$$
$$= \exp\left(iV_0\int_0^\infty f(t)e^{-\lambda t}dt\right)\exp\left(\int_0^\infty \psi_{L_\lambda}\left(\omega_{\lambda s}\int_0^\infty f(t+s)e^{-\lambda t}dt\right)ds\right).$$

Example Again, in the special case $f(t)=\theta$, we obtain the characteristic function

$$\mathbb{E}\left(\exp\left(if\bullet\int_0^t V_s ds\right)\middle|\,(\omega_s)_{s\geq 0}, V_0\right) = \exp\left(iV_0\frac{\theta}{\lambda}\right)\exp\left(\int_0^\infty \psi_{L_\lambda}\left(\omega_{\lambda s}\frac{\theta}{\lambda}\right)ds\right).$$

A.2.5 Characteristic functionals for Model 2

Just for comparison, we present the corresponding results for Model 2, i.e. the non–Gaussian Ornstein–Uhlenbeck process which is driven by a time–changed Lévy subordinator. Recall that Model 2 is given by

$$dU_t = -\lambda U_t dt + dL_{\tau(\lambda t)},$$

where $\lambda>0$ and where $L=(L_t)_{t\geq 0}$ and $\tau=(\tau_t)_{t\geq 0}$ denote independent (Lévy) subordinators. Then

$$U_t = U_0 e^{-\lambda t} + \int_0^t e^{-\lambda(t-s)}dL_{\tau(\lambda s)} = \int_{-\infty}^t e^{-\lambda(t-s)}dL_{\tau(\lambda s)}.$$

Let f denote an arbitrary function. By applying Fubini's theorem, we obtain

$$\begin{aligned} f\bullet U_t &= \int_0^\infty f(t)dU_t = -\lambda\int_0^\infty f(t)U_t dt + \int_0^\infty f(t)dL_{\tau(\lambda t)} \\ &= -\lambda U_0\int_0^\infty e^{-\lambda t}f(t)dt + -\lambda\int_0^\infty f(t)\int_0^t e^{-\lambda t}e^{\lambda s}dL_{\tau(\lambda s)}dt + \int_0^\infty f(t)dL_{\tau(\lambda t)} \\ &= -\lambda U_0\int_0^\infty e^{-\lambda t}f(t)dt + \int_0^\infty\left(e^{\lambda s}\int_s^\infty e^{-\lambda t}f(t)dt + f(s)\right)dL_{\tau(\lambda s)}. \end{aligned}$$

Remark Revuz & Yor (2001, p. 9, Proposition): Let A be of finite variation and u a continuous nondecreasing function on $[a,b]$. For any nonnegative function f on $[u(a),u(b)]$, we have

$$\int_{[a,b]} f(u(s))dA(u(s)) = \int_{[u(a),u(b)]} f(t)dA_t.$$

In the following, we assume that τ is *strictly non-decreasing and continuous*, e.g.

$$\tau_t = \int_0^t \xi_s ds.$$

Then, the inverse function τ^{-1} exists and $\tau(\tau^{-1}(x)) = x$ and $\tau^{-1}(\tau(x)) = x$. Then, we have

$$\begin{aligned} f \bullet U_t &= -\lambda U_0 \int_0^\infty e^{-\lambda t} f(t)dt + \int_0^\infty \left(e^{\frac{1}{\lambda}\tau^{-1}(\tau(\lambda s))} \int_{\frac{1}{\lambda}\tau^{-1}(\tau(s))}^\infty e^{-\lambda t} f(t)dt + f(\frac{1}{\lambda}\tau^{-1}(\tau(\lambda s))) \right) dL_{\tau(\lambda s)} \\ &= -\lambda U_0 \int_0^\infty e^{-\lambda t} f(t)dt + \int_0^\infty \left(e^{\frac{1}{\lambda}\tau^{-1}(u)} \int_{\frac{1}{\lambda}\tau^{-1}(u)}^\infty e^{-\lambda t} f(t)dt + f\left(\frac{1}{\lambda}\tau^{-1}(u)\right) \right) dL_u. \end{aligned}$$

Let ψ denote the characteristic exponent of Z. Then we have

$$\begin{aligned} &\mathbb{E}\left(\exp(if \bullet U_t)|U_0, \tau\right) \\ &= \exp\left(-\lambda i U_0 \int_0^\infty e^{-\lambda t} f(t)dt\right) \exp\left(\int_0^\infty \psi\left(e^{\tau^{-1}(u)} \int_{\frac{1}{\lambda}\tau^{-1}(u)}^\infty e^{-\lambda t} f(t)dt + f\left(\frac{1}{\lambda}\tau^{-1}(u)\right)\right) du\right). \end{aligned}$$

For $U_t^+ = \int_0^t U_s ds$, we get

$$\begin{aligned} f \bullet U^+ &= \int_0^\infty f(t)U(t)dt = U_0 \int_0^\infty e^{-\lambda t} f(t)dt + \int_0^\infty f(t) \int_0^t e^{-\lambda(t-s)} dL_{\tau(\lambda s)} dt \\ &= U_0 \int_0^\infty e^{-\lambda t} f(t)dt + \int_0^\infty \int_s^\infty e^{-\lambda t} f(t)dt e^{\lambda s} dL_{\tau(\lambda s)} \\ &= U_0 \int_0^\infty e^{-\lambda t} f(t)dt + \int_0^\infty \int_{\frac{1}{\lambda}\tau^{-1}(\tau(\lambda s))}^\infty e^{-\lambda t} f(t)dt e^{\tau^{-1}(\tau(\lambda s))} dL_{\tau(\lambda s)} \\ &= U_0 \int_0^\infty e^{-\lambda t} f(t)dt + \int_0^\infty \int_{\frac{1}{\lambda}\tau^{-1}(u)}^\infty e^{-\lambda t} f(t)dt e^{\tau^{-1}(u)} dL_u. \end{aligned}$$

Hence, we get

$$\begin{aligned} &\mathbb{E}\left(\exp(f \bullet U^+)|U_0, \tau\right) \\ &= \exp\left(iU_0 \int_0^\infty e^{-\lambda t} f(t)dt\right) \exp\left(\int_0^\infty \psi\left(\int_{\frac{1}{\lambda}\tau^{-1}(u)}^\infty e^{-\lambda t} f(t)dt\right) e^{\tau^{-1}(u)} du\right). \end{aligned}$$

A.3 Superposition of generalised OU processes

Finally, we will address the aspect of allowing for long memory in the volatility process. Clearly, also the generalised non-Gaussian Ornstein–Uhlenbeck process has an exponentially fast decaying autocorrelation function and is therefore unable to capture long memory. However, as in Barndorff-Nielsen & Shephard (2001, 2002), we can consider a superposition of $J \in \mathbb{N}$ generalised Ornstein–Uhlenbeck processes $V^J = \left(V_t^J\right)_{t \geq 0}$, which we define by

$$\begin{aligned} V_t^J &= \sum_{i=1}^J w_j V_t^{(j)}, \quad w_j \geq 0, \quad \sum_{j=1}^J w_j = 1, \\ dV_t^{(j)} &= -\lambda^{(j)} V_t^{(j)} dt + v_{\lambda^{(j)} t}^{(j)} dL_{\lambda^{(j)} t}^{(j)}, \quad j = 1, \ldots, J. \end{aligned}$$

for $\lambda^{(j)} > 0$; the stochastic volatility process $v^{(j)}$ is assumed to be nonnegative, stationary and independent of the Lévy subordinators $L^{(j)}$, for $j = 1, \ldots, J$. For simplicity, we will assume that the $L^{(j)}$ are independent.

Remark Note that if we want to allow for correlation between $v^{(j)}$ and the Brownian motion W in the asset price (1) for at least one $j \in \{1, \ldots, J\}$, we cannot assume that the J generalised Ornstein–Uhlenbeck processes in the superposition are independent. This will obviously make computations in such a framework slightly more difficult than under the independence assumption.

B Proofs of the results from Section A

B.1 (Conditional) moments of V

Proof of Proposition 9 The conditional mean in given by

$$\mathbb{E}\left(V_t|\, V_0\right) = \mathbb{E}\left(V_0 e^{-\lambda t} + \int_0^t e^{-\lambda(t-s)}\omega_{\lambda s} dL_{\lambda s} \middle| V_0\right) = e^{-\lambda t}V_0 + \kappa_1(L_1)\mathbb{E}(\omega_0)\left(1 - e^{-\lambda t}\right).$$

Next, we compute the conditional second moment. From Itô's formula, we obtain

$$V_t^2 - V_0^2 = 2\int_0^t V_{s-}dV_s + [V]_t = -2\lambda\int_0^t V_{s-}^2 ds + 2\int_0^t V_{s-}\omega_{\lambda s}dL_{\lambda s} + \int_0^t \omega_{\lambda s}^2 d[L]_{\lambda s}.$$

So, we deduce that the conditional second moment of V satisfied the following first order ordinary differential equation:

$$\frac{d}{dt}\mathbb{E}\left(V_t^2\middle|\, V_0\right) = -2\lambda\mathbb{E}\left(V_t^2\middle|\, V_0\right) + 2\lambda\kappa_1(L_1)\mathbb{E}\left(V_t\omega_{\lambda t}|\, V_0\right) + \lambda\kappa_2(L_2)\mathbb{E}\left(\omega_0^2\right).$$

Note here that the joint moment of V and ω is given by

$$\begin{aligned}
\mathbb{E}\left(V_t\omega_{\lambda t}|\, V_0\right) &= \mathbb{E}\left(\omega_{\lambda t}\left(V_0 e^{-\lambda t} + \int_0^t e^{-\lambda(t-s)}\omega_{\lambda s-}dL_{\lambda s}\right)\middle| V_0\right)\\
&= V_0 e^{-\lambda t}\mathbb{E}\left(\omega_0\right) + \lambda\kappa_1(L_1)\int_0^t e^{-\lambda(t-s)}\mathbb{E}\left(\omega_{\lambda t}\omega_{\lambda s}\right)ds\\
&= V_0 e^{-\lambda t}\mathbb{E}\left(\omega_0\right) + \lambda\kappa_1(L_1)\int_0^t e^{-\lambda(t-s)}\gamma(\lambda(t-s))ds\\
&+ \lambda\kappa_1(L_1)\left(\mathbb{E}\left(\omega_0\right)\right)^2\int_0^t e^{-\lambda(t-s)}ds\\
&= V_0 e^{-\lambda t}\mathbb{E}\left(\omega_0\right) + \lambda\kappa_1(L_1)\int_0^t e^{-\lambda(t-s)}\gamma(\lambda(t-s))ds\\
&+ \kappa_1(L_1)\left(\mathbb{E}\left(\omega_0\right)\right)^2\left(1 - e^{-\lambda t}\right).
\end{aligned}$$

Therefore, we obtain the following ODE:

$$\frac{d}{dt}\mathbb{E}\left(V_t^2\middle|\, V_0\right) + 2\lambda\mathbb{E}\left(V_t^2\middle|\, V_0\right) = q(t),$$

where

$$\begin{aligned}
q(t) = \lambda\kappa_2(L_2)\mathbb{E}\left(\omega_0^2\right) + 2\lambda\kappa_1(L_1)\Bigg(V_0 e^{-\lambda t}\mathbb{E}\left(\omega_0\right) + \lambda\kappa_1(L_1)\int_0^t e^{-\lambda(t-s)}\gamma(\lambda(t-s))ds\\
+ \kappa_1(L_1)\left(\mathbb{E}\left(\omega_0\right)\right)^2\left(1 - e^{-\lambda t}\right)\Bigg).
\end{aligned}$$

From solving the differential equation, we get

$$\begin{aligned}
\mathbb{E}\left(V_t^2\middle|\, V_0\right) &= e^{-2\lambda t}\left(V_0^2 + \int_0^t e^{2\lambda s}q(s)ds\right)\\
&= e^{-2\lambda t}V_0^2 + \frac{1}{2}\kappa_2(L_2)\mathbb{E}\left(\omega_0^2\right)\left(1 - e^{-2\lambda t}\right) + 2\kappa_1(L_1)V_0\mathbb{E}\left(\omega_0\right)\left(e^{-\lambda t} - e^{-2\lambda t}\right)\\
&+ 2\lambda^2\kappa_1^2(L_1)\int_0^t\int_0^s e^{-2\lambda t}e^{\lambda(s+u)}\gamma(\lambda(s-u))duds\\
&+ \kappa_1^2(L_1)\left(\mathbb{E}\left(\omega_0\right)\right)^2\left(1 - 2e^{-\lambda t} + e^{-2\lambda t}\right).
\end{aligned}$$

Also note that from change of variables with $x = s - u$, $u = s - x$, $du = -dx$ and the Fubini theorem, we get

$$\int_0^t \int_0^s e^{\lambda(u+s)}\gamma(\lambda(s-u))duds = \int_0^t \int_0^s e^{\lambda(2s-x)}\gamma(\lambda x)dxds = \int_0^t \int_x^t e^{\lambda(2s-x)}\gamma(\lambda x)dsdx$$
$$= \int_0^t e^{-\lambda x}\gamma(\lambda x)\int_x^t e^{2\lambda s}dsdx = \int_0^t e^{-\lambda x}\gamma(\lambda x)\frac{1}{2\lambda}\left(e^{2\lambda t} - e^{2\lambda x}\right)dx$$
$$= \frac{1}{2\lambda^2}\int_0^{\lambda t}\gamma(y)\left(e^{2\lambda t - y} - e^y\right)dy.$$

Altogether, we obtain the following expression for the conditional variance

$$Var\left(V_t|\, V_0\right) = \frac{1}{2}\kappa_2(L_2)\mathbb{E}\left(\omega_0^2\right)\left(1 - e^{-2\lambda t}\right) + \kappa_1^2(L_1)\int_0^{\lambda t}\gamma(y)\left(e^{-y} - e^{-2\lambda t + y}\right)dy.$$

Finally, we compute the conditional covariation. Clearly,

$$Cov\left(V_t, V_{t+h}|\, V_0\right) = \mathbb{E}\left(V_t^2\middle|\, V_0\right) + \mathbb{E}\left(V_t(V_{t+h} - V_t)|\, V_0\right) - \mathbb{E}\left(V_t|\, V_0\right)\mathbb{E}\left(V_{t+h}|\, V_0\right).$$

The only still unknown quantity is $\mathbb{E}\left(V_t(V_{t+h} - V_t)|\, V_0\right)$. Note that

$$V_t(V_{t+h} - V_t) = V_t V_0 e^{-\lambda t}\left(e^{-\lambda h} - 1\right) + V_t\int_0^t e^{-\lambda(t-s)}\left(e^{-\lambda h} - 1\right)\omega_{\lambda s}dL_{\lambda s}$$
$$+ V_t\int_t^{t+h} e^{-\lambda(t+h-s)}\omega_{\lambda s}dL_{\lambda s}$$
$$= \left(e^{-\lambda h} - 1\right)V_t^2 + V_t\int_t^{t+h} e^{-\lambda(t+h-s)}\omega_{\lambda s}dL_{\lambda s}.$$

Hence

$$Cov\left(V_t, V_{t+h}|\, V_0\right) = e^{-\lambda h}\mathbb{E}\left(V_t^2\middle|\, V_0\right) + \mathbb{E}\left(V_t\int_t^{t+h} e^{-\lambda(t+h-s)}\omega_{\lambda s}dL_{\lambda s}\middle|\, V_0\right)$$
$$- \mathbb{E}\left(V_t|\, V_0\right)\mathbb{E}\left(V_{t+h}|\, V_0\right). \quad (8)$$

So we only have to compute the second term on the right hand side of equation (8).

$$\mathbb{E}\left(V_t\int_t^{t+h} e^{-\lambda(t+h-s)}\omega_{\lambda s}dL_{\lambda s}\middle|\, V_0\right)$$
$$= V_0\kappa_1(L_1)\mathbb{E}\left(\omega_0\right)e^{-\lambda t}\left(1 - e^{-\lambda h}\right) + \mathbb{E}\left(\int_0^t e^{-\lambda(t-u)}\omega_{\lambda u}dL_{\lambda u}\int_t^{t+h} e^{-\lambda(t+h-s)}\omega_{\lambda s}dL_{\lambda s}\middle|\, V_0\right)$$
$$= V_0\kappa_1(L_1)\mathbb{E}\left(\omega_0\right)e^{-\lambda t}\left(1 - e^{-\lambda h}\right) + \kappa_1^2(L_1)\left(\mathbb{E}\left(\omega_0\right)\right)^2\left(1 - e^{-\lambda t}\right)\left(1 - e^{-\lambda h}\right)$$
$$+ \lambda^2\kappa_1^2(L_1)\int_t^{t+h}\int_0^t e^{-\lambda(t+h-s)}e^{-\lambda(t-u)}\gamma(\lambda(u-s))duds.$$

Note that

$$\int_t^{t+h}\int_0^t e^{\lambda(u+s)}\gamma(\lambda(u-s))dsdu = \int_t^{t+h}\int_{u-t}^u e^{\lambda(2u-x)}\gamma(\lambda x)dxdu.$$

Altogether, we obtain the following

$$Cov\left(V_t, V_{t+h}|\, V_0\right)$$
$$= e^{-2\lambda t}e^{-\lambda h}V_0^2 + \frac{1}{2}\kappa_2(L_2)\mathbb{E}\left(\omega_0^2\right)\left(1 - e^{-2\lambda t}\right)e^{-\lambda h} + 2\kappa_1(L_1)V_0\mathbb{E}\left(\omega_0\right)\left(e^{-\lambda t} - e^{-2\lambda t}\right)e^{-\lambda h}$$

$$
\begin{aligned}
&+ 2\lambda^2\kappa_1^2(L_1)\int_0^t\int_0^s e^{-2\lambda t}e^{\lambda(s+u)}\gamma(\lambda(s-u))duds e^{-\lambda h}\\
&+ \kappa_1^2(L_1)\left(\mathbb{E}\left(\omega_0\right)\right)^2\left(1-2e^{-\lambda t}+e^{-2\lambda t}\right)e^{-\lambda h}\\
&+ V_0\kappa_1(L_1)\mathbb{E}\left(\omega_0\right)e^{-\lambda t}\left(1-e^{-\lambda h}\right)+\kappa_1^2(L_1)\left(\mathbb{E}\left(\omega_0\right)\right)^2\left(1-e^{-\lambda t}\right)\left(1-e^{-\lambda h}\right)\\
&+ \lambda^2\kappa_1^2(L_1)\int_t^{t+h}\int_0^t e^{-\lambda(t+h-s)}e^{-\lambda(t-u)}\gamma(\lambda(u-s))dsdu\\
&- e^{-2\lambda t-\lambda h}V_0^2-\kappa_1(L_1)\mathbb{E}\left(\omega_0\right)V_0\left(e^{-\lambda t}\left(1-e^{-\lambda(t+h}\right)+e^{-\lambda(t+h)}\left(1-e^{-\lambda t}\right)\right)\\
&- \kappa_1^2(L_1)\left(\mathbb{E}\left(\omega_0\right)\right)^2\left(1-e^{-\lambda t}\right)\left(1-e^{-\lambda(t+h)}\right)\\
=&\frac{1}{2}\kappa_2(L_2)\mathbb{E}\left(\omega_0^2\right)\left(1-e^{-2\lambda t}\right)e^{-\lambda h}+\kappa_1^2(L_1)e^{-\lambda h}\int_0^{\lambda t}\left(e^{-y}-e^{-2\lambda t+y}\right)\gamma(y)dy\\
&+ \lambda^2\kappa_1^2(L_1)e^{-2\lambda t-\lambda h}\int_t^{t+h}\int_{u-t}^{u}e^{2\lambda u-\lambda x}\gamma(\lambda x)dxdu.
\end{aligned}
$$

□

Proof of Proposition 11 We work with the following representation of V:

$$V_t = \int_{-\infty}^t e^{-\lambda(t-s)}\omega_{\lambda s}dL_{\lambda s}.$$

Then,

$$\mathbb{E}\left(V_t\right|\omega) = \lambda\kappa_1(L_1)\int_{-\infty}^t e^{-\lambda(t-s)}\omega_{\lambda s}ds.$$

Note that for $n \in \mathbb{N}$

$$\Delta V_t^n = V_t^n - V_{t-}^n = \left(\Delta V_t + V_{t-}\right)^n - V_{t-}^n = \sum_{k=1}^n \binom{n}{k}\left(\Delta V_t\right)^k V_{t-}^{n-k}.$$

From Itô's formula, we get

$$
\begin{aligned}
V_t^n - V_0^n &= n\int_0^t V_{s-}^{n-1}dV_s + \sum_{0<s\le t}\left(\Delta V_s^n - nV_{s-}^{n-1}\Delta V_s\right)\\
&= -\lambda n\int_0^t V_{s-}^n ds + n\int_0^t V_{s-}^{n-1}\omega_{\lambda s}dL_{\lambda s} + \sum_{0<s\le t}\sum_{k=2}^n\binom{n}{k}\left(\Delta V_s\right)^k V_{s-}^{n-k}\\
&= -\lambda n\int_0^t V_{s-}^n ds + n\int_0^t V_{s-}^{n-1}\omega_{\lambda s}dL_{\lambda s} + \sum_{k=2}^n\binom{n}{k}\sum_{0<s\le t}\left(\Delta V_s\right)^k V_{s-}^{n-k}.
\end{aligned}
$$

Hence

$$
\begin{aligned}
\frac{d}{dt}\mathbb{E}\left(V_t^n\right|\omega) &= -\lambda n\mathbb{E}\left(V_t^n\right|\omega) + n\lambda\kappa_1(L_1)\mathbb{E}\left(V_t^{n-1}\omega_{\lambda t}\right|\omega)\\
&\quad + \sum_{k=2}^n\binom{n}{k}\lambda\kappa_k(L_1)\mathbb{E}\left(\omega_{\lambda t}^k V_t^{n-k}\ |\omega\right),\\
\mathbb{E}\left(V_t^n\right|\omega) &= e^{-\lambda nt}\left\{\mathbb{E}\left(V_0^n\right|\omega) + \int_0^t g(u)e^{\lambda nu}du\right\},
\end{aligned}
$$

where

$$g(u) = n\lambda\kappa_1(L_1)\mathbb{E}\left(V_u^{n-1}\omega_{\lambda u}\middle|\,\omega\right) + \sum_{k=2}^{n}\binom{n}{k}\lambda\kappa_k(L_1)\mathbb{E}\left(\omega_{\lambda u}^k V_u^{n-k}\middle|\,\omega\right)$$
$$= n\lambda\kappa_1(L_1)\mathbb{E}\left(V_u^{n-1}\middle|\,\omega\right)\omega_{\lambda u} + \sum_{k=2}^{n}\binom{n}{k}\lambda\kappa_k(L_1)\mathbb{E}\left(V_u^{n-k}\middle|\,\omega\right)\omega_{\lambda u}^k.$$

So, for the conditional second moment, we get

$$\mathbb{E}\left(V_t^2\middle|\,\omega\right) = 2\lambda^2\kappa_1^2(L_1)e^{-2\lambda t}\int_{-\infty}^{t} e^{\lambda u}\omega_{\lambda u}\int_{-\infty}^{u} e^{\lambda s}\omega_{\lambda s}dsdu + \lambda\kappa_2(L_1)e^{-2\lambda t}\int_{-\infty}^{t} e^{2\lambda u}\omega_{\lambda u}^2 du.$$

The conditional variance is hence given by

$$Var\left(V_t\middle|\,\omega\right) = \lambda\kappa_2(L_1)e^{-2\lambda t}\int_{-\infty}^{t} e^{2\lambda u}\omega_{\lambda u}^2 du.$$

Finally, we compute the conditional covariance.

$$\mathbb{E}\left(V_tV_{t+h}\middle|\,\omega\right) = \mathbb{E}\left(V_t^2\middle|\,\omega\right) + \mathbb{E}\left(V_t(V_{t+h}-V_t)\middle|\,\omega\right).$$

Note that

$$V_{t+h} - V_t = \left(e^{-\lambda h} - 1\right)\int_{-\infty}^{t} e^{-\lambda(t-s)}\omega_{\lambda s}dL_{\lambda s} + \int_{t}^{t+h} e^{-\lambda(t+h-s)}\omega_{\lambda s}dL_{\lambda s},$$

hence, we deduce that

$$\mathbb{E}\left(V_tV_{t+h}\middle|\,\omega\right) = e^{-\lambda h}\mathbb{E}\left(V_t^2\middle|\,\omega\right) + \mathbb{E}\left(\int_{-\infty}^{t} e^{-\lambda(t-u)}\omega_{\lambda u}dL_{\lambda u}\int_{t}^{t+h} e^{-\lambda(t+h-s)}\omega_{\lambda s}dL_{\lambda s}\middle|\,\omega\right)$$
$$= \lambda^2\kappa_1^2(L_1)e^{-\lambda h}e^{-2\lambda t}\int_{-\infty}^{t} e^{\lambda u}\omega_{\lambda u}du\int_{-\infty}^{t+h} e^{\lambda s}\omega_{\lambda s}ds$$
$$+ \lambda\kappa_2(L_1)e^{-\lambda h}e^{-2\lambda t}\int_{-\infty}^{t} e^{2\lambda u}\omega_{\lambda u}^2 du.$$

Hence the covariance is given by

$$Cov\left(V_t, V_{t+h}\middle|\,\omega\right) = \lambda\kappa_2(L_1)e^{-\lambda h}e^{-2\lambda t}\int_{-\infty}^{t} e^{2\lambda u}\omega_{\lambda u}^2 du.$$

□

Proof of Proposition 12 The result for the mean follows immediately from Proposition 9 for $t \to \infty$. Also, we carry out very similar computations as before in order to compute the (co–)variance. For the variance, we get

$$\mathbb{E}\left(V_t^2\right) = e^{-2\lambda t}\left(\mathbb{E}\left(V_0^2\right) + \int_0^t e^{2\lambda u}\left(2\mathbb{E}\left(V_{u-}\omega_{\lambda u}\right)\lambda\kappa_1(L_1) + \lambda\kappa_2(L_1)\mathbb{E}\left(\omega_{\lambda u}^2\right)\right)\right)du.$$

Note that

$$\mathbb{E}\left(V_t\omega_{\lambda t}\right) = \mathbb{E}\left(\int_{-\infty}^{t} e^{-\lambda(t-s)}\omega_{\lambda s-}\omega_{\lambda t}dL_{\lambda s}\right)$$
$$= \lambda\kappa_1(L_1)\int_{-\infty}^{t} e^{-\lambda(t-s)}\left(\gamma(\lambda(t-s)) + \left(\mathbb{E}\left(\omega_0\right)\right)^2\right)ds$$

$$= \lambda\kappa_1(L_1)\int_{-\infty}^{t} e^{-\lambda(t-s)}\gamma(\lambda(t-s))ds + \kappa_1(L_1)\left(\mathbb{E}\left(\omega_0\right)\right)^2,$$

and

$$\mathbb{E}\left(V_0^2\right) = \frac{1}{2}\kappa_2(L_1)\mathbb{E}\left(\omega_0^2\right) + 2\lambda^2\kappa_1^2(L_1)\int_{-\infty}^{0}\int_{-\infty}^{u} e^{\lambda(s+u)}\mathbb{E}\left(\omega_{\lambda s}\omega_{\lambda u}\right)duds.$$

Hence,

$$\mathbb{E}\left(V_t^2\right) = 2\lambda^2\kappa_1^2(L_1)e^{-2\lambda t}\int_{-\infty}^{t}\int_{-\infty}^{u} e^{\lambda(u+s)}\gamma(\lambda(u-s))dsdu + \kappa_1^2(L_1)\left(\mathbb{E}\left(\omega_0\right)\right)^2 \\ + \frac{1}{2}\kappa_2(L_1)\mathbb{E}\left(\omega_0^2\right),$$

and

$$Var(V_t) = \frac{1}{2}\kappa_2(L_1)\mathbb{E}\left(\omega_0^2\right) + 2\lambda^2\kappa_1^2(L_1)e^{-2\lambda t}\int_{-\infty}^{t}\int_{-\infty}^{u} e^{\lambda(u+s)}\gamma(\lambda(u-s))dsdu.$$

Finally, an application of Fubini's theorem leads to

$$Var(V_t) = \frac{1}{2}\kappa_2(L_1)\mathbb{E}\left(\omega_0^2\right) + \kappa_1^2(L_1)\int_{0}^{\infty} e^{-y}\gamma(y)dy.$$

For the covariance, we have

$$\begin{aligned}
&\mathbb{E}\left(V_tV_{t+h}\right) = \mathbb{E}\left(V_t^2\right) + \mathbb{E}\left(V_t(V_{t+h}-V_t)\right) \\
&= e^{-\lambda h}\mathbb{E}\left(V_t^2\right) + e^{-2\lambda t}e^{-\lambda h}\lambda^2\kappa_1^2(L_1)\int_{-\infty}^{t}\int_{t}^{t+h} e^{\lambda(u+s)}\mathbb{E}\left(\omega_{\lambda s}\omega_{\lambda u}\right)dsdu \\
&= \frac{1}{2}\kappa_2(L_1)\mathbb{E}\left(\omega_0^2\right)e^{-\lambda h} + \kappa_1^2(L_1)\left(\mathbb{E}\left(V_t^2\right)\right)^2 \\
&+ \lambda^2\kappa_1^2(L_1)e^{-\lambda h}e^{-2\lambda t}\left(2\int_{-\infty}^{t}\int_{-\infty}^{u} e^{\lambda(u+s)}\gamma(\lambda(u-s))dsdu + \int_{-\infty}^{t}\int_{t}^{t+h} e^{\lambda(u+s)}\gamma(\lambda(s-u))dsdu\right) \\
&= \frac{1}{2}\kappa_2(L_1)\mathbb{E}\left(\omega_0^2\right)e^{-\lambda h} + \kappa_1^2(L_1)\left(\mathbb{E}\left(V_t^2\right)\right)^2 + \lambda^2\kappa_1^2(L_1)e^{-\lambda h}e^{-2\lambda t}\int_{-\infty}^{t}\int_{-\infty}^{t+h} e^{\lambda(u+s)}\gamma(\lambda|s-u|)dsdu
\end{aligned}$$

Hence, we deduce that

$$\begin{aligned}
&Cov(V_t, V_{t+h}) \\
&= e^{-\lambda h}\left(\frac{1}{2}\kappa_2(L_1)\mathbb{E}\left(\omega_0^2\right) + \lambda^2\kappa_1^2(L_1)e^{-2\lambda t}\int_{-\infty}^{t}\int_{-\infty}^{t+h} e^{\lambda(u+s)}\gamma(\lambda|s-u|)dsdu\right) \\
&= e^{-\lambda h}\left(\frac{1}{2}\kappa_2(L_1)\mathbb{E}\left(\omega_0^2\right) + \kappa_1^2(L_1)\left(\int_{0}^{\infty} e^{-y}\gamma(y)dy + \lambda^2e^{-2\lambda t}\int_{-\infty}^{t}\int_{t}^{t+h} e^{\lambda(u+s)}\gamma(\lambda|s-u|)dsdu\right)\right).
\end{aligned}$$

Since V is stationary, the expression above simplifies to

$$\begin{aligned}
&Cov(V_t, V_{t+h}) \\
&= e^{-\lambda h}\left(\frac{1}{2}\kappa_2(L_1)\mathbb{E}\left(\omega_0^2\right) + \lambda^2\kappa_1^2(L_1)\int_{-\infty}^{0}\int_{-\infty}^{h} e^{\lambda(u+s)}\gamma(\lambda|s-u|)dsdu\right) \\
&= e^{-\lambda h}\left(\frac{1}{2}\kappa_2(L_1)\mathbb{E}\left(\omega_0^2\right) + \kappa_1^2(L_1)\left(\int_{0}^{\infty} e^{-y}\gamma(y)dy + \lambda^2\int_{-\infty}^{0}\int_{0}^{h} e^{\lambda(u+s)}\gamma(\lambda|s-u|)dsdu\right)\right).
\end{aligned}$$

□

B.2 (Conditional) moments of V^+

Proof of Proposition 1 Recall that

$$dV_t = -\lambda V_t dt + \omega_{\lambda t} dL_{\lambda t}, \qquad \text{and} \qquad V_t = V_0 e^{-\lambda t} + \int_0^t e^{-\lambda(t-s)} \omega_{\lambda s} dL_{\lambda s}.$$

Therefore we obtain for $\epsilon(t,\lambda) = \frac{1}{\lambda}\left(1 - e^{-\lambda t}\right)$:

$$\begin{aligned}
\int_0^t dV_s &= V_t - V_0 = -\lambda \int_0^t V_s ds + \int_0^t \omega_{\lambda s} dL_{\lambda s}, \\
\int_0^t V_s ds &= \frac{1}{\lambda}\left(V_0 - V_t + \int_0^t \omega_{\lambda s} dL_{\lambda s}\right) \\
&= \frac{1}{\lambda}\left(\left(1 - e^{-\lambda t}\right) V_0 + \int_0^t \left(1 - e^{-\lambda(t-s)}\right) \omega_{\lambda s} dL_{\lambda s}\right) \\
&= \epsilon(t,\lambda) V_0 + \int_0^t \epsilon(t-s,\lambda) \omega_{\lambda s} dL_{\lambda s}.
\end{aligned}$$

□

Proof of Proposition 14 The conditional mean in given by

$$\begin{aligned}
\mathbb{E}\left(V_t^+ \middle| V_0\right) &= \frac{1}{\lambda}\left(1 - e^{-\lambda t}\right) V_0 + \frac{1}{\lambda}\mathbb{E}\left(\int_0^t \left(1 - e^{-\lambda(t-s)}\right) \omega_{\lambda s} dL_{\lambda s}\right) \\
&= \frac{1}{\lambda}\left(1 - e^{-\lambda t}\right) V_0 + \kappa_1(L_1)\mathbb{E}\left(\omega_0\right) \frac{1}{\lambda}\left(e^{-\lambda t} + \lambda t - 1\right) \\
&= \frac{1}{\lambda}\left(1 - e^{-\lambda t}\right)\left(V_0 - \kappa_1(L_1)\mathbb{E}\left(\omega_0\right)\right) + \kappa_1(L_1)\mathbb{E}\left(\omega_0\right) t.
\end{aligned}$$

Next, we compute the second moment. Note that

$$\left(V_t^+\right)^2 = 2\int_0^t V_{s-}^+ dV_s^+ + [V^+]_t = 2\int_0^t V_s \int_0^s V_u du ds = 2\int_0^t \int_0^s V_s V_u du ds.$$

Hence, the conditional second moment is given by

$$\begin{aligned}
\mathbb{E}\left(\left(V_t^+\right)^2 \middle| V_0\right) &= \int_0^t \int_0^s \mathbb{E}\left(V_s V_u \middle| V_0\right) du ds = 2\int_0^t \int_s^t \mathbb{E}\left(V_s V_u \middle| V_0\right) du ds \\
&= 2\int_0^t \int_s^t \left(Cov\left(V_s, V_u \middle| V_0\right) + \mathbb{E}\left(V_s \middle| V_0\right) \mathbb{E}\left(V_u \middle| V_0\right)\right) du ds \\
&= \kappa_2(L_1)\mathbb{E}\left(\omega_0^2\right) \int_0^t \int_s^t e^{-\lambda(u-s)} \left(1 - e^{-2\lambda s}\right) du ds \\
&\quad + 2\kappa_1^2(L_1) \int_0^t \int_s^t e^{-\lambda(u-s)} \int_0^{\lambda s} \left(e^{-y} - e^{-2\lambda s + y}\right) \gamma(y) dy du ds \\
&\quad + 2\lambda^2 \kappa_1^2(L_1) \int_0^t \int_s^t e^{-\lambda(u+s)} \int_u^s \int_{x-u}^x e^{2\lambda x - \lambda y} \gamma(\lambda y) dy dx du ds \\
&\quad + 2\int_0^t \int_s^t \mathbb{E}\left(V_s \middle| V_0\right) \mathbb{E}\left(V_u \middle| V_0\right) du ds \\
&= \kappa_2(L_1)\mathbb{E}\left(\omega_0^2\right) \frac{1}{\lambda^2}\left(-\frac{1}{2}e^{-2\lambda t} + 2e^{-\lambda t} + \lambda t - \frac{3}{2}\right) \\
&\quad + 2\kappa_1^2(L_1) \int_0^t \int_s^t e^{-\lambda(u-s)} \int_0^{\lambda s} \left(e^{-y} - e^{-2\lambda s + y}\right) \gamma(y) dy du ds
\end{aligned}$$

$$+ 2\lambda^2\kappa_1^2(L_1)\int_0^t\int_s^t e^{-\lambda(u+s)}\int_u^s\int_{x-u}^x e^{2\lambda x-\lambda y}\gamma(\lambda y)dydxduds$$
$$+ 2\int_0^t\int_s^t \mathbb{E}\left(V_s|\, V_0\right)\mathbb{E}\left(V_u|\, V_0\right)duds.$$

Hence, the conditional variance is given by

$$\begin{aligned}
Var\left(V_t^+\middle|\, V_0\right) &= \kappa_2(L_1)\mathbb{E}\left(\omega_0^2\right)\frac{1}{\lambda^2}\left(-\frac{1}{2}e^{-2\lambda t}+2e^{-\lambda t}+\lambda t-\frac{3}{2}\right)\\
&+ 2\kappa_1^2(L_1)\int_0^t\int_s^t e^{-\lambda(u-s)}\int_0^{\lambda s}\left(e^{-y}-e^{-2\lambda s+y}\right)\gamma(y)dyduds\\
&+ 2\lambda^2\kappa_1^2(L_1)\int_0^t\int_s^t e^{-\lambda(u+s)}\int_u^s\int_{x-u}^x e^{2\lambda x-\lambda y}\gamma(\lambda y)dydxduds.
\end{aligned}$$

□

Proof of Proposition 15 For the mean, we get:

$$\mathbb{E}\left(V_t^+\right) = \int_0^t \mathbb{E}\left(V_s\right)ds = \kappa_1(L_1)\mathbb{E}\left(\omega_0\right)t.$$

Next, we compute the second moment, by using Fubini's theorem.

$$\begin{aligned}
\mathbb{E}\left(V_t^+\right)^2 &= 2\int_0^t\int_s^t \mathbb{E}\left(V_sV_u\right)duds = 2\int_0^t\int_s^t\left(Cov(V_s,V_u)+\left(\mathbb{E}\left(V_0\right)\right)^2\right)duds\\
&= \kappa_2(L_1)\mathbb{E}\left(\omega_0^2\right)\int_0^t\int_s^t e^{-\lambda(u-s)}duds + 2\int_0^t\int_s^t G(u-s)duds + \left(\kappa_1(L_1)\mathbb{E}\left(\omega_0\right)t\right)^2\\
&= \kappa_2(L_1)\mathbb{E}\left(\omega_0^2\right)\frac{1}{\lambda^2}\left(e^{-\lambda t}-1+\lambda t\right) + 2\int_0^t\int_s^t G(u-s)duds + \left(\kappa_1(L_1)\mathbb{E}\left(\omega_0\right)t\right)^2,\\
Var\left(V_t^+\right) &= \kappa_2(L_1)\mathbb{E}\left(\omega_0^2\right)\frac{1}{\lambda^2}\left(e^{-\lambda t}-1+\lambda t\right) + 2\int_0^t\int_s^t G(u-s)duds.
\end{aligned}$$

Finally, we compute the covariance.

$$\begin{aligned}
&Cov\left(V_t^+, V_{t+h}^+\right)\\
&= \mathbb{E}\left(V_t^+V_{t+h}^+\right) - \mathbb{E}\left(V_t^+\right)\mathbb{E}\left(V_{t+h}^+\right)\\
&= \mathbb{E}\left(\int_0^t V_sds\int_0^{t+h}V_udu\right) - \left(\kappa_1(L_1)\mathbb{E}\left(\omega_0\right)\right)^2\left(t^2+th\right)\\
&= \mathbb{E}\left(V_t^+\right)^2 + \int_0^t\int_t^{t+h}\mathbb{E}\left(V_sV_u\right)duds - \left(\kappa_1(L_1)\mathbb{E}\left(\omega_0\right)\right)^2\left(t^2+th\right)\\
&= \kappa_2(L_1)\mathbb{E}\left(\omega_0^2\right)\frac{1}{\lambda^2}\left(e^{-\lambda t}-1+\lambda t\right) + 2\int_0^t\int_s^t G(u-s)duds - \left(\kappa_1(L_1)\mathbb{E}\left(\omega_0\right)\right)^2 th\\
&\quad + \int_0^t\int_t^{t+h}\left(\frac{1}{2}\kappa_2(L_1)\mathbb{E}\left(\omega_0^2\right)e^{-\lambda(u-s)} + G(u-s) + \left(\kappa_1(L_1)\mathbb{E}\left(\omega_0\right)\right)^2\right)duds,\\
&= \frac{1}{2}\kappa_2(L_1)\mathbb{E}\left(\omega_0^2\right)\frac{1}{\lambda^2}\left(e^{-\lambda(t+h)}+e^{-\lambda t}-e^{-\lambda h}-1+2\lambda t\right) + \int_0^t\int_t^{t+h}G(u-s)duds.
\end{aligned}$$

□

B.3 Characteristic functionals

Proof or Proposition 16

$$
\begin{aligned}
&\mathbb{E}(\exp(i\theta V_t)|V_0,\omega_s,0\le s\le\lambda t)\\
&=\exp\left(i\theta e^{-\lambda t}V_0\right)\mathbb{E}\left(\exp\left(i\theta e^{-\lambda t}\int_0^{\lambda t}e^s\omega_s dL_s\right)\Big| |V_0,\omega_s,0\le s\le\lambda t)\right)\\
&=\exp\left(i\theta e^{-\lambda t}V_0\right)\exp\left(\int_0^{\lambda t}\psi_L\left(\theta e^{-\lambda t}e^s\omega_s\right)ds\right).
\end{aligned}
$$

□

Proof of Proposition 17 From Fubini's theorem, we get:

$$
\begin{aligned}
f\bullet V_T=\int_0^T f(t)dV_t&=-\lambda\int_0^T f(t)V_t dt+\int_0^T f(t)\omega_{\lambda t}dL_{\lambda t}\\
&=-\lambda\int_0^T f(t)\left(V_0e^{-\lambda t}+\int_0^t e^{-\lambda(t-s)}\omega_{\lambda s}dL_{\lambda s}\right)dt+\int_0^T f(t)\omega_{\lambda t}dL_{\lambda t}\\
&=-\lambda V_0\int_0^T f(t)e^{-\lambda t}dt-\lambda\int_0^T\left(\int_s^T f(t)e^{-\lambda(t-s)}dt\right)\omega_{\lambda s}dL_{\lambda s}+\int_0^T f(t)\omega_{\lambda t}dL_{\lambda t}\\
&=-\lambda V_0\int_0^T f(t)e^{-\lambda t}dt-\lambda\int_0^T\left(\int_0^{T-s} f(t+s)e^{-\lambda t}dt\right)\omega_{\lambda s}dL_{\lambda s}+\int_0^T f(t)\omega_{\lambda t}dL_{\lambda t}\\
&=-\lambda V_0\int_0^T f(t)e^{-\lambda t}dt-\lambda\int_0^T\left(-\frac{1}{\lambda}f(s)+\left(\int_0^{T-s} f(t+s)e^{-\lambda t}dt\right)\right)\omega_{\lambda s}dL_{\lambda s}.
\end{aligned}
$$

Hence, the characteristic functional is given by

$$
\begin{aligned}
&\mathbb{E}\left(\exp\left(if\bullet V_T\right)\right|V_0,\omega_{\lambda s},0\le s\le T)\\
&=\mathbb{E}\left(\exp\left(i\int_0^T f(t)dV_t\right)\Big|V_0,\omega_{\lambda s},0\le s\le T\right)\\
&=\exp\left(-i\lambda V_0\int_0^T f(t)e^{-\lambda t}dt\right)\cdot\\
&\quad\mathbb{E}\left(\exp\left(-i\lambda\int_0^T\left(-\frac{1}{\lambda}f(s)+\left(\int_0^{T-s}f(t+s)e^{-\lambda t}dt\right)\right)\omega_{\lambda s}dL_{\lambda s}\right)\Big|\omega_{\lambda s},0\le s\le t\right)\\
&=\exp\left(-i\lambda V_0\int_0^T f(t)e^{-\lambda t}dt\right)\exp\left(\int_0^T\psi_L\left(\left(f(s)-\lambda\int_0^{T-s}f(t+s)e^{-\lambda t}dt\right)\omega_{\lambda s}\right)ds\right).
\end{aligned}
$$

□

Proof of Proposition 19 Note that we can write

$$
\int_0^t V_s ds=\frac{\left(1-e^{-\lambda t}\right)V_0}{\lambda}+\int_0^{\lambda t}\frac{\left(1-e^{-(\lambda t-s)}\right)}{\lambda}\omega_s dL_s
$$

$$
\begin{aligned}
&\mathbb{E}\left(\exp\left(i\theta\int_0^t V_u du\right)\Big|V_0\right)\\
&=\exp\left(\frac{i\theta V_0}{\lambda}\left(1-\exp(-\lambda t)\right)\right)\mathbb{E}\left(\exp\left(\frac{i\theta}{\lambda}\int_0^{\lambda t}\left(1-e^{-(\lambda t-s)}\right)\omega_s dL_s\right)\right)
\end{aligned}
$$

$$= \exp\left(\frac{i\theta V_0}{\lambda}(1-\exp(-\lambda t))\right)\mathbb{E}\left(\mathbb{E}\left(\exp\left(\frac{i\theta}{\lambda}\int_0^{\lambda t}\left(1-e^{-(\lambda t-s)}\right)\omega_s dL_s\right)\middle|\omega_s, 0\le s\le \lambda t\right)\right)$$
$$= \exp\left(\frac{i\theta V_0}{\lambda}(1-\exp(-\lambda t))\right)\mathbb{E}\left(\exp\left(\int_0^{\lambda t}\psi_L\left(\frac{\theta}{\lambda}\left(1-e^{-(\lambda t-s)}\right)\omega_s\right)ds\right)\right),$$

where $\psi(u)$ is the characteristic exponent of L_λ. So, we get

$$\mathbb{E}\left(\exp\left(i\theta\int_0^t V_u du\right)\middle| V_0\right)$$
$$= \phi\left(\exp\left(\frac{\theta}{\lambda}(1-\exp(-\lambda t))\right)\right)\mathbb{E}\left(\exp\left(\int_0^{\lambda t}\psi_L\left(\frac{\theta}{\lambda}\left(1-e^{-(\lambda t-s)}\right)\omega_s ds\right)\right)\right),$$

where ϕ is the characteristic function of V_0.

□

Proof of Proposition 20 We obtain from Fubini's theorem:

$$f\bullet V_t^+ = f\bullet\int_0^t V_s ds = \int_0^\infty f(t)V_t dt = \int_0^\infty f(t)\left(V_0 e^{-\lambda t} + \int_0^t e^{-\lambda(t-s)}\omega_{\lambda s} dL_{\lambda s}\right)dt$$
$$= V_0\int_0^\infty f(t)e^{-\lambda t}dt + \int_0^\infty f(t)e^{-\lambda t}\int_0^t e^{\lambda s}\omega_{\lambda s} dL_{\lambda s} dt$$
$$= V_0\int_0^\infty f(t)e^{-\lambda t}dt + \int_0^\infty\left(\int_s^\infty f(t)e^{-\lambda(t-s)}dt\right)\omega_{\lambda s} dL_{\lambda s}$$
$$= V_0\int_0^\infty f(t)e^{-\lambda t}dt + \int_0^\infty\left(\int_0^\infty f(t+s)e^{-\lambda t}dt\right)\omega_{\lambda s} dL_{\lambda s}.$$

Therefore, the conditional characteristic functional of the integrated process is given by

$$\mathbb{E}\left(\exp\left(if\bullet\int_0^t V_s ds\right)\middle|(\omega_s)_{s\ge 0}, V_0\right)$$
$$= \exp\left(iV_0\int_0^\infty f(t)e^{-\lambda t}dt\right)\cdot$$
$$\mathbb{E}\left(\exp\left(i\int_0^\infty\left(\int_0^\infty f(t+s)e^{-\lambda t}dt\right)\omega_{\lambda s} dL_{\lambda s}\right)\middle|(\omega_s)_{s\ge 0}, V_0\right)$$
$$= \exp\left(iV_0\int_0^\infty f(t)e^{-\lambda t}dt\right)\exp\left(\int_0^\infty\psi_{L_\lambda}\left(\omega_{\lambda s}\int_0^\infty f(t+s)e^{-\lambda t}dt\right)ds\right).$$

So, the conditional cumulant functional is given by

$$\log\left(\mathbb{E}\left(\exp\left(if\bullet\int_0^t V_s ds\right)\middle|(\omega_s)_{s\ge 0}, V_0\right)\right)$$
$$= iV_0\int_0^\infty f(t)e^{-\lambda t}dt + \int_0^\infty\psi_{L_\lambda}\left(\omega_{\lambda s}\int_0^\infty f(t+s)e^{-\lambda t}dt\right)ds.$$

□

C Proof of the consistency

Proof of Proposition 8: We know from Mancini (2006), Jacod (2008) that the truncated realised variance is a consistent estimator of integrated variance:

$$\sum_{i=1}^{\lfloor t/\delta_n\rfloor}(\delta_i^n Y)^2\,\mathbb{I}_{\{|\delta_i^n Y|\le c\delta_n^\epsilon\}}\overset{ucp}{\to}\int_0^t\sigma_s^2 ds,\quad \text{as } n\to\infty,$$

for $c > 0$ and $\epsilon \in (0, 1/2)$, where the convergence is uniform on compacts in probability (ucp). Then we choose K_n such that $K_n \to \infty$ and $K_n\delta_n \to 0$ as $n \to \infty$ and

$$\widehat{\sigma}^2(\delta_n)_{i\Delta_n} = \frac{1}{K_n\delta_n} \sum_{j=\lfloor i\Delta_n/\delta_n\rfloor - K_n}^{\lfloor i\Delta_n/\delta_n\rfloor} (\delta_i^n Y)^2 \, \mathbb{I}_{\{|\delta_i^n Y| \le c\delta_n^\epsilon\}}.$$

We know from Mancini (2006), Aït-Sahalia & Jacod (2009) that the above locally averaged truncated realised variance is a consistent estimator of the spot variance. In fact, it approximates the integral $\frac{1}{K_n\delta_n}\int_{s-K_n\delta_n}^{s} \sigma_x^2 dx$. And we know that for $s \in [0, T]$,

$$\frac{1}{K_n\delta_n}\int_{s-K_n\delta_n}^{s} \sigma_x^2 dx \overset{a.s.}{\to} \sigma_s^2, \quad \text{as } n \to \infty. \tag{9}$$

Clearly, the rate at which the convergence in (9) happens depends crucially on the degree of smoothness of the stochastic variance σ^2.

We obtain from standard arguments that $RV_t(\sigma^2, \Delta_n) \to [\sigma^2]_t$, as $n \to \infty$, where the convergence is uniformly on compacts in probability. Hence, it only remains to show that

$$N_t^n := RV_t\left(\widehat{\sigma}^2(\delta_n), \Delta_n\right) - RV_t\left(\sigma^2, \Delta_n\right) = \sum_{i=1}^{\lfloor t/\Delta_n\rfloor} \left(\left(\Delta_i^n \widehat{\sigma}^2(\delta_n)\right)^2 - \left(\Delta_i^n \sigma^2\right)^2\right)$$

converges to 0 in probability as $n \to \infty$. For all $i \in \{0, 1, \ldots, \lfloor t/\Delta_n\rfloor\}$, we can write

$$\left(\left(\Delta_i^n \widehat{\sigma}^2(\delta_n)\right)^2 - \left(\Delta_i^n \sigma^2\right)^2\right) = \left(\Delta_i^n \widehat{\sigma}^2(\delta_n) - \Delta_i^n \sigma^2\right)\left(\Delta_i^n \widehat{\sigma}^2(\delta_n) + \Delta_i^n \sigma^2\right).$$

The first term in this product satisfies

$$A_i^n := \Delta_i^n \widehat{\sigma}^2(\delta_n) - \Delta_i^n \sigma^2 = \left(\widehat{\sigma}^2(\delta_n)_{i\Delta_n} - \sigma_{i\Delta_n}^2\right) - \left(\widehat{\sigma}^2(\delta_n)_{(i-1)\Delta_n} - \sigma_{(i-1)\Delta_n}^2\right),$$

and the second term is given by

$$B_i^n := \Delta_i^n \widehat{\sigma}^2(\delta_n) + \Delta_i^n \sigma^2 = 2\Delta_i^n \sigma^2 + A_i^n,$$

and $N_t^n = \sum_{i=1}^{\lfloor t/\Delta_n\rfloor} A_i^n B_i^n$. Clearly,

$$|N_t^n| \le \sup_{i\in\{0,1,\ldots,\lfloor t/\Delta_n\rfloor\}} |B_i^n| \sum_{i=1}^{\lfloor t/\Delta_n\rfloor} |A_i^n| \le \sup_{i\in\{0,1,\ldots,\lfloor t/\Delta_n\rfloor\}} |B_i^n| \left\lfloor \frac{t}{\Delta_n} \right\rfloor \sup_{i\in\{0,1,\ldots,\lfloor t/\Delta_n\rfloor\}} |A_i^n|.$$

Now, we proceed by proving that the following two equalities hold:

$$\sup_{i\in\{0,1,\ldots,\lfloor t/\Delta_n\rfloor\}} \frac{1}{\Delta_n} |A_i^n| = o_P(1), \tag{10}$$

$$\sup_{i\in\{0,1,\ldots,\lfloor t/\Delta_n\rfloor\}} |B_i^n| = O_P(1), \tag{11}$$

which implies that $N_t^n = o_P(1)$. Actually, since $B_i^n = 2\Delta_i^n \sigma^2 + A_i^n$ and σ is càdlàg, (11) follows immediately from (10). So, it remains to prove that (10) holds. Clearly, we can write

$$|A_i^n| \le \left|\widehat{\sigma}^2(\delta_n)_{i\Delta_n} - \sigma_{i\Delta_n}^2\right| + \left|\widehat{\sigma}^2(\delta_n)_{(i-1)\Delta_n} - \sigma_{(i-1)\Delta_n}^2\right|,$$

which are terms of the same stochastic order. So, it will be sufficient to prove the uniform convergence for just one of those terms. Note that for any $i \in \{0, 1, \ldots, \lfloor t/\Delta_n\rfloor\}$, we have

$$\widehat{\sigma}^2(\delta_n)_{i\Delta_n} - \sigma_{i\Delta_n}^2 = \left(\widehat{\sigma}^2(\delta_n)_{i\Delta_n} - \frac{1}{K_n\delta_n}\int_{i\Delta_n - K_n\delta_n}^{i\Delta_n} \sigma_x^2 dx\right) + \left(\frac{1}{K_n\delta_n}\int_{i\Delta_n - K_n\delta_n}^{i\Delta_n} \sigma_x^2 dx - \sigma_{i\Delta_n}^2\right).$$

From Jacod (2008, Theorem 2.4 (iii) and Theorem 2.10), we can deduce that

$$\sup_{i\in\{0,1,\dots,\lfloor t/\Delta_n\rfloor\}}\left(\widehat{\sigma}^2(\delta_n)_{i\Delta_n}-\frac{1}{K_n\delta_n}\int_{i\Delta_n-K_n\delta_n}^{i\Delta_n}\sigma_x^2dx\right)=O_P\left(\frac{1}{\sqrt{K_n}}\right),$$

if Jacod (2008, Assumption (L-s)) holds for $s\leq\frac{4\epsilon-1}{2\epsilon}$. So basically, that means that the Blumenthal–Getoor index of J has to be strictly smaller than $\frac{4\epsilon-1}{2\epsilon}$, where ϵ is chosen as in Definition 7. For the second term, we get

$$\sup_{i\in\{0,\dots,\lfloor t/\Delta_n\rfloor\}}\left|\frac{1}{K_n\delta_n}\int_{i\Delta_n-K_n\delta_n}^{i\Delta_n}\sigma_x^2dx-\sigma_{i\Delta_n}^2\right|\leq\sup_{i\in\{0,\dots,\lfloor t/\Delta_n\rfloor\}}\sup_{\{x\in[i\Delta_n-K_n\delta_n,i\Delta_n]\}}\left|\sigma_x^2-\sigma_{i\Delta_n}^2\right|=:\mathcal{A}_t^n.$$

In order for the central limit theorem to hold, we need to derive restrictions on B and C to ensure that

$$\sqrt{K_n}\mathcal{A}_t^n=o_{a.s.}(\alpha_n), \tag{12}$$

for a sequence $(\alpha_n)_{n\in\mathbb{N}}$ which converges to 0 as $n\to\infty$. Then, we have

$$\sup_{i\in\{0,1,\dots,\lfloor t/\delta_n\rfloor\}}\left|\widehat{\sigma}^2(\delta_n)_{i\Delta_n}-\frac{1}{K_n\delta_n}\int_{i\Delta_n-K_n\delta_n}^{i\Delta_n}\sigma_x^2dx\right|=O_P\left(\frac{1}{\sqrt{K_n}}\right),$$

and

$$\sup_{i\in\{0,\dots,\lfloor t/\Delta_n\rfloor\}}\sup_{x\in[i\Delta_n-K_n\delta_n,i\Delta_n]}\left|\sigma_x^2-\sigma_{i\Delta_n}^2\right|=o_{a.s.}\left(\frac{\alpha_n}{\sqrt{K_n}}\right).$$

Finally, we have to impose further restrictions on the parameters B,C to ensure that both

$$\frac{1}{\Delta_n}O_P\left(\frac{1}{\sqrt{K_n}}\right)=o(1), \tag{13}$$

and

$$\text{and}\quad\frac{1}{\Delta_n}\mathcal{A}_t^n=\frac{1}{\Delta_n}o_{a.s.}\left(\frac{\alpha_n}{\sqrt{K_n}}\right)=o(1). \tag{14}$$

Remark Note that condition (14) is weaker than condition (12), since $\sqrt{K_n}$ grows faster to ∞ than $\frac{1}{\Delta_n}$:

$$\frac{\sqrt{K_n}}{\Delta_n}=O\left(\delta_n^{B/2-C}\right)=\infty,\quad\text{for }\frac{B}{2}<C,$$

which is always satisfied for $C>0$ and $B<0$. Hence, we will only impose conditions (12), (13) in the following.

Now, we study the cases where the volatility model satisfies assumption (V1) or assumption (V2) separately.

Case 1: σ^2 satisfies assumption (V1)

Now, we assume (V1). I.e. $d\sigma_t^2=b_tdt+\gamma_tdB_t$, for càdlàg, square–integrable processes $b=(b_t)_{t\geq0},(\gamma_t)_{t\geq0}$ and a standard Brownian motion $B=(B_t)_{t\geq0}$. (If we assume that σ rather than σ^2 is a Brownian semimartingale, the proof is analogous after applying the mean value theorem.)

Let $\kappa>0$ denote a constant, which might change from line to line throughout the proof. Since b and γ are bounded, we deduce from Lévy's modulus of continuity and the Dubins-Schwarz theorem (Protter (2004)) for a new Brownian motion $\tilde{B}=(\tilde{B}_t)_{t\geq0}$ that

$$\sup_{i\in\{1,\dots,\lfloor t/\Delta_n\rfloor\}}\sup_{x\in[i\Delta_n-K_n\delta_n,i\Delta_n]}\left|\sigma_x^2-\sigma_{i\Delta_n}^2\right|$$

$$
\begin{aligned}
&= \sup_{i\in\{1,\dots,\lfloor t/\Delta_n\rfloor\}} \sup_{x\in[i\Delta_n-K_n\delta_n,i\Delta_n]} \left|\int_x^{i\Delta_n} b_s ds + \int_x^{i\Delta_n} \gamma_s dB_s\right| \\
&= \sup_{i\in\{1,\dots,\lfloor t/\Delta_n\rfloor\}} \sup_{x\in[i\Delta_n-K_n\delta_n,i\Delta_n]} \left|\int_x^{i\Delta_n} b_s ds + \tilde{B}_{\int_x^{i\Delta_n}\gamma_s^2 ds}\right| \\
&\overset{L}{\leq} \sup_{i\in\{1,\dots,\lfloor t/\Delta_n\rfloor\}} \sup_{x\in[i\Delta_n-K_n\delta_n,i\Delta_n]} |b_x|K_n\delta_n + \kappa \sup_{i\in\{1,\dots,\lfloor t/\Delta_n\rfloor\}} \sup_{x\in[i\Delta_n-K_n\delta_n,i\Delta_n]} \left|\tilde{B}_{i\Delta_n-x}\right| \\
&= o_{a.s.}(K_n\delta_n) + O_{a.s.}\left(\sqrt{2(K_n\delta_n)\log((K_n\delta_n)^{-1})}\right).
\end{aligned}
$$

In order to simplify this further, we use $\Delta_n = O(\delta_n^C)$ for some $C > 0$ and $K_n = O(\delta_n^B)$ for some $-1 < B < 0$. Furthermore, note that $\lim_{x\to 0} \frac{\sqrt{x\log(1/x)}}{x^\xi} = 0$ if and only if $0 < \xi < 1/2$. So, we formulate the following sufficient conditions for our consistency result:

(i) Condition (12) is satisfied if $\sqrt{K_n}\sqrt{K_n\delta_n\log((K_n\delta_n)^{-1})} = o(1)$, which is equivalent to

$$
\frac{\sqrt{\delta_n^{B+1}\log(\delta_n^{-(B+1)})}}{\left(\delta_n^{B+1}\right)^{\frac{-B/2}{B+1}}} = o(1) \Leftrightarrow \qquad \frac{-B/2}{B+1} < \frac{1}{2} \Leftrightarrow \qquad B > -\frac{1}{2},
$$

and if

$$
\sqrt{K_n}K_n\delta_n = o(1), \Leftrightarrow \quad \delta_n^{B/2+B+1} = o(1), \Leftrightarrow \quad 3/2B+1>0, \Leftrightarrow \quad B > -2/3.
$$

(ii) Condition (13) is satisfied if

$$
\frac{1}{\Delta_n\sqrt{K_n}} = o(1) \Leftrightarrow \quad \delta_n^{-(C+B/2)} = o(1) \Leftrightarrow \quad -(C+B/2) > 0 \Leftrightarrow \quad C < -\frac{B}{2},
$$

and if

$$
\frac{K_n\delta_n}{\Delta_n} = o(1), \Leftrightarrow \quad \delta_n^{B+1-C} = o(1), \Leftrightarrow \quad B+1-C>0, \Leftrightarrow \quad C < B+1.
$$

So, altogether, we get the following restrictions on B and C:

$$
-\frac{1}{2} < B < 0, \qquad \text{and} \qquad 0 < C < -\frac{B}{2}.
$$

Case 2: σ^2 satisfies assumption (V2)

Now, we turn our attention to jump driven stochastic volatility models. Clearly, the degree of smoothness of σ^2 in the spot volatility estimation will not be the same as in the Brownian motion case. So we cannot work with Lévy 's modulus of continuity. We assume that (V2) holds, i.e. $d\sigma_t^2 = b_t dt + \gamma_t dL_t$, for càdlàg processes $b = (b_t)_{t\geq 0}, (\gamma_t)_{t\geq 0}$ and a Lévy subordinator $L = (L_t)_{t\geq 0}$ possibly with drift, which we denote by $\tilde{b}$, and with Lévy measure ν. We know from Pruitt (1981), that $\limsup_{t\to 0} \frac{\sup_{0\leq s\leq t}|L_s|}{t^{1/\alpha}} = 0$ a.s. for $\alpha > \beta$, where $\beta = \inf\{\alpha : \limsup_{x\to 0} x^\alpha h(x) = 0\}$, and

$$
h(x) = \int_{\{|y|>x\}} y d\nu(y) + \frac{1}{x^2}\int_{\{|y|\leq x\}} y^2 d\nu(y) + \frac{1}{x}\left|\tilde{b} + \int_{\{|y|\leq x\}} \frac{yy^2}{1+y^2} - \int_{\{|y|>x\}} \frac{y}{1+y^2} d\nu(y)\right|.
$$

So, β is essentially the Blumenthal–Getoor index of L. Since L is a subordinator, its increments are always positive and, hence, the following inequality holds.

$$
\sup_{i\in\{1,\dots,\lfloor t/\Delta_n\rfloor\}} \sup_{x\in[i\Delta_n-K_n\delta_n,i\Delta_n]} \left|\sigma_x^2 - \sigma_{i\Delta_n}^2\right|.
$$

$$\leq \sup_{i\in\{1,\ldots,\lfloor t/\Delta_n\rfloor\}} \sup_{x\in[i\Delta_n-K_n\delta_n,i\Delta_n]} |b_x|K_n\delta_n + \sup_{x\in[0,t]}|\gamma_x| \sup_{i\in\{1,\ldots,\lfloor t/\Delta_n\rfloor\}} \sup_{x\in[i\Delta_n-K_n\delta_n,i\Delta_n]} |L_{i\Delta_n}-L_x|$$
$$= o_{a.s.}(K_n\delta_n) + o_{a.s.}\left((K_n\delta_n)^{1/\alpha}\right), \quad \text{for } \alpha > \beta.$$

So, we formulate the following sufficient conditions for our consistency result:

(i) Condition (12) is satisfied if $\sqrt{K_n}\,(K_n\delta_n)^{1/\alpha} = o(1)$ which is equivalent to

$$\delta_n^{B/2+B/\alpha+1/\alpha} = o(1) \Leftrightarrow \qquad B/2 + B/\alpha + 1/\alpha > 0 \Leftrightarrow \qquad B > -\frac{2}{2+\alpha},$$

and if

$$\sqrt{K_n}\,(K_n\delta_n) = o(1), \Leftrightarrow \qquad \delta_n^{B/2+(B+1)} = o(1), \Leftrightarrow \qquad B > -2/3.$$

(ii) Condition (13) is satisfied if

$$\frac{1}{\Delta_n\sqrt{K_n}} = o(1) \Leftrightarrow \qquad \delta_n^{-(C+B/2)} = o(1) \Leftrightarrow \qquad -(C+B/2) > 0 \Leftrightarrow \qquad C < -\frac{B}{2},$$

and if

$$\frac{K_n\delta_n}{\Delta_n} = o(1), \Leftrightarrow \qquad \delta_n^{B+1-C} = o(1), \Leftrightarrow \qquad C < B+1.$$

So, if the volatility process satisfies assumption (V2), we assume for $\alpha > \max\{1,\beta\}$ that

$$-\frac{2}{2+\alpha} < B < 0, \qquad \text{and} \qquad 0 < C < -\frac{B}{2}.$$

□

Research Papers 2009

CREATES
Center for Research in Econometric Analysis of Time Series

2009-11:	Lasse Bork: Estimating US Monetary Policy Shocks Using a Factor-Augmented Vector Autoregression: An EM Algorithm Approach
2009-12:	Konstantinos Fokianos, Anders Rahbek and Dag Tjøstheim: Poisson Autoregression
2009-13:	Peter Reinhard Hansen and Guillaume Horel: Quadratic Variation by Markov Chains
2009-14:	Dennis Kristensen and Antonio Mele: Adding and Subtracting Black-Scholes: A New Approach to Approximating Derivative Prices in Continuous Time Models
2009-15:	Charlotte Christiansen, Angelo Ranaldo and Paul Söderlind: The Time-Varying Systematic Risk of Carry Trade Strategies
2009-16:	Ingmar Nolte and Valeri Voev: Least Squares Inference on Integrated Volatility and the Relationship between Efficient Prices and Noise
2009-17:	Tom Engsted: Statistical vs. Economic Significance in Economics and Econometrics: Further comments on McCloskey & Ziliak
2009-18:	Anders Bredahl Kock: Forecasting with Universal Approximators and a Learning Algorithm
2009-19:	Søren Johansen and Anders Rygh Swensen: On a numerical and graphical technique for evaluating some models involving rational expectations
2009-20:	Almut E. D. Veraart and Luitgard A. M. Veraart: Stochastic volatility and stochastic leverage
2009-21:	Ole E. Barndorff-Nielsen, José Manuel Corcuera and Mark Podolskij: Multipower Variation for Brownian Semistationary Processes
2009-22:	Giuseppe Cavaliere, Anders Rahbek and A.M.Robert Taylor: Co-integration Rank Testing under Conditional Heteroskedasticity
2009-23:	Michael Frömmel and Robinson Kruse: Interest rate convergence in the EMS prior to European Monetary Union
2009-24:	Dominique Guégan: A Meta-Distribution for Non-Stationary Samples
2009-25:	Ole E. Barndorff-Nielsen and Almut E. D. Veraart: Stochastic volatility of volatility in continuous time