Testing for non-correlation between price and volatility jumps

Jean Jacod[*] Claudia Klüppelberg[†] Gernot Müller[‡]

Abstract

We consider a log-price process X_t, which is observed at discrete times $0, \Delta_n$, $2\Delta_n, \ldots$, and the process has a stochastic squared volatility σ_t^2. Assuming that the price process as well as the volatility process have common jumps, we suggest tests for non-correlation between log-price and squared volatility jumps, or functions of such jumps. Our tests have a prescribed asymptotic level, as the mesh Δ_n tends to 0 and the observation time T_n tends to ∞. The finite sample performance of our test is studied using simulations. We finally apply our tests to real data, and the test rejects the non-correlation hypothesis for the combination of squared log-price jumps and the moduli of the jumps of the squared volatility. This sheds new light on economically motivated statements on causality between price and volatility jumps and on econometric modeling.

MSC2010 Subject Classification: Primary 91G70, 62G10 , Secondary 62M02, 60H30

Keywords: common jumps, discrete sampling, high-frequency data, Itô semimartingale, statistical test, stochastic volatility model

1 Introduction

Models for financial data involving stochastic volatility and allowing for sample path discontinuities in the volatility as well as in the underlying asset price (stock, index, exchange rate) have become popular in recent years. Since Merton's paper [24] there is an increasing number of publications showing empirical evidence of jumps in the asset prices: from the distribution of log-returns viewpoint as in [1, 12, 15], or from a non-parametric statistical viewpoint as in [5, 10, 16] using multipower variations, or [2, 22] with other methods, see also the references in these papers. Naturally, if the asset price process has jumps, one might suppose that the volatility process also exhibits jumps, which is e.g. modeled by a continuous-time GARCH model [21]. However, as in [9], the volatility process can exhibit jumps, although price is continuous. The presence of

---
[*]Institut de Mathématiques de Jussieu, 4 Place Jussieu, 75 005 Paris, France (CNRS – UMR 7586, and Université Pierre et Marie Curie - P6), Email: jean.jacod@upmc.fr

[†]Centre for Mathematical Sciences, and Institute for Advanced Study, Technische Universität München, Boltzmannstrasse 3, 85748 Garching, Germany, Email: cklu@ma.tum.de

[‡]Centre for Mathematical Sciences, Technische Universität München, Boltzmannstrasse 3, 85748 Garching, Germany, Email: mueller@ma.tum.de

jumps, for price and/or volatility, implies that the model is incomplete (a stochastic volatility usually also implies incompleteness). The consequences of this fact for risk management, portfolio optimization and derivatives pricing, starting again with [24], have been extensively discussed already in the last decades, see for example [14] and many others.

Another important topic is the (statistical) connection between price and volatility, besides the mere fact that "$dX_t^c = \sigma_t\,dW_t$" for the "continuous martingale part" X^c of the log-price process X, where σ_t is the spot volatility and W the Brownian motion driving X. In continuous models, this connection often takes the name of *leverage* and is usually understood as being the correlation between the two Brownian motions driving X and σ. Most empirical studies conclude to a significant negative correlation, typically between -0.3 and -0.6. Since [11] and [13], there have been numerous economical interpretations or justifications of empirical findings of negative correlation between log-returns and functions of the volatility, and the fact that leads to the name "leverage" is that, when price goes down, the asset becomes (relatively) more risky, hence volatility goes up.

Now, once statistical evidence for the presence of jumps is established, it is important to investigate in a next step a possible relationship between price jumps and volatility jumps. It might of course happen that price and volatility never jump together; however, this seems ruled out by empirical evidence, see e.g. [19]. If common jumps occur, the same economical justification as for the (continuous) leverage to be negative would suggest the sizes of a log-price jump and a volatility jump (or some function of it) occurring simultaneously to be negatively correlated, and in any case this (positive or negative) correlation can be viewed as an assessment of the "jump leverage".

More generally, and even if the above correlation is found to be essentially zero, knowing the joint law – or at least some of its properties – of the two components of a "co-jump" is of uttermost importance for modeling purposes.

Of course, there are many possibilities for the joint law of co-jumps, ranging from an extreme case where there is a functional relationship between the two sizes, to the other extreme where they are independent. A precise functional relationship is – not surprisingly – ruled out by empirical evidence. On the other hand, testing for independence *stricto sensu*, especially in a non-parametric setting, is always a difficult task, and completely out of reach here with the relatively scarce large jumps one observes in finance, even with high-frequency data. So the more modest question addressed in this paper is to determine whether, for log-price and (squared) volatility co-jumps, the two jump sizes are correlated or not; or, more generally, if some functions of these are correlated or not. An important point is that we completely disentangle the jumps from the continuous part of our processes, unlike in [6] where a "global" leverage which includes both the leverage for the continuous part and a kind of correlation between jumps, is evaluated and found significantly negative in empirical studies.

In this paper we consider a single, liquid, asset. The price is observed at discrete and regularly spaced times $i\Delta_n$, for some small time lag Δ_n, whereas the final horizon T_n is large: so the asymptotic is $\Delta_n \to 0$ *and* $T_n \to \infty$, and in the empirical study we use $\Delta_n = 1$ minute and $T_n = 7$ years. We also assume that the microstructure noise is

negligible (probably a reasonable assumption when $\Delta_n = 1$ minute), that is, prices at times $i\Delta_n$ are exactly observed.

Price jumps, though, are not observed exactly, but the sizes of the "big" jumps are actually known up to a small – essentially negligible – error. As for the volatility, which is not directly observed *per se*, we use spot volatility estimators such as those in [3, 19], based on rolling windows of size about 1 hour when Δ_n is 1 minute; as we will see on Monte-Carlo experiments, this allows us to get "reasonable" estimators for the volatility jumps, although by far not as good as those for price jumps. The assumption $T_n \to \infty$, rather unusual in a high-frequency setting, is necessary here because otherwise one would only have a finite number of big jumps (or of jumps of any size when they have finite activity), and of course asymptotic inference is impossible when a fixed amount of data is observed.

The empirical analysis of co-jumps and their joint statistical properties has of course been done by a number of authors, although it typically concerns daily returns and the corresponding daily moves of volatility: see for example [7] and the references therein. The usual finding is that, as for the continuous leverage, a negative correlation is found for those daily returns large enough to be interpreted as jumps. In this paper we consider higher frequency observations and take into consideration intra-day (large enough) jumps. The volatility jumps are estimated by taking differences of spot volatility estimators based on windows of size almost one hour, and from our simulation studies these spot volatility estimators seem to behave well enough to obtain reasonably accurate estimators for the actual volatility jumps. This analysis of actual jumps rather than daily returns leads us to different empirical conclusions: the correlation between co-jumps is not statistically different from 0, although there seems to be some statistically significant dependence between the price and volatility co-jumps. However, from a mere visual inspection (which is of course not compelling in any sense) of the plots in Figure 3, for example, it looks like the dependence is indeed quite weak.

The paper is organized as follows: In Section 2 we set the stage by formulating the (log) price process and stochastic volatility models in its full generality of a bivariate Itô semimartingale. We also describe the kind of statistical hypotheses which we subject to testing. The theoretical results are given in Section 3, in the case of finite activity jumps for the price process, and extended to the general case in Section 4. Section 5 provides some Monte Carlo experiments. Section 6 shows the empirical results for SPDR S&P 500 ETF (SPY) high-frequency data during 2005-2011. Two more data sets have been analyzed and results are reported in appendix [17]. Section 7 concludes. Finally, proofs are given in Section 8.

2 Model framework and statistical hypotheses

Our (non-parametric) model is for the pair of the single log-price X and the corresponding squared volatility $c = \sigma^2$. The basic structural assumption is that (X, c) is an Itô semimartingale on some filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$. Their jumps are inte-

grable (since our main interest is their correlation, they should at least have finite second moment anyway). So we may write them as

$$
\begin{array}{rcl}
X_t & = & X_0 + \int_0^t b_s ds + \int_0^t \sigma_s \, dW_s + \int_0^t \int_E \delta(s,z) \, (\mu - \nu)(ds, dz) \\
c_t & = & c_0 + \int_0^t \widetilde{b}_s ds + \int_0^t \widetilde{\sigma}_s dW_s + \int_0^t \widehat{\sigma}_s dW'_s + \int_0^t \int_E \widetilde{\delta}(s,z) \, (\mu - \nu)(ds, dz)
\end{array}
\tag{1}
$$

(see [18] for all unexplained, but classical, notions or notation). Here, W and W' are independent standard Brownian motions, $\mu = \mu(\omega, dt, dz)$ is a Poisson random measure on $\mathbb{R}_+ \times E$, and the predictable compensator (or intensity measure) of μ is $\nu(dt, dz) = dt \otimes \lambda(dz)$ for some σ-finite measure λ on $(E, \mathcal{E})$. Further $(E, \mathcal{E})$ is an auxiliary Polish space (there is a lot of freedom for choosing the driving measure μ and λ, and even the space E: it is always possible to take $E = \mathbb{R}$ for example, and λ to be the Lebesgue measure).

This formulation is general enough to accommodate a large variety of situations: cases without or with leverage for the "continuous part", according to whether $\widetilde{\sigma}_t$ is identically 0 or not: cases where X and c do not jump together, if the product $\delta\widetilde{\delta}$ vanishes identically; cases where they *always* jump together, if $\delta \neq 0$ if and only if $\widetilde{\delta} \neq 0$. And, of course, all intermediate cases for simultaneous and non-simultaneous jumps.

The "coefficients" of the model are the processes $b_t, \widetilde{b}_t, \widetilde{\sigma}_t$ and $\widehat{\sigma}_t$, which are progressively measurable, and the predictable functions δ and $\widetilde{\delta}$ on $\Omega \times \mathbb{R}_+ \times E$. They should be such that the integrals in (1) are meaningful, and also that $c_t \geq 0$ for all $t \geq 0$. In our calculations below we will make two simplifying assumptions: one is that all moments of c_t are bounded in t and that those of X_t are all finite (although not bounded in t); second, we assume that the jumps of X have *finite activity* (we do not assume the same for c, because this property is without consequences on the mathematical treatment below). The first assumption is reasonable, at least for stock and foreign exchange markets; the second one is probably less so, but in Section 4 we explain, how it can be relaxed. The precise assumption is thus as follows:

Assumption (A) a) There are functions $\Gamma, \widetilde{\Gamma} \geq 0$ on E and processes $\gamma_t, \widetilde{\gamma}_t$ such that $|\delta(\omega, t, z)| \leq \gamma_t(\omega)\Gamma(z)$ and $|\widetilde{\delta}(\omega, t, z)| \leq \widetilde{\gamma}_t(\omega)\widetilde{\Gamma}(z)$, with $\int_E (1_{\{\Gamma(z)>0\}} \vee \Gamma(z)^p) \, \lambda(dz) < \infty$ and $\int_E (\widetilde{\Gamma}(z)^2 \vee \widetilde{\Gamma}(z)^p) \, \lambda(dz) < \infty$ for all $p \geq 2$.
  b) For all $p > 0$ we have

$$
\sup_{t \geq 0} \mathbb{E}\left[\sup_{s \in [t, t+1]} \left(|b_s| + |\widetilde{b}_s| + |\widetilde{\sigma}_s| + |\widehat{\sigma}_s| + \gamma_s + \widetilde{\gamma}_s + c_s \right)^p \right] < \infty.
\tag{2}
$$

Under (A-a) the process X_t has finite activity jumps, and finite moments of all order, but those are typically *not* bounded in time.

(A-b) replaces the weaker requirement that all processes $b_t, \widetilde{b}_t, \cdots$ are locally bounded, which is more common in the literature on high-frequency data with fixed time horizon: this property is "localized" and replaced by boundedness, but of course localization is impossible when the time horizon goes to infinity. This is needed because the number of jumps of X and c should go to infinity if we wish to have independence or non-correlation

4

tests to be consistent, and this cannot be achieved if the time horizon is kept fixed because we would only observe finitely many jumps of size larger than ε for any $\varepsilon > 0$.

With $\Delta Y_t = Y_t - Y_{t-}$ denoting the jump size at time t of any càdlàg process Y_t, we set

$$N_t = \sum_{s \leq t} 1_{\{\Delta X_s \neq 0\}}, \qquad S_1, S_2, \cdots \text{ are the successive jump times of } X. \qquad (3)$$

Then (A-a) implies that N_t is finite-valued, and under Assumption (B) below we will have $S_n \uparrow \infty$, and of course $N_t = \sum_{n \geq 1} 1_{\{S_n \leq t\}}$.

Remark 1 (i) The signs of $\sigma_t, \widetilde{\sigma}_t, \widehat{\sigma}_t$ in (1) are irrelevant (the sign of the product $\sigma_t \widetilde{\sigma}_t$ is relevant for the continuous leverage effect, but we are not considering this question here) and, for example, one may take $\sigma_t = \sqrt{c_t}$, which gives a càdlàg version for σ_t, and is even an Itô semimartingale if c_t and c_{t-} never vanish.(i)

(ii) We have required $c_t \geq 0$ above. Now, in virtually all models featuring a standard Brownian motion W in the formula for X, and of which we are aware, we have $c_t > 0$ and $c_{t-} > 0$ for all t. In this case one can also consider the log-volatility $c_t' = \log c_t$, which has been considered in [11, 13]. The process c' satisfies an equation of the form

$$c_t' = c_0' + \int_0^t \widetilde{b}_s' ds + \int_0^t \widetilde{\sigma}_s' dW_s + \int_0^t \widehat{\sigma}_s' dW_s' + \int_0^t \int_E \widetilde{\delta}'(s, z) (\mu - \nu)(ds, dz) \qquad (4)$$

for suitable coefficients $\widetilde{b}', \widetilde{\sigma}', \widehat{\sigma}', \widetilde{\delta}'$. For example, the connection between $\widetilde{\delta}$ and $\widetilde{\delta}'$ is

$$\widetilde{\delta}'(\omega, t, z) = \log \left(1 + \frac{\widetilde{\delta}(\omega, t, z)}{c_{t-}(\omega)} \right), \qquad \widetilde{\delta}(\omega, t, z) = c_{t-}(\omega) \left(e^{\widetilde{\delta}'(\omega, t, z)} - 1 \right). \qquad (5)$$

When c_t and c_{t-} do not vanish, it is theoretically equivalent to model c_t or its logarithm c_t'. In practice, modeling c_t' (which can take positive and negative values) is often easier than modeling c_t (which does not take negative values), when there are jumps, for the following reason: in (1) we need $\widetilde{\delta}(\omega, t, z) \geq -c_{t-}(\omega)$ for all z in the support of λ; so, unless we take $\delta \geq 0$ identically (implying positive volatility jumps only), this requirement prevents us to use a non-random function $\widetilde{\delta}(\omega, t, z) = \widetilde{\delta}(t, z)$. In the version (4) nothing prevents us from taking a non-random $\widetilde{\delta}'$.

(iii) If we use (4) we need assumptions on the coefficients to ensure that (A) is satisfied. For example, considering the conditions on $\widetilde{\delta}$ and $\widetilde{\delta}'$, if for instance $|\widetilde{\delta}'(\omega, t, z)| \leq \widetilde{\Gamma}'(z)$ for a function $\widetilde{\Gamma}'$ satisfying $\int \left(\widetilde{\Gamma}'(z)^2 \wedge e^{p\widetilde{\Gamma}'(z)} \right) \lambda(dz) < \infty$ for all $p \geq 0$, then (A) is satisfied by δ with $\widetilde{\gamma}_t = c_t$ and $\widetilde{\Gamma} = e^{\widetilde{\Gamma}'} - 1$. This is the reason why in (A) we have required $|\widetilde{\delta}(\omega, t, z)| \leq \widetilde{\gamma}_t(\omega) \widetilde{\Gamma}(z)$ instead of the – seemingly – more natural condition $|\widetilde{\delta}(\omega, t, z)| \leq \widetilde{\Gamma}(z)$. $\qquad \square$

Our general (albeit unreachable) aim is to test for "independence" of the sizes of simultaneous jumps of X and c or, rather, for "conditional independence" knowing the past before the jumps: this is what is really meaningful in a dynamical system and, thus,

it is what we want to check for modeling purposes. The conditional distribution of jumps, knowing the past, is easily obtained in terms of the coefficients δ and $\widetilde{\delta}$. Namely, for any two measurable functions f, g of polynomial growth we can define the process (with the convention $\frac{0}{0} = 0$)

$$V(f,g)_t = \frac{\int_E 1_{\{\delta(t,z)\neq 0\}} f(\delta(t,z)) \, g(\widetilde{\delta}(t,z)) \, \lambda(dz)}{\int_E 1_{\{\delta(t,z)\neq 0\}} \lambda(dz)}, \tag{6}$$

and we have

$$V(f,g)_{S_n} = \mathbb{E}\big[f(\Delta X_{S_n}) \, g(\Delta c_{S_n}) \mid \mathcal{F}_{S_n-}\big] \quad \text{on the set } \{S_n < \infty\}. \tag{7}$$

In other words, the above-mentioned conditional independence amounts to

$$V(f,g)_{S_n} = V(f,1)_{S_n} \, V(1,g)_{S_n} \quad \forall n \geq 1 \tag{8}$$

for all pairs f, g of functions, and by a density argument it suffices to check this, when f and g belong to the class $\mathcal{L}_{pol}$ of all *functions h of polynomial growth* satisfying $|h(x + y) - h(x)| \leq K|y|(1 + |x|^p + |y|^p)$ for some $K, p \geq 0$ (taking bounded functions would even be sufficient).

Now, testing (8) simultaneously for all pairs f, g is impossible. So we will choose a pair f, g and propose a test of the property (8) for this particular choice. One can then repeat the same test for several pairs, thus "approaching" an independence test. When

$$f(x) = g(x) = x, \tag{9}$$

and because the difference of the two sides of (8) is then the (conditional) covariance of the two jump sizes ΔX_{S_m} and Δc_{S_m}, testing (8) amounts to testing for non-correlation, but other choices for f and g may be as useful as (9).

For testing (8), and even if the variables $V(f,g)_{S_n}$ were observable (which they are not), one needs a large number of them, and also a nice behavior when n tends to infinity. This leads us to introduce the following assumption:

Assumption (B) a) We have convergence in probability: $\frac{1}{n} N_{nt} \xrightarrow{\mathbb{P}} A_t$ as $n \to \infty$ for all t (recall (3) for N_t), where the process A is continuous and $A_1 > 0$.

b) For all functions $f, g \in \mathcal{L}_{pol}$ we have (non-random) numbers $v(f,g)$ such that as $n \to \infty$

$$\frac{1}{n} \sum_{j=1}^{[nt]} V(f,g)_{S_j} \xrightarrow{\mathbb{P}} v(f,g) \, t \quad \text{for all } t. \tag{10}$$

Under (B-a) the jump times S_n are not too "unevenly distributed" through time, and in particular they are all finite. This is a rather mild condition, clearly satisfied when X is a discontinuous Lévy process or, more generally, when X has *jumps of Lévy type*, meaning that the function $\delta(\omega, t, z) = \delta(z)$ depends on z only. In this case the S_n form a Poisson process.

Remark 2 One could weaken (B-a) so as to be $\gamma_n^{-1} N_{nt} \xrightarrow{\mathbb{P}} A_t$ for a sequence $\gamma_n \to \infty$: the forthcoming results still hold, with an appropriate modification of the condition $T_n \Delta_n^{1/2-\zeta} \to 0$ in Theorem 1 by $T_n \Delta_n^{2/3-\zeta} \to 0$, and the extension at the end of Section 5. But it is difficult to imagine realistic models satisfying this with γ_n other than n. □

Assumption (B-b) is a kind of "ergodic" hypothesis for the discrete-time process $V(f,g)_{S_m}$. Assumption (B) is satisfied when the process c is ergodic and X has Lévy jumps. It is also satisfied when the pair (X,c) is ergodic.

Finally, besides being in $\mathcal{L}_{pol}$, the pair of functions f, g is supposed to satisfy:

Assumption (C) a) We have $v(f^2,1) > v(f,1)^2$ and $v(1,g^2) > v(1,g)^2$.
   b) We have $V(f,1)_{S_m} = v(f,1)$ for all $m \geq 1$.

(C-a) is a non-degeneracy condition, basically stating that the variables $f(\Delta X_{S_m})$ and $g(\Delta c_{S_m})$ are not constant (if they were, (8) would be trivially satisfied and there would be nothing to test), and also implying implicitly that there are infinitely many jumps of c occurring together with those of X. As for (C-b), it looks, and is, quite restrictive, and we can only think of three situations in which this can reasonably be true:

(S1) X has jumps of Lévy type: then $V(f,1)_t$ does *not* depend on t and (C-b) holds for all f.

(S2) The distribution of the sizes of X-jumps, conditionally on the past, is symmetrical about 0. In this case, any *odd* function f satisfies (C-b), because then $V(f,1)_{S_m} = 0$ for all m, and also $v(f,1) = 0$. In practice, many models for (discontinuous) prices are assumed to have jumps satisfying this symmetry property.

(S3) Even more models have the weaker property that the jumps of X, conditionally on the past, are centered. In this case the function $f(x) = x$ satisfies (C-b).

At this stage, we can state the statistical hypotheses, for any given (fixed) pair f, g of functions in $\mathcal{L}_{pol}$, under (A, B, C). The null and alternative hypotheses are, respectively,

$$(\mathrm{H}_0): \ V(f,g)_{S_n} = V(f,1)_{S_n} V(1,g)_{S_n} \text{ for all } n, \text{ hence also } v(f,g) = v(f,1)v(1,g)$$
$$(\mathrm{H}_1): \ v(f,g) \neq v(f,1)v(1,g). \tag{11}$$

These two hypotheses are disjoint, but there is a gap between them: it is possible that both fail. (H_0) is genuinely the null in which we are interested (the conditional independence (8)). Note that, within the sub-class of our model for which the two-dimensional discrete-time process $(\Delta X_{S_n}, \Delta c_{S_n})$ is stationary ergodic, we have $v(f,g) = \mathbb{E}[f(\Delta X_{S_n})g(\Delta c_{S_n})]$ for all n, hence (H_1) amounts to the fact that the random variables $f(\Delta X_{S_n})$ and $g(\Delta c_{S_n})$ are (unconditionally) correlated.

3 The test procedure

We start with the observation scheme. For $n \geq 1$ the process X is observed at discrete times $i\Delta_n$ for all integers i between 0 and T_n/Δ_n, where the time span T_n is a multiple of the time lag Δ_n, and below we assume that $T_n \to \infty$ (long span asymptotic) *and* $\Delta_n \to 0$ (high-frequency in-fill asymptotic) as $n \to \infty$. The starting point X_0 is irrelevant for us, so the observation amounts to the returns (or log-returns):

$$\Delta_i^n X = X_{i\Delta_n} - X_{(i-1)\Delta_n}, \tag{12}$$

and later the same notation $\Delta_i^n Y$ is used for the increments of any other process Y.

As explained at the beginning of Chapter 9 of [20], large increments $\Delta_i^n X$ are due to jumps, most usually to a single one (recalling finite activity), whereas small increments reflect the variation of the continuous (Brownian) part of X. Since we are interested in the jumps of X we need to disentangle them from the continuous part, and for this we use decreasing truncation levels $u_n > 0$. To "estimate" the spot volatility we also use time windows of size $k_n\Delta_n$ for an increasing sequence k_n of integers. We must have $u_n \to 0$, but more slowly than $\sqrt{\Delta_n}$, and $k_n \to \infty$, but more slowly than $1/\sqrt{\Delta_n}$. These convergence rates hold, if we choose exponents ϖ and ρ such that for some constant $A \in (1, \infty)$

$$\frac{1}{A} \leq \frac{u_n}{\Delta_n^\varpi} \leq A, \quad \frac{1}{A} \leq k_n \Delta_n^\rho \leq A, \quad \text{with } \varpi, \rho \in \left(0, \frac{1}{2}\right). \tag{13}$$

The choice of u_n and k_n in practice will be discussed at the end of this section.

Based on the truncation method initiated by Mancini [23], the next quantities serve as "local estimators" of the squared spot volatility $c_t = \sigma_t^2$:

$$\widehat{c}_i^n = \frac{1}{k_n\Delta_n} \sum_{j=1}^{k_n} |\Delta_{i+j}^n X|^2 \, 1_{\{|\Delta_{i+j}^n X| \leq u_n\}}, \tag{14}$$

so that $\widehat{c}_{[t/\Delta_n]}^n$ and $\widehat{c}_{[t/\Delta_n]-k_n}^n$ will be used as estimators of c_t and c_{t-}, respectively. Taking only returns smaller than u_n into account allows us to eliminate the jumps of X, which otherwise would introduce a positive bias into the volatility estimator.

We will construct the tests under the assumptions (A,B,C). In view of (7), it is only natural that the following process plays a key role:

$$U(f,g)_t = \sum_{m \geq 1} f(\Delta X_{S_m}) \, g(\Delta c_{S_m}) \, 1_{\{S_m \leq t\}}. \tag{15}$$

It is not observable, but its value at time t is consistently estimated by the following (observable) statistics (below, empty sums are set to 0):

$$U(f,g)_t^n = \sum_{i=k_n+1}^{[t/\Delta_n]-k_n} f(\Delta_i^n X) \, g(\widehat{c}_i^n - \widehat{c}_{i-k_n-1}^n) \, 1_{\{|\Delta_i^n X| > u_n\}}. \tag{16}$$

8

The test statistics itself will be as follows (below, $U(1, g)_t^n$, for example, stands for $U(f, g)_t^n$ when $f(x) = 1$ for all x, and $U(1, 1)_t^n$ counts the number of returns larger than u_n in the estimation window):

$$\Psi_n = \sqrt{U(1, 1)_{T_n}^n} \, \frac{\Upsilon_n}{\sqrt{\Phi_n}},$$

where

$$
\begin{aligned}
\Phi_n &= (U(1, 1)_{T_n}^n)^3 U(f^2, g^2)_{T_n}^n + U(1, 1)_{T_n}^n \, (U(f, 1)_{T_n}^n)^2 \, U(1, g^2)_{T_n}^n \\
&\quad + U(1, 1)_{T_n}^n \, (U(1, g)_{T_n}^n)^2 \, U(f^2, 1)_{T_n}^n + 4 U(1, 1)_{T_n}^n \, U(1, g)_{T_n}^n \, U(f, 1)_{T_n}^n \, U(f, g)_{T_n}^n \\
&\quad - 2 U(1, 1)_{T_n}^n \, U(f, 1)_t \, U(f, g^2)_{T_n}^n - 2 U(1, 1)_{T_n}^n \, U(1, g)_{T_n}^n \, U(f^2, g)_{T_n}^n \\
&\quad - 3 (U(f, 1)_{T_n}^n)^2 \, (U(1, g)_{T_n}^n)^2 \\
\Upsilon_n &= U(1, 1)_{T_n}^n \, U(f, g)_{T_n}^n - U(f, 1)_{T_n}^n \, U(1, g)_{T_n}^n.
\end{aligned}
\tag{17}
$$

In words, Ψ_n is the Studentized version of Υ_n, which in turn is the empirical correlation between proxies of the jumps of X (replaced by the observed returns) and those of c (replaced by the increments of the spot volatility estimators, over non-overlapping adjacent windows). Let us also mention that the functionals (16) are of the same type as those introduced for example in [19]. However, the conditions on f and g are different, and also $T_n \to \infty$, so that the next result needs a new proof.

Before describing the tests, we give the asymptotic behavior of these statistics (recall that $\xrightarrow{\mathbb{P}}$ denotes convergence in probability, and $\xrightarrow{\mathcal{L}}$ convergence in law):

Theorem 1 *Assume (A,B,C) for the stochastic volatility model* (1) *and the test functions f and g, and let $T_n \to \infty$ and $\Delta_n \to 0$ be such that $T_n \Delta_n^{1/2-\eta} \to 0$ for some $\eta \in (0, \frac{1}{2})$. We also choose u_n and k_n as in* (13)*, with $\rho \in \left[\frac{1}{2} - \eta, \frac{1}{2}\right)$. Then, as $n \to \infty$,*

$$
\begin{aligned}
\text{under } (\mathrm{H}_0): \quad \Psi_n &\xrightarrow{\mathcal{L}} \mathcal{N}(0, 1) \\
\text{under } (\mathrm{H}_1): \quad |\Psi_n| &\xrightarrow{\mathbb{P}} \infty.
\end{aligned}
\tag{18}
$$

This theorem readily provides a way to construct a test with a given *asymptotic level* $\alpha \in (0, 1)$ for the null hypothesis (H_0). We consider the symmetric α-quantile y_α of the standard normal law, such that $\mathbb{P}(|U| > y_\alpha) = \alpha$ when U is $\mathcal{N}(0, 1)$. The critical regions are

$$\mathcal{C}_n = \left\{ |\Psi_n| > y_\alpha \right\}. \tag{19}$$

In our non-parametric setting, the asymptotic size of the tests above is defined

$$\sup \left\{ \limsup_{n \to \infty} \mathbb{P}(\mathcal{C}_n) : \mathbb{P} \text{ satisfies } (\mathrm{H}_0) \right\},$$

and this family of critical regions is called *consistent for the alternative*, if $\mathbb{P}(\mathcal{C}_n) \to 1$ whenever $\mathbb{P}$ satisfies (H_1).

Theorem 2 *Under the assumptions of Theorem 1 the critical regions $\mathcal{C}_n$ have the asymptotic size α for testing the null hypothesis (H_0) and are consistent for the alternative (H_1).*

Practical considerations. Apart from the natural choice (9), which allows us to test for non-correlation of the jump sizes of X and c, the choice of other pairs f, g, for which to conduct the test, really depends on the type of model one wants to use or which one believes to be more or less realistic, and also on how far one wants to go towards testing "complete" conditional independence. In our empirical studies we use a few pairs, but other choices are of course possible.

The procedure also requires to choose the truncation levels u_n and the window k_n. For u_n, the aim is to eliminate the Brownian parts in (16), and to eliminate the jumps and keep most of the Brownian parts in (14). This type of truncation, which originates in [23], is quite often used in similar context, and is for example discussed in [19], or in [2] for u_n only. Taking the threshold u_n fixed over the whole time period has the effect that we investigate more jumps during periods with high volatility than in periods with low volatility. Therefore we will use an adaptive threshold, computed locally in time on the basis of a few days. As for k_n, saying that (13) holds with ρ almost equal to $1/2$ does not say much. We can take $k_n = [k^* \Delta_n^{-.49}]$ as suggested in [19] with $k^* = 4$. However, taking k_n to be between 50 and 100 for a "typical" number and span of observations seems to be a reasonable choice according to our Monte Carlo study, which mimics the empirical analysis where 1 minute data is used.

One may also wonder whether $T_n \Delta_n^{1/2}$ is really "small enough" to suppose that we are in the asymptotic regime. As usual in this context, there is no reliable rule about this, apart from conducting the test for various values of Δ_n (that is, doing sub-sampling with various ranges) and check for significant differences.

Finally, our test completely disregards microstructure noise, and is clearly *not* robust against it. Therefore, it should be used with care, when very high-frequency data are employed. This depends, of course, on the nature of the data: for stock prices, sampling more frequently than 1 minute for the most liquid ones and 5 minutes for the others is probably unreasonable; for exchange rates or indices, frequency might be higher. It would indeed be possible to take care of microstructure noise by using a "pre-averaged" version of log-prices and spot volatility estimators; however, this requires another tuning parameter (the length of the pre-averaging window), and matters become even more complicated, both theoretically and practically, and less transparent.

Remark 3 The above procedure is designed to test for correlation between two functions $f(\Delta X_t)$ and $g(\Delta c_t)$ for a co-jump at time t, under the conditional law knowing the past before t. One might prefer to test for correlation between $f(\Delta X_t)$ and $g(\Delta h(c)_t)$ for some function h, such as $h(x) = \sqrt{x}$ (so $\Delta h(c)_t$ would be the jump $\Delta \sigma_t$ of the volatility and not the squared volatility), or $h(x) = \log x$ (so $\Delta h(c)_t$ would be the jump of the log-volatility).

As indicated in Remark 1(ii), this is possible, under the additional assumption that the process c_t satisfies $\inf_s c_s > 0$ (a rather weak assumption, satisfied by virtually all models for stochastic volatility). The jump $\Delta h(c)_t$ would then be estimated by the quantity $h(\widehat{c}_i^n) - h(\widehat{c}_{i-k_n}^n)$, where $\widehat{c}_i^n = \widehat{c}_i^n \vee \rho_n$ for some appropriately chosen sequence ρ_n of positive numbers, decreasing sufficiently slow to 0. With this modification, the previous procedure works in exactly the same way.

Below, we do not implement these slightly different tests explicitly. If we are interested in whether there is independence or not, this property holds for the pair $(\Delta X_t, \Delta c_t)$ if and only if it holds for the pair $(\Delta X_t, \Delta h(c)_t)$, as soon as h is continuous and strictly monotonic. $\qquad\square$

4 Extension to the infinite activity case

Assumption (A) is quite weak, *except* for the requirement that jumps of X have finite activity. Relaxing this amounts to replace (A) by the following, for some number $r \in [0, 2]$:

Assumption (A-r) We have (A), except that the integrability condition on Γ is replaced by $\int_E (\Gamma(z)^r \vee \Gamma(z)^p) \, \lambda(dz) < \infty$, for all $p \geq 2$.

The number r is an upper bound for the activity index, or *Blumenthal-Getoor index*, of the jumps of X, and imposing (A-0) is the same as adopting (A).

Testing for the independence of the infinitely many very small jumps of X with those of c seems out of practical reach. We can however fix a level $\varepsilon > 0$ and decide that only jumps of X with absolute size bigger than ε are tested. If desired, we can repeat the same test for a decreasing (finite) sequence $\varepsilon_1 > \varepsilon_2 > \cdots > \varepsilon_m$. One should also make sure that X has a.s. no jump such that $|\Delta X_t| = \varepsilon$; this is true for all $\varepsilon > 0$ except countably many values, and also for all $\varepsilon > 0$ as soon as the Lévy measure of X has a density.

In this situation we have to replace (3) and (6) by

$$N_t = \sum_{s \leq t} 1_{\{|\Delta X_s| > \varepsilon\}}; \quad S_1, S_2, \cdots \text{ are the jump times of } X \text{ with size bigger than } \varepsilon,$$

$$V(f, g)_t = \frac{\int_E 1_{\{|\delta(t,z)| > \varepsilon\}} f(\delta(t,z)) \, g(\widetilde{\delta}(t,z)) \, \lambda(dz)}{\int_E 1_{\{|\delta(t,z)| > \varepsilon\}} \, \lambda(dz)}.$$

Assumptions (B) and (C) are not formally modified, although of course N_t, hence A_t, and also all $v(f, g)$, depend implicitly on the value ε. In particular (C-b) is satisfied in the situations (S1) (for all f) and (S2) (for all odd f) described just after Assumption (C), but not in situation (S3) for $f(x) = x$, in general.

The two statistical hypotheses (H_0) and (H_1) are the same as in (11), and again implicitly depend on ε. The testing procedure works in exactly the same way, provided we replace (16) by

$$U(f, g)_t^n = \sum_{i=k_n+1}^{[t/\Delta_n]-k_n} f(\Delta_i^n X) \, g(\widehat{c}_i^n - \widehat{c}_{i-k_n-1}^n) \, 1_{\{\{|\Delta_i^n X| > \varepsilon\}}.$$

However, due to the presence of many small jumps, the spot volatility estimators (14) are not as good as in the finite activity case, and the quality deteriorates as r increases. This results in more stringent assumptions in Theorems 1 and 2. Namely, instead of $T_n \Delta_n^{1/2-\eta} \to 0$ for some $\eta \in (0, \frac{1}{2})$, and when e.g. $r \geq \frac{4}{3}$, we need $T_n \Delta_n^{\frac{2-r}{r}-\eta} \to 0$. The proofs, though, are much more involved, and we will not pursue this topic here.

5   Monte Carlo results

To assess the behavior of the test in a realistic scenario, we conduct a Monte Carlo study using a two-factor stochastic volatility model with diurnal effect and mid-range changing volatility level. We consider 16 simulation scenarios to investigate the sensitivity of the test on the jump frequency, on the length of the data set, and on the threshold u_n and the length k_n of the window for the volatility estimation, cf. Eq. (14) and Eq. (16).

5.1   Two-factor stochastic volatility model

For our experiments we use the following model:

$$
\begin{aligned}
dX_t &= \sqrt{V_t + ae^{Y_t}}\sqrt{D_t}\sqrt{M_t}dW_t + dJ_t^X &(20)\\
dV_t &= \kappa_1(\theta - V_t)dt + \tau\sqrt{V_t}dB_t &(21)\\
dY_t &= -\kappa_2 Y_t dt + dH_t + dJ_t^Y &(22)\\
D_t &= 4.8((t - \lfloor t \rfloor) - 0.5)^2 + 0.6 &(23)
\end{aligned}
$$

Note that a similar two-factor model has been used in [19], and regarding the choice of the parameters we closely follow the setup given there. The process X is the log-price process, and the processes V and Y are two stochastic components of the volatility process. Since $D_t = D_{t+1}$ for all t and we identify the time unit with one day, D represents a deterministic diurnal effect on the volatility. Note that the process D mimics quite accurately the diurnal volatility pattern found in [25] for the SPY equity (cf. Figure 2 in [25]) between 9:30 am and 4:00 pm. The process M is not modeled explicitly, but is the 20-day moving average volatility level (standardized to an average level of 1 over the full period), which we observe in the empirical analysis in Section 6. This way we model realistic volatility fluctuations over long periods. The four volatility components V, Y, D and M are linked to the volatility process σ, given by

$$
\sigma_t = \sqrt{V_t + ae^{Y_t}}\sqrt{D_t}\sqrt{M_t}, \tag{24}
$$

in Eq. (20). The processes W and B are assumed to be independent standard Brownian motions, driving the processes X and V. We conduct a similar Monte Carlo study with leverage effect (i.e. with W and B being correlated Brownian motions) in the appendix [17]. The *common* jumps of the log-price process and the squared volatility process are generated by the bivariate compound Poisson process (J^X, J^Y) with intensity λ_J. Note that σ can jump upwards and downwards, if the jump distribution of J^Y allows for positive and negative jumps. The process H is assumed to be a univariate compound Poisson process with intensity λ_H generating jumps occurring *only* in the volatility process.

5.2   Simulation scenarios

We consider 16 different simulation scenarios, which all have some parameter setting in common: $\theta = 0.4$, $\kappa_1 = 0.02$, $\kappa_2 = 0.5$, $\tau = 0.04$, $a = 0.6$, and $\lambda_H = 0.5$, $\Delta_n^{-1} = 390$.

Table 1: The 16 simulation scenarios, differing in the common jump intensity λ, in the number T of days in the observation period, and in the threshold parameters k^* (3 and 5) and u^* (3.89 and 3.72).

Scenario	λ_J	T	k^*	u^*
A.1250.k^*.a	0.68	1250		3.89
B.1250.k^*.a	1.36	1250		3.89
A.2500.k^*.a	0.68	2500	3 / 5	3.89
B.2500.k^*.a	1.36	2500		3.89
A.1250.k^*.b	0.68	1250		3.72
B.1250.k^*.b	1.36	1250		3.72
A.2500.k^*.b	0.68	2500	3 / 5	3.72
B.2500.k^*.b	1.36	2500		3.72

We further assume that the jump distribution of H is uniform on $(-0.4, 0.4)$, so that H generates additional small jumps in the volatility without simultaneous jumps in X. The number of observations per day, Δ_n^{-1}, corresponds to our empirical analysis in Section 6. The parameters fit high-frequency financial data very well, cf. the choice of the parameters in the two-factor model in [19]. Note, however, that in contrast to the model in [19] we additionally employ the diurnal volatility pattern D, the 20-day moving average volatility level M, and the process H, which all lead to higher volatility fluctuations (intraday and over longer periods). Moreover, we use the exponential function in (20) in combination with possibly negative values in Y to handle both upward and downward volatility jumps generated by J^Y. Having in mind that V is mean-reverting to θ, Y is mean reverting to 0, the parameter $a = 0.6$, and that the processes D and M have mean level 1, the volatility level for the Brownian part in (20) varies around 1.

The scenarios differ in the (common) jump intensity λ_J, in the length of the observation period $[0, T]$, and in the constants u^* and k^* which are used as factors determining the threshold u_n and the length k_n. The parameter settings can be found in Table 1.

For the 16 scenarios, we simulate 1000 data sets each, and take $T \in \{1250, 2500\}$, which corresponds to 5 and 10 years with 250 trading days each. To assess the sensitivity of the test on the threshold u_n and the length k_n of the window for the volatility estimation (cf. Eq. (14) and Eq. (16)), we apply the test always using subsequently two different values of k^* and two different values of u^*, cf. Table 1. More precisely, as in Section 6 we set $u_n = u^*\sqrt{\Delta_n}\sqrt{M_t^* D_t}$, where $\sqrt{M_t^*}$ is the 20-day average volatility level (estimated from the simulated data set at hand), and D_n is the diurnal effect from (23). Finally, $k_n = [k^*\Delta_n^{-0.49}]$.

One main goal of this simulation study is to learn about the sensitivity of the test statistics Ψ_n in Theorem 1 on the level of correlation between the log-price jumps and the squared volatility jumps. To this end, we choose the joint jump distribution of (J^X, J^Y)

to be a truncated bivariate normal distribution with correlation ρ_{sim}; more precisely,

$$\begin{pmatrix} \Delta J^X \\ \Delta J^Y \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0.5 & \rho_{\text{sim}}\sqrt{0.9} \\ \rho_{\text{sim}}\sqrt{0.9} & 1.8 \end{pmatrix} \right). \tag{25}$$

For comparison, [19] employ a uniform log-price jump distribution in $[-1.0420, -0.1] \cup [0.1, 1.0420]$ (for a jump intensity of 0.5), and a uniform volatility jump distribution in $[0.04, 0.76]$. Hence, our choice leads to similar price jumps, but to volatility jumps which can be larger in magnitude and both upwards and downwards. Furthermore, using this non-truncated bivariate normal distribution we produce effectively not only significant common jumps, but also significant price jumps paired with invisibly small volatility jumps and vice versa (and some common jumps which can hardy be detected, neither in the price nor in the volatility). Finally, our choice is in line with two criteria which are important for our purpose: First, as we will see later, using these parameters the (random) number of jumps taken into consideration by our test corresponds (at least in certain setups) quite well to the empirical analysis in Section 6 (there, in the main analysis 330 jumps were taken into consideration). Second, the fraction of maximal and minimal squared volatility over the whole time period should be at least of the magnitude of the empirical analysis, where this fraction is about 63.

Due to the model structure the correlation parameter ρ_{sim} of (25) does not represent the real correlation between the jumps in X and the jumps in c. After simulating a data set, we therefore always compute the empirical correlation between the jumps in X and the jumps in c. We denote this empirical correlation simply by ρ. Over the 1000 data sets in a simulation scenario we vary the correlation parameter ρ_{sim} uniformly over the interval $[-0.6, 0.6]$; then we compare the resulting test statistics Ψ_n to the corresponding empirical correlation ρ.

Always the test is performed for $f = g = \text{id}$ (where id denotes the identity function $\text{id}(x) = x$), as in Eq. (9). For details on the specific selection of jumps (concerning the selection of volatility jumps, handling of overnight data frictions, etc., we refer to Section 6, where we proceed completely analogously.

5.3   Results

The 16 plots in Figure 1 show the values of the test statistics Ψ_n against the empirical correlation ρ computed from the corresponding simulated data sets. Information about the number of jumps analyzed by the test and about the fraction of maximal and minimal squared volatility can be found in the appendix [17].

The similarity of the plots to data from linear models is obvious. Moreover, the slope of the regression lines varies over the simulation scenarios. Of course, in order to have a test which is able to detect also small correlations between log-price and squared volatility jumps, the (absolute) values of the test statistics should be large as soon as the empirical correlation ρ moves away from 0. Hence, a larger slope indicates that the test has more power in detecting small correlations.
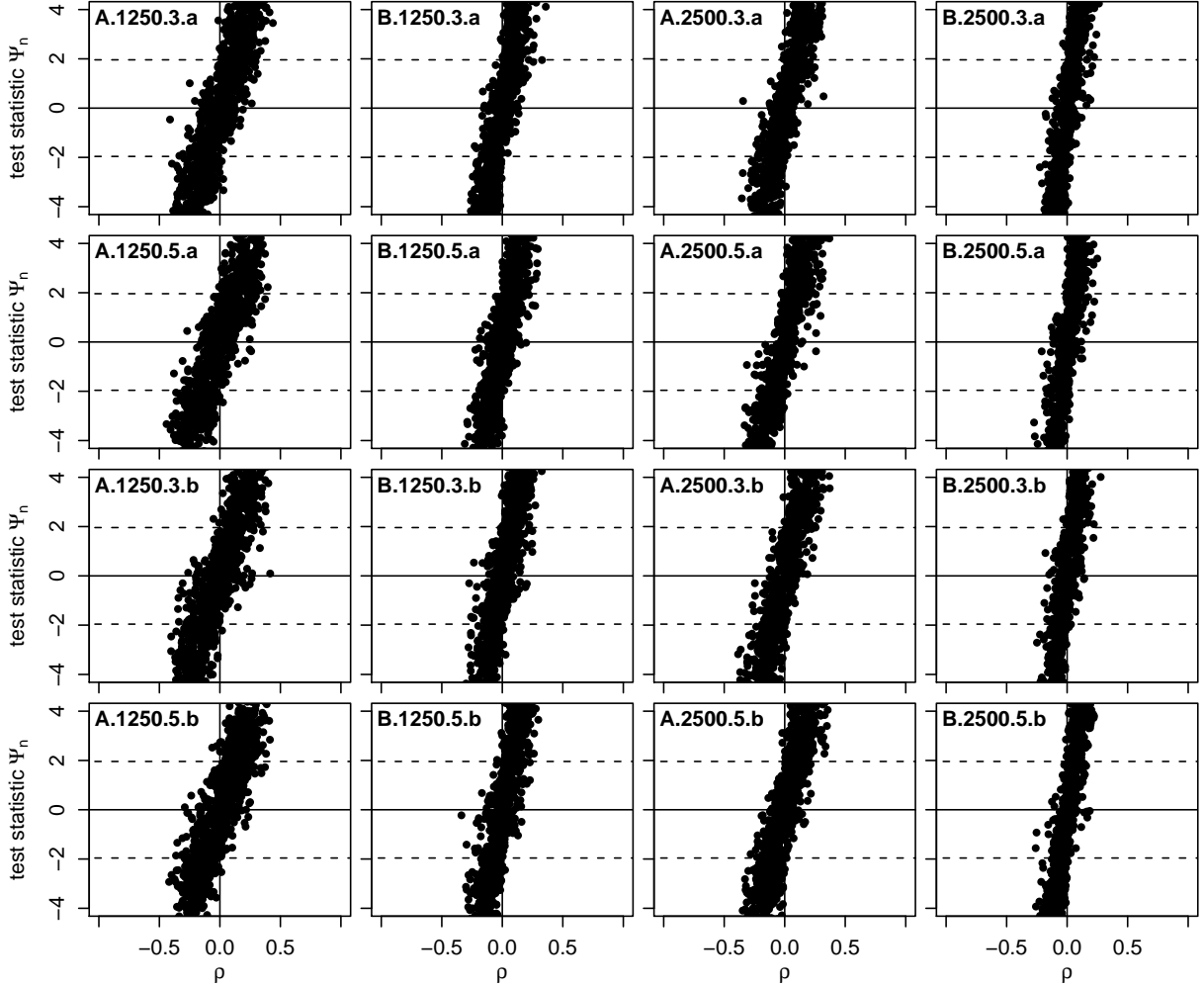
Figure 1: Monte Carlo results for the 16 simulation scenarios of Table 1. The plots show the values of the test statistics Ψ_n against the empirical correlation ρ computed from the corresponding simulated data sets. The null *no correlation* is rejected on the level α if $|\Psi_n| > y_\alpha$, cf. Eq. (19). For illustration, the horizontal dashed lines indicate the 2.5% and 97.5% quantiles of the standard normal distribution, used for a test level of 5%.

To make this precise, we conduct regression analysis for all 16 plots of Figure 1; details on the results can be found in [17]. As Figure 1 suggests, scenario B2500.3.b has the highest power. Moreover, there is a strong connection between the number of investigated price jumps and the value of the slope: the more jumps are investigated, the larger the slope is. Obviously, a higher *jump intensity* as well as a longer *observation interval* increases the number of jumps (using the same parameter values otherwise, in particular u^*). Here it is useful to compute from Table 1 that the total expected number of common jumps is 850 in scenarios A.1250, 1700 in scenarios B1250 and A.2500, and 3400 in scenarios B.2500 (note, however, that usually less jumps are investigated by the test, cf. the histograms in [17]). As to the *influence of the choice of u^** we do not observe
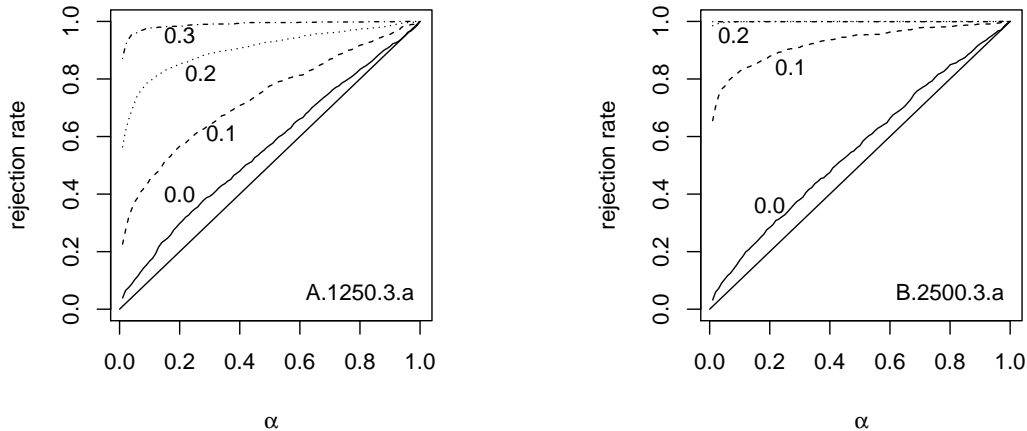
Figure 2: Rejection rates against test level α, for scenarios A.1250.3.a and B.2500.3.a, and for $\rho^* \in \{0.0, 0.1, 0.2, 0.3\}$. Each curve is based on 1000 data sets.

significant differences in the slopes for $u^* = 3.89$ and $u^* = 3.72$. Of course, taking a higher value for u^* protects better against a misspecification of jumps due to larger movements of the Brownian motion. The *influence of the choice of k^** on the slope is marginal, although in our simulations we get slightly better results for $k^* = 3$.

Next we investigate rejection rates in the different scenarios. Figure 2 exemplarily shows rejection rates against test levels $\alpha \in (0, 1)$ for scenarios A.1250.3.a and B.2500.3.a. Here we pool data sets with similar empirical correlation ρ, e.g. $\rho \in (\rho^* - 0.01, \rho^* + 0.01)$ as having correlation *approximately* ρ^* for $\rho^* \in \{0.0, 0.1, 0.2\}$. To reduce the Monte Carlo error, we conducted additional simulations for these two scenarios, so that each curve in Figure 2 is based on 1000 data sets. Due to the smaller number of large jumps, the test has only small power in scenario A.1250.3.a to distinguish between *no correlation* and a *correlation of about 0.1*, whereas it has the power to distinguish very precisely between *no correlation* and a *correlation of about 0.3*. In scenario B.2500.3.a, where more jumps are available, the test has more power and is able to detect very precisely correlations of 0.2. Moreover, we learn from Figure 2, that the test is slightly over-rejecting in general, but this effect is only marginal for the usual levels of 5% of 10%.

Summarizing our results, we conclude that the more jumps are considered, the more power the test has in order to distinguish between *no correlation* and *correlation*; this, of course, coincides with our expectations from asymptotics in Theorem 1, where $T_n \to \infty$ and $\Delta_n \to 0$. The sensitivity on both k^* and u^* is quite weak.

6   Empirical results

As an application we consider trade data from the SPDR S&P 500 ETF (SPY), an exchange traded fund that tracks the S&P 500 index. Our data comes from *TickData* and contains all trades of SPY between January 3, 2005 and December 30, 2011, at 15

different U.S. exchanges. However, we only retained trades from NASDAQ, where SPY was traded more often than at all other 14 exchanges.

Before calculating returns we applied a procedure for cleaning high-frequency data (cf., e.g., Section 3 in [8]). In particular, we used median prices if multiple transactions with identical time stamp occurred. Since SPY is the most liquid equity in U.S. markets with a very small spread, market microstructure noise is very modest (cf. Section 6 in [25]), and we can frequently sample from the price process. Since we calculated returns after every 60 seconds for the period 9:30 am to 4:00 pm, we have 390 observations on a full day, hence we set $\Delta_n = 1/390$. Moreover, we deleted days, where periods of more than 60 consecutive seconds without trades occur. Applying this rule, we retain 1266 days in the data set, corresponding to 493 740 one minute observations.

According to the recommendations of Section 3 and the conclusions from our Monte Carlo experiments, we choose $k_n = 56$. To account for changing volatility levels - intraday as well as over the period of 7 years - we use a time-varying threshold $u_{n,i}$. To this end we first estimate the current volatility level by computing a classical volatility estimate $v_{n,i}$ based on a time window of 3900 returns before and 3899 returns after observation i (i.e. the estimate $v_{n,i}$ is based on 7800 observations). Second we multiply this estimate by the diurnal volatility pattern given in (23); recall that this pattern was found in [25] for the SPY equity. Finally we multiply by $u^* = 3.89$ (the 99.995% quantile of the standard normal), which gives us $u_{n,i}$.

After specifying the thresholds $u_{n,i}$, we get a preliminary set of jump candidates by selecting those i for which $|\Delta X_i| > u_{n,i}$. Recall that for the volatility jump estimation we use returns within k_n minutes before and k_n minutes after the corresponding price jump. In consequence, due to the overnight discontinuities in the data, the estimation of the volatility jumps may not be accurate in the first k_n minutes of the trading day as well as in the last k_n minutes before closing. Accordingly, we remove all jumps before 10:26 am and after 3:04 pm from the preliminary set of jump candidates. Similarly, also clusters of price jumps may deteriorate the quality of the volatility jump estimation. Hence, we investigate only isolated common jumps; i.e. we remove all jumps from the current set of jump candidates where other price jumps occur within k_n minutes before or after.

We improve the quality of the squared volatility estimators further by adapting them to the intraday volatility pattern. To this end, we replace the estimator from (14) by a weighted mean,

$$\frac{\sum_k 1_{\{|\Delta_{i+k}^n X| \leq u_n\}}}{k_n \Delta_n \sum_k [D_{i+k} 1_{\{|\Delta_{i+k}^n X| \leq u_n\}}]} \sum_{j=1}^{k_n} \frac{1}{D_{i+j}} |\Delta_{i+j}^n X|^2 1_{\{|\Delta_{i+j}^n X| \leq u_n\}}, \tag{26}$$

where D_i is the value of the diurnal component for observation i according to (23).

Finally we use additional thresholds to identify volatility jumps occurring simultaneously with price jumps, in order to ensure that only relevant jumps are retained. In the end we only retain those price jumps, where the corresponding volatility jump is at least 10% upwards or at least 9% downwards. This way, we finally investigate 330 common jumps in total.
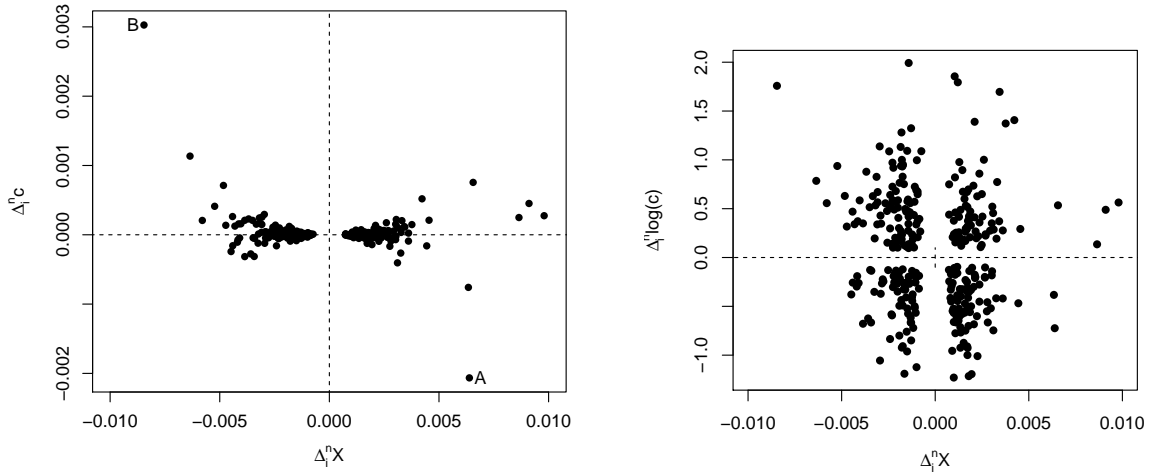
Figure 3: Extreme returns $\Delta_i^n X$ (on the horizontal axis) together with corresponding estimated jumps $\Delta_i^n c$ (left) and estimated jumps $\Delta_i^n \log(c)$ (right), on the vertical axis.

These common jumps are shown in Figure 3 with log-price jumps on the horizontal axis and corresponding estimated jumps $\Delta_i^n c := \hat{c}_i^n - \hat{c}_{i-k_n-1}^n$ (left plot) and estimated jumps $\Delta_i^n \log(c) := \log(\hat{c}_i^n) - \log(\hat{c}_{i-k_n-1}^n)$ (right plot), respectively, on the vertical axis. The two most extreme variance jumps (which are marked in Figure 3 by A and B) were observed on October 10, 2008, at 12:43 pm and on October 29, 2008, 2:19 pm, respectively. Note from Figure 3 that we estimate both upward and downward volatility jumps; moreover, the empirical correlation between the estimated variance jumps and the log-price jumps is -0.23. This indicates that the correlation between log-price and variance jumps is not as strong as the overall correlation between returns and variance increments, where values of up to around -0.6 are reported (e.g., [4]). Our result for the common jumps is in line with [7], who find a smaller leverage of -0.11 for the co-jumps (of log-price and volatility σ instead of σ^2) than in total (-0.37) for a similar data set. Removing the points A and B in our data set decreases the empirical correlation even to -0.06. From the right-hand side of Figure 3 one does not detect any correlation between the jumps of the log-price and the log-variance at all, in the sense of *linear dependence*. However, given the plot on the left-hand side of Figure 3) log-price and squared volatility jumps are nevertheless *not independent*.

For our data set, values of the volatility estimates of up to 0.1002 are observed for the last minutes of October 9, 2008, and the minimal value is 0.0016 for April 16, 2007, at 12:45 pm. These dates correspond also to the largest and smallest values of the time-varying threshold $u_{n,i}$: here, the smoothest period was around April 20, 2007, and the most turbulent period around October 17, 2008.

Table 2 shows the values of the test statistics Ψ_n for the SPY data set for three different choices of the functions f and g, cf. Eq. (6) and thereafter. We see that one of the nine test statistics has a value greater than 1.96: for $f(x) = x^2$ and $g(x) = |x|$ the test statistic Ψ_n is 2.0007. Accordingly, we reject the null hypothesis of *no correlation* between $(\Delta X)^2$ and $|\Delta c|$ on a 5% level. On a 10% level the null hypothesis is also rejected for $|\Delta X|$ versus

Table 2: Values of the test statistics Ψ_n for the SPY data set, for different choices of f and g, and for testing $f(\Delta X)$ vs. $g(\Delta c)$. The test statistics are asymptotically standard normal and, hence, are compared to the corresponding quantiles. For large values (positive or negative) the null (non-correlation) is rejected.

	$f(\Delta X)$ vs. $g(\Delta c)$		
	$f(x) = x$	$f(x) = |x|$	$f(x) = x^2$
$g(x) = x$	-1.4213	0.8859	1.1421
$g(x) = |x|$	-0.3561	1.7766	2.0007
$g(x) = x^2$	-0.6019	1.0949	1.2920

$|\Delta c|$.

At 5% level the rejection is not very substantial; however, as seen in the appendix [17], for the Intel and Microsoft data we also reject under the same circumstances. These decisions should be reasonably reliable since we have learnt from our simulation study that, even if only about 330 jumps are investigated, the error of *wrongly rejecting the null* is quite well controlled by the chosen level of the test: for small levels (e.g. 5%) the test only very slightly overrejects.

The appendix [17] contains a sensitivity analysis, where we vary the constants u^* and k^* as well as the length and the position of the window used for the moving average estimation of the (mid-range) volatility level M. We always come to the same conclusion, at least at a 10% level.

7 Conclusions

For a stochastic volatility model given as a bivariate Itô semimartingale (X, σ^2) (the log-price and squared volatility) observed on a discrete grid with grid size Δ_n we have proposed a test for correlation of the common jumps or functions of common jumps with prominent examples being the identity, the modulus or the squares of the jumps. We have also indicated the necessary modifications to test for correlation, for common jumps of $(X, h(\sigma^2))$ for functions h like squareroot or logarithm. We concentrate on large price jumps (larger than a threshold u_n) and estimate a possible jump of the spot volatility at the time of such large price jumps. The estimation of the latent spot volatility is of course crucial and we use a window of size $k_n \Delta_n$.

For a two-factor stochastic volatility model with diurnal effect we have investigated the small sample properties of the test in a simulation study. In the paper we present 16 different simulation scenarios taken from the fit of high-frequency data investigating the influence of the (common) jump intensity, the length of the observation period, the choice of the threshold u_n for the size of the log-price jumps, and the sensitivity of the procedure with respect to the choice of k_n. Some details are provided in the appendix [17].

A plot of the values of the test statistics versus the correlation reveals that the procedure works well in the Monte-Carlo study; i.e. it rejects H_0 : *the jumps in squared*

volatility and log-price process are uncorrelated: for a sufficiently large correlation. Not surprisingly, more jumps (larger jump intensity or larger observation interval) lead to a better test performance. As for the influence of the choice of the threshold u_n and of the choice of k_n for the window size, it turns out that the test is not very sensitive on these parameters. However, choosing u_n relatively small can be dangerous, since in this case big increments of the Brownian part could wrongly be identified as jumps.

As an application we consider 1-minute returns of the SPDR S&P 500 ETF (SPY) from NASDAQ between January 3, 2005 and December 30, 2011. After cleaning the data according to an established procedure we have finally $493\,740$ data points. Our test procedure is then based on 330 price jumps. We consider 9 tests based on jumps in log-price and squared volatility taking simply their values, the jump moduli and the squared jumps. On a 5% level, the test does *not* reject the non-correlation hypothesis for the log-price and squared volatility co-jumps, however it rejects the non-correlation hypothesis for the combination of squared log-price jumps and the moduli of the jumps of the squared volatility; i.e., we can conclude that $(\Delta X_t)^2$ and $|\Delta\sigma_t^2|$ are correlated, and thus the two components of a co-jump are indeed dependent. The same empirical findings holds for the stock prices analyzed in the appendix [17].

8 The proofs

We shall need the notion of stable convergence in law. This kind of convergence is slightly stronger than convergence in law. For details and extensions see [20], Section 2.2.1. A formal definition is as follows: let V_n be random variables on $(\Omega, \mathcal{F}, \mathbb{P})$, and V be a variable defined on an extension $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}})$ of this space, that is $\widetilde{\Omega} = \Omega \times \Omega'$ and $\widetilde{\mathcal{F}} = \mathcal{F} \otimes \mathcal{F}'$ and the Ω-marginal of the measure $\widetilde{\mathbb{P}}$ on $(\widetilde{\Omega}, \widetilde{\mathcal{F}})$ is $\mathbb{P}$. Then V_n converges to V stably in law $(V_n \overset{\mathcal{L}-(s)}{\longrightarrow} V)$ if $\mathbb{E}[Y\phi(Z_n)] \to \widetilde{\mathbb{E}}[Y\phi(Z)]$ as $n \to \infty$ for all bounded $\mathcal{F}$-measurable random variables Y and all bounded Lipschitz functions ϕ.

8.1 Some auxiliary processes

We start with two lemmas which are about some processes related to the jumps of X and c, but not to the actual observations. The pair f, g of functions of polynomial growth $(f, g \in \mathcal{L}_{pol})$ is fixed. Throughout, K is a positive constant which changes from line to line, and it is written as K_p if it depends on an additional parameter p. We also need the following expression

$$\begin{aligned} a(f,g) \;=\;\; & v(f^2, g^2) + v(f,1)^2\, v(1,g^2) + v(1,g)^2\, v(f^2,1) + 4v(f,1)\, v(1,g)\, v(f,g) \\ & -2v(f,1)\, v(f,g^2) - 2v(1,g)\, v(f^2,g) - 3v(f,1)^2\, v(1,g)^2. \end{aligned} \tag{27}$$

An important property is that $a(f,g) > 0$ when f, g satisfies (C). To see this, we observe that the family $v(f,g)$ when f, g vary can be interpreted as $v(f,g) = \mathbb{E}[f(Y)g(Y')]$ for a pair of random variables (Y, Y'), and then $a(f,g) = \mathbb{E}\big[(f(Y) - \mathbb{E}(f(Y)))^2\, (g(Y') - $

$\mathbb{E}(g(Y')))^2\big]$, whereas (C) implies $\mathbb{E}\big[(f(Y) - \mathbb{E}(f(Y)))^2\big] > 0$ and $\mathbb{E}\big[(f(Y') - \mathbb{E}(f(Y')))^2\big] > 0$, hence $a(f, g) > 0$

Lemma 1 *(a) Under (B) we have as $n \to \infty$,*

$$D^n(f, g)_t := \frac{1}{n} \sum_{i=1}^{[nt]} f(\Delta X_{S_i}) \, g(\Delta c_{S_i}) \quad \xrightarrow{u.c.p.} \quad v(f, g) \, t \qquad (28)$$

(convergence in probability, uniform in t over each compact interval).
  (b) Under (B,C), and the null hypothesis (H$_0$) relative to f, g, the processes

$$
\begin{aligned}
Z^n_t &= \sqrt{n} \left(D^n(f, g)_t - v(f, 1) \, D^n(1, g)_t - v(1, g) \, D^n(f, 1)_t + v(f, 1) v(1, g) \, t \right) \\
\overline{Z}^n_t &= \sqrt{n} \left(D^n(f, 1)_t - v(f, 1) t \right)
\end{aligned}
$$

converge stably in law to $\sqrt{a(f, g)}\, Z$ and $\sqrt{v(f^2, 1) - v(f, 1)^2}\, Z$ as $n \to \infty$, respectively, where Z is a standard Brownian motion defined on an extension of the space $(\Omega, \mathcal{F}, \mathbb{P})$ and independent of $\mathcal{F}$.

There is also joint convergence of Z^n and $\overline{Z}^n$ towards two correlated Brownian motions, but we do not need this later on.

Proof. (a) For $i \geq 1$ set $\zeta(f, g)_i := f(\Delta X_{S_i}) \, g(\Delta c_{S_i})$ and $\mathcal{G}_i := \mathcal{F}_{S_i-}$. By (7) and (10), we have

$$\frac{1}{n} \sum_{i=1}^{[nt]} \mathbb{E}[\zeta(f, g)_i \mid \mathcal{G}_{i-1}] \quad \xrightarrow{\mathbb{P}} \quad v(f, g) \, t \quad \text{for all } t. \qquad (29)$$

Since the limit is continuous, this convergence is locally uniform in t, when $f, g \geq 0$. Upon decomposing f and g as differences of positive functions, we deduce that the convergence in (29) is locally uniform in t for all $f, g \in \mathcal{L}_{pol}$. Moreover, $\zeta(f, g)_i$ is $\mathcal{G}_i$-measurable, so by well known results on triangular arrays, (28) will follow if we prove that for all t,

$$\frac{1}{n^2} \sum_{i=1}^{[nt]} \mathbb{E}[\zeta(f, g)_i^2 \mid \mathcal{G}_{i-1}] \quad \xrightarrow{\mathbb{P}} \quad 0.$$

This follows from (29) applied with the functions f^2 and g^2.

  (b) The proof is the same for Z^n and $\overline{Z}^n$, so we only consider Z^n_t, which equals $\sum_{i=1}^{[nt]} \xi^n_i + v(f, 1)v(1, g)\sqrt{n}\left(\frac{[nt]}{n} - t\right)$, where

$$\xi^n_i = \frac{1}{\sqrt{n}} \left(f(\Delta X_{S_i}) g(\Delta c_{S_i}) - v(f, 1) g(\Delta c_{S_i}) - v(1, g) f(\Delta X_{S_i}) + v(f, 1)v(1, g) \right).$$

In view of (C) and (H$_0$) one easily checks that $\mathbb{E}[\xi^n_i \mid \mathcal{G}_{i-1}] = 0$. Then, since $\sqrt{n}\left(\frac{[nt]}{n} - t\right)$ goes to 0 locally uniformly in t, and by classical convergence results (see e.g. [18]), in order

to prove the convergence in law $Z^n \xrightarrow{\mathcal{L}} \sqrt{a(f,g)}\, Z$, it is enough to show the following two convergences, for all $t > 0$:

$$\sum_{i=1}^{[nt]} \mathbb{E}[(\xi^n)^2 \mid \mathcal{G}_{i-1}] \xrightarrow{\mathbb{P}} a(f,g)\, t, \qquad \sum_{i=1}^{[nt]} \mathbb{E}[(\xi_i^n)^4 \mid \mathcal{G}_{i-1}] \xrightarrow{\mathbb{P}} 0. \qquad (30)$$

For this we observe that (use again (C) and (H_0)):

$$\begin{aligned}
\mathbb{E}[(\xi_i^n)^2 \mid \mathcal{G}_{i-1}] \;=\; \tfrac{1}{n} \Big(&V(f^2, g^2)_{S_i} + v(f,1)^2\, V(1, g^2)_{S_i} + v(1,g)^2\, V(f^2, 1)_{S_i} \\
-&v(f,1)^2\, v(1,g)^2 - 2v(f,1)\, V(f, g^2)_{S_i} - 2v(1,g)\, V(f^2, g)_{S_i} + 2v(f,1)^2\, v(1,g)\, V(1,g)_{S_i} \Big)
\end{aligned}$$

whereas $\mathbb{E}[(\xi_i^n)^4 \mid \mathcal{G}_{i-1}] \le \frac{K}{n^2}\big(1 + V(f^8, 1)_{S_i} + +V(1, g^8)_{S_i}\big)$. Then (30) follows from (29), plus the property $v(f,g) = v(f,1)v(1,g)$, which holds under (H_0).

Now, we need to prove that this convergence in law is indeed stable, which amounts to proving

$$\mathbb{E}[L\,\phi(Z^n)] \;\to\; \mathbb{E}[L]\,\mathbb{E}[\phi(\sqrt{a(f,g)}\, Z)] \qquad (31)$$

for any bounded $\mathcal{F}$-measurable random variable L and any bounded Lipschitz function ϕ on the Skorokhod space $\mathbb{D}(\mathbb{R}_+, \mathbb{R})$ endowed with a bounded Skorokhod distance ρ (cf. [20], Section 2.2.2 for details and references). It suffices to prove (31) for all L in a set of variables which is dense in $\mathbb{L}^1$, hence to consider an $\mathcal{F}_{S_j}$-measurable variable L, where j is an arbitrarily large integer (because $S_j \to \infty$ as $j \to \infty$). On the one hand, as $n \to \infty$ we have $Z^n_{t \wedge \frac{j}{n}} \to 0$ uniformly in t, so $\rho(Z^n, Z(j)^n) \to 0$ if $Z(j)^n_t = Z^n_t - Z^n_{t \wedge \frac{j}{n}}$, and thus

$$\big|\mathbb{E}[L\,\phi(Z^n)] - \mathbb{E}[L\,\phi(Z(j)^n)]\big| \;\le\; K\mathbb{E}[\rho(Z^n, Z(j)^n)] \;\to\; 0.$$

On the other hand $Z(j)^n$ has the same structure as Z^n, except that the sum is taken from $j+1$ instead of 1, hence the same proof as above shows that the sequence $Z(j)^n$ converges in law to $\sqrt{a(f,g)}\, Z$ as well, under $\mathbb{P}$, and also under the conditional probability $\mathbb{P}(\cdot \mid \mathcal{F}_{S_i})$. Thus

$$\mathbb{E}[\phi(Z(j)^n) \mid \mathcal{F}_{S_i}] \;\to\; \mathbb{E}[\phi(\sqrt{a(f,g)}\, Z)].$$

Since $\mathbb{E}[L\phi(Z(j)^n)] = \mathbb{E}\big[L\,\mathbb{E}(\phi(Z(j)^n) \mid \mathcal{F}_{S_j})\big]$, (31) follows from what precedes. $\qquad\square$

Lemma 2 *Assume* (B,C), *and recall* U *from* (15) *with* $U(1,1)_t = N_t = \sum_{m \ge 1} 1_{\{S_m \le t\}}$. *Set*

$$\begin{aligned}
\overline{\Phi}_t =\; & (N_t)^3\, U(f^2, g^2)_t + N_t(U(f,1)_t)^2\, U(1, g^2)_t \\
& + N_t(U(1,g)_t)^2\, U(f^2, 1)_t + 4N_t U(f,1)_t\, U(1,g)_t\, U(f,g)_t \\
& - 2N_t U(f,1)_t\, U(f, g^2)_t - 2N_t U(1,g)_t\, U(f^2, g)_t - 3U(f,1)_t^2\, U(1,g)_t^2 \\
\overline{\Upsilon}_t =\; & N_t\, U(f,g)_t - U(f,1)_t\, U(1,g)_t.
\end{aligned} \qquad (32)$$

and define $\overline{\Psi}_t = \overline{\Upsilon}_t / \sqrt{\overline{\Phi}_t / N_t}$. *Then, as* $T_n \to \infty$, *we have*

$$\overline{\Phi}_{T_n}/N_{T_n}^4 \xrightarrow{\mathbb{P}} a(f,g), \qquad \overline{\Upsilon}_{T_n}/N_{T_n}^2 \xrightarrow{\mathbb{P}} v(f,g) - v(f,1)v(1,g), \qquad (33)$$

22

and

$$\text{under (H}_0): \quad \overline{\Psi}_{T_n} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$$
$$\text{under (H}_1): \quad |\overline{\Psi}_{T_n}| \xrightarrow{\mathbb{P}} \infty. \tag{34}$$

Proof. 1) We set $z_n = [T_n]$, and the aim of this step is to show that, as $n \to \infty$,

$$A_t^n := \frac{1}{z_n} N_{T_n t} \xrightarrow{\text{u.c.p.}} A_t. \tag{35}$$

Since both A^n and A are increasing in t, and A_t is continuous, it suffices to show that $A_t^n \xrightarrow{\mathbb{P}} A_t$ for any fixed t. We have $\frac{1}{z_n} N_{z_n t} \le A \le \frac{1}{z_n} N_{(z_n+1)t}$, whereas $\frac{z_n+1}{z_n} \to 1$, so (35) readily follows from (B-a).

2) Next, we recall a classical result about the convergence of time-changed processes, see for example [18], Chapter VI. Let Y^n and Y be (possibly multi-dimensional) processes, with Y continuous. Then, since A^n is increasing and its limit A is continuous, we have

- $Y^n \xrightarrow{\mathbb{P}} Y$ (functional convergence) $\quad \Rightarrow \quad Y^n_{A_1^n} \xrightarrow{\mathbb{P}} Y_{A_1}$,

- $Y^n \xrightarrow{\mathcal{L}-(s)} Y$ (functional convergence) $\quad \Rightarrow \quad Y^n_{A_1^n} \xrightarrow{\mathcal{L}-(s)} Y_{A_1}$. $\tag{36}$

Even though we are interested in the convergence of the random variables $Y^n_{A_1^n}$ at time $t = 1$ only, we need the "functional" convergence for the above to be true. We also remark that in the second part Y is defined on an extension of the space.

We also mention a key point, which follows from the definition in (28), namely

$$U(f,g)_{T_n t} = z_n D^{z_n}(f,g)_{A_t^n}. \tag{37}$$

3) Let Y^n be the 8-dimensional process with components

$$Y^{n,1} = D^{z_n}(f^2,g^2), \quad Y^{n,2} = D^{z_n}(f,g^2), \quad Y^{n,3} = D^{z_n}(f^2,g), \quad Y^{n,4} = D^{z_n}(f^2,1)$$
$$Y^{n,5} = D^{z_n}(1,g^2), \quad Y^{n,6} = D^{z_n}(f,1), \quad Y^{n,7} = D^{z_n}(1,g), \quad Y^{n,8} = D^{z_n}(f,g)$$

so by (37) a simple calculation shows

$$\overline{\Phi}_{T_n} = z_n^4 \Big((A_1^n)^3 \, Y^{n,1}_{A_1^n} + A_1^n \, (Y^{n,6}_{A_1^n})^2 \, Y^{n,5}_{A_1^n} + A_1^n \, (Y^{n,7}_{A_1^n})^2 \, Y^{n,4}_{A_1^n} + 4 A_1^n \, Y^{n,6}_{A_1^n} \, Y^{n,7}_{A_1^n} \, Y^{n,8}_{A_1^n}$$
$$- 2 A_1^n \, Y^{n,6}_{A_1^n} \, Y^{n,2}_{A_1^n} - 2 A_1^n \, Y^{n,7}_{A_1^n} \, Y^{n,3}_{A_1^n} - 3 (Y^{n,6}_{A_1^n})^2 \, (Y^{n,7}_{A_1^n})^2 \Big) \tag{38}$$
$$\overline{\Upsilon}_{T_n} = z_n^2 \Big(A_1^n \, Y^{n,8}_{A_1^n} - Y^{n,6}_{A_1^n} \, Y^{n,7}_{A_1^n} \Big), \qquad \overline{\Psi}_{T_n} = \overline{\Upsilon}_{T_n} / \sqrt{\overline{\Phi}_{T_n}/(z_n A_1^n)}.$$

Combining Lemma 1 and (35) yields $(A^n, Y^n) \xrightarrow{\mathbb{P}} (A,Y)$ in the functional sense, where

$$Y_t = \big(v(f^2,g^2)t, v(f,g^2)t, v(f^2,g)t, v(f^2,1)t, v(1,g^2)t, v(f,1)t, v(1,g)t, v(f,g)t \big).$$

Then the first part of (36) and $N_{T_n} = z_n A_1^n$ and $A_1 > 0$ yield (33). Since further $N_{T_n} \to \infty$, whereas $v(f,g) \ne v(f,1)v(1,g)$ under (H$_1$), we deduce the second part of (34).

4) It remains to prove the first part of (34), and we thus assume (H_0). With the notation $Z^n, \overline{Z}^n$ of Lemma 1, we can rewrite $\overline{\Upsilon}_{T_n}$ as

$$\overline{\Upsilon}_{T_n} = z_n^{3/2} \left(A_1^n Z_{A_1^n}^{z_n} - \overline{Z}_{A_1^n}^{z_n} \left(Y_{A_1^n}^{n,7} - v(1,g) A_1^n \right) \right).$$

As seen before, $Y_{A_1^n}^{n,7} - v(1,g) A_1^n \overset{\mathbb{P}}{\longrightarrow} 0$, whereas the second part of (36) with $Y^n = Z^{z_n}$ and with $Y^n = \overline{Z}^{z_n}$, plus Lemma 1, imply that $Z_{A_1^n}^{z_n} \overset{\mathcal{L}-(s)}{\longrightarrow} \sqrt{a(f,g)} Z_{A_1}$ and $\overline{Z}_{A_1^n}^{z_n} \overset{\mathcal{L}-(s)}{\longrightarrow} \sqrt{v(f^2,1) - v(f,1)^2} Z_{A_1}$. Plugging this in the expression for $\overline{\Upsilon}_{T_n}$, we deduce that $\overline{\Upsilon}_{T_n}/z_n^{3/2} \overset{\mathcal{L}-(s)}{\longrightarrow} \sqrt{a(f,g)} A_1 Z_{A_1}$. Coming back to (38), and since $a(f,g) > 0$ and upon using (35), it follows that $\overline{\Psi}_{T_n} \overset{\mathcal{L}-(s)}{\longrightarrow} Z_{A_1}/\sqrt{A_1}$. Since Z is a standard Brownian motion independent of A_1, we see that this limit is $\mathcal{N}(0,1)$, and the proof is complete. $\square$

8.2 Fixed-time approximations

Here we compare the process $U(f,g)$ of (15) with its approximation $U(f,g)^n$ given by (16). It should be emphasized that our Assumption (A) is *not* directly comparable to the boundedness or local boundedness assumptions on the coefficients which are made in the book [20]. However, the estimates are still valid, and thus so is for example Theorem 9.5.1 of [20], which asserts that $U(f,g)^n$ converges in probability, for the Skorokhod topology, to $U(f,g)$.

Here we need precise moment estimates on the difference $U(f,g)_t^n - U(f,g)_t$ for any given time t, and for this we need some notation. We set $\overline{X}_t = X_t - \sum_{s \leq t} \Delta X_s = X_t - \sum_{m \geq 1} \Delta X_{S_m} 1_{\{S_m \leq t\}}$, and we denote by $S_1', S_2', \cdots$ the successive jump times of the Poisson process $N_t' = \int_0^t \int_E 1_{\{\Gamma(z) > 0\}} \mu(ds, dz) = \mu((0,t] \times \{z : \Gamma(z) > 0\})$, and by $i(n,m)$ the unique integer such that $(i(n,m) - 1)\Delta_n < S_m' \leq i(n,m)\Delta_n$. Next, $\mathcal{G}_t$ denotes the σ-field generated by $\mathcal{F}_t$ and all random times S_m'.

We set $v_n = k_n \Delta_n$, and it is no restriction to assume $v_n \leq 1/2$. Let $\Omega_{n,t}$ be the set on which $|\Delta_i^n \overline{X}| \leq u_n$ for all $i \leq [t/\Delta_n]$, and $\Omega_{n,t}'$ be the set on which all S_m' smaller than t are more than $2v_n$ apart from one another and from t itself. Observe that $\mathbb{E}[|\Delta_i^n \overline{X}|^q] \leq K_q \Delta_n^{q/2}$ for any non-random i (use (A) and classical estimates). Then by Markov's inequality and $u_n \asymp \Delta_n^\varpi$ with $\varpi < \frac{1}{2}$ on the one hand (cf. (13) for the precise formulation), and by the Poisson property of the times S_m' on the other hand, we get

$$\mathbb{P}(\Omega_{n,t}) \geq 1 - Kt\Delta_n, \qquad \mathbb{P}(\Omega_{n,t}') \geq 1 - Ktv_n. \tag{39}$$

Below, q is an arbitrary number in $[2, \infty)$. We set

$$M_t := 1 + |b_t| + |\widetilde{b}_t| + |\widetilde{\sigma}_t| + |\widehat{\sigma}_t| + \gamma_t + \widetilde{\gamma}_t + \sqrt{c_t}, \quad \overline{M}(q,t) := \mathbb{E}\left[\sup_{s \in [0,t]} (M_s)^q \mid \mathcal{G}_0 \right].$$

Observing that $\mathbb{E}[\overline{M}(q,t)] \leq \sum_{m=0}^{[t]} \mathbb{E}\left[\sup_{s \in [m,m+1]} (M_s)^q \right]$, we readily deduce from (A-b) that

$$\mathbb{E}[\overline{M}(q,t)] \leq K_q(t+1). \tag{40}$$

By (A-a) we have $\mathbb{E}[|\Delta X_{S'_m}|^q \mid \mathcal{G}_{S'_m-}] \leq K_q \gamma^q_{S'_m}$ and a similar estimate for $\Delta c_{S'_m}$, hence

$$S'_m \leq t \quad \Rightarrow \quad \mathbb{E}[|\Delta X_{S'_m}|^q \mid \mathcal{G}_0] + \mathbb{E}[|\Delta c_{S'_m}|^q \mid \mathcal{G}_0] \leq K_q \overline{M}(q, t). \tag{41}$$

By (A) and the arguments in Section 2.1.6 of [20], and since $i(n, m)$ is $\mathcal{G}_0$-measurable,

$$\mathbb{E}[|\Delta^n_{i(n,m)} \overline{X}|^q \mid \mathcal{G}_{(i(n,m)-1)\Delta_n}] \leq K_q \Delta_n^{q/2} \mathbb{E}\Big[\sup_{s \in ((i(n,m)-1)\Delta_n, i(n,m)\Delta_n]} (M_s)^q \mid \mathcal{G}_{(i(n,m)-1)\Delta_n}\Big],$$

which yields

$$S'_m \leq t \quad \Rightarrow \quad \mathbb{E}[|\Delta^n_{i(n,m)} \overline{X}|^q \mid \mathcal{G}_0] \leq K_q \Delta_n^{q/2} \overline{M}(q, t+1). \tag{42}$$

We also set

$$\bar{c}^n_i := \frac{1}{k_n \Delta_n} \sum_{j=1}^{k_n} (\Delta^n_{i+j} \overline{X})^2, \tag{43}$$

which is the estimated spot volatility based upon the continuous process $\overline{X}$. By standard estimates of the same type as above,

$$\mathbb{E}[|\bar{c}^n_{i(n,m)} - c_{S'_m}|^q \mid \mathcal{G}_{S'_m}] \leq K_q \Big(v_n + \frac{1}{k_n^{q/2}}\Big) \mathbb{E}\Big[\sup_{s \in (S'_m, S'_m + v_n + \Delta_n]} (M_s)^{2q} \mid \mathcal{G}_{i(n,m)\Delta_n}\Big]$$

and similarly for $\mathbb{E}[|\bar{c}^n_{i(n,m)-k_n-1} - c_{S'_m-}|^q \mid \mathcal{G}_{(i(n,m)-k_n-1)\Delta_n}]$. Then, as for (42), we get

$$S'_m \leq t \Rightarrow \mathbb{E}[|\bar{c}^n_{i(n,m)-k_n-1} - c_{S'_m-}|^q + |\bar{c}^n_{i(n,m)} - c_{S'_m}|^q \mid \mathcal{G}_0] \leq K_q \Big(v_n + \frac{1}{k_n^{q/2}}\Big) \overline{M}(2q, t+1). \tag{44}$$

The properties of f and g yield $|f(x+u)g(y+v) - f(x)g(y)| \leq K(1 + |x|^p + |u|^p + |y|^p + |v|^p)(|u| + |v|)$ for some $K, p \geq 1$. Therefore, on the set $\Omega_{n,t} \cap \Omega'_{n,t}$ we have

$$|U(f,g)^n_t - U(f,g)_t| \leq K \sum_{m \geq 1: S'_m \leq t} \zeta^n_m \overline{\zeta}^n_m, \quad \text{where}$$
$$\zeta^n_m = 1 + |\Delta X_{S'_m}|^p + |\Delta^n_{i(n,m)} \overline{X}|^p + |\Delta c_{S'_m}|^p + |\bar{c}^n_{i(n,m)} - c_{S'_m}|^p + |\bar{c}^n_{i(n,m)-k_n-1} - c_{S'_m-}|^p$$
$$\overline{\zeta}^n_m = |\Delta^n_{i(n,m)} \overline{X}| + |\bar{c}^n_{i(n,m)} - c_{S'_m}| + |\bar{c}^n_{i(n,m)-k_n-1} - c_{S'_m-}|.$$
$$\tag{45}$$

We have $\overline{M}(p, s) \leq \overline{M}(q, s)^{p/q} \leq \overline{M}(q, s)$ if $q \geq p$, by Hölder's inequality and $\overline{M}(q, s) \geq 1$, so (41), (42) and (44) and Hölder's inequality again give us, with $N'_t = \sum_{m \geq 1} 1_{\{S'_m \leq t\}}$:

$$\mathbb{E}\Big[|U(f,g)^n_t - U(f,g)_t| 1_{\Omega_{n,t} \cap \Omega'_{n,t}}\Big] \leq K \sum_{m \geq 1} \mathbb{E}\Big[1_{\{S'_m \leq t\}} \mathbb{E}(\zeta^n_m \overline{\zeta}^n_m \mid \mathcal{G}_{S'_{m-1}})\Big]$$

$$\leq K\big(\sqrt{\Delta_n} + \sqrt{v_n} + \frac{1}{\sqrt{k_n}}\big) \mathbb{E}\big[\overline{M}(4p, t+1) N'_t\big]$$

$$\leq K\big(\sqrt{v_n} + \frac{1}{\sqrt{k_n}}\big) \mathbb{E}\big[\overline{M}(4pq, t+1)\big]^{1/q} \mathbb{E}[(N'_t)^r]^{1/r}$$

25

for every conjugate pair $q, r \in (1, \infty)$ (i.e. $1/q + 1/r = 1$). Then (40) and the fact that N'_t is a Poisson variable with parameter Ct, implying $\mathbb{E}[(N'_t)^r] \leq Kt^r$, allow us to deduce

$$\mathbb{E}\left[|U(f,g)^n_t - U(f,g)_t| \, 1_{\Omega_{n,t} \cap \Omega'_{n,t}} \right] \leq K_\varepsilon \big(\sqrt{v_n} + \frac{1}{\sqrt{k_n}} \big) (1 + t^{1+\varepsilon}), \qquad (46)$$

for any $\varepsilon > 0$, where K_ε depends on ε, and also on the functions f and g, arbitrary in $\mathcal{L}_{pol}$. Note that, when $f = g = 1$, the left side of (46) vanishes, whereas $U(1,1)_t = N_t$.

8.3 Proof of Theorems 1 and 2

Theorem 2 is an immediate consequence of Theorem 1, which we now proceed to prove. In view of Lemma 2, we need to compare Ψ_n and $\overline{\Psi}_{T_n}$. Note that $T_n \Delta_n^{1/2-\eta} \to 0$ and $\rho \in \left[\frac{1}{2} - \eta, \frac{1}{2} \right)$ imply the existence of some $\varepsilon > 0$ such that

$$T_n v_n \to 0, \quad T_n^{1+2\varepsilon} w_n^2 \to 0, \quad \text{where} \ \ w_n = \sqrt{v_n} + 1/\sqrt{k_n}. \qquad (47)$$

The inequalities (39) and the above yield that $\Omega_n = \Omega_{n,T_n} \cap \Omega'_{n,T_n}$ satisfies $\mathbb{P}(\Omega_n) \to 1$. By (35) and $[T_n] \sim T_n$ we have $N_{T_n}/T_n \xrightarrow{\mathbb{P}} A_1 > 0$. Moreover, combining (28), (35) and (37), we see that $U(f', g')_{T_n}/N_{T_n} \xrightarrow{\mathbb{P}} v(f', g')$ for every $f', g' \in \mathcal{L}_{pol}$. Thus, the above and (46) applied with an ε satisfying (47) imply the existence of a $[1, \infty)$-valued variable ζ, of a sequence ξ_n of variables bounded in probability, and of subsets $\Omega'_n \subset \Omega_n$ with $\mathbb{P}(\Omega'_n) \to 1$, such that on Ω'_n we have

$$|U(f^j, g^k)^n_{T_n} - U(f^j, g^k)_{T_n}| \leq \xi_n \, w_n \, T_n^{1+\varepsilon}, \quad |U(f^j, g^k)_{T_n}| \leq \zeta \, T_n, \quad T_n \leq \zeta \, N_{T_n} \leq \zeta^2 T_n \tag{48}$$

for all powers $j, k = 0, 1, 2$. Recalling $U(1,1)_t = N_t$ and (17) and (32), we deduce that, on the set Ω'_n,

$$|\Phi_n - \overline{\Phi}_{T_n}| \leq K(\xi_n^4 \, w_n^4 \, T_n^{43\varepsilon} + \xi_n \, w_n \, T_n^\varepsilon \, \zeta^3) T_n^4, \quad |\Upsilon_n - \overline{\Upsilon}_{T_n}| \leq K(\xi_n^2 \, w_n^2 \, T_n^{2\varepsilon} + \xi_n \, w_n \, T^\varepsilon \, \zeta) T_n^2. \tag{49}$$

Therefore, we deduce from (33) and $w_n T_n^\varepsilon \to 0$ that the variables Φ_n and Γ_n also satisfy the same properties (33): the second part of (18) is then proved as in Lemma 2.

For the first part, we assume (H_0). It suffices to show that $(\Psi_n - \overline{\Psi}_{T_n}) 1_{\Omega'_n} \xrightarrow{\mathbb{P}} 0$. As Φ_n satisfies the first property (33), it is enough to show that $\big((\Upsilon_n - \overline{\Upsilon}_{T_n})/(N_{T_n})^{3/2} \big) 1_{\Omega'_n} \xrightarrow{\mathbb{P}} 0$, and by (48) and (49) this trivially holds under (47).

References

[1] Aït-Sahalia, Y. (2002). Telling from discrete data whether the underlying continuous-time model is a diffusion. *J. Finance* **57** 2075–2112.

[2] Aït-Sahalia, Y. and Jacod, J. (2009). Testing for jumps in a discretely observed process. *Ann. Stat.* **37** 184–222.

[3] Alvarez, A., Panloup F., Pontier M. and Savy, N. (2012). Estimation of the instantaneous volatility. *Statistical Inference for Stochastic Processes* **15** 27–59.

[4] Andersen, T.G., Benzoni, L. and Lund, J. (2002). An empirical investigation of continuous-time equity return models. *J. Finance* **57** 1239–1284.

[5] Andersen, T.G., Bollerslev, T. and Diebold, F.X. (2007). Roughing it up: including jump components in the measurement, modeling, and forecasting of return volatility. *Rev. Econ. Stat.* **89** 701–720.

[6] Bandi, F. and Renò, R. (2012). Time-varying leverage effect. *Journal of Econometrics*, **169**, 94-113.

[7] Bandi, F. and Renò, R. (2013). Price and volatility co-jumps. Working paper, Johns Hopkins University and Università di Siena.

[8] Barndorff-Nielsen, O.E., Hansen, P.R., Lunde, A. and Shephard, N. (2009). Realised kernels in practice: trades and quotes. *Econometrics J.* **12** 1–32.

[9] Barndorff-Nielsen, O.E. and Shephard, N. (2001). Non–Gaussian Ornstein–Uhlenbeck–based models and some of their uses in financial economics (with discussion). *J. R. Statist. Soc. Ser. B* **63** 167–241.

[10] Barndorff-Nielsen, O.E. and Shephard, N. (2006). Econometrics of testing for jumps in financial economics using bipower variation. *J. Financial Econometrics* **4** 1–30.

[11] Black, F. (1976). Studies of stock price volatility changes. *Proceedings of the Business and Economic Section of the American Statistical Association*, 177-181.

[12] Carr, P., Geman, H., Madan, D. and Yor, M. (2003). The fine structure of asset returns: an empirical investigation. *J. of Business* **75** 305–333.

[13] Christie, A. (1982). The stochastic behaviour of common stock variances: value, leverage, and interest rate effects. *Journal of Financial Economics*, **3**, 407-431.

[14] Duffie, D., Pan, J. and Singleton, K. (2000). Transform analysis and asset pricing for affine jump diffusions. *Econometrica* **68** 1343–1376.

[15] Eberlein, E. and Keller U. (1995). Hyperbolic distributions in finance. *Bernoulli* **1** 281–299.

[16] Huang, X. and Tauchen, G. (2006). The relative contribution of jumps to total price variance. *J. Financial Econometrics* **4** 456–499.

[17] Jacod, J., Klüppelberg, C. and Müller, G. (2015). Testing for non-correlation between price and volatility jumps: appendix. Working Paper, Université Paris VI and Technische Universität München.

[18] Jacod, J. and Shiryaev, A.N. (2003). *Limit Theorems for Stochastic Processes*, 2nd ed. Springer-Verlag, Berlin.

[19] Jacod, J. and Todorov, V. (2010). Do price and volatility jump together? *Annals of Applied Probability* **20** 1425–1469.

[20] Jacod, J. and Protter P. (2011). *Discretization of Processes*. Springer-Verlag, Berlin.

[21] Klüppelberg, C., Lindner, A. and Maller, R. (2004). A continuous time GARCH process driven by a Lévy process: stationarity and second order behavior. *J. Appl. Probab.* **41** 601–622.

[22] Lee, S.S. and Mykland, P.A. (2008). Jumps in financial markets: a new nonparametric test and jump dynamics. *Rev. Financial Stud.* **21** 2535–2563.

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

[24] Merton, R.C. (1976). Option pricing when underlying stock returns are discontinuous. *J. Fin. Econ.* **3** 125–44.

[25] Mykland, P.A., Shephard, N. and Sheppard, K. (2012). Efficient and feasible inference for the components of financial variation using blocked multipower variation. Working Paper, University of Chicago and University of Oxford.