Global estimation of realized instantaneous volatility in the presence of price jumps

Wale Dare*

Department of Mathematics and Statistics
University of St.Gallen
Switzerland

PRELIMINARY AND INCOMPLETE. PLEASE DO NOT CIRCULATE.

Abstract

We propose a non-parametric procedure for estimating the realized spot volatility of a price process described by an Itô semimartingale with Lévy jumps. The procedure integrates the threshold jump elimination technique of Mancini (2009) with a frame (Gabor) expansion of the realized trajectory of spot volatility. We show that the procedure converges in probability in $L^2([0,T])$ for a wide class of spot volatility processes, including those with discontinuous paths. Our analysis assumes the time interval between price observations tends to zero; as a result, the intended application is the analysis of high frequency financial data.

Volatility estimation using discretely observed asset prices has received a great deal of attention recently, however, much of that effort has been focused on estimating the *integrated* volatility and, to a lesser extent, the *spot* volatility at a given point in time. Notable contributions to the literature on volatility estimation include the papers by Foster & Nelson (1996), Fan & Wang (2008), Florens-Zmirou (1993), and Barndorff-Nielsen & Shephard (2004). In these studies, the object of interest is local in nature: spot volatility at a given point in time or integrated volatility up to a terminal point in

*wale.dare@unisg.ch

time. In contrast, estimators which aim to obtain spot volatility estimates for entire time windows have received much less coverage. These are the so-called "global" spot volatility estimators. These estimators derive their name from the fact that the objects of interest are not localized. Typically, a global estimator would be a random elements whose realizations are elements of some function space.

There are potential benefits to adopting global estimators of spot volatility. Given a consistent global estimate of spot volatility σ^2 over an interval $[0, T]$, the integrated volatility at any point t within $[0, T]$ may be consistently estimated by integrating σ^2 over the interval $[0, t]$. In fact, by the continuous mapping theorem, consistent estimates of continuous transformations of σ^2 are immediately available. Hence, integrated powers of spot volatility, $\int_0^t \sigma_s^p \, \mathrm{d}s$, $p > 0$, the running maximum of spot volatility, $\sigma_t^* := \sup_{s \leq t} |\sigma_s|$, and volatility in excess of a given threshold, $\sigma_t^a := \sigma_t I_{\{|\sigma_t| > a\}}$, $a > 0$, to name just a few, are easily obtained via the obvious transformation of the estimated global spot volatility. This flexibility is one of the more appealing features of this class of estimators.

The estimator by Genon-Catalot et al. (1992) is an early contribution to the study of the realized trajectory of spot volatility. Working within the context of continuous asset prices and deterministic spot volatility, the authors decribed an estimator of the realized trajectory of spot volatility using wavelet projection methods. This basic framework has since been extended by Hoffmann et al. (2012). The authors proposed adaptive estimators of spot volatility in the framework of asset prices subject to market microstructure noise.

Another important contribution to the global spot volatility estimation literature is the estimator proposed by Malliavin & Mancino (2002), which relies on Fourier methods to estimate the realized path of spot volatility in the context of continuous prices. In their procedure, Fourier coefficients of the realized price path are first estimated and the used to derive expressions for the Fourier coefficients of the realized path of spot volatility.

In the current work, we extend the study of the realized path of spot volatility to situations where the price process or the volatility coefficient itself cannot be assumed to be continuous. That is we describe a procedure for consistently estimating càdlàg volatility paths in the presence of price jumps. By employing Gabor frames in our analysis we are able to leverage the excellent time-frequency localization property of Gabor frames to obtain the sparsest representation for the realized trajectories of spot volatility.

The rest of this paper is organized as follows: in Section 1 we introduce notation and give general description of the dynamics of observed prices. In Section 2 we introduce Gabor frames and review the basic theory required for our subsequent analysis. We present our main results in Section 3 and 4, where we specify the estimator and give proof of its consistency. Section 5 describes simulation exercises that lend further support to the theoretical analysis. Section 6 contains concluding remarks.

1 Prices

We fix a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, P)$. We recall the definition of an Itô semimartingale with Lévy jumps.

1.1 Definition *An $\mathbb{R}$-valued process X is an Itô semimartingale with Lévy jumps if it admits the representation:*

$$X_t = X_0 + \int_0^t b_s ds + \int_0^t \sigma_s dW_s + xI_{\{|x|>1\}} * \mu_t + xI_{\{|x|\leq 1\}} * (\mu - \nu)_t \tag{1.1}$$

with

$$\nu(dt, dx) = F(dx)dt,$$

$t \geq 0$, *where W is a Brownian motion, σ and b are $\mathbb{R}$-valued progressively measurable processes, μ is an integer-valued measure induced by the jumps of X, ν is its Lévy system, and $F(dx)$ is a deterministic σ-finite measure on $\mathbb{R}$.*

1.1 Remark Generally, Itô semimartingales are those with characteristic triplet that is absolutely continuous with respect to the Lebesgue measure. Here, we further restrict the Lévy system ν to be deterministic. This assumption ensures the jump measure μ is a Poisson random measure.

We assume prices are observed in the fixed time interval $[0, 1]$ at discrete, equidistant times $t_i = i\Delta_n, i = 0, 1, \cdots, n$, where

$$\Delta_n = 1/n = t_{i+1} - t_i, \qquad i = 0, \cdots, n-1. \tag{1.2}$$

Given the finite sequence $\{X_{t_i}, i = 0, 1, 2, \cdots, n\}$, our aim is to estimate the spot variance σ^2 in the time interval $[0, 1]$ by nonparametric methods. Note that our objective is not an approximation of a point but rather the approximation of an entire function. Thus an estimator of the spot variance may be viewed as a random element (function), as opposed to a random variable, that must converge in some sense to the spot variance, which itself is a random element. We approach this task by estimating the expansion of the spot variance using collections of Gabor frame elements.

2 Frames

Frames generalize the notion of orthonormal bases in Hilbert spaces. If $\{f_k\}_{k\in\mathbb{N}}$ is a frame for a separable Hilbert space $\mathcal{H}$ then every vector $f \in \mathcal{H}$ may be expressed as a linear combination of the frame elements, i.e.

$$f = \sum_{k\in\mathbb{N}} c_k f_k. \tag{2.3}$$

This is similar to how elements in a Hilbert space may be expressed in terms of orthonormal basis; but unlike orthonormal basis, the representation in (2.3) need not be unique, and the frame elements need not be orthogonal. Loosely speaking, frames contain redundant elements. The absence of uniqueness in the frame representation is by no means a shortcoming; on the contrary, we are afforded a great deal of flexibility and stability as a result. In fact, given a finite data sample, the estimated basis expansion coefficients are likely to be imprecise. This lack of precision can create significant distortions when using an orthonormal basis. These distortions are somewhat mitigated when using frames because of the built-in redundancy of frame elements.

Furthermore, if $\{f_k\}_{k\in\mathbb{N}}$ is a frame for $\mathcal{H}$, then surjective, bounded transformations of $\{f_k\}_{k\in\mathbb{N}}$ also constitute frames for $\mathcal{H}$, e.g. $\{f_k + f_{k+1}\}_{k\in\mathbb{N}}$ is a frame. So, once we have a frame, we can generate an arbitrary number of them very easily. We may then obtain estimates using each frame and compare results. If our results using the different frames fall within a tight band, then we are afforded some indication of the robustness of the computations.

Our discussion of frame theory will be rather brief; we only mention concepts needed for our specification of the volatility estimator. For a more detailed treatment see the book by Christensen (2008). In the sequel if z is a complex number then we shall denote respectively by $\bar{z}$ and $|z|$ the complex conjugate and magnitude of z. Let $L^2(\mathbb{R})$ denote the space of complex-valued functions defined on the real line with finite norm given by

$$\|f\| := \left(\int_{\mathbb{R}} f(t)\overline{f(t)}\,\mathrm{d}t \right)^{1/2} < \infty, \qquad f \in L^2(\mathbb{R}).$$

Define the inner product of two elements f and g in $L^2(\mathbb{R})$ as $\langle f, g\rangle := \int_{\mathbb{R}} f(t)\overline{g(t)}\,\mathrm{d}t$.

Denote by $\ell^2(\mathbb{N})$ the set of complex-valued sequences defined on the set of natural numbers $\mathbb{N}$ with finite norm given by

$$\|c\| := \left(\sum_{k\in\mathbb{N}} c_k \overline{c_k} \right)^{1/2} < \infty, \qquad c \in \ell^2(\mathbb{N}),$$

where c_k is the k-th component of c. The inner product of two sequences c and d in $\ell^2(\mathbb{N})$ is $\langle c, d\rangle := \sum_{k\in\mathbb{N}} c_k \overline{d_k}$. Now we may give a definition for frames:

2.1 Definition *A sequence $\{f_k\}_{k\in\mathbb{N}} \subset L^2(\mathbb{R})$ is a frame if there exists positive constants C_1 and C_2 such that*

$$C_1\|f\|^2 \le \sum_{k\in\mathbb{N}} |\langle f, f_k\rangle|^2 \le C_2\|f\|^2, \qquad f \in L^2(\mathbb{R}).$$

The constants C_1 and C_2 are called *frame bounds*. If $C_1 = C_2$ then $\{f_k\}_{k\in\mathbb{N}}$ is said to be *tight*. Because an orthonormal basis satisfies Parseval's equality

$$\mathcal{A}f = c \qquad \text{where} \qquad c_k = \langle f, f_k\rangle, \qquad k \in \mathbb{N}. \tag{2.4}$$

Because $\mathcal{A}$ takes a function defined on a continuum ($\mathbb{R}$) to a sequence, which is a function defined on the discrete set $\mathbb{N}$, $\mathcal{A}$ is known as the *analysis* operator associated with the frame $\{f_k\}_{k\in\mathbb{N}}$. The boundedness of the analysis operator follows from the frame bounds in Definition (2.1). Now $\mathcal{A}^*$, the adjoint of $\mathcal{A}$, is well-defined and takes sequences in $\ell^2(\mathbb{N})$ to functions in $L^2(\mathbb{R})$. Using the fact that $\mathcal{A}^*$ must satisfy the equality $\langle \mathcal{A}f, c\rangle = \langle f, \mathcal{A}^*c\rangle$ for all $f \in L^2(\mathbb{R})$ and $c \in \ell^2(\mathbb{N})$, it may be deduced that

$$\mathcal{A}^*c = \sum_{k\in\mathbb{N}} c_k f_k, \qquad c \in \ell^2(\mathbb{N}),$$

where c_k is the k-th component of the sequence c. The adjoint, $\mathcal{A}^*$, may be thought of as reversing the operation or effect of the analysis operator; for this reason it is known as the *synthesis* operator.

Now an application of the operator $(\mathcal{A}^*\mathcal{A})^{-1}$ to every frame element f_k yields a sequence $\{\tilde{f}_k := (\mathcal{A}^*\mathcal{A})^{-1}f_k\}_{k\in\mathbb{N}}$, which is yet another frame for $L^2(\mathbb{R})$. The frame $\{\tilde{f}_k\}_{k\in\mathbb{N}}$ is known as the *canonical dual* of $\{f_k\}_{k\in\mathbb{N}}$. Denoting the analysis operator associated with the canonical dual by $\tilde{\mathcal{A}}$, it may be shown[1] that

$$\mathcal{A}^*\tilde{\mathcal{A}} = \tilde{\mathcal{A}}^*\mathcal{A} = \mathcal{I}, \tag{2.5}$$

where $\mathcal{I}$ is the identity operator and $\tilde{\mathcal{A}}^*$ is the adjoint of the analysis operator of the canonical dual. Furthermore, Proposition 3.2.3 of Daubechies (1992) shows that $\tilde{\mathcal{A}}$ satisfies

$$\tilde{\mathcal{A}} = \mathcal{A}(\mathcal{A}^*\mathcal{A})^{-1}, \tag{2.6}$$

[1] See for example Daubechies (1992, Proposition 3.2.3)

so that the analysis operator of the canonical dual frame is fully characterized by $\mathcal{A}$ and its adjoint. It is easily seen that (2.5) yields a representation result since if $f \in L^2(\mathbb{R})$ then

$$f = \tilde{\mathcal{A}}^* \mathcal{A} f = \mathcal{A}^* \tilde{\mathcal{A}} f = \sum_{k \in \mathbb{N}} \langle f, \tilde{f}_k \rangle f_k. \tag{2.7}$$

Thus, in a manner reminiscent of orthonormal basis representations, every function in $L^2(\mathbb{R})$ is expressible as a linear combination of the frame elements, with the frame coefficients given by $\langle f, \tilde{f}_k \rangle$, the correlation between the function and the elements of the dual frame. It follows from the first equality in (2.5) and the commutativity of the duality relationship that functions in $L^2(\mathbb{R})$ may also be written as linear combinations of the elements in $\{\tilde{f}_k\}_{k \in \mathbb{N}}$, with coefficients given by $\langle f, f_k \rangle$, i.e. $f = \sum_{k \in \mathbb{N}} \langle f, f_k \rangle \tilde{f}_k$.

2.1 Gabor frames

Next, we specialize the discussion to Gabor frames. The analysis of Gabor frames involves two operators: the *translation* operator $\mathcal{T}$ and the *modulation* operator $\mathcal{M}$ defined as follows:

$$\mathcal{T}_b f(t) := f(t - b), \qquad b \in \mathbb{R}, f \in L^2(\mathbb{R}), \tag{2.8}$$

$$\mathcal{M}_a f(t) := \mathrm{e}^{2\pi \mathrm{i} a t} f(t), \qquad a \in \mathbb{R}, f \in L^2(\mathbb{R}), \tag{2.9}$$

where i is the imaginary number, i.e. $\mathrm{i} = \sqrt{-1}$. Both $\mathcal{T}$ and $\mathcal{M}$ are shift operators: $\mathcal{T}$ is a shift or translation operator on the time axis, whereas $\mathcal{M}$ performs shifts on the frequency axis. A Gabor system is constructed by fixing $a, b \in \mathbb{R}$, and performing shifts of a single nontrivial function $g \in L^2(\mathbb{R})$ in time-frequency space. For example, if a and b are real numbers then the sequence of functions

$$\{\mathcal{M}_{ha} \mathcal{T}_{kb} g\}_{h,k \in \mathbb{Z}},$$

constitutes a Gabor system.

2.2 Definition *Let $g \in L^2(\mathbb{R})$, and let $a > 0$, $b > 0$ be positive real numbers. Define for $t \in \mathbb{R}$*

$$g_{h,k}(t) := \mathrm{e}^{2\pi \mathrm{i} hat} g(t - kb), \qquad h, k \in \mathbb{Z}.$$

If the sequence $\{g_{h,k}\}_{h,k \in \mathbb{Z}}$ constitutes a frame for $L^2(\mathbb{R})$, then it is called a Gabor frame.[2]

[2] It is also sometimes referred to as a *Weyl-Heisenberg* frame.

The fixed function g is known as the *Gabor frame generator*[3]; a is known as the *modulation parameter*; and b is known as the *translation parameter*. In order to obtain sharp asymptotic rates, we require g and its dual $\tilde{g}$ (see (2.7)) to be continuous and compactly supported. The following result, stated in Christensen (2006, Lemma 1.2) and in Zhang (2008, Proposition 2.4), tells us how to construct such dual pairs.

2.1 Lemma *Let $[r, s]$ be a finite interval, let $a > 0$, $b > 0$ be positive constants, and let g be a continuous function. If $g(t) \neq 0$ when $t \in (r, s)$; $g(t) = 0$ when $t \notin (r, s)$; and a, b satisfy: $a < 1/(s - r)$, $0 < b < s - r$; then $\{g, \tilde{g}\}$ is a pair of dual Gabor frame generators, with the dual Gabor generator given by*

$$\tilde{g}(t) := g(t)/G(t), \text{ where} \tag{2.10}$$

$$G(t) := \sum_{k \in \mathbb{Z}} |g(t - kb)|^2/a. \tag{2.11}$$

Furthermore,

$$\tilde{g}_{h,k}(t) := \mathrm{e}^{2\pi i h a t} \tilde{g}(t - kb), \qquad h, k \in \mathbb{Z} \tag{2.12}$$

is compactly supported.

In the sequel, we assume the Gabor frame setup in Lemma 2.1.

3 Volatility estimation: continuous prices

In this section we specify a consistent estimator of spot volatility within a framework of continuous prices. That is, we simplify the general setup of (1.1) to:

$$X_t = X_0 + \int_0^t b_s \, \mathrm{d}s + \int_0^t \sigma_s \, \mathrm{d}W_s, \qquad t \geq 0. \tag{3.13}$$

We further restrict the processes b and σ as follows:

3.1 Assumption

1. The drift b is progressively measurable, whereas the diffusion coefficient σ is adapted and càdlàg.

[3] It is referred to elsewhere as the *window function*.

2. There is a sequence of stopping times $\{T_m\}$ tending to infinity almost surely such that

$$E(\sup_{0\leq s\leq T_m} |b_s - b_0|^4) + E(\sup_{0\leq s\leq T_m} |\sigma_s - \sigma_0|^4) < \infty,$$

for all m.

3.1 Remark These assumptions are satisfied by a wide range of practically relevant processes; these include continuous Lévy and additive processes with càdlàg volatility coefficients. Also included are continuous solutions of stochastic differential equations; indeed all processes with locally bounded b and σ satisfy these requirements.

Let $\{g, \tilde{g}\}$ be a pair of dual Gabor frame generators constructed as in Lemma 2.1, then σ^2 admits a Gabor frame expansion given by:

$$\sigma^2(t) = \sum_{h,k\in\mathbb{Z}} c_{h,k}\, g_{h,k}(t), \text{ where} \tag{3.14}$$

$$c_{h,k} = \langle \sigma^2, \tilde{g}_{h,k}\rangle. \tag{3.15}$$

Note that both σ^2 and $\tilde{g}$ have compact support. Indeed σ^2 has support in $[0, 1]$, whereas $\tilde{g}$ has support in $[s, r]$. So, $c_{h,k} \neq 0$ only if the supports of σ^2 and $\tilde{g}_{h,k}$ overlap. Furthermore, we note from (2.12) that $\tilde{g}_{h,k+1}$ is simply $\tilde{g}_{h,k}$ shifted by b units; so, $c_{h,k} = 0$ if $|k| \geq K_0$ with

$$K_0 := \lceil (1 + |s| + |r|)/b \rceil, \tag{3.16}$$

where $\lceil x\rceil$, $x \in \mathbb{R}$, is the least integer that is greater than or equal to x. Thus σ^2 admits a representation of the form:

$$\sigma^2(t) = \sum_{\substack{(h,k)\in\mathbb{Z}^2\\ |k|\leq K_0}} c_{h,k}\, g_{h,k}(t),$$

and for sufficiently large positive integer H,

$$\sigma^2(t) \approx \sum_{\substack{|h|\leq H\\ |k|\leq K_0}} c_{h,k}\, g_{h,k}(t).$$

Now, suppose n observations of the price process are available, and let

$$\Theta_n := \{(h, k) \in \mathbb{Z}^2 : |h| \leq H_n \text{ and } |k| \leq K_0\}, \tag{3.17}$$

where H_n is an increasing sequence in n. We propose the following estimator of the volatility coefficient:

$$v_n(X,t) := \sum_{(h,k)\in\Theta_n} \hat{c}_{h,k}\, g_{h,k}(t), \qquad t \in [0,1], \text{ where} \tag{3.18}$$

$$\hat{c}_{h,k} := \sum_{i=0}^{n-1} \overline{\tilde{g}_{h,k}(t_i)}(X_{t_{i+1}} - X_{t_i})^2. \tag{3.19}$$

So $|\Theta_n|$ is the number of frame elements included in the expansion. Specifically, $|\Theta_n| = (2K_0+1)(2H_n+1)$; and since K_0 is a finite quantity, it follows that $|\Theta_n| = O(H_n)$, i.e. the number of estimated coefficients is proportional to H_n, and therefore, will grow with the number of observations, n. In the next section we show that the estimator converges to σ^2 on $[0,1]$ in a mean integrated square error sense.

3.1 Proposition *Suppose the price process is specified as in* (3.13) *and satisfies the conditions of Assumption 3.1. Let $\{g, \tilde{g}\}$ be pair of dual Gabor generators satisfying the conditions of Lemma 2.1 such that g is Lipschitz continuous on the unit interval. If $H_n \uparrow \infty$ satisfies*

$$(H_n)^2 \Delta_n^{1/2} = o(1),$$

then $v_n(X,t)$, defined in (3.18)*, converges in $L^2[0,1]$ to σ^2 in probability.*

Proof. We begin by noting that

$$\begin{aligned} v_n(X,t) - \sigma^2(t) = &\sum_{(h,k)\in\Theta_n} (\hat{c}_{h,k} - c_{h,k})\, g_{h,k}(t) \\ &- \sum_{(h,k)\notin\Theta_n} c_{h,k}\, g_{h,k}(t), \end{aligned} \tag{3.20}$$

where

$$\hat{c}_{h,k} = \sum_{i=0}^{n-1} \overline{\tilde{g}_{h,k}(t_i)}(X_{t_{i+1}} - X_{t_i})^2 \text{ and}$$

$$c_{h,k} = \int_0^1 \overline{\tilde{g}_{h,k}(s)}\sigma^2(s)\,\mathrm{d}s.$$

We tackle the summands in (3.20) in turn starting with the first one. But first let

$$M_i := \int_{t_i}^{t_{i+1}} b(s)\,\mathrm{d}s, \quad \text{and} \quad S_i := \int_{t_i}^{t_{i+1}} \sigma(s)\,\mathrm{d}W_s,$$

and note that since $X_{t_{i+1}} - X_{t_i} = M_i + S_i$, it follows that

$$(X_{t_{i+1}} - X_{t_i})^2 = M_i^2 + 2M_iS_i + S_i^2.$$

So, (3.20) may be written as

$$v_n(X,t) - \sigma^2(t) = B_{1,n}(t) + B_{2,n}(t) + B_{3,n}(t) + B_{4,n}(t),$$

where

$$\begin{aligned}
B_{1,n}(t) &:= \sum_{(h,k)\in\Theta_n} g_{h,k}(t)\left(\sum_{i=0}^{n-1}\overline{\tilde{g}_{h,k}(t_i)}S_i^2 - c_{h,k}\right),\\
B_{2,n}(t) &:= 2\sum_{(h,k)\in\Theta_n} g_{h,k}(t)\left(\sum_{i=0}^{n-1}\overline{\tilde{g}_{h,k}(t_i)}S_iM_i\right),\\
B_{3,n}(t) &:= \sum_{(h,k)\in\Theta_n} g_{h,k}(t)\left(\sum_{i=0}^{n-1}\overline{\tilde{g}_{h,k}(t_i)}M_i^2\right),\\
B_{4,n}(t) &:= -\sum_{(h,k)\notin\Theta_n} g_{h,k}(t)c_{h,k}.
\end{aligned} \tag{3.21}$$

We start by recalling the well-known fact that frame expansions converge unconditionally in $L^2[0,1]$, that is, the expansion converges regardless of the order of summation (Christensen, 2008, Theorem 5.1.7), so that

$$\|B_{4,n}\|_{L^2[0.1]} = o_{a.s.}(1).$$

We now obtain an estimate for $B_{3,n}(t)$. Suppose without loss of generality that $b_0 = \sigma = 0$ and let $\{T_m\}$ be a localizing sequence for b and σ. Then, by Jensen's inequality

$$\begin{aligned}
E\left(\int_{t_i}^{t_{i+1}} b_{s\wedge T_m}\,\mathrm{d}s\right)^2 &\le \Delta_n E\left(\int_{t_i}^{t_{i+1}} b_{s\wedge T_m}^2\,\mathrm{d}s\right)\\
&\le \Delta_n\int_{t_i}^{t_{i+1}} E(b_{s\wedge T_m}^2)\,\mathrm{d}s\\
&\le \Delta_n\int_{t_i}^{t_{i+1}} E(\sup_{u\le T_m} b_u^4)^{1/2}\,\mathrm{d}s\\
&\le c\Delta_n^2,
\end{aligned} \tag{3.22}$$

where the change in the order of integration is justified by Fubini's theorem. Set $M_i^m = \int_{t_i}^{t_{i+1}} b^2_{s\wedge T_m}\,\mathrm{d}s$ and

$$B^m_{3,n}(t) = \sum_{(h,k)\in\Theta_n} g_{h,k}(t)\left(\sum_{i=0}^{n-1}\overline{\tilde{g}_{h,k}(t_i)}(M_i^m)^2\right)$$

and note that given $\eta > 0$,

$$P(\sup_{t\in[0,1]}|B_{3,n}(t)| > \eta) \leq P(T_m \leq 1) + P(\sup_{t\in[0,1]}|B^m_{3,n}(t)| > \eta).$$

Since $T_m \uparrow \infty$ a.s., the first term on the right becomes arbitrarily small as m tends to infinity. Now since $g_{h,k}$ and $\tilde{g}_{h,k}$ are bounded independently of h and k, and $n\Delta_n = 1$, it follows by Markov's inequality and (3.22) that

$$P(\sup_{t\in[0,1]}|B^m_{3,n}(t)| > \eta) \leq cH^n\Delta_n.$$

Hence,

$$\sup_{t\in[0,1]}|B_{3,n}(t)| = o_P(1). \tag{3.23}$$

Now set $S_i^m := \int_{t_i}^{t_{i+1}} \sigma_{s\wedge T_m} dW_s$. Note the following

$$\begin{aligned}
E((S_i^m)^2) &= E\left(\int_{t_i}^{t_{i+1}} \sigma^2_{s\wedge T_m}\,\mathrm{d}s\right)\\
&= \int_{t_i}^{t_{i+1}} E(\sigma^2_{s\wedge T_m})\,\mathrm{d}s\\
&= \int_{t_i}^{t_{i+1}} (E(\sup_{u\wedge T_m}\sigma_u^4)^{1/2})\,\mathrm{d}s\\
&\leq c\Delta_n. \qquad (3.24)
\end{aligned}$$

By Hölder's inequality, (3.22), and (3.24), we have

$$\begin{aligned}
E(M_i^m S_i^m) &= (E(M_i^m)^2 E(S_i^m)^2)^{1/2}\\
&\leq c\Delta_n^{3/2}. \qquad (3.25)
\end{aligned}$$

Set

$$B^m_{2,n}(t) := 2\sum_{(h,k)\in\Theta_n} g_{h,k}(t)\left(\sum_{i=0}^{n-1}\overline{\tilde{g}_{h,k}(t_i)}S_i^m M_i^m\right).$$

Then for each m, because $g_{h,k}$ and $\tilde{g}_{h,k}$ are bounded independently of h and k, and $n\Delta_n = 1$, we conclude by an appeal to Markov's inequality that $P(\sup_{t\in[0,1]} |B^m_{2,n}(t)| > \eta) \leq cH^n\Delta_n^{1/2}$. By the previously used localization argument,

$$\sup_{t\in[0,1]} |B_{2,n}(t)| = o_P(1). \tag{3.26}$$

Now we tackle the final piece $B_{1,n}(t)$. Let

$$A^n := \sum_{i=0}^{n-1} \overline{\tilde{g}_{h,k}(t_i)} S_i^2 - \int_0^1 \sigma^2(s)\overline{\tilde{g}_{h,k}(s)}\,\mathrm{d}s. \tag{3.27}$$

We will first obtain an upper bound for A^n; we proceed by adding and subtracting $\sum_{i=0}^{n-1}\int_{t_i}^{t_{i+1}} \overline{\tilde{g}_{h,k}(t_i)}\sigma^2(s)\,\mathrm{d}s$ from A to yield:

$$\begin{aligned} A^n &= \sum_{i=0}^{n-1} \overline{\tilde{g}_{h,k}(t_i)} \left(S_i^2 - \int_{t_i}^{t_{i+1}} \sigma^2(s)\,\mathrm{d}s \right) \\ &\quad + \sum_{i=0}^{n-1} \left(\int_{t_i}^{t_{i+1}} \sigma^2(s)\{\overline{\tilde{g}_{h,k}(t_i)} - \overline{\tilde{g}_{h,k}(s)}\}\,\mathrm{d}s \right) \\ &=: A_1^n + A_2^n. \end{aligned}$$

We obtain estimates in turn for the summands. By Assumption 3.1, σ is càdlàg so that it is almost surely bounded on $[0,1]$; by the continuity of $\tilde{g}_{h,k}$ and Lemma (7.1), we have

$$\begin{aligned} A_2^n &= \sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} \sigma^2(s)\{\overline{\tilde{g}_{h,k}(t_i)} - \overline{\tilde{g}_{h,k}(s)}\}\,\mathrm{d}s \\ &\leq c\bar{\omega}(\tilde{g}_{h,k}, \Delta_n), \quad a.s., \end{aligned}$$

where $\bar{\omega}(\tilde{g}_{h,k}, \Delta_n)$ is the modulus of continuity of $\tilde{g}_{h,k}$ on an interval of length Δ_n. By the Lipschitz continuity of g we have,

$$A_2^n = O_{a.s.}(\bar{\omega}(g, \Delta_n)) = O_{a.s.}(\Delta_n).$$

Now, we obtain an estimate for A_1^n. First, let $D_i^n : \Omega \times [0,1] \to \mathbb{R}$ for $i = 0, \cdots, n-1$ be defined as follows:

$$D_i^n(t) := \overline{\tilde{g}_{h,k}(t_i)} \left(\int_{t_i}^t \sigma_{u\wedge T_m}\,\mathrm{d}W_u \right) \mathbb{1}_{(t_i,t_{i+1}]}(t). \tag{3.28}$$

$$D_0^n(0) := 0. \tag{3.29}$$

So, $D_i^n(t)$ is 0 on $[0,1]$ except when t is in $(t_i, t_{i+1}]$. Moreover,

$$D_i^n(t)D_j^n(t) = 0, \qquad i \neq j.$$

Now, for $t \in (t_i, t_{i+1}]$,

$$\begin{aligned}
E(D_i^n(t)^4) &= \overline{\tilde{g}_{h,k}(t_i)}^4 \mathbb{1}_{(t_i,t_{i+1}]}(t) E\left(\left(\int_{t_i}^t \sigma_{u\wedge T_m}\,\mathrm{d}W_u\right)^4\right) \\
&\leq c\mathbb{1}_{(t_i,t_{i+1}]}(t) E\left(\left(\int_{t_i}^t \sigma^2_{u\wedge T_m}\,\mathrm{d}u\right)^2\right) && \text{B.D.G} \\
&\leq c(t-t_i)\mathbb{1}_{(t_i,t_{i+1}]}(t) E\left(\int_{t_i}^t \sigma^4_{u\wedge T_m}\,\mathrm{d}u\right) && \text{Jensen} \\
&\leq c\Delta_n \mathbb{1}_{(t_i,t_{i+1}]}(t) \int_{t_i}^{t_{i+1}} E\left(\sigma^4_{u\wedge T_m}\right)\mathrm{d}u && \text{Fubini} \\
&\leq c\mathbb{1}_{(t_i,t_{i+1}]}(t)\Delta_n^2 && (3.30)
\end{aligned}$$

where the application of Fubini's theorem (Halmos, 1950, Theorem VII.36.B) is justified by the fact that σ^4 is non negative and measurable with respect to the product σ-algebra on $[0,1] \times \Omega$. Now, using Itô's integration by parts formula, we may write

$$\begin{aligned}
E((A_1^n)^2) &= E\left(2\sum_{i=0}^{n-1}\int_{t_i}^{t_{i+1}} \overline{\tilde{g}_{h,k}(t_i)}\left(\int_{t_i}^s \sigma_{u\wedge T_m}\,\mathrm{d}W_u\right)\sigma_{s\wedge T_m}\,\mathrm{d}W_s\right)^2 \\
&= 4E\left(\int_0^1 \sum_{i=0}^{n-1} D_i^n(s)\sigma_{s\wedge T_m}\,\mathrm{d}W_s\right)^2 \\
&\leq c\int_0^1 \sum_{i=0}^{n-1} E(D_i^n(s)^2\sigma^2_{s\wedge T_m})\,\mathrm{d}s \\
&\leq c\int_0^1 \sum_{i=0}^{n-1} E(D_i^n(s)^4)^{1/2}E(\sigma^4_{s\wedge T_m})^{1/2}\,\mathrm{d}s \\
&\leq c\int_0^1 \sum_{i=0}^{n-1} \mathbb{1}_{(t_i,t_{i+1}]}(s)\Delta_n\,\mathrm{d}s \\
&\leq c\Delta_n.
\end{aligned}$$

By Chebyshev's inequality and the previously used stopping time argument, we have $A^n = O_P(\Delta_n)$. By the boundedness of $g_{h,k}$, we have

$$\sup_{t\in[0,1]} |B_{1,n}(t)| = o_P(1).$$

Hence, $B_{j,n}(t)$ for $j = 1, \cdots, 4$, tends to zero in $L^2[0,1]$ in probability. $\square$

4 Volatility estimation: discontinuous prices

In this section we specify a global spot volatility estimator for possibly discontinuous Itô semimartingale price processes. That is, for $t \geq 0$,

$$X_t = X_0 + \int_0^t b_s ds + \int_0^t \sigma_s dW_s + xI_{\{|x|>1\}} * \mu_t + xI_{\{|x|\leq 1\}} * (\mu - \nu)_t$$

with $\nu(dt, dx) = F(dx)dt$ for a determinsistic and constant-in-time σ-finite measure F. We assume σ and b satisfy the requirements of Assumption 3.1, and we further restrict the Lévy system of X as follows:

4.1 Assumption The Lévy measure F satisfies the following condition $(x^2 I_{\{|x|\leq u\}}) * \nu_t = \int_0^t \int_{-u}^u x^2 F(dx)dt = O(u)$ as $u \to 0$.

4.1 Remark The requirement is satisfied if F is absolutely continuous with bounded density f, as is the case with the Gaussian; more generally, it is satisfied if $f(x) = O(x^{-2})$ as $x \to 0$.

As in the preceeding section, we observe a realization of the price process at $n + 1$ equidistant points t_i, $i = 0, 1, \cdots, n$. The observation interval is normalized to $[0, 1]$ with practically no loss of generality. The estimator proposed in the previous section, where there is no jump activity, will not do here. It is inconsistent on account of the presence of jumps; its quality deteriorates as a function of how active the jumps of X are. We will counter this phenomenon with a modified spot variance estimator, but first we introduce the following notation. Let $\Delta_i X$ denote $X_{t_{i+1}} - X_{t_i}$ for $i = 0, 1, \cdots, n-1$, and let u_n be a positive decreasing sequence such that

$$u_n = O(\Delta_n^\beta), \text{ where } \quad 0 < \beta < 1. \tag{4.31}$$

We specify the jump-robust global estimator of spot volatility as follows:

$$V_n(X,t)(t) := \sum_{(h,k)\in\Theta_n} \hat{a}_{h,k}\, g_{h,k}(t), \qquad \forall t \in [0,1], \text{ where} \tag{4.32}$$

$$\hat{a}_{h,k} := \sum_{i=0}^{n-1} \overline{\tilde{g}_{h,k}(t_i)}(\Delta_i X)^2 I_{\{(\Delta_i X)^2 \le u_n\}}, \tag{4.33}$$

where $\{g_{h,k}, \tilde{g}_{h,k}\}$ is a pair of dual Gabor frames constructed as in Lemma (2.1); Θ_n retains its meaning from (3.17); and $I_{\{(\Delta_i X)^2 \le u_n\}}$ is one if $(\Delta_i X)^2$ is less than or equal to u_n and zero otherwise.

There are obvious similarities between $v_n(X,t)$, defined at (3.18), and $V_n(X,t)$ with the key difference being that $V_n(X,t)$ discards realized squared increments over intervals that likely contain jumps; u_n determines the threshold for what is included in the computation and what is not. This determination becomes more accurate as the observation interval becomes infinitessimally small. Clearly it makes sense to use $v_n(X,t)$ if we have reason to believe that the price process is not subject to jumps; $v_n(X,t)$ will always employ all available data and therefore may be assumed to produce more accurate results.

We now proceed to prove the consistency of the estimator. First we introduce the following notation and prove an intermediate lemma. Set

$$\begin{aligned} X_t^c &:= X_0 + \int_0^t b_s \,\mathrm{d}s + \int_0^t \sigma_s \,\mathrm{d}W_s, \\ A_t &:= x I_{\{|x|>1\}} * \mu_t, \\ X_t^f &:= X_t^c + A_t. \end{aligned} \tag{4.34}$$

We now prove the following:

4.1 Lemma *Let X^f be specified as in (4.34) with σ and b satisfying Assumption 3.1. Let $\{g, \tilde{g}\}$ be pair of dual Gabor generators satisfying the conditions of Lemma (2.1) such that g is Lipschitz continuous on the unit interval. If $u_n = O(\Delta_n^\beta), 0 < \beta < 1$, and $H_n \uparrow \infty$ are sequences satisfying*

$$u_n^{-1/2}(H^n)^2 \Delta_n^{1/2} = o(1),$$

then $V_n(X^f,t)$ as defined in (4.32) converges in $L^2[0,1]$ in probability to σ^2.

Proof. We have

$$\begin{aligned} V_n(X^f,t) - \sigma^2(t) = &\{V_n(X^f,t) - V_n(X^c,t)\} + \{V_n(X^c,t) - v_n(X^c,t)\} \\ &+ \{v_n(X^c,t) - \sigma^2(t)\}. \end{aligned} \tag{4.35}$$

That the third summand on the right converges to 0 in $L^2[0,1]$ in probability is the content of Proposition 3.1. Set $\hat{b}_{h,k} := \sum_{i=0}^{n-1} \overline{\tilde{g}_{h,k}(t_i)}(\Delta_i X^c)^2 I_{\{(\Delta_i X^c)^2 \le u_n\}}$ and $\hat{d}_{h,k} := \sum_{i=0}^{n-1} \overline{\tilde{g}_{h,k}(t_i)}(\Delta_i X^c)^2$. Now note that $V_n(X^c,t) - v_n(X^c,t) = \sum_{(h,k)\in\Theta_n} (\hat{b}_{h,k} - \hat{d}_{h,k})\, g_{h,k}(t)$ with

$$\begin{aligned}\hat{b}_{h,k} - \hat{d}_{h,k} &= \sum_{i=0}^{n-1} \overline{\tilde{g}_{h,k}(t_i)}\{(\Delta_i X^c)^2 I_{\{(\Delta_i X^c)^2 \le u_n\}} - (\Delta_i X^c)^2\} \\ &= \sum_{i=0}^{n-1} \overline{\tilde{g}_{h,k}(t_i)}(\Delta_i X^c)^2 I_{\{(\Delta_i X^c)^2 > u_n\}}.\end{aligned}$$

Without loss of generality, suppose $b_0 = \sigma_0 = 0$; let $\{T_m\}$ be a localizing sequence for b and σ. Set $\Delta_i M_m := \int_{t_i}^{t_{i+1}} \sigma_{s\wedge T_m}\, \mathrm{d}W_s$, $\Delta_i S_m := \int_{t_i}^{t_{i+1}} b_{s\wedge T_m}\, \mathrm{d}s$, and $\Delta_i X_m^c := \Delta_i M_m + \Delta_i S_m$. Define $\hat{b}_{h,k}^m - \hat{d}_{h,k}^m$ as above by substituting $\Delta_i X_m^c$ for $\Delta_i X^c$. Now note the following

$$\begin{aligned}E(|\hat{b}_{h,k}^m - \hat{d}_{h,k}^m|) &\le cnE((\Delta_i X_m^c)^2 I_{\{(\Delta_i X_m^c)^2 > u_n\}}) \\ &\le cnE((\Delta_i X_m^c)^4)^{1/2} P((\Delta_i X_m^c)^2 > u_n)^{1/2} \\ &\le cnu_n^{-1/2} E((\Delta_i X_m^c)^4)^{1/2} E((\Delta_i X_m^c)^2)^{1/2}.\end{aligned}$$

Arguing as in Proposition 3.1, it is easily verified that $E((\Delta_i X_m^c)^4) \le c(\Delta_n^4 + \Delta_n^3 + \Delta_n^2)$ and $E((\Delta_i X_m^c)^2) \le c(\Delta_n^2 + \Delta_n^{3/2} + \Delta_n)$. Hence, $E(|\hat{b}_{h,k}^m - \hat{d}_{h,k}^m|) \le cnu_n^{-1/2}\Delta_n^{3/2} = cu_n^{-1/2}\Delta_n^{1/2}$. Because $\tilde{g}_{h,k}$ is bounded, this allows us to conclude by way of Markov's inequality that given $\eta > 0$,

$$P(\sup_{t\in[0,1]} |V_n(X^c,t) - v_n(X^c,t)| > \eta) \le P(T_m \le 1) + cu_n^{-1/2} H^n \Delta_n^{1/2},$$

which becomes arbitrarily small as m and n tend to infinity simultaneously.

To obtain an estimate for the first summand in (4.35), denote $\hat{e}_{h,k} := \sum_{i=0}^{n-1} \overline{\tilde{g}_{h,k}(t_i)}(\Delta_i X^f)^2 I_{\{(\Delta_i X^f)^2 \le u_n\}}$ and observe that $V_n(X^f,t) - V_n(X^c,t) = \sum_{(h,k)\in\Theta_n} (\hat{e}_{h,k} - \hat{b}_{h,k})\, g_{h,k}(t)$ with

$$\hat{e}_{h,k} - \hat{b}_{h,k} = \sum_{i=0}^{n-1} \overline{\tilde{g}_{h,k}(t_i)}\{(\Delta_i X^f)^2 I_{\{(\Delta_i X^f)^2 \le u_n\}} - (\Delta_i X^c)^2 I_{\{(\Delta_i X^c)^2 \le u_n\}}\}.$$

By definition $X^f = X^c + A$, where A represents the jumps of X in excess of 1. We may write $(\Delta_i X^f)^2 I_{\{(\Delta_i X^f)^2 \le u_n\}} - (\Delta_i X^c)^2 I_{\{(\Delta_i X^c)^2 \le u_n\}} = \gamma_i^1 + \gamma_i^2 + \gamma_i^3$ with

$$\begin{aligned}\gamma_i^1 &:= (\Delta_i X^c)^2 (I_{\{(\Delta_i X^f)^2 \le u_n\}} - I_{\{(\Delta_i X^c)^2 \le u_n\}}), \\ \gamma_i^2 &:= (\Delta_i X^c \Delta_i A) I_{\{(\Delta_i X^f)^2 \le u_n\}}, \\ \gamma_i^3 &:= (\Delta_i A)^2 I_{\{(\Delta_i X^f)^2 \le u_n\}}. \qquad (4.36)\end{aligned}$$

Because, X is càdlàg, there is at most a finite number of jumps in excess of 1 per outcome in $[0,1]$. For sufficiently large n, each interval $(t_i, t_{i+1}]$ contains at most one jump. If the i-th interval does not contains a jump then $\gamma_i^2 = \gamma_i^3 = 0$ because $\Delta_i A = 0$. If the i-th interval contains a jump, we have

$$|\Delta_i X^f| = |\Delta_i A + \Delta_i X^c| \geq 1 - |\Delta_i X^c|. \tag{4.37}$$

Now observe that because X^c has continuous paths, it is uniformly continuous on the compact domain $[0,1]$, so that as n tends to infinity, $1-\sup_{i<n} |\Delta_i X^c| \uparrow 1$; meanwhile, $u_n^{1/2} \downarrow 0$. Hence, for n large enough, we have $|\Delta_i X^f| > u_n^{1/2}$ so that, almost surely, γ_i^2 and γ_i^3, for all i, are uniformly eventually zero.

To pin down γ_i^1, we introduce the following events

$$\begin{aligned}
\Omega_n^1 &:= \{\omega : \mu(\omega, (t_i, t_{i+1}] \times \{|x| > 1\}) \leq 1, \text{ for all } i < n\}, & n \in \mathbb{N}, \\
\Omega_n^2 &:= \{\omega : |\Delta_i X^c(\omega)| < 1 - u_n^{1/2}, \text{ for all } i < n\}, & n \in \mathbb{N}, \\
\Omega_k &:= \{\omega : \mu(\omega, [0,1] \times \{|x| > 1\}) \leq k\}. & k \in \mathbb{N}.
\end{aligned}$$

Set $\Omega_n := \Omega_n^1 \cap \Omega_n^2$. As previously argued (see (4.37)), $P(\Omega_n^2) \to 1$ as $n \to \infty$. Because X is càdlàg, $\mu([0,1] \times \{|x| > 1\})$ is almost surely finite, so that $P(\Omega_n^1) \to 1$ as $n \to \infty$. Hence, $P(\Omega_n) \to 1$ as $n \to \infty$. It is also the case that $P(\Omega_k) \to 1$ as $k \to \infty$ since X is càdlàg and the number of jumps larger than one in any bounded interval must be finite almost surely. Now, recall that $\{T_m\}$ is a localizing sequence for b and σ; set $\Omega(m,n,k) := \Omega_n \cap \Omega_k \cap \{T_m > 1\}$ and note that $P(\Omega(m,n,k)) \to 1$ as $n, m, k \to \infty$. Thus, on $\Omega(m,n,k)$ there is at most k jumps larger than one with no more than one jump per interval; the increments of X^c are small enough to ensure the increments of X^f exceed $u_n^{1/2}$; and the processes σ^4 and b^4 are integrable.

Set $\gamma_i^1(n,m,k) = \gamma_i^1 I_{\Omega(m,n,k)}$ and denote $G_i := \{|\Delta_i A| > 0\}$. By triangle inequality, $E(|\gamma_i^1(n,m,k)|) \leq E(|\gamma_i^1(n,m,k) I_{G_i}|) + E(|\gamma_i^1(n,m,k) I_{G_i^c}|)$. Clearly, $\gamma_i^1(n,m,k) = 0$ on G_i^c so that

$$\begin{aligned}
\sum_{i=0}^{n-1} \overline{\tilde{g}_{h,k}(t_i)} E(|\gamma_i^1(n,m,k)|) &\leq \sum_{i=0}^{n-1} \overline{\tilde{g}_{h,k}(t_i)} E(|\gamma_i^1(n,m,k) I_{G_i}|) \\
&= \sum_{i=1}^{k} \overline{\tilde{g}_{h,k}(t_i)} E((\Delta_i X_m^c)^2 I_{\{(\Delta_i X_m^c)^2 \leq u_n\}} I_{G_i}) \\
&\leq \sum_{i=1}^{k} \overline{\tilde{g}_{h,k}(t_i)} E((\Delta_i X_m^c)^2) \\
&\leq ck\Delta_n.
\end{aligned}$$

Hence, given $\eta > 0$,

$$P(\sup_{t\in[0,1]} |V_n(X^f,t) - V_n(X^c,t)| > \eta) \leq P(\Omega(m,n,k)^c) + cH^n k\Delta_n.$$

By taking m, n, k large enough, the first term can be made as small as required; for fixed m, k, letting $n \to \infty$ will make the second term as small as desired. This completes the proof. $\square$

We now prove consistency for the estimator when the price process admits both large and small jumps. That is

$$X_t = X_0 + X_t^c + J_t^l + J_t^s.$$

where $J_t^l := (xI_{\{|x|>1\}}) * \mu_t$ and $J_t^s := (xI_{\{|x|\leq 1\}}) * (\mu - \nu)_t$. We now give the main result of the paper.

4.1 Proposition *Let the price process X be specified as in* (1.1). *We assume that the requirements of Assumption 3.1 and 4.1 are met. Let $\{g, \tilde{g}\}$ be pair of dual Gabor generators satisfying the conditions of Lemma* (2.1) *such that g is Lipschitz continuous on the unit interval. Suppose the sequences $u_n = O(\Delta_n^\beta), 0 < \beta < 1$, and $H_n \uparrow \infty$ satisfy*

$$u_n^{-1/2}(H^n)^2\Delta_n^{1/2} = o(1)$$
$$H^n u_n^{1/2} = o(1)$$

Then $V_n(X,t)$, defined in (4.32), *converges in $L^2[0,1]$ in probability to σ^2.*

Proof. We argue along the lines of Theorem 4 of Mancini (2009). First, consider the following decomposition of the process X:

$$X = X^f + J^s, \tag{4.38}$$
$$X^f = X^c + J^l, \tag{4.39}$$

where $X^c = \int_0^t b_s \,\mathrm{d}s + \int_0^t \sigma_s \,\mathrm{d}W_s$, $J^l = (xI_{|x|>1}) * \mu$, and $J^s = (xI_{|x|\leq 1}) * (\mu - \nu)$. By localization, it is enough to assume σ^4, b^4, and $x^2 * \mu_t$ are integrable. Let t be a point in the unit interval, then

$$V_n(X,t) - \sigma_t^2 = \sum_{(h,k)\in\Theta_n} (\hat{a}_{h,k} - c_{h,k}) g_{h,k}(t) - \sum_{(h,k)\notin\Theta_n} c_{h,k} g_{h,k}(t), \tag{4.40}$$

with $\hat{a}_{h,k}$ and $c_{h,k}$ defined by (4.32) and (3.15), respectively. The last term tends to zero, almost surely, in $L^2[0,1]$ as $n \to \infty$ because Gabor frames converge unconditionally.

To obtain a bound on the first item on the right of (4.40), we may use (4.38) to write

$$\sum_{(h,k)\in\Theta_n} (\hat{a}_{h,k} - c_{h,k}) g_{h,k}(t) = \sum_{(h,k)\in\Theta_n} (w_{h,k} + x_{h,k} + y_{h,k} + z_{h,k}) g_{h,k}(t), \tag{4.41}$$

where

$$\begin{aligned}
w_{h,k} &:= \sum_{i=0}^{n-1} \overline{\tilde{g}_{h,k}(t_i)} (\Delta_i X^f)^2 I_{\{(\Delta_i X^f)^2 \le 4u_n\}} - \int_0^1 \sigma^2(s) \overline{\tilde{g}_{h,k}(s)} \, \mathrm{d}s \\
x_{h,k} &:= \sum_{i=0}^{n-1} \overline{\tilde{g}_{h,k}(t_i)} (\Delta_i X^f)^2 (I_{\{(\Delta_i X)^2 \le u_n\}} - I_{\{(\Delta_i X^f)^2 \le 4u_n\}}) \\
y_{h,k} &:= 2 \sum_{i=0}^{n-1} \overline{\tilde{g}_{h,k}(t_i)} \Delta_i X^f \Delta_i J^s I_{\{(\Delta_i X)^2 \le u_n\}} \\
z_{h,k} &:= \sum_{i=0}^{n-1} \overline{\tilde{g}_{h,k}(t_i)} (\Delta_i J^s)^2 I_{\{(\Delta_i X)^2 \le u_n\}}.
\end{aligned} \tag{4.42}$$

By Lemma (4.1), if $\delta > 0$ then $P(\sup_{t\in[0,1]} |\sum_{(h,k)\in\Theta_n} w_{h,k} g_{h,k}(t)| > \delta) \to 0$ as n tends to infinity. It remains to show that the last three terms on the right of (4.41) converge to zero in probability. Starting with the second summand, denote $A_i := \{(\Delta_i X)^2 \le u_n\}$, $B_i := \{(\Delta_i X^f)^2 \le 4u_n\}$ and note that $I_{A_i} - I_{B_i} = I_{A_i \cap B_i^c} - I_{A_i^c \cap B_i}$. By the reverse triangle inequality, $A_i \cap B_i^c \subset \{|\Delta_i J^s| > u_n^{1/2}\}$. Hence,

$$\begin{aligned}
(\Delta_i X^f)^2 I_{A_i \cap B_i^c} &\le (\Delta_i X^f)^2 I_{\{(\Delta_i J^s)^2 > u_n\}} && (4.43) \\
&\le 2(\Delta_i X^c)^2 I_{\{(\Delta_i J^s)^2 > u_n\}} + 2(\Delta_i J^l)^2 I_{\{(\Delta_i J^s)^2 > u_n\}} \\
&=: v_i + w_i. && (4.44)
\end{aligned}$$

We have,

$$\begin{aligned}
E(v_i) &= 2E((\Delta_i X^c)^4)^{1/2} P((\Delta_i J^s)^2 > u_n)^{1/2} \\
&= 2u_n^{-1/2} E((\Delta_i X^c)^4)^{1/2} E((\Delta_i J^s)^2)^{1/2} \\
&\le c u_n^{-1/2} \Delta_n^{3/2}.
\end{aligned} \tag{4.45}$$

Hence, by Markov's inequality and the boundedness of $g_{h,k}$

$$\sup_{t\in[0,1]} \sum_{(h,k)\notin\Theta_n} \sum_{i=0}^{n-1} \overline{\tilde{g}_{h,k}(t_i)} v_i g_{h,k}(t) = O_P(u_n^{-1/2} H^n \Delta_n^{1/2}). \tag{4.46}$$

Now note that if $\eta > 0$ then

$$
\begin{aligned}
P(\sup_{t\in[0,1]} | \sum_{(h,k)\notin\Theta_n} \sum_{i=0}^{n-1} \overline{\tilde{g}_{h,k}(t_i)} w_i g_{h,k}(t)| > \eta) \\
&\leq P\left(\cup_i\{I_{\{|x|>1\}} * \mu((t_i, t_{i+1}] \times \mathbb{R}) > 0, (\Delta_i J^s)^2 > u_n\}\right) \\
&\leq nP(\mu([0, 1/n] \times \{|x| > 1\}) > 0)E((\Delta_{1/n} J^s)^2)u_n^{-1} \\
&\leq c\Delta_n u_n^{-1},
\end{aligned}
$$

which clearly tends to zero in n. Now, define

$$
\begin{aligned}
\Omega_n^1 &:= \{\omega : |\Delta_i X^f(\omega)| > 2u_n^{1/2}, \forall i < n\}, \\
\Omega_n^2 &:= \{\omega : \mu(\omega, (t_i, t_{i+1}] \times \{|x| > 1\}) \leq 1, \forall i < n\},
\end{aligned}
$$

$\forall n \in \mathbb{N}$. These sets are clearly measurable. Denote $\Omega_n := \Omega_n^1 \cap \Omega_n^2$. Since there can be at most a finite number of jumps larger than 1, in absolute value, per outcome on $[0, 1]$ and $u_n \downarrow 0$ while X^c is uniformly continuous on $[0, 1]$, it follows that $P(\Omega_n) \to 1$ as $n \to \infty$ (See (4.37)). Now note that

$$
\begin{aligned}
A_i^c \cap B_i \cap \Omega_n &\subset \{(\Delta_i X^c + \Delta_i J^s)^2 > u_n\} \\
&\subset \{(\Delta_i X^c)^2 > u_n/4\} \cup \{(\Delta_i J^s)^2 > u_n/4\}.
\end{aligned}
$$

Hence, by successive applications of Hölder and Markov inequaities,

$$
\begin{aligned}
E((\Delta_i X^f)^2 I_{A_i^c\cap B_i\cap\Omega_n}) &= E((\Delta_i X^c)^2 I_{A_i^c\cap B_i\cap\Omega_n}) \\
&\leq E((\Delta_i X^c)^2 I_{\{(\Delta_i X^c)^2>u_n/4\}}) + E((\Delta_i X^c)^2 I_{\{(\Delta_i J^s)^2>u_n/4\}}) \\
&\leq c\Delta_n^{3/2} u_n^{-1/2}.
\end{aligned}
$$

Let η be a given positive number; put $y_i := (\Delta_i X^f)^2 I_{A_i^c\cap B_i}$. Then it is clear that

$$
P(\sup_{t\in[0,1]} | \sum_{(h,k)\in\Theta_n} \sum_{i=0}^{n-1} \overline{\tilde{g}_{h,k}(t_i)} y_i g_{h,k}(t)| > \eta) \leq P(\Omega_n^c) + cu_n^{-1/2} H^n \Delta_n^{1/2},
$$

which tends to zero in n. This completes the proof that

$$
P(\sup_{t\in[0,1]} | \sum_{(h,k)\in\Theta_n} x_{h,k} g_{h,k}(t)| > \eta) \to 0. \tag{4.47}
$$

Now we obtain a bound for the third summand in (4.41). First, denote $C_i := \{(\Delta_i J^s)^2 \leq 4u_n\}$, $p_{h,k} := 2\sum_{i=0}^{n-1} \overline{\tilde{g}_{h,k}(t_i)} \Delta_i X^f \Delta_i J^s I_{A_i\cap C_i}$, and $q_{h,k} :=$

$2\sum_{i=0}^{n-1}\overline{\tilde{g}_{h,k}(t_i)}\Delta_i X^f\Delta_i J^s I_{A_i\cap C_i^c}$. Clearly, $\sum_{(h,k)\notin\Theta_n} y_{h,k}g_{h,k}(t)=\sum_{(h,k)\notin\Theta_n}(p_{h,k}+q_{h,k})g_{h,k}(t)$. By the reverse triangle inequality,

$$\begin{aligned}A_i\cap C_i^c &\subset \{u_n^{1/2}<|\Delta_i X^f|\}\\ &\subset \{u_n^{1/2}/2<|\Delta_i X^c|\}\cup\{u_n^{1/2}/2<|\Delta_i J^l|\}\\ &=: G_i^1\cup G_i^2.\end{aligned}$$

So that

$$\begin{aligned}\Delta_i X^f\Delta_i J^s I_{A_i\cap C_i^c} &\le \Delta_i X^f\Delta_i J^s(I_{G_i^1}+I_{G_i^2})\\ &\le \Delta_i X^c\Delta_i J^s(I_{G_i^1}+I_{G_i^2})+\Delta_i J^l\Delta_i J^s(I_{G_i^1}+I_{G_i^2})\\ &=: \gamma_i^1+\gamma_i^2+\gamma_i^3+\gamma_i^4.\end{aligned}$$

Now,

$$\begin{aligned}E(\gamma_i^1) &\le E((\Delta_i X^c I_{G_i^1})^2)^{1/2}E((\Delta_i J^s)^2)^{1/2}\\ &\le E((\Delta_i X^c)^4)^{1/4}E(I_{G_i^1})^{1/4}E((\Delta_i J^s)^2)^{1/2}\\ &\le c\Delta_n^{1/2}(u_n^{-1/2}\Delta_n^{1/2})\Delta_n^{1/2}\\ &\le cu_n^{-1/2}\Delta_n^{3/2}. \qquad (4.48)\end{aligned}$$

Hence, given positive η,

$$P(\sup_{t\in[0,1]}|\sum_{(h,k)\notin\Theta_n}\sum_{i=0}^{n-1}\overline{\tilde{g}_{h,k}(t_i)}\gamma_i^1 g_{h,k}(t)|>\eta)\le cH^n(u_n^{-1}\Delta_n)^{1/2}.$$

Similarly, we have

$$\begin{aligned}E(\gamma_i^2) &= E(\Delta_i X^c\Delta_i J^s I_{G_i^2})\\ &\le E((\Delta_i X^c)^2 I_{G_i^2})^{1/2}E((\Delta_i J^s)^2)^{1/2}\\ &\le E((\Delta_i X^c)^4)^{1/4}P(\Delta_i J^l>u_n^{1/2}/2)^{1/4}E((\Delta_i J^s)^2)^{1/2}\\ &\le cu_n^{-1/8}\Delta_n^{5/4}.\end{aligned}$$

So that given positive η,

$$P(\sup_{t\in[0,1]}|\sum_{(h,k)\notin\Theta_n}\sum_{i=0}^{n-1}\overline{\tilde{g}_{h,k}(t_i)}\gamma_i^2 g_{h,k}(t)|>\eta)\le cH^n(u_n^{-1/2}\Delta_n)^{1/4}.$$

Observe that,

$$
\begin{aligned}
&P(\sup_{t\in[0,1]} |\sum_{(h,k)\notin\Theta_n}\sum_{i=0}^{n-1}\overline{\tilde{g}_{h,k}(t_i)}\gamma_i^3 g_{h,k}(t)| > \eta) \\
&\qquad \le P(\cup_i\{\mu((t_i,t_{i+1}]\times\{|x|>1\})>0,(\Delta_i X^c)^2 > u_n/4\}) \\
&\qquad \le c\Delta_n u_n^{-1}
\end{aligned}
$$

and

$$
\begin{aligned}
&P(\sup_{t\in[0,1]} |\sum_{(h,k)\notin\Theta_n}\sum_{i=0}^{n-1}\overline{\tilde{g}_{h,k}(t_i)}\gamma_i^4 g_{h,k}(t)| > \eta) \\
&\qquad \le P(\cup_i\{\mu((t_i,t_{i+1}]\times\{|x|>1\})>0,(\Delta_i J^s)^2 > u_n/4\}) \\
&\qquad \le c\Delta_n u_n^{-1}.
\end{aligned}
$$

Now, let $\Psi_n := \{\omega : |\Delta_i X^c(\omega)| > u_n^{1/2} \text{ for some } i < n\}$. Then

$$
\begin{aligned}
P(\Psi_n) &\le \sum_{i=0}^{n-1} P(|\Delta_i X^c| > u_n^{1/2}) \\
&\le u_n^{-3/2(1-\beta)}\sum_{i=0}^{n-1} E((\Delta_i X^c)^{3/(1-\beta)}) \\
&\le c\Delta_n^{1/2}. \qquad (4.49)
\end{aligned}
$$

Hence, $P(\Psi_n) \to 0$. On $A_i\cap C_i\cap\Psi_n^c$, it is easily seen that $|\Delta_i J^l| - |\Delta_i X^c + \Delta_i J^s| < |\Delta_i X| \le u_n^{1/2}$, so that $|\Delta_i J^l| \le u_n^{1/2} + |\Delta_i X^c| + |\Delta_i J^s|$, and $|\Delta_i J^l| = O(u_n^{1/2})$. Let $r_{h,k} := 2\sum_{i=0}^{n-1}\overline{\tilde{g}_{h,k}(t_i)}\Delta_i X^c\Delta_i J^s I_{A_i\cap C_i\cap\Psi_n^c}$ and $s_{h,k} := 2cu_n^{1/2}\sum_{i=0}^{n-1}\overline{\tilde{g}_{h,k}(t_i)}\Delta_i J^s I_{A_i\cap C_i\cap\Psi_n^c}$. Then for $\delta > 0$,

$$
\begin{aligned}
&P(|\sum_{(h,k)\in\Theta_n} p_{h,k}g_{h,k}(t)| > \delta/2) \\
&\qquad \le P(\Psi_n) + P(|\sum_{(h,k)\in\Theta_n} r_{h,k}g_{h,k}(t)| > \delta/4) + P(|\sum_{(h,k)\in\Theta_n} s_{h,k}g_{h,k}(t)| > \delta/4).
\end{aligned}
$$

Now consider that $\sum_{(h,k)\in\Theta_n} r_{h,k}g_{h,k}(t) \le cH_n\sum_{i=0}^{n-1}\Delta_i X^c\Delta_i J^s I_{A_i\cap C_i}$, which implies that

$$
\begin{aligned}
&P(|\sum_{(h,k)\in\Theta_n} r_{h,k}g_{h,k}(t)| > \delta/4) \le P(cH_n|\sum_{i=0}^{n-1}\Delta_i X^c\Delta_i J^s I_{A_i\cap C_i}| > \delta/4) \\
&\qquad \le P\left(\left(\sum_{i=0}^{n-1}(\Delta_i X^c)^2\right)^{1/2}\left(\sum_{i=0}^{n-1}(\Delta_i J^s I_{A_i\cap C_i})^2\right)^{1/2} > \delta(4H_n c)^{-1}\right).
\end{aligned}
$$

It is a well known fact that $\sum_{i=0}^{n-1}(\Delta_i X^c)^2(t)$ converges to $\int_0^t \sigma^2(s)\,\mathrm{d}s$ in probability uniformly on the unit interval. Hence, there is a sufficiently large N such that if $n > N$ then $P(|(\sum_{i=0}^{n-1}(\Delta_i X^c)^2)^{1/2} - (\int_0^1 \sigma^2(s)\,\mathrm{d}s)^{1/2}| > \delta) \leq \varepsilon/4$, and because integrated volatility is almost surely finite, there is a sufficiently large K satisfying $K/2 > \delta$ such that $P(\int_0^1 \sigma^2(s)\,\mathrm{d}s > K/2) \leq \varepsilon/4$. Hence, we may write

$$
\begin{aligned}
&P(|\sum_{(h,k)\in\Theta_n} r_{h,k} g_{h,k}(t)| > \delta/4) \\
&\quad\leq P\left(\sum_{i=0}^{n-1}(\Delta_i J^s I_{A_i\cap C_i})^2 > \delta^2(4KH_n c)^{-2}\right) + \varepsilon/2 \\
&\quad\leq P\left((x^2 I_{\{|x|\leq 1\wedge 2u_n^{1/2}\}}) * \mu_1 > \delta^2(4KH_n c)^{-2}\right) + \varepsilon/2 \\
&\quad\leq \delta^{-2}(4KH_n c)^2 E\left((x^2 I_{\{|x|\leq 1\wedge 2u_n^{1/2}\}}) * \mu_1\right) + \varepsilon/2 \\
&\quad\leq \delta^{-2}(4KH_n c)^2 (x^2 I_{\{|x|\leq 1\wedge 2u_n^{1/2}\}}) * \nu_1 + \varepsilon/2
\end{aligned}
$$

which for sufficiently large n is less than ε. Now it is easily seen that for sufficiently large c

$$
\begin{aligned}
&P(|\sum_{(h,k)\in\Theta_n} s_{h,k} g_{h,k}(t)| > \delta/4) \leq P(\sum_{i=0}^{n-1}\Delta_i J^s I_{A_i\cap C_i} > (8cu_n^{1/2})^{-1}\delta) \\
&\quad\leq (64c^2 u_n)\delta^{-2} E\left((x^2 I_{\{|x|\leq 1\wedge 2u_n^{1/2}\}}) * \mu_1\right) \\
&\quad\leq (64c^2 u_n)\delta^{-2} (x^2 I_{\{|x|\leq 1\wedge 2u_n^{1/2}\}}) * \nu_1 \qquad (4.50)
\end{aligned}
$$

which, as above, is less than $\varepsilon/4$ for sufficiently large n. Hence,

$$
P(\sup_{t\in[0,1]} |\sum_{(h,k)\in\Theta_n} y_{h,k} g_{h,k}(t)| > \delta) \to 0.
$$

Next, write $z_{h,k} = a_{h,k} + b_{h,k}$ where $a_{h,k} := \sum_{i=0}^{n-1} \overline{\tilde{g}_{h,k}(t_i)}(\Delta_i J^s)^2 I_{A_i\cap C_i}$ and $b_{h,k} := \sum_{i=0}^{n-1} \overline{\tilde{g}_{h,k}(t_i)}(\Delta_i J^s)^2 I_{A_i\cap C_i^c}$. Then

$$
\begin{aligned}
&P(|\sum_{(h,k)\in\Theta_n} z_{h,k} g_{h,k}(t)| > \delta) \\
&\quad\leq P(|\sum_{(h,k)\in\Theta_n} a_{h,k} g_{h,k}(t)| > \delta/2) + P(|\sum_{(h,k)\in\Theta_n} b_{h,k} g_{h,k}(t)| > \delta/2).
\end{aligned}
$$

Note,

$$
\begin{aligned}
P(|&\sum_{(h,k)\in\Theta_n} b_{h,k}g_{h,k}(t)| > \delta/2) \\
&\leq P\left(\cup_i\{I_{\{|x|>1\}} * \mu((t_i, t_{i+1}] \times \mathbb{R}) > 0, (\Delta_i J^s)^2 > 4u_n\}\right) \\
&\leq nP(I_{\{|x|>1\}} * \mu([0, 1/n] \times \mathbb{R}) > 0)E((\Delta_{1/n}J^s)^2)(4u_n)^{-1} \\
&\leq cn^{-1}\kappa u_n^{-1}. \qquad (4.51)
\end{aligned}
$$

which can be made as small as desired. Now consider

$$
\begin{aligned}
P(|&\sum_{(h,k)\in\Theta_n} a_{h,k}g_{h,k}(t)| > \delta/2) \\
&\leq P(\sum_{i=0}^{n-1}(\Delta_i J^s)^2 I_{\{|\Delta_i J^s|\leq 2u_n^{1/2}\}} > \delta(2cH_n)^{-1}) \\
&\leq \delta^{-1}(2cH_n)E\left(x^2 I_{\{|x|\leq 1\wedge 2u_n^{1/2}\}} * \mu_1\right) \\
&\leq \delta^{-1}(2cH_n)(x^2 I_{\{|x|\leq 1\wedge 2u_n^{1/2}\}} * \nu)_1 \\
&\leq cH^n u_n^{1/2}
\end{aligned}
$$

which can be made arbitrarily small by the constraints on H^n. Hence,

$$
P(\sup_{t\in[0,1]} |\sum_{(h,k)\in\Theta_n} z_{h,k}g_{h,k}(t)| > \delta) \to 0. \qquad (4.52)
$$

□

5 Simulation

5.1 Continuous prices

In this section, we confirm via simulations the results established analytically. We will first focus on the continuous case to mirror Proposition (3.1). Specifically, we will demonstrate that the mean integrated square error (MISE), the square bias, and the variance of the frame-based estimator tends to zero as the number of obervations increases. We use prices generated by 4 commonly used models of asset prices, namely, the arithmetic Brownian motion (ABM), the Ornstein-Uhlenbeck process (OU), the geometric Brownian motion (GBM), and the Cox-Ingersoll-Ross (CIR) process.

We simulate prices using the following stochastic differential equations:

$$X_t = 0.8 + 0.5t + 0.2W_t, \tag{ABM}$$

$$X_t = 0.8 - \int_0^t 4X_s \,\mathrm{d}s + \int_0^t 0.2 \,\mathrm{d}W_s, \tag{OU}$$

$$X_t = 0.8 + \int_0^t 0.5X_s \,\mathrm{d}s + \int_0^t 0.2X_s \,\mathrm{d}W_s, \tag{GBM}$$

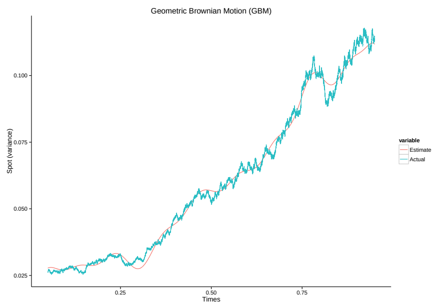
$$X_t = 0.8 + \int_0^t (0.1 - 0.5X_s) \,\mathrm{d}s + \int_0^t 0.2\sqrt{X_s} \,\mathrm{d}W_s, \tag{CIR}$$

where W_t is a standard Brownian motion. For convenience, the observation interval is set to the unit interval $[0, 1]$. In all 4 cases, $X_0 = 0.8$. For each price model, we obtain estimates for the MISE, the square bias, and the variance of the estimator when the number of observations are 500, 5000, and 50000, respectively. In a high-frequency framework, 500 observations for an actively traded stock is likely too small; 5,000 is about right, but 50,000 is not entirely unheard of. At any rate, our objective is not to capture the average number of trades of any particular security, but rather, to obtain support for our asymptotic results by showing an inverse relationship between the number of observations and the MISE, and thereby gain a better understanding of the finite sample behavior of the estimator.

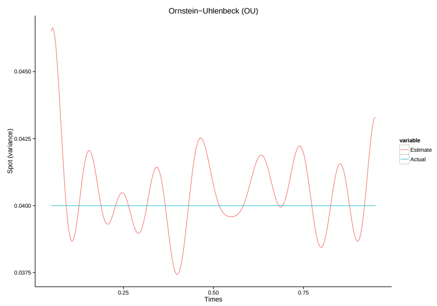
The starting point for constructing the estimator is to fix a generator for the Gabor frame. We have denoted the generator and its dual by g and $\tilde{g}$, respectively. For our purposes, any continuous and compactly supported function would work.

Figure 1: Estimated vs. actual spot volatility

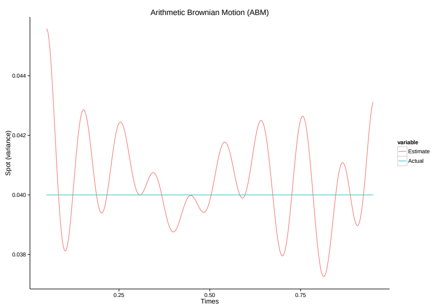
(a) GBM



(b) OU



(c) ABM



(d) CIR

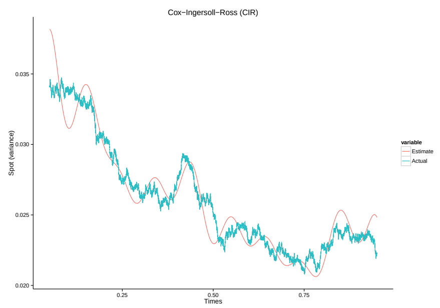


Table 1: Mean integrated square error (MISE) of $v_n(X, t)$.

	ABM			OU		
n	MISE	Sq. Bias	Var	MISE	Sq. Bias	Var
500	1.30×10^{-4}	2.86×10^{-6}	1.27×10^{-4}	1.43×10^{-4}	1.19×10^{-5}	1.31×10^{-4}
5000	1.41×10^{-5}	1.11×10^{-6}	1.30×10^{-5}	1.45×10^{-5}	1.62×10^{-6}	1.28×10^{-5}
50000	2.32×10^{-6}	1.02×10^{-6}	1.30×10^{-6}	2.36×10^{-6}	1.12×10^{-6}	1.23×10^{-6}

	GBM			CIR		
n	MISE	Sq. Bias	Var	MISE	Sq. Bias	Var
500	2.18×10^{-4}	4.18×10^{-6}	2.14×10^{-4}	6.26×10^{-5}	8.51×10^{-7}	6.17×10^{-5}
5000	2.33×10^{-5}	1.58×10^{-6}	2.17×10^{-5}	6.82×10^{-6}	6.00×10^{-7}	6.22×10^{-6}
50000	4.66×10^{-6}	1.02×10^{-6}	3.64×10^{-6}	1.46×10^{-6}	6.06×10^{-7}	8.52×10^{-7}

Note: The mean of the integrated square errors are obtained by taking an average over 100 sample paths generated for each model/number of observations pair.

From an implementation perspective, using a B-spline makes the construction of a *dual* frame generator a trivial matter. This is a consequence of Theorems 2.2 and 2.7 in Christensen (2006), which together specify a very simple rule for constructing dual pairs: Let $a > 0$ and $b > 0$ denote translation and modulation parameters, and let h be a B-spline of order p. Define the dilation operator $\mathcal{D}_c$ as follows:

$$\mathcal{D}_c f(x) = c^{-1/2} f(x/c). \tag{5.53}$$

If $0 < ab \leq 1/(2p-1)$ then $\{\mathcal{D}_a h, \mathcal{D}_a \tilde{h}\}$, where

$$\tilde{h}(x) = abh(x) + 2ab \sum_{n=1}^{p-1} h(x+n), \qquad x \in \mathbb{R}, \tag{5.54}$$

is a pair of dual Gabor frame generators. So if we start with a B-spline h then the dual generator will be a finite linear combination of scaled translates of h; consequently, the dual generator will be a spline, with similar regularity properties. For our simulation, we used a third-order B-spline. Our choice of the third order B-spline is motivated by a desire for a generator with a Fourier transform that decays like a quadratic polynomial. Specifically, we set

$$h(x) = \begin{cases} x^2/2 & x \in (1,0] \\ (-2x^2 + 6x - 3)/2 & x \in (2,1] \\ (3 - x^2)/2 & x \in (3,2] \\ 0 & x \notin (3,0] \end{cases}, \tag{5.55}$$

with $\tilde{h}$ computed as in (5.54) above. Our choice of the modulation and translation parameters is rather arbitrary. The only constraint is that $0 < ab \leq 1/(2p-1) = 1/5$; from our experimentation with different values, performance seems to be about the same for different choices satisfying the inequality; we settled on $a = 1/5$ and $b = 1/3$. Ideally H_n, the order of the number of frequency domain shifts, would be selected optimally to minimize MISE while balancing integrated variance and integrated square bias; this is an open research question. For the time being we set H_n naively equal to 50.

The simulation results indicate that the Gabor frame estimator performs satisfactorily. Figure 1 displays, for each of the 4 price models (ABM, OU, GBM, and CIR), simulated spot variance sample paths plotted against spot variance paths produced by the Gabor frame estimator. A visual inspection shows that the estimator produces a relatively good fit even with the naive selection of H_n. This claim is further corroborated by the analysis of the the integrated mean aquare error (MISE), the integrated square bias, and the integated variance summarized in Table 1.We found that the variance, estimated in the foregoing manner, is only approximately the difference between the MISE and the integrated square bias. The reported figures for variance are in fact the difference between the MISE and the integrated square bias. The discrepancy is rather slight and does not materially change the result. In all 4 model, an inverse relation between MISE, square bias, and variance may be read off from the table. As was established mathematically, we expect MISE to vanish if the number of price observations were made to grow without bound.

5.2 Prices with jumps

We continue our investigation by simulating prices with jumps.

$$X_t = 0.8 + 0.5t + 0.2W_t + \sum_{i=1}^{N} Y_i, \tag{ABM + JMP}$$

$$X_t = 0.8 - \int_0^t 4X_s \,\mathrm{d}s + \int_0^t 0.2 \,\mathrm{d}W_s + \sum_{i=1}^{N} Y_i, \tag{OU + JMP}$$

$$X_t = 0.8 + \int_0^t 0.5X_s \,\mathrm{d}s + \int_0^t 0.2X_s \,\mathrm{d}W_s + \sum_{i=1}^{N} Y_i, \tag{GBM + JMP}$$

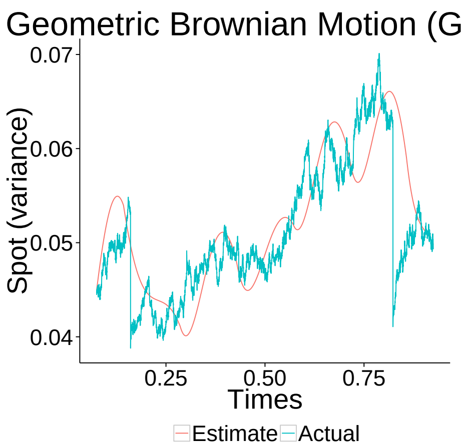
$$X_t = 0.8 + \int_0^t (0.1 - 0.5X_s) \,\mathrm{d}s + \int_0^t 0.2\sqrt{X_s} \,\mathrm{d}W_s + \sum_{i=1}^{N} Y_i, \tag{CIR + JMP}$$

wher N is a Poisson random variable with intensity 5 and Y_i, $1 \leq i \leq N$, is a normal random variable with mean zero and standard deviation 0.4.

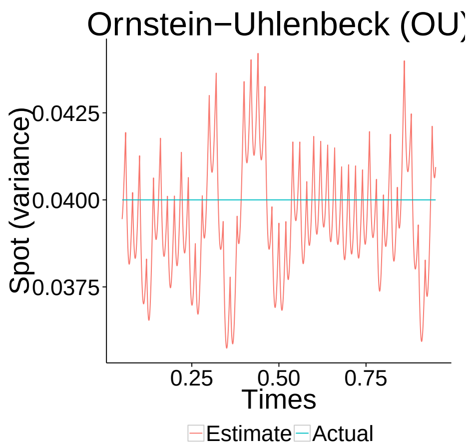
We construct the dual Gabor frames as in the previous subsection using the third order B-Spline specified in (5.55). With the introduction of jumps into the simulation, we found out that better results may be obtained by varying tha parameters a, b, and H_n. We settled on $a = 1/7$, $b = 1/25$, and $H_n = 50$. The jump threshold is obtained by setting $u_n = n^{\alpha}$, where $\alpha = -0.9$. The results of the simulations are recorded in Table 2. We also produce a graph of a single observations (paths) in Figure 2.

Figure 2: Estimated vs. actual spot volatility

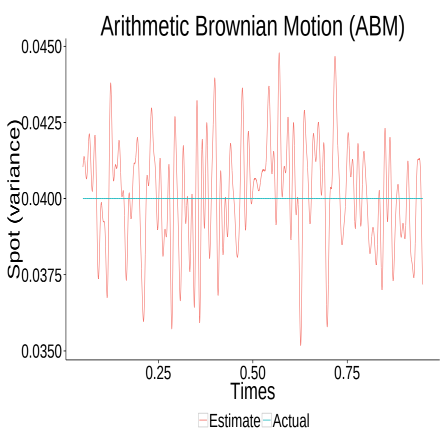
(a) GBM + JMP



(b) OU + JMP



(c) ABM + JMP



(d) CIR + JUMP

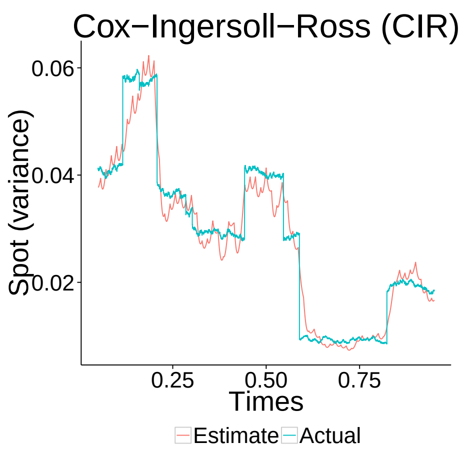


Table 2: Mean integrated square error (MISE) of $V_n(X,t)$.

	ABM + JMP			OU + JMP		
n	MISE	Sq. Bias	Var	MISE	Sq. Bias	Var
500	1.53×10^{-4}	8.95×10^{-6}	1.44×10^{-4}	8.51×10^{-4}	1.31×10^{-4}	7.20×10^{-4}
5000	2.19×10^{-5}	2.27×10^{-6}	1.96×10^{-5}	5.48×10^{-5}	9.76×10^{-6}	4.50×10^{-5}
50000	2.13×10^{-6}	9.00×10^{-8}	2.04×10^{-6}	6.61×10^{-6}	2.65×10^{-6}	3.97×10^{-6}

	GBM + JMP			CIR + JMP		
n	MISE	Sq. Bias	Var	MISE	Sq. Bias	Var
500	6.13×10^{-3}	8.70×10^{-4}	5.26×10^{-3}	3.74×10^{-4}	2.32×10^{-4}	1.43×10^{-4}
5000	3.42×10^{-4}	4.07×10^{-5}	3.02×10^{-4}	1.12×10^{-5}	8.29×10^{-6}	2.95×10^{-6}
50000	7.11×10^{-5}	6.36×10^{-6}	6.47×10^{-5}	7.05×10^{-6}	5.64×10^{-6}	1.40×10^{-6}

Note: The mean of the integrated square errors are obtained by taking an average over 50 sample paths generated for each model/number of observations pair.

6 Conclusion

We have investigated estimators of the instantaneous volatility of asset prices for entire time windows based on Gabor frame expansions of the realized trajectory of spot volatility. The main practical advantage of this type of estimator is their versatility. Once an estimate obtained various functionals of instantaneous volatility such as the ubiquitous integrated volatility are obtained immediately. We derived our estimators of global instantaneous volatility under the assumption that the price process is an Itô semimartingale with Lévy jumps. We have also assumed that the densities of the first and second predictable characteristics belong to the localized class of processes with finite fourth moment.

We proposed a preliminary version of the estimator to be used in situations where the assumption of continuous asset prices hold. Under the assumption that observations of the asset price occur at discrete equidistant intervals with a mesh tending to zero within a fixed time interval, we have shown using standard arguments that the estimator converges in probability in $L^2[0,1]$. In the case of asset prices with discontinuous prices, we modified the basic estimator to require the computation of the Gabor frame coefficients to depend on a threshold. The threshold itself is allowed to shrink to zero at a sufficiently slow rate to ensure consistency of the estimator in $L^2[0,1]$.

References

Barndorff-Nielsen, Ole E. and Shephard, Neil (2004) "Econometric analysis of realized covariation: High frequency based covariance, regression, and correlation in financial economics", *Econometrica*, Vol. 72, No. 3, pp. 885–925.

Christensen, Ole (2006) "Pairs of dual Gabor frame generators with compact support and desired frequency localization", *Applied Computational Harmonic Analysis*, Vol. 20, pp. 403–410.

——— (2008) *Frames and bases: An introductory course*, Boston: Birkhauser.

Daubechies, Ingrid (1992) *Ten lectures on wavelets*: CBMS-NSF Series in Applied Mathematics, SIAM.

Fan, Jianqing and Wang, Yazhen (2008) "Spot volatility estimation for high-frequency data", *Statistics and its Interface*, Vol. 1, pp. 279–288.

Florens-Zmirou, Danielle (1993) "On estimating the diffusion coefficicient from discrete observations", *Journal of Applied Probability*, Vol. 30, No. 4, pp. 790–804.

Foster, Dean P. and Nelson, Dan B. (1996) "Continuous record asymptotics for rolling sample variance estimators", *Econometrica*, Vol. 64, No. 1, pp. 139–174.

Genon-Catalot, V., Laredo, C., and Picard, D. (1992) "Non-parametric estimation of the diffusion coefficient by wavelets methods", *Scandinavian Journal of Statistics*, Vol. 19, No. 4, pp. 317–335.

Halmos, Paul R. (1950) *Measure Theory*: Springer-Verlag New York.

Hoffmann, M., Munk, A., and Schmidt-Hieber (2012) "Adaptive wavelet estimation of the diffusion coefficient under additive error measurements", *Annales de l'institut Henry Poincaré, Probabilités et Statistiques*, Vol. 48, No. 4, pp. 1186–1216.

Malliavin, Paul and Mancino, Maria Elvira (2002) "Fourier series methods for measurement of multivariate volatilities", *Finance and Stochastics*, Vol. 6, No. 1, pp. 49–61.

Mancini, Cecilia (2009) “Non-parametric threshold estimation for models with stochastic diffusion coefficient and jumps”, *Scandinavian Journal of Statistics*, Vol. 36, pp. 270 – 296.

Zhang, Zhihua (2008) “Convergence of Weyl-Heisenberg frame series”, *Indian Journal of Pure and Applied Mathematics*, Vol. 39, No. 2, pp. 167–175.

7 appendix

7.1 Lemma *Let the dual Gabor frame generator* $\tilde{g}$ *be constructed as in* (2.10). *If* $\bar{\omega}(g,\delta)$ *denotes the modulus of continuity of* g*, i.e.* $\bar{\omega}(g,\delta) := \sup\{|g(t)-g(t')| : t,t' \in \mathbb{R}$ *and* $|t-t'|<\delta\}$*, then*

$$\bar{\omega}(\tilde{g}_{j,k},\delta) \leq C\bar{\omega}(g,\delta) \qquad h,k \in \mathbb{Z},$$

where C *is a positive constant.*

Proof. G is bounded away from zero. To see this, note that since g has support in $[r,s]$, the series on the left hand side of (2.11) has finitely many terms for each t. In addition, it is straight forward to verify that $G(t) = G(t+b)$ for all t; so, G is periodic with period b. It is also clear that because g is continuous, so is G. It follows that G attains its min and max on any interval of length b. Let I_b denote the interval $[(s+r-b)/2, (s+r+b)/2]$, then

$$\begin{aligned}\min_{t\in\mathbb{R}} G(t) &= \min_{t\in I_b} G(t)\\ &\geq a^{-1}\min_{t\in I_b}|g(t)|^2.\end{aligned}$$

Because g is continuous and g doesn't vanish in (r,s), we conclude that $G_* := \min_{t\in\mathbb{R}} G(t) > 0$. It is also straight forward that $G^* := \max_{t\in\mathbb{R}} G(t) < \infty$. Now, let $t,t' \in \mathbb{R}$, $t > t'$, such that $|t-t'| \leq \delta$, then

$$\begin{aligned}|\tilde{g}(t)-\tilde{g}(t')| &= |(G(t)G(t'))^{-1}(g(t)G(t')-g(t')G(t))|\\ &\leq (G_*^{-2})\{|g(t)||G(t)-G(t')| + |G(t)||g(t)-g(t')|\}.\end{aligned} \tag{7.56}$$

For a real number x, denote $\lfloor x \rfloor$ the largest integer less than or equal to x and $\lceil x \rceil$ the smallest integer that is greater than or equal to x. Now, Let A denote the set of integers i such that $r < t-ib < s$. By definition of g, $g(t-jb) = 0$, whenever $j \notin A$. Since $b > 0$, A contains at most $\lceil (1+|s|+|r|)/b \rceil$ number of elements. Let $\tau := \min\{t-ib : i \in A\}$, i.e. τ is the smallest $t-ib$ such that $i \in A$. Because A contains at most a finite number of elements, there exists an integer k such that $\tau = t-kb$. Set $\tau' := t'-kb$.

It is straight forward to verify that $|\tau-\tau'| \leq \delta$ and

$$\begin{aligned}a|G(t)-G(t')| &\leq \sum_{j=0}^{\lceil (1+|s|+|r|)/b\rceil} |g(\tau+jb)^2 - g(\tau'+jb)^2|\\ &\leq \sum_{j=0}^{\lceil (1+|s|+|r|)/b\rceil} |g(\tau+jb)-g(\tau'+jb)||g(\tau+jb)+g(\tau'+jb)|\\ &\leq 2\lceil (1+|s|+|r|)/b\rceil g^*\bar{\omega}(g,\delta),\end{aligned} \tag{7.57}$$

where $g^* := \max_{t\in\mathbb{R}} |g(t)|$. Returning to (7.56), we see that

$$|\tilde{g}(t) - \tilde{g}(t')| \le C_{\tilde{g}}\bar{\omega}(g,\delta),$$

where $C_{\tilde{g}} = G_*^2(2a(\lceil(1+|s|+|r|)/b\rceil)(g^*)^2 + G^*)$. Now let $h, k \in \mathbb{Z}$, then

$$\begin{aligned} |\tilde{g}_{h,k}(t) - \tilde{g}_{h,k}(t')| &= |\mathrm{e}^{2\pi \mathrm{i} hat}(\tilde{g}(t-kb) - \tilde{g}(t'-kb))| \\ &\le |\tilde{g}(t-kb) - \tilde{g}(t'-kb)| \le C_{\tilde{g}}\bar{\omega}(g,\delta). \end{aligned} \tag{7.58}$$

The last inequality follows because translating a function leaves its modulus of continuity unchanged.

□