PhD thesis proposal:
On nonparametric volatility estimation, market microstructure noise, and fixed-income market stability.

Wale Dare

July 25, 2015

Contents

Chapter 1

Nonparametric spot volatility estimation by Gabor frames methods

Volatility estimation for discretely observed asset prices has received a great deal of attention recently, but much of that effort has been focused on integrated volatility and, to a lesser extent, the spot volatility at a given point in time. In both cases, the object of interest is local: spot volatility at a given point in time or integrated volatility up to a given terminal time. Notable contributions to the spot and integrated volatility estimation literature include the papers by Foster & Nelson (1996), Fan & Wang (2008), Florens-Zmirou (1993), and Barndorff-Nielsen & Shephard (2004). On the other hand, global approaches, which aim to obtain volatility estimates not just at a point in time but for an entire time window, have received much less coverage. As a result, very little is known about the behaviour of these estimators; in fact, what is known is confined to the case where both the price and volatility processes have continuous paths.

So far, two types of global spot volatility estimators have been proposed. The first of these is the Fourier-based volatility estimator proposed by Malliavin & Mancino (2002) and Malliavin & Mancino (2009). This estimator is built up by first obtaining an estimate of the Fourier series expansion of the price process. The estimated Fourier coefficients of the price process are then used to obtain estimates for the Fourier coefficients of the volatility function. Even though Malliavin & Mancino show that estimates of the individual coefficients in the Fourier expansion of the volatility function converge in a mean square sense, they stopped short of provoding an explicit rate of convergence for the entire volatiltity function. The second class of global estimators is the wavelet-based estimators proposed by Genon-Catalot et al. (1992) and Hoffmann et al. (2012). A lot more is known about this class of estimators; for instance, uniform and integrated mean square rates are well known. In both the Fourier and the wavelet approaches, there is a reliance on orthonormal bases: the Fourier and wavelet orthonormal bases, respectively. Using orthonormal bases is optimal provided the individual coefficients in the orthonormal basis expansion of the volatility function can be estimated with good precision. This may not always be the case when using a finite number of data points to estimate the basis coefficients. The global spot volatility estimator we propose is aimed squarely at this problem by employing a Gabor frame methodology. Frames are very flexible and yield robust estimates when the individual coefficients cannot be estimated with high precision; which is probably the case when using a finite amount of data. Additionally, Gabor frames possess highly efficient noise reduction capabilities. This point becomes important when price measurements are subject to market microstructure noise. We elaborate on these points further below.

The rest of this paper is organized as follows: Section 1.1 gives a description of the price dynamics and observed prices; Section 1.2 reviews the part of the theory of Gabor frames that is pertinent to our study; Section 1.3 gives a specification of the Gabor frame based estimator; Section 1.4 contains a discussion of the convergence of the estimator in a MISE sense; Section 1.5 contains a simulation exercise providing further support for the estimator; in Section 1.6 we use the proposed estimator to gain insight into the diurnal pattern of intra day volatility in the bond market; Section 1.8 concludes and briefly discusses future work. All proofs are relegated to the Appendix.

1.1 Model

Let $\{X_t\}_{t\geq 0}$ be log prices with dynamics given by

$$\mathrm{d}X_t = \mu(t, X_t)\mathrm{d}t + \sigma(t)\mathrm{d}W_t, \qquad X_0 = x, \tag{1.1}$$

where $\{W_t\}$ is a standard Brownian motion with respect to the filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}\}_t, \mathrm{P})$ satisfying the usual conditions; the initial price $x \in \mathbb{R}$ is known; the non-stochastic functions μ and σ are as yet unknown, but assumed to satisfy the Lipschitz and growth conditions sufficient for the existence of a strong solution. We assume prices are observed in the fixed time interval $[0, T]$ at discrete, equidistant times $t_i := i\Delta_n$, where $i = 0, 1, \cdots, n$ and $\Delta_n = T/n$. 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 $[0, T]$ by means of projection methods. We approach this task by estimating the projection of the spot variance in the finite dimensional subspace spanned by finite Gabor frame elements.

1.2 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.

$$f = \sum_{k\in\mathbb{N}} c_k f_k. \tag{1.2}$$

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 (1.2) 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 redundancies they contain. Of course, we end up computing more coefficients but there is no hard limit on the number of coefficients we should compute; we use the same n data points whether we compute k or $k + 10$ coefficients.

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 our computations.

Another reason frames might be a good idea is that high-frequency financial data is seldom without market microstructure noise, while Fourier and wavelet methods have noise reduction capabilities, Gabor frames are particularly efficient in this regards. As a result, Gabor frames can potentially yield much sparser representations of the volatility process when working in a noisy environment. We will not deal explicitly with market microstructure noise here, we will do so in a second paper. Our discussion of frame theory is rather brief, we only mention concepts needed for our specification of the volatility estimator, for a 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

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

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 $\ell^2(\mathbb{Z})$ the set of complex-valued sequences defined over the set of integers $\mathbb{Z}$ with finite norm given by

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

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

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

$$c\|f\|^2 \leq \sum_{k\in\mathbb{N}} |\langle f, f_k\rangle|^2 \leq C\|f\|^2, \qquad \forall f \in L^2(\mathbb{R}).$$

The constants c and C are called *frame bounds*. Now if $\{f_k\}$ is a frame, we may associate with it a bounded operator A that maps every function f in $L^2(\mathbb{R})$ to a sequence c in $\ell^2(\mathbb{N})$ in the following way:

$$Af = c \qquad \text{where} \qquad c_k = \langle f, f_k\rangle.$$

On account of the fact that A takes a function defined on a continuum ($\mathbb{R}$) to a sequence, which is a function defined on the discrete set $\mathbb{N}$, A is known as the *analysis* operator associated with the frame $\{f_k\}_{k\in\mathbb{N}}$. That the analysis operator is bounded follows from the frame bounds in Definition (1.1). Now, the adjoint of A, A^*, is well-defined and takes sequences in $\ell^2(\mathbb{N})$ to $L^2(\mathbb{R})$. Using the fact that A^* must satisfy the equality $\langle Af, c\rangle = \langle f, A^*c\rangle$ for all f and c in $L^2(\mathbb{R})$ and $\ell^2(\mathbb{N})$, respectively, it may be deduced that

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

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

Now an application of the operator $(A^*A)^{-1}$ to every frame element f_k yields a sequence $\{\tilde{f}_k := (A^*A)^{-1}f_k\}_{k\in\mathbb{N}}$, 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 $\tilde{A}$, it is may be shown that $A^*\tilde{A} = \tilde{A}^*A = I$ (the identity operator) and that $\tilde{A}A^* = A\tilde{A}^*$ is the orthogonal projection operator of $\ell^2(\mathbb{N})$ into the range of A, $R(A)$ (Daubechies, 1992, Proposition 3.2.3).

1.2.1 Why frames?

The main reason we might be interested in frame methods for estimating volatility is robustness to coefficient noise; by this we mean the imprecision that may result by virtue of the fact that in practice the frame coefficients may not be known with precision and must be estimated. Coefficient error has many sources including the fact that we have to rely on finite data; we have to store the coefficients with rounding, and market microstructure noise.

1. Frame coefficients can be stored or estimated with low precision, and still result in a reconstruction of that is or realtively higher precision

2. Daubechies (1992) gives the following heuristic explanation.

1.2.2 Gabor frames

Next, we specialize the discussion to Gabor frames. The analysis of Gabor frames involves two operators T and M, called translation and modulation operators, respectively. (T as used here will not be confused with the upper bound of the observation interval $[0, T]$, as the meaning of T will be clear from the context). If $f \in L^2(\mathbb{R})$ then

$$T_b f(t) := f(t - b),$$
$$M_a f(t) := e^{2\pi iat} f(t),$$

for $a, b \in \mathbb{R}$, where $i^2 = -1$. Both T and M are shift operators: T is a shift or translation operator on the time axis, whereas M performs shifts on the frequency axis. A Gabor system is constructed by preforming time-frequency shifts on a single function $g \in L^2(\mathbb{R})$, i.e.

$$\{M_h T_k g\}_{h,k\in\mathbb{Z}}$$

is a Gabor system. A Gabor system need not be a frame.

1.2 Definition *Let $g \in L^2(\mathbb{R})$ and $a, b > 0$ and, for all $t \in \mathbb{R}$, define*

$$g_{h,k}(t) := e^{ihat} g(t - kb), \qquad \forall h, k \in \mathbb{Z},$$

where $i = \sqrt{-1}$. 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 or a Weyl-Heisenberg frame.

The fixed function g is called the *generator* or the *window function*. In order to obtain sharp asymptotic rates, we require g and its dual $\tilde{g}$ to be continuous and compactly supported. The following Lemma taken from Christensen (2006) and Zhang (2008) tells us exactly how to construct such dual pairs.

1.1 Lemma *Let $[r, s]$ be a finite interval, a and b positive constants, and g a continuous function. If $g(t) \neq 0$ when $t \in (r, s)$; $g(t) = 0$ when $t \notin (r, s)$; $a < 2\pi/(s - r)$; and $0 < b < s - r$; then $\{g, \tilde{g}\}$ is a pair of dual Gabor frame generators with*

$$\tilde{g}(t) := g(t)/G(t), \text{ where} \tag{1.3}$$

$$G(t) := (2\pi/a) \sum_{k\in\mathbb{Z}} |g(t - kb)|^2. \tag{1.4}$$

Furthermore,

$$\tilde{g}_{h,k}(\cdot) := e^{ihat} \tilde{g}(\cdot - kb), \qquad \forall h, k \in \mathbb{Z} \tag{1.5}$$

is compactly supported.

Next, we establish that the dual generator $\tilde{g}$ also inherits the continuity properties of g.

1.2 Lemma *Let the dual Gabor frame generator $\tilde{g}$ be constructed as in (1.3). If $\omega_g(\delta)$ denotes the modulus of continuity of g, i.e. $\omega_g(\delta) := \sup\{|g(t) - g(t')| : t, t' \in \mathbb{R}$ and $|t - t'| < \delta\}$, then*

$$\omega_{\tilde{g}_{j,k}}(\delta) = C\omega_g(\delta) \qquad \forall h, k \in \mathbb{Z},$$

where C is a positive constant.

Proof. See Appendix A. □

In the sequel, we assume the Gabor frame setup in Lemma (1.1).

1.3 Volatility estimation

We make the following assumptions about the drift and volatility coefficients explicit.

1.1 Assumption

1. The volatility function is strictly positive, bounded, and continuous.
2. The modulus of continuity of the volatility coefficient, $\omega_\sigma(\Delta_n)$, is equal to $o(1/\log(1/\Delta_n))$ as $n \to \infty$.
3. There is $0 < C_T < \infty$ such that $|\mu(t,x)| \leq C_T(1+|x|)$, for all $t \in [0,T]$ and $x \in \mathbb{R}$.

Since σ^2 may not necessarily be defined on the entire real line, we proceed as in Genon-Catalot et al. (1992) by constructing an extension over the real line that has support in $[0,T]$; this way we are able to apply the Hilbert space machinery in $L^2(\mathbb{R})$. We let $\bar{\sigma}^2$ denote the extension of σ^2, i.e. $\bar{\sigma}^2$ coincides exactly with σ^2 on $[0,T]$ but vanishes outside of $[0,T]$. We summarize these notions as follows:

1.2 Assumption $\bar{\sigma}^2$ is in $L^2(\mathbb{R})$, has support in $[0,T]$, and coincides with σ^2 on $[0,T]$.

With this substitution, we end up with a new process $\bar{X}$ coinciding with X on $[0,T]$ such that

$$\mathrm{d}\bar{X} = \mu(t,\bar{X}_t)\mathrm{d}t + \bar{\sigma}(t)\mathrm{d}W_t, \qquad \bar{X}_0 = x.$$

Now we may avail ourselves of the Gabor frame representation on $L^2(\mathbb{R})$. Let $\{g,\tilde{g}\}$ be a dual Gabor pair constructed as in Lemma (1.1), then $\bar{\sigma}^2$ admits a Gabor frame expansion given by:

$$\bar{\sigma}^2(t) = \sum_{h,k\in\mathbb{Z}} c_{h,k}\; g_{h,k}(t), \text{ where} \tag{1.6}$$

$$c_{h,k} = \langle \bar{\sigma}^2, \tilde{g}_{h,k}\rangle. \tag{1.7}$$

Note that both $\bar{\sigma}^2$ and $\tilde{g}$ have compact support. Indeed $\bar{\sigma}^2$ has support in $[0,T]$, whereas $\tilde{g}$ has support in $[s,r]$. So, $c_{h,k} \neq 0$ only if the supports of $\bar{\sigma}^2$ and $\tilde{g}_{h,k}$ overlap. Furthermore, we note from (1.5) that $\tilde{g}_{h,k+1}$ is simply $\tilde{g}_{h,k}$ shifted by b units; so, $c_{h,k} = 0$ if $|k| > K_0$ with

$$K_0 := \lceil (T + |s| + |r|)/b \rceil, \tag{1.8}$$

where $\lceil x \rceil$ is the smallest positive integer larger than $x \in \mathbb{R}$. Thus $\bar{\sigma}^2$ admits a representation of the form:

$$\bar{\sigma}^2(t) = \sum_{\substack{(h,k)\in\mathbb{Z}^2 \\ |k|\le K_0}} c_{h,k}\, g_{h,k}(t). \tag{1.9}$$

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

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

where H_n is an increasing sequence in n. We propose the following estimator of the volatility coefficient in $[0,T]$:

$$\hat{\sigma}_n^2(t) = \sum_{(h,k)\in\Theta_n} \hat{c}_{h,k}\, g_{h,k}(t), \text{ where} \tag{1.10}$$

$$\hat{c}_{h,k} = \sum_{i=0}^{n-1} \overline{\tilde{g}_{h,k}(t_i)} (X_{t_{i+1}} - X_{t_i})^2. \tag{1.11}$$

In the next section we show that the estimator converges to σ^2 on $[0,T]$ in a mean integrated square error sense.

1.4 Asymptotic properties

Let R_n denote the average integrate square deviation of $\hat{\sigma}_n^2$ from $\bar{\sigma}^2$, i.e.

$$R_n = \mathrm{E} \int_{\mathbb{R}} \{\hat{\sigma}_n^2(t) - \bar{\sigma}^2(t)\}^2 \lambda(t) \mathrm{d}t, \tag{1.12}$$

where λ is a positive and continuous weight function with support in $(0,T)$. The weight function allows us to emphasis different time windows when estimating the volatility. For instance, we may wish to emphasize the recent past in real-time applications. We show that R_n tends to 0 as a function of the sample size, n. Note that R_n is the sum of a bias and a variance component, which we write as follows:

$$R_n = B_n^2 + V_n,$$

where

$$B_n^2 := \int_{\mathbb{R}} (\mathrm{E}[\hat{\sigma}_n^2(t) - \bar{\sigma}^2(t)])^2 \lambda(t) \mathrm{d}t$$

$$V_n := \int_{\mathbb{R}} \mathrm{E}[\{\hat{\sigma}_n^2(t) - \mathrm{E}[\hat{\sigma}_n^2(t)]\}^2] \lambda(t) \mathrm{d}t.$$

1.1 Proposition *Let $\{g, \tilde{g}\}$ be pair of dual Gabor generators constructed as in Lemma* (1.1)*. Suppose the conditions in Assumptions* (1.1) *and* (1.2) *hold. If $H_n^2\Delta_n, H_n\omega_g(\Delta_n)$, and $\omega_{\sigma^2}(1/H_n)\log H_n \to 0$, then the mean integrated square error R_n tends to 0 as n tends to infinity, with*

$$\begin{aligned} &B_n^2 = O(H_n^2\Delta_n + \{H_n\omega_g(\Delta_n)\}^2 + \{\omega_{\sigma^2}(1/H_n)\log H_n\}^2) \\ &V_n = O(H_n^2\Delta_n), \end{aligned} \tag{1.13}$$

where $\omega_g(\delta) := \sup\{|g(t) - g(t')| : t, t' \in \mathbb{R}$ and $|t - t'| < \delta\}$.

Proof. See Appendix A.

□

1.1 Remark

1. First, the above bounds are remarkably similar to those achievable using an orthonormal basis such as wavelets (Genon-Catalot et al., 1992). The variance component is slower by a factor of H_n. This comes about because the vectors in a frame need not be orthogonal. The bias term is slower by a logarithmic factor. Intuitively, the logarithmic term shows up because we are expanding $\bar{\sigma}^2$ using a frame, which may be thought of as containing some redundant term. In practical implementations, this may be a small price to pay for the added flexibility and robustness gained by using frames.

2. Second, this result shows that the variance component of the MISE does not depend on the smoothness properties of either σ^2 and g.

1.5 Simulation study

1.6 Volatility in the bond market

1.7 Extensions

1.7.1 Estimating the spot co-volatility matrix

We propose the following extension to $\hat{\sigma}_n^2$ from Chapter 1 to a multivariate setting. Let $\{X\}_{t\geq 0}$ be a d-dimensional vector of log-prices satisfying the stochastic integral equation:

$$X_t = X_0 + \int_0^t \mu(t, X_s)\mathrm{d}s + \int_0^t \sigma(t)\mathrm{d}W_s, \qquad t \geq 0,$$

where $\{W_s\}_{t\geq 0}$ is an m-dimensional standard Brownian motion; μ is $\mathbb{R}^d$-valued, continuous, and locally bounded in both variables; σ is $\mathbb{R}^d \times \mathbb{R}^m$-valued, continuous and locally bounded in time. Let

$$\Sigma(t) := \sigma(t)\sigma'(t), \qquad t \geq 0$$

where $\sigma'(t)$ is the transpose of $\sigma(t)$. Our aim is to obtain an estimate for Σ on the basis of n discretely and synchronously observed price vectors $\{X_1, \cdots, X_n\}$ in some fixed time interval $[0, T]$, where $T < \infty$. With very little loss of generality we assume that the prices are observed at equidistant intervals given by

$$\Delta_n := T/n. \tag{1.14}$$

Let $\{g_{h,k}, \tilde{g}_{h,k}\}_{h,k\in\mathbb{Z}}$ be a pair of dual Gabor frames generated as in Lemma (1.1), and $\Theta_n := \{(h,k) \in \mathbb{Z}^2 : |h| \leq H_n \text{ and } |k| \leq K_0\}$ with H_n an increasing sequence n and $K_0 = \lceil (T + |s| + |r|)/b \rceil$. We propose to estimate the spot co-volatility matrix in $[0, T]$ using $\hat{\Sigma}_n$ defined component-wise for $1 \leq u, v \leq d$ as follows:

$$\hat{\Sigma}_n^{u,v}(t) = \sum_{(h,k)\in\Theta_n} \hat{c}_{h,k}^{u,v}\, g_{h,k}(t), \text{ where} \tag{1.15}$$

$$\hat{c}_{h,k}^{u,v} = \sum_{i=0}^{n-1} \overline{\tilde{g}_{h,k}(t_i)}(X_u(t_{i+1}) - X_u(t_i))(X_v(t_{i+1}) - X_v(t_i)). \tag{1.16}$$

We conjecture that $\hat{\Sigma}_n$ is consistent for Σ and that it converges in the mean integrated error sense at the same rate of convergence as $\hat{\sigma}_n^2$ (more precisely, the same order of convergence. So actual rate modulo a constant factor which we conjecture to be equal to the number m of driving Brownian motions). A rigorous proof of this conjecture will be given and further substantiated with simulations.

1.8 Conclusion

We proposed an estimator for the spot volatility function using Gabor frame methods. We showed that the estimator converges in a MISE sense and obtained an explicit convergence rate. The evidence for the validity of the proposed estimator will be further reinforced in a simulation study. We will also take the estimator to task using data from the Forex and bond market.

Chapter 2

Market microstructure noise and spot volatility estimation

The need to estimate the volatility of an asset is pervasive in finance. Volatility is the key component in portfolio selection, option pricing, and risk management. Without further restriction on the price process the estimation of the volatility coefficient would be all but impossible. Thankfully financially theory provides some guidance in this regard. At relatively long observation intervals such as a week or a month, it is generally agreed that observed prices may be thought of as discrete realizations from a semimartigales. The theory of semimartingales provides a complete answer on what form we should expect the volatility coefficient of the price process to take. Furthermore, the theory tells us that the usual realized variance estimator converges to the quadratic variation process as the observation interval shrinks to zero (Protter, 2004, Theorem 23).

Unfortunately, this general concensus on the semimartingale status of observed prices breaks down when we consider prices sampled at higher frequencies. The problem is that, at high frequencies, it is hard to justify the assumption that the efficient price process, which may be a semimartingale, is directly observable. Instead, what is observable is part efficient price and part noise resulting from the established processes in the market. The noisy component of observed prices has several sources; these include the so-called bid-ask bound, the release of asynchronous information, and rounding error resulting from discrete prices etc. We elaborate on this sources below. The term *market microstructure noise* was coined by Garman (1976) to describe this type of price contamination. In short, due to the presence of market microstructure noice in high frequency prices, it may no longer be justifiable to assume that observed prices are the discrete realizations of a semimartigale.

There is a large literature dealing with nonparametric estimation of the volatility process in the presence of market microstructure noise. Here too the focus of these efforts has been on obtaining estimates for the integrated volatility. Some of the proposed approaches such as the two-time scale estimator of Zhang et al. (2005) and the pre-averging estimator of Podolskij & Vetter (2007) have been extended to cover local spot volatility estimation; see for example Zu & Boswijk (2014). On the other hand, the study of market-microstructure-robust global volatility estimators has thus far been rather scant. A notable contribution, is the wavelet-based estimator proposed by Hoffmann et al. (2012), which works essentially by estimating the wavelet coefficient using pre-averaged market data. In this chapter, we propose an alternative estimator which combines the Gabor frame estimator from the previous chapter with the multi-time scale procedure popularized by Zhang et al. (2005). We believe that the Gabor frame type estimators truly come into their own when dealing with high frequency data that is corrupted with market microstructure noise. Their highly developed time-frequency localization properties ensures that market microstructure noise occurring as either an independent or dependent process can be effectively tracked and eliminated by a simple thresholding approach.

The rest of this chapter is organized as follows: In Section 2.1 we specify the price process and the structure of the noise process. In Section 2.2 we give a specification of the noise robust spot volatility estimator. In Section 2.3 we conduct a simulation exercise to verify the validity of the proposed estimator. In section 2.4 we use the estimator to gain insight into the diurnal pattern of volatility in the bond market; Section 2.5 concludes the paper. A rigorous proof of the consistency of the estimator will be given in the Appendix.

2.1 Model

We consider the problem of making inference on the spot volatility of a security price process using market data sampled discretely in the presence of market microstructure noise. The presence of the noise component implies that the price process of interest is unobservable directly; instead what we have are discrete transaction or bid and ask price data with market microstructure noise contamination. The usual way this market setup is modeled in the literature is via the Additive Market Microstructure (AMN) model, which as the name implies, states that prices are affected additively by microstructure noise. That is, for $i = 0, 1, \cdots, n$ and $0 = t_0 \leq t_i \leq t_n = T$, the i-th observed price at time t_i may be modeled as:

$$Y_{t_i} = X_{t_i} + \varepsilon_{t_i}, \tag{2.1}$$

where ε_{t_i} is the i-th coordinate in an i.i.d sequence of market microstructure noise. The noise component is assumed to be independent of the efficient price process. The unobserved efficient price of interest X is the unique solution to the following stochastic differential equation:

$$\mathrm{d}X_t = \mu(t, X_t)\mathrm{d}t + \sigma(t)\mathrm{d}W_t, \tag{2.2}$$

where W_t is a Brownian motion, σ is strictly positive, and both μ and σ are continuous and bounded from above. The functions μ and σ are referred to respectively as the drift and volatility functions. This setup is similar to the microstructure environment considered by Zhang et al. (2005) in their estimation of the integrated volatility.

2.1.1 Market microstructure noise

According to Garman (1976), market microstructure noise arises from the moment-to-moment aggregate exchange behavior. Some major sources of market microstructure noise include:

1. *The bid-ask spread.* The price at which an investor can buy an asset, at any fixed point in time, is almost always greater than the price at which he may sell the asset. The *real* or efficient price of the asset is somewhere in between (in some cases, it could be outside the range if there is private information not available to the other participants in the market)

2. *The price impact of trade.* The idea is that each transaction releases information about the underlying asset. For instance, a buyer-initiated transaction tells the market that the asset is more valuable than its current price to somebody. Now, a really big buyer-initiated transaction tells the market that someone with a lot of money and, with no doubt a sophisticated knowledge of the market, thinks the asset is more valuable in the future than its current price. This type of information release can lead to a domino effect where the market goes through several rounds overbidding the price of the asset even though the fundamentals of the asset may not have changed. A pioneering work in the theoretical study of the price impact of trades is the paper by Roll (1984); Hasbrouck (1991) provides an empirically-oriented treatment.

3. *Price round-off* Suppose the market valuation of IBM stock is CHF 19.95666, but because markets price are quoted up to a certain decimal place, the security may be exchanged at say CHF 19.95. Economically,

this seems like a small matter, but implementation-wise this a problem for any statistical procedure relying on the assumption that prices satisfy some form of *recurrence or mixing* property. This is because with prices rounded at 2 decimal places, it is no longer the case that any possible value in the continuous range of the asset price will eventually show up in the data, given enough time. Thus the vast majority of price data in fact will *never* be reported.

4. *Human error* This is especially a problem for prices resulting from trading pits. The chaos of the trading pit almost surely guarantees the occurence of data entry errors throughout the trading day.

2.2 Noise-robust estimator

We start by highlighting the difficulty with which we are faced when trying to estimate the volatility function in the presence of market microstructure noise. The main difficulty in Chapter 1 is that we are attempting to estimate an unobservable or latent volatility function on the basis of discretely observable price process; in the present setting, not even the price process X is observable. What we have a our disposal is data Y, which is part efficient price X and part noise ε. We have no idea how much of the observe price Y is noise and how much is efficient price. It is thus safe to assume that we need to do something different in order to take care of the contamination. Motivated by the ideas in Zhang et al. (2005), we propose to divide the sample n into R_n subsamples containing m_n data points so that $R_n m_n$ is approximately n or $R_n m_n \sim n$. We compute the coefficients of the Gabor frame expansion over these subsamples, average the coefficients out, and use the coefficient estimates over the entire sample to bias-correct the resulting coefficients. More rigorously, we propose the following:

$$\hat{\sigma}^2_{n,b}(t) = \sum_{(h,k)\in\Theta_n} \hat{c}^b_{h,k}\, g_{h,k}(t), \tag{2.3}$$

where

$$\hat{c}^b_{h,k} = \hat{c}^R_{h,k} - (m_n/n)\hat{c}_{h,k} \tag{2.4}$$

$$\hat{c}^R_{h,k} = (1/R_n) \sum_{i=0}^{n-R_n} \overline{\tilde{g}_{h,k}(t_i)}(Y_{t_{i+R_n}} - Y_{t_i})^2 \tag{2.5}$$

$$\hat{c}_{h,k} = \sum_{i=0}^{n-1} \overline{\tilde{g}_{h,k}(t_i)}(Y_{t_{i+1}} - Y_{t_i})^2. \tag{2.6}$$

Here as in Chapter 1, $\{g_{h,k}, \tilde{g}_{h,k}\}_{h,k\in\mathbb{Z}}$ is a pair of dual Gabor frames generated as in Lemma (1.1), and $\Theta_n := \{(h,k) \in \mathbb{Z}^2 : |h| \leq H_n \text{ and } |k| \leq K_0\}$ with H_n an increasing sequence n and $K_0 = \lceil (T + |s| + |r|)/b \rceil$. Any similarities between $\hat{\sigma}_n^2$ from Chapter 1 and $\hat{\sigma}_{n,b}^2$ are at best superficial. Note that $\hat{\sigma}_n^2$ is constructed using the actual efficient price process X, whereas in the present context we have to make do with corrupted market data Y. Also, note that the coefficients $\hat{c}_{h,k}$ play a secondary role here; they merely serve as a devise for bias correction. On the other hand there are strong similarities between the two time scale estimators of Zhang et al. (2005) and Zu & Boswijk (2014). To see this note that we may express (2.5) as follows:

$$\hat{c}_{h,k}^R = (1/R_n) \sum_{i=1}^{R_n-1} \sum_{j=1}^{m_n} \overline{\tilde{g}_{h,k}(t_{i+(j-1)R_n\Delta_n})} (Y_{i+jR_n\Delta_n} - Y_{i+(j-1)R_n\Delta_n})^2.$$

Now it is clear that $\hat{c}_{h,k}^R$ is the average coefficient taken over the R_n coefficients estimates obtained using the R_n subsamples. In the sequel we show using a simulation study the validity of the proposed estimator. We also apply the estimator compute and study diurnal patterns in intra day volatility in the bond market. A rigorous proof of the consistency of the estimator will be given in the Appendix.

2.3 Simulation study

2.4 Diurnal pattern in the bond market revisited

2.5 Conclusion

Chapter 3

Detecting changes in bond market stability

It is a well-known tenet in empirical finance that a large portions of the variability in financial markets results from a handful of risk-factors. In particular, it has been known since, at least, the paper by Litterman & Scheinkman (1991) that fixed-income securities can be explained by no more than four factors. This observation has been further substantiated by the empirical study by Bouchaud et al. (1999). Our aim in this chapter is to formulate and test the predictive power of a test statistic that tracks the evolution of the principal eigen vector of the spot co-volatility matrix.

3.1 Forward rates and bond prices

We fix notation by going over the basics of continuous-time fixed-income pricing. For a detailed presentation see Carmona & Tehranchi (2006) or Heath et al. (1992). Let $P(t, \tau)$ denote the time-t price of a zero-coupon bond (bond) with maturity date τ and unit nominal value, where $0 \leq t \leq \tau < \infty$. Under the assumption that the bond price $P(t, \tau)$ is a smooth function of the maturity date τ, the forward rate $f(t, \tau)$ is given by the instantaneous rate of return on a loan contracted at time t to take effect starting at time τ; it is defined as:

$$f(t, \tau) = -\partial \log P(t, \tau)/\partial T. \tag{3.1}$$

The above definition yields a formula for the unit price of a bond at time t with maturity τ; this is achieved by integrating and then taking the exponential of both sides of (3.1) to yield:

$$P(t,\tau) = \exp\left(-\int_t^\tau f(t,s)\mathrm{d}s\right). \tag{3.2}$$

It is clear from (3.2) that the forward rate fully characterizes bond prices. Specializing to the case where the contract and effective dates coincides yields a characterization of the spot interest rate in terms of the forward rate function. Thus, denoting the spot interest rate by $r(t)$, we have

$$r(t) = f(t,t).$$

It is clear from the above equality that a discussion of the spot rate may be subsumed in a discussion of the forward rate. This is the point of departure of the model of fixed-income markets proposed by Heath et al. (1992) (HJM). We highlight the salient point of the HJM perspective in the next section.

3.2 HJM model of fixed-income markets

The HJM model provides a model for the dynamics of the entire term structure. It is assumed that there is continuum of bonds indexed by maturity date τ taking values in some closed and bounded interval. Bond prices and forward rates are denoted by $P(t,\tau)$ and $f(t,\tau)$, respectively. The forward rate evolution in time is subject to uncertainty arising from m-dimensional standard Brownian vector $(W_1,\cdots,W_d)$, with $m<\infty$, defined on the filtered probability space $\{\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\geq 0},\mathrm{P}\}$, satisfying the usual conditions. To recap, while there are uncountably many bonds and consequently forward rates - one for each maturity date - there are only m risk factors affecting the time evolution of the entire term structure. Given a fixed maturity date τ, the forward rate $f(t,\tau)$ satisfies the following:

$$f(t,\tau) = f(0,\tau) + \int_0^t \mu(t,\tau)\mathrm{d}s + \int_0^t \sum_{j=1}^m \sigma_j(s,\tau)\mathrm{d}W_j(s), \tag{3.3}$$

where $0\leq t\leq\tau$; the drift function is locally integrable and the volatility coefficient is locally square integrable. Heath et al. (1992) showed that the no-arbitrage condition restricts the drift coefficient according to the following:

$$\mu(t,\tau) = \sum_{j=1}^m \sigma_j(t,\tau)\int_t^\tau \sigma_j(t,s)\mathrm{d}s.$$

So, the drift function is fully characterized by the volatility coefficient.

3.3 Realized Spectra

Now, in practice we do not have market data for a continuum of maturities. Most fixed-income instruments are contracted for maturities that are multiples of 3 months, up to a mximum of 30 years; so, in practice we have about 120 maturities. Based on the empirical evidence (Bouchaud et al., 1999; Litterman & Scheinkman, 1991), the principal risk factors, those accounting for the vast majority of the variability, say 90%, are less than 4; so, 120 maturities is probably more than we need to start teasing out what these principal risk factors are, and to monitor how they change or evolve over time. We start by fixing $1 < d < \infty$ maturities for which we have high-frequency data available. We let $[0, T]$ denote the observation interval. We assume the d forward rates are observed synchronously at equidistant intervals. Let $\Sigma(t)$ denote the time-t $d \times d$ spot co-volatility matrix of the d forward rates. Since, $\Sigma(t)$ is symmetric and positive definite, its spectral decomposition is given by:

$$\Sigma(t) = \sum_{i=1}^{d} \lambda_i(t)\{v_i(t) \otimes v_i(t)\}, \qquad t \geq 0,$$

where $\{\lambda(t)_i, v_i(t)\}$ are time-t eigen pairs with $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_d$. Given the empirical evidence aluded to previously, there is $d' << d$ such that

$$\Sigma(t) \approx \sum_{i=1}^{d'} \lambda_i(t)\{v_i(t) \otimes v_i(t)\}, \qquad t \geq 0.$$

Now, because Σ is not observable, we propose using $\hat{\Sigma}_n$ specified in Equation (1.15) of Chapter 1 to estimate Σ. Now given, $\hat{\Sigma}_n$, we propose to estimate the spectrum of Σ by means of the spectrum of $\hat{\Sigma}_n$, which we denote by $\{\hat{\lambda}_{n,j}, \hat{v}_{n,j}\}$ for j between 1 and d. Now let

$$\hat{\chi}_n(t_i) = \frac{\langle \hat{v}_{n,1}(t_i), \hat{v}_{n,1}(t_i - 1)\rangle}{\|\hat{v}_{n,1}(t_i)\| \|\hat{v}_{n,1}(t_i - 1)\|}, \tag{3.4}$$

where $\langle\cdot\rangle$ and $\|\cdot\|$ are, respectively, the Euclidean inner product and norm. The statistic $\hat{\chi}_n(t_i)$ is an empirical construct measuring the cosine between successive realizations of the most important eigen vectors $\hat{v}_{n,1}(t_i)$ and $\hat{v}_{n,1}(t_i - 1)$. The empirical envidence (Carmona & Tehranchi, 2006) suggests that the principal or most important factor can account for a disproportionately portion of total variability; so, it makes sense to track just the most principal vector. Since this eigen vector is a description of the composition of the most important or principal risk factor, we may view $\hat{\chi}_n(t_i)$ as telling us how the direction or composition of the principal risk factor changes over time. Knowledge of this time evolution, if tracked in real-time, could provide actionable information to traders about market direction and stability. Furthermore, Malliavin et al. (2007) showed empirically that the Fourier-based counterpart of $\hat{\chi}_n(t_i)$ is fairly stable when it is business-as-usual in the market; in crisis periods this statistic can exhibit substantial fluctuations. Our aim is to replicate this empirical result and to answer a question that Malliavin et al. (2007) have left unanswered: is the realized velocity of the eigen vector a leading indicator for market returns? Intuitively, changes in volatility need not be reflected immediately in prices since the Brownian motions driving rates may offset the change in volatility for some time before they begin to show in return data. So, we conjecture the question may be answered in the affirmative, but we will have to conduct an empirical study to settle the question. To test this hypothesis, we propose the following simple model:

$$f(t_i, \tau) = \alpha + \sum_{j=1}^{q} \beta_j \hat{\chi}_n(t_{i-j}) + \sum_{j=1}^{p} \gamma_j f(t_{i-j}, \tau) + \eta_i, \tag{3.5}$$

where eta_i is i.i.d error with constant variance and zero mean; p and q are positive integers; and α, $\{\beta_j\}_{j=1}^{q}$ and $\{\gamma_j\}_{j=1}^{p}$ are constants to be determined and tested for statistical significance.

3.4 Empirical study

3.5 Conclusion

Bibliography

Applebaum, David (2009) *Lévy processes and stochastic calculus*: Cambridge University Press, 2nd edition.

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.

Bouchaud, Jean-Philippe, Sagna, Nicolas, Cont, Rama, El-Karoui, Nicole, and Potters, Marc (1999) "Phenomenology of the interest rate curve", *Applied Mathematical Finance*, Vol. 6, pp. 209–232.

Carmona, René and Tehranchi, Michael (2006) *Interest rate models: an infinite dimensional stochastic analysis perspective*, Berlin Heidelberg: Springer-Verlag.

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.

Garman, Mark B. (1976) "Market microstructure", *Journal of Financial Economics*, Vol. 3, No. 3, pp. 257–275.

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.

Hasbrouck, Joel (1991) "Measuring the information content of stock trades", *The Journal of Finance*, Vol. 46, No. 1, pp. 179–207.

Heath, David, Jarrow, Robert, and Morton, Andrew (1992) "Bond pricing and the term structure of interest rates: A new methodology for continuous claims valuation", *Econometrica*, Vol. 60, No. 1, pp. 77–105.

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.

Litterman, Robert B. and Scheinkman, José (1991) "Common factors affecting bond market returns", *The Journal of Fixed Income*, Vol. 1, No. 1, pp. 54–61.

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.

——— (2009) "A Fourier transform method for nonparametric estimation of multivariate volatility", *Annals of Statistics*, Vol. 37, No. 4, pp. 1983 – 2010.

Malliavin, Paul, Mancino, Maria Elvira, and Recchioni, Maria Cristina (2007) "A non-parametric calibration of the hjm geometry: an application of itô calculus to financial statistics", *Japanese Journal of Mathematics*, Vol. 2, pp. 55–77.

Podolskij, Mark and Vetter, Matthias (2007) "Estimation of the volatility functional in the simultaneous presence of microstructure noise and jumps", *CREATES Research Papers, School of Economics and Management, Aarhus University.*

Protter, Philip E. (2004) *Stochastic Integration and Differential Equations*: Springer, 2nd edition.

Roll, Richard (1984) "A simple implicit measure of the effective bid-ask spread in an efficient market", *Journal of Finance*, Vol. 39, No. 4, pp. 1127–1139.

Zhang, Lan, Mykland, Per A., and Aït-Sahalia, Yacine (2005) "A tale of two time scales: Determining integrated volatility with noisy high-frequency data", *Journal of the American Statistical Association*, Vol. 100, No. 472, pp. 1394 – 1411.

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

Zu, Yang and Boswijk, Peter H. (2014) "Estimating spot volatility with high-frequency financial data", *Journal of Econometrics*, Vol. 181, No. 2, pp. 117–135.

Appendices

Appendix A

Proofs

We now give the proof of Lemma (1.2).

Proof. Note that G is bounded away from zero. Indeed, since g has support in $[r, s]$, the series on the left hand side of (1.4) 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 $I_b := [(s + r - b)/2, (s + r + b)/2]$, then

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

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 $t, t' \in \mathbb{R}$ such that $|t - t'| \leq \delta$, then

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

Let $\tau := r + (t \bmod b)$, and $\tau' := r + (t' \bmod b)$. It is straight forward to verify that if $|\tau - \tau'| \leq \delta$, then

$$\begin{aligned}|G(t) - G(t')| &\leq \sum_{j=0}^{\lfloor (s+r)/b\rfloor} |g(\tau + jb)^2 - g(\tau' + jb)^2|\\ &\leq \sum_{j=0}^{\lfloor (s+r)/b\rfloor} |g(\tau + jb) - g(\tau' + jb)||g(\tau + jb) + g(\tau' + jb)|\\ &\leq 2\lceil (s + r)/b\rceil g^* \omega_g(\delta), \quad \text{(A.2)}\end{aligned}$$

where $g^* := \max_{t\in\mathbb{R}} |g(t)|$. On the other hand, if $|\tau - \tau'| > \delta$, then

$$\begin{aligned}
&|G(t) - G(t')| \\
&\quad \leq |g(\tau')^2 - g(r)^2| + |g(s)^2 - g(\tau + c)^2| \\
&\qquad + \sum_{j=1}^{\lfloor (s+r)/b \rfloor} \{|g(\tau + (j-1)b)^2 - g(\tau' + jb)^2|\}.
\end{aligned}$$

where $c = \lfloor (s + r)/b \rfloor b$. It follows as above that

$$|G(t) - G(t')| \leq 2(\lceil (s + r)/b\rceil + 1)g^* \omega_g(\delta, T). \tag{A.3}$$

Returning to (A.1), we see that

$$|\tilde{g}(t) - \tilde{g}(t')| \leq C_{\tilde{g}} \omega_g(\delta),$$

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

$$\begin{aligned}
|\tilde{g}_{h,k}(t) - \tilde{g}_{h,k}(t')| &= |e^{ihat}(\tilde{g}(t - kb) - \tilde{g}(t' - kb)| \\
&\leq |(\tilde{g}(t - kb) - \tilde{g}(t' - kb)| \leq C_{\tilde{g}} \omega_g(\delta).
\end{aligned} \tag{A.4}$$

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

□

Next, we provide a proof of Proposition (1.1).

Proof. We begin with B_n^2, the bias component of the integrated mean square error. Using Itô's product formula (Applebaum, 2009, p. 257), we may express, using Itô's isometry theorem, the average deviation at time t as follows:

$$\mathrm{E}[\hat{\sigma}_n^2(t) - \bar{\sigma}^2(t)] = B_{1,n}(t) + B_{2,n}(t) + B_{3,n}(t),$$

where

$$\begin{aligned}
B_{1,n}(t) &:= \sum_{(h,k)\in\Theta_n} g_{h,k}(t) \left\{ \sum_{i=0}^{n-1} \overline{\tilde{g}_{h,k}(t_i)} \int_{t_i}^{t_{i+1}} \sigma^2(s)\mathrm{d}s - c_{h,k} \right\}, \\
B_{2,n}(t) &:= 2 \sum_{(h,k)\in\Theta_n} g_{h,k}(t) \left\{ \sum_{i=0}^{n-1} \overline{\tilde{g}_{h,k}(t_i)} \mathrm{E}\left[\int_{t_i}^{t_{i+1}} (X_s - X_{t_i})\mu(s, X_s)\mathrm{d}s \right] \right\}, \\
B_{3,n}(t) &:= - \sum_{(j,k)\notin\Theta_n} g_{h,k}(t) c_{h,k}.
\end{aligned} \tag{A.5}$$

The first two components, $B_{1,n}(t)$ and $B_{2,n}(t)$, result from the fact that X is being observed discretely; whereas $B_{3,n}$ results because only a finite number of the frame elements are being used in the approximation. We refer to Theorem 4.1 in Zhang (2008) for an estimate of $B_{3,n}$:

$$B_{3,n} = O(\omega_{\bar{\sigma}^2}(1/H_n)\log H_n).$$

So, the smoother the volatility coefficient the smaller the number of frame elements needed to obtain a decent approximation. We obtain a bound on $B_{1,n}$ by noting that

$$\begin{aligned}\sum_{i=0}^{n-1}\overline{\tilde{g}_{h,k}(t_i)}\int_{t_i}^{t_{i+1}}\sigma^2(s)\mathrm{d}s - c_{h,k} &= \sum_{i=0}^{n-1}\int_{t_i}^{t_{i+1}}\sigma^2(s)\{\overline{\tilde{g}_{h,k}(t_i)} - \overline{\tilde{g}_{h,k}(s)}\}\mathrm{d}s\\ &\le C_\beta\omega_{\tilde{g}_{h,k}}(\Delta_n),\end{aligned}$$

where $C_\beta = \int_0^T \sigma^2(s)\mathrm{d}s < \infty$. It follows from Lemma (1.2) that

$$B_{1,n}(t) \le C_B H_n \omega_g(\Delta_n)$$

with $C_B = 2(2K_0+1)g^*\tilde{g}^*C_\beta C_{\tilde{g}}$. Next, we estimate $B_{2,n}(t)$. Note that

$$X_s - X_{t_i} = \int_{t_i}^{s}\mu(u, X_u)\mathrm{d}u + \int_{t_i}^{s}\sigma(u)\mathrm{d}W_u.$$

So, we may write

$$\begin{aligned}&\mathrm{E}\left[\int_{t_i}^{t_{i+1}}(X_s - X_{t_i})\mu(s, X_s)\mathrm{d}s\right]\\ &\quad= \mathrm{E}\left[\int_{t_i}^{t_{i+1}}\left(\int_{t_i}^{s}\mu(u, X_u)\mathrm{d}u\right)\mu(s, X_s)\mathrm{d}s\right]\\ &\qquad+ \mathrm{E}\left[\int_{t_i}^{t_{i+1}}\left(\int_{t_i}^{s}\sigma(u)\mathrm{d}W_u\right)\mu(s, X_s)\mathrm{d}s\right]\\ &\quad=: \beta_{2,i} + \beta_{3,i}.\end{aligned}$$

By Fubinni's theorem, the Cauchy-Schwarz inequality, Itô's isometry theorem, and the linear growth condition on the drift, we have

$$\begin{aligned}\beta_{3,i} &\le C_T\int_{t_i}^{t_{i+1}}\left(\int_{t_i}^{s}\sigma^2(u)\mathrm{d}u\right)^{1/2}\mathrm{E}[(1+|X_s|)^2]^{1/2}\mathrm{d}s\\ &\le C_3\Delta_n^{3/2}.\end{aligned}\tag{A.6}$$

Similarly, it may be verified that $\beta_{2,i} \leq C_2\Delta_n^2$. Now we may write

$$
\begin{aligned}
B_{2,n}(t) &\leq \Delta_n^{1/2} \sum_{(h,k)\in\Theta_n} g_{h,k}(t) \left\{ \sum_{i=0}^{n-1} \overline{\tilde{g}_{h,k}(t_i)} C_4 \Delta_n \right\} \\
&\leq ((2K_0+1)C_4 g^* \tilde{g}^* T)(2H_n+1)\Delta_n^{1/2} \\
&= O(H_n \Delta_n^{1/2}). \qquad \text{(A.7)}
\end{aligned}
$$

Thus the square bias is bounded as follows:

$$
\begin{aligned}
B_n^2 &= \int_R (B_{1,n}(t) + B_{2,n}(t) + B_{3,n}(t))^2 \lambda(t) \mathrm{d}t \\
&= O(H_n^2 \Delta_n + H_n^2 \omega_g^2(\Delta_n) + \omega_{\tilde{\sigma}^2}^2(1/H_n) \log^2 H_n). \qquad \text{(A.8)}
\end{aligned}
$$

Next, we obtain a bound for the variance term. Note that V_n may be expressed as follows:

$$
V_n = V_{1,n} + V_{2,n}
$$

with

$$
\begin{aligned}
V_{1,n} &:= \sum_{(h,k)\in\Theta_n} \mathrm{var}[\hat{c}_{h,k}] \gamma_{h,k}^2, \text{ and} \\
V_{2,n} &:= \sum_{\substack{(h,k),(h',k')\in\Theta_n \\ (h,k)\neq(h',k')}} \mathrm{cov}[\hat{c}_{h,k}, \hat{c}_{h',k'}] \gamma_{h,k} \gamma_{h',k'},
\end{aligned}
$$

where $\gamma_{h,k} := \int_0^T g_{h,k}(t)\mathrm{d}t$. We start with $V_{1,n}$. If we set

$$
\begin{aligned}
Y_i &:= \left(\int_{t_i}^{t_{i+1}} \sigma(s) \mathrm{d}W_s \right)^2 \qquad \text{and} \\
Z_i &:= \left(\int_{t_i}^{t_{i+1}} \mu(s, X_s) \mathrm{d}s \right)^2 + 2 \left(\int_{t_i}^{t_{i+1}} \mu(s, X_s) \mathrm{d}s \right) \left(\int_{t_i}^{t_{i+1}} \sigma(s) \mathrm{d}W_s \right),
\end{aligned}
$$

then we may write

$$
\hat{c}_{h,k} = \sum_{i=0}^{n-1} g_{h,k}(t_i)(Y_i + Z_i).
$$

Furthermore, setting

$$
\alpha_{1,i} := \sum_{(h,k)\in\Theta_n} g_{h,k}^2(t_i) \gamma_{h,k}^2,
$$

allows us to write

$$V_{1,n} := \sum_{i=0}^{n-1} \alpha_{1,i}(\mathrm{var}[Y_i] + \mathrm{var}[Z_i] + 2\mathrm{cov}[Y_i, Z_i]).$$

Using the Cauchy-Schwarz inequality, we obtain

$$\mathrm{var}[Y_i] = \left(\int_{t_i}^{t_{i+1}} \sigma^2(t_i)\mathrm{d}s\right)^2 = O(\Delta_n^2); \tag{A.9}$$

$$\mathrm{var}[Z_i] \le \mathrm{E}[Z_i^2] = O(\Delta_n^3); \tag{A.10}$$

$$\mathrm{cov}[Y_i, Z_i] \le (\mathrm{var}[Z_i]\mathrm{var}[Y_i])^{1/2} = O(\Delta_n^{5/2}). \tag{A.11}$$

Since $\alpha_{1,i} = O(H_n)$, we conclude that $V_{1,n} = O(H_n\Delta_n)$. It is clear that $V_{2,n}$ may be bounded in a similar fashion. Indeed, let

$$\alpha_{2,i} := \sum_{\substack{(h,k)\in\Theta_n \\ (h,k)\neq(h',k')}} g_{h',k'}(t_i)g_{h,k}(t_i)\gamma_{h',k'}\gamma_{h,k},$$

then we may write

$$V_{2,n} := \sum_{i=0}^{n-1} \alpha_{2,i}(\mathrm{var}[Y_i] + \mathrm{var}[Z_i] + 2\mathrm{cov}[Y_i, Z_i]).$$

Since $\alpha_{2,i} = O(H_n^2)$, it follows that $V_{2,n} := O(H_n^2\Delta_n)$.

$\square$