10-2017

Estimation of a noisy subordinated Brownian motion via two-scales power variations

José E. Figueroa-López
Washington University in St. Louis, figueroa-lopez@wustl.edu

Kiseop Lee

Estimation of a noisy subordinated Brownian Motion via two-scale power variations

Jose E. Figueroa-López[†] **and Kiseop Lee**[‡*]

[†]*Department of Statistics*
Purdue University
W. Lafayette, IN 47907, USA
figueroa@purdue.edu

[‡]*Department of Mathematics*
University of Louisville
Louisville, KY, 40292, USA
kiseop.lee@louisville.edu

August 1, 2014

Abstract: High frequency based estimation methods for a pure-jump subordinated Brownian motion exposed to a small additive microstructure noise are developed building on the two-scale realized variations approach developed by Zhang *et al.* (2005) for the estimation of a continuous Itô process. The proposed estimators are shown to be robust against the noise and to attain better rates of convergence than those of standard method of moment estimators *even in the absence of noise*. Our main results give approximate optimal values for the number K of regular sparse subsamples to be used, which is a crucial tune-up parameter of the method. Finally, a two-step data-driven procedure is devised to implement the proposed estimators with the optimal K-value. The superior finite-sample performance and empirical robustness of the resulting estimators are illustrated by Monte Carlo simulations and a real high-frequency data application.

Keywords and phrases: Geometric Lévy Models, Kurtosis and Volatility Estimation, Power Variation Estimators, Microstructure Noise, Robust Estimation Methods.

1. Introduction

In this paper, we propose some estimation methods for a subordinated Brownian motion (SBM), whose sampling observations have been contaminated by a small additive noise ε. In addition to a "volatility" type parameter σ, which accounts for the variability of a time series of the process' increments, a SBM is endowed with an additional parameter, denoted by κ, which accounts for the tail heaviness of the increments' distribution. Therefore, κ determines the proneness of the process to produce extreme increment observations. Such a measure is clearly of critical relevance in many applications such as risk management. Mathematically, σ measures the variance of the increments divided by the time span δ of the increments, while κ measures the excess kurtosis of the increments multiplied by the time span δ, so that increments at higher-frequency exhibit larger kurtosis and, hence, heavier distribution tails. It is important to emphasize that the models considered here are pure-jump Lévy models and σ is not the volatility of a continuous

[*]The first author's research is partially supported by the NSF grant DMS-1149692.

component. Nevertheless, given that σ determines the variability of a time series of log-returns under an exponential SBM, it is natural to call σ the volatility parameter of the model.

As in the context of a regression model, the additive noise ε, typically called microstructure noise, can be seen as a modeling artifact to account for the deviation of the actual price formation process from the SBM. In some circumstances, the noise can also be given some concrete interpretation based on the actual trading mechanism such as in the case of bid/aks bounce effects (cf. Roll (1984)). At low frequencies the microstructure noise is relatively negligible compared to the SBM's observations but at high-frequency the noise is significant and heavily tilts any estimates that do not account for it. The aim is then to develop inference methods that are robust against potential microstructure noise components.

The literature of statistical estimation methods under microstructure noise has grown extensively during the last decade. See, for instance, Aït-Sahalia *et al.* (2005), Zhang *et al.* (2005), Hansen & Lunde (2006), Bandi & Russell (2008), Mykland & Zhang (2012), to mention just a few. Among these, the problem of estimating the integrated variance has received a great deal of attention. However, to the best of our knowledge, there is almost no paper that focus on the estimation of the volatility and kurtosis parameters of a pure-jump Lévy model in the presence of a high-frequency noise component as it is the case in this paper. The empirical performance of some standard parametric methods, in the absence of a microstructure noise, has been studied in a few works such as and Seneta (2004), Ramezani & Zeng (2007), Behr & Pötter (2009), and Figueroa-López *et al.* (2011).

We begin our analysis with a derivation of approximate Method of Moment Estimators (MME) for σ^2 and κ, respectively denoted by $\hat{\sigma}^2_{n,T}$ and $\hat{\kappa}_{n,T}$ throughout the remainder of the introduction. MMEs and related estimators are widely used in high-frequency data analysis due to their simplicity, computational efficiency, and known robustness against potential correlation between observations. In fact, moment type estimators are arguably the only possible estimators that can efficiently be applied for high-frequency data due to the extremely high computational and numerical burden associated with the volume of such data.

Once the MMEs have been introduced, we analyze the behavior of the estimators when $n \to \infty$ (infill asymptotics) and when $T \to \infty$ (long-run asymptotics) both in the absence and presence of microstructure noise. We identify the order $O(T^{-1})$, when $T \to \infty$, as the rate of convergence of the estimators (in the mean-squared error sense) under the ideal situation of absence of noise. Hence, the goal is to develop estimators that are able to achieve at least this rate of convergence in the presence of microstructure noise. The asymptotic analysis of the estimators in the presence of noise allows to qualitatively characterize the behavior of $\hat{\sigma}^2_{n,T}$ and $\hat{\kappa}_{n,T}$. In particular, we found that $\hat{\sigma}^2_{n,T} \to \infty$ and $\hat{\kappa}_{n,T} \to 0$, as $n \to \infty$, both of which are stylized facts validated using real high-frequency observations (see Section 6.3 below). Furthermore, denoting δ_n the time span between observations, it turns out that $\delta_n \hat{\sigma}^2_{n,T}$ and $\delta_n^{-1} \hat{\kappa}_{n,T}$ converge to the second moment and the excess kurtosis of the microstructure noise, respectively. The latter properties will be useful to devise noise-robust estimators for σ^2 and κ.

In order to develop estimators that are robust to a microstructure noise component, we borrow ideas from Zhang *et al.* (2005)'s seminal approach based on combining the realized quadratic variations at two-scales or frequencies. More concretely, the idea consists of the three steps. First, we break up the high-frequency sampling observations in K groups of observations taken at a lower frequency (sparse subsampling). Second, the relevant estimators (say, realized quadratic variations) are applied to each group and the resulting K point estimates are averaged. Finally, a bias correction step is necessary for which one typically uses the estimators at the highest possible frequency.

A fundamental problem in the approach described in the previous paragraph is how to determine the number of subgroups, K, which strongly affects the performance of the estimators. We propose a method to find approximate optimal values for K. For the estimator of σ^2, it is found that the optimal K takes the form

$$K_\sigma^* := n^{\frac{2}{3}} \left(\frac{6 \left(\mathbb{E}\varepsilon^4 + (\mathbb{E}\varepsilon^2)^2 \right)}{T^2 \sigma^4} \right)^{\frac{1}{3}}, \tag{1.1}$$

where ε represents the additive microstructure noise associated to one observation of the SBM. It is also found that the mean-squared error (MSE) of the resulting estimator (using K as above) attains a rate of convergence $C_\sigma \left(\mathbb{E}\varepsilon^4 + (\mathbb{E}\varepsilon^2)^2 \right)^{\frac{1}{3}} n^{-\frac{1}{3}} T^{-\frac{2}{3}}$ (up to a constant C_σ), which, since $T/n \to 0$, shows the surprising fact that the estimator converges at a rate of $o(T^{-1})$, which is faster than the rate attained by the MMEs in the absence of noise. For the estimation of κ, it is found that the optimal K takes the form

$$K_\kappa^* = n^{\frac{4}{5}} \left(\frac{5 \operatorname{Var}\left((\varepsilon_2 - \varepsilon_1)^4 \right)}{3^3 2^4 T^4 \sigma^8} \right)^{\frac{1}{5}}, \tag{1.2}$$

while the mean-squared error of the resulting estimator converges at the rate of $C_\kappa \operatorname{Var}\left((\varepsilon_2 - \varepsilon_1)^4 \right)^{\frac{3}{5}} n^{-\frac{3}{5}} T^{-\frac{2}{5}}$, up to constant C_κ. Here, ε_1 and ε_2 represent the microstructure noise associated to two different observations of the SBM. In particular, we again infer that the resulting estimator attains a better MSE performance than the plain MME in the absence of noise.

In order to implement the estimators with the corresponding optimal choices of K^*, we propose an iterative procedure in which an initial reasonable guess for σ^2 is used to find K^*, which in turn is used to improve the initial guess of σ, and so forth. The superior finite-sample and empirical performance of the resulting estimators are illustrated by simulation and real high-frequency stock data. In particular, the estimators don't exhibit the lack of robustness as the sampling frequency increases as their MME counterparts.

The rest of the paper is organized as follows. In Section 2, we give the model and the estimation framework. Section 3 introduces the method of moment estimators. Their in-fill and long-run asymptotic behavior are analyzed in Section 4. Section 5 introduces the estimators for σ and κ that are robust to a microstructure noise component together with bias corrected versions of these with optimal selection of K. Section 6 shows the finite-sample performance of the proposed estimators via simulations as well as their empirical robustness using real high-frequency transaction data. Finally, the proofs of the paper are deferred to the Appendix.

2. The model and the sampling scheme

In this section, we introduce the model used throughout the paper. We consider a subordinated Brownian motion model for the price process $\{S_t\}_{t\geq 0}$ of a risky asset. Concretely, given some constants $\sigma, \kappa > 0, \theta, b \in \mathbb{R}$, the log return process $X_t := \log(S_t/S_0)$ of the asset is assumed to take the form

$$X_t = \sigma W(\tau_t) + \theta \tau_t + bt, \tag{2.1}$$

where $W := \{W(t)\}_{t\geq 0}$ is a Wiener process and $\{\tau_t\}_{t\geq 0} := \{\tau(t;\kappa)\}_{t\geq 0}$ is an independent subordinator (i.e., a non-decreasing Lévy process) satisfying the following conditions:

$$\text{(i) } \mathbb{E}\tau_t = t, \qquad \text{(ii) } \mathrm{Var}(\tau_t) = \kappa t, \qquad \text{(iii) } \mathbb{E}\tau_1^j < \infty, \quad j \geq 3. \tag{2.2}$$

The condition (2.2-iii) is imposed so that X_t admits finite moments of arbitrary order. In the formulation (2.1), τ plays the role of a random clock aimed at incorporating variations in business activities through time. It is well known that the resultant process X is a Lévy process (see, e.g, Sato (1999)). Hereafter, ν will denote the Lévy measure of X.

Prototypical examples of (2.1) are the Variance Gamma (VG) and the Normal Inverse Gaussian (NIG) Lévy processes, which were proposed by Carr *et al.* (1998) and Barndorff-Nielsen (1998), respectively. In the VG model, $\tau(t;\kappa)$ is Gamma distributed with scale parameter $\beta := \kappa$ and shape parameter $\alpha := t/\kappa$, while in the NIG model $\tau(t;\kappa)$ follows an Inverse Gaussian distribution with mean $\mu = 1$ and shape parameter $\lambda = 1/(t\kappa)$.

As seen from the formulas for their moments (see (3.1) below), the model's parameters have the following interpretation:

1. σ dictates the overall variability of the log returns of the asset; in the "symmetric" case ($\theta = 0$), σ^2 is the variance of log returns divided by the time span of the returns;
2. κ controls the kurtosis or the tail's heaviness of the log return distribution; in the symmetric case ($\theta = 0$), κ is the excess kurtosis of log returns multiplied by the time span of the returns;
3. b is a drift component in the calendar time;
4. θ is a drift component in the business time and controls the skewness of log returns;

Throughout the paper, we also assume that the log return process $\{X_t\}_{t\geq 0}$ is sampled during a time interval $[0,T]$ at evenly spaced times:

$$t_i := i\delta_n, \quad i = 1,\dots,n, \qquad \text{where} \qquad \delta_n := \frac{T}{n}. \tag{2.3}$$

This sampling scheme is sometimes called *calendar time sampling* (c.f. Oomen (2006)). Under the assumption of independence and stationarity of increments, we have at our disposal a random sample

$$\Delta_i^n X := X_{i\delta_n} - X_{(i-1)\delta_n}, \quad i = 1,\dots,n, \tag{2.4}$$

of size n of the density of X_{δ_n}.

In real markets, log returns at high-frequency exhibit certain stylized features which cannot be accurately explained by efficient models such as (2.4). There are different approaches to model these features, widely termed as *microstructure noise*. Microstructure noises may come from different sources, such as clustering noises, non-clustering noises such as bid/ask bounce effects, and roundoff errors (cf. Campbell *et al.* (1997), Zeng (2003)). In what follows, we adopt a popular approach due to Zhang *et al.* (2005), where the net effect of the market microstructure is incorporated as an additive noise to the observed log-return process:

$$\widetilde{X}_t := \widetilde{X}(t) := X_t + \varepsilon_t, \tag{2.5}$$

where $\{\varepsilon_t\}_{t\geq 0}$ is assumed to be a centered process, independent of X. In particular, under this setup, the log return observations at a frequency δ_n are given by

$$\Delta_i^n \widetilde{X} := \widetilde{X}_{i\delta_n} - \widetilde{X}_{(i-1)\delta_n} = \left(X_{i\delta_n} - X_{(i-1)\delta_n}\right) + \left(\varepsilon_{i\delta_n} - \varepsilon_{(i-1)\delta_n}\right) =: \Delta_i^n X + \tilde{\varepsilon}_{i,n}, \qquad i = 1, \ldots, n. \tag{2.6}$$

In the simplest case, the noise $\{\varepsilon_t\}_{t\geq 0}$ is a white noise; i.e., the variables $\{\varepsilon_t\}_{t\geq 0}$ are independent identically distributed with mean 0.

It is well known (and not surprising) that standard statistical methods do not perform well when applied to high-frequency observations if the microstructure noise is not incorporated in the derivation of the estimators. A standing problem is then to derive inference methods that are robust against a wide range of microstructure noises. In Section 5, we proposed an approach to address the latter problem, borrowing ideas from the seminal two-scale correction technique of Zhang *et al.* (2005) applied to Method of Moment Estimators (MME). Before that, we first introduce the considered MMEs and carry on a simple infill asymptotic analysis of the estimators both in the absence and presence of the microstructure noise.

3. Method of Moment Estimators

The Method of Moment Estimators (MME) are widely used to deal with high-frequency data due to their simplicity, computational efficiency, and known robustness against potential correlation between observations. For the general subordinated Brownian model (2.2)-(2.1), the central moments are given in closed forms as follows:

$$\begin{aligned}
\mu_1(X_\delta) &:= \mathbb{E}(X_\delta) = (\theta + b)\delta, \\
\mu_2(X_\delta) &:= \mathrm{Var}(X_\delta) = (\sigma^2 + \theta^2\kappa)\delta, \\
\mu_3(X_\delta) &:= \mathbb{E}(X_\delta - \mathbb{E}X_\delta)^3 = \left(3\sigma^2\theta\kappa + \theta^3 c_3(\tau_1)\right)\delta, \\
\mu_4(X_\delta) &:= \mathbb{E}(X_\delta - \mathbb{E}X_\delta)^4 = \left(3\sigma^4\kappa + 6\sigma^2\theta^2 c_3(\tau_1) + \theta^4 c_4(\tau_1)\right)\delta + 3\mu_2(X_\delta)^2,
\end{aligned} \tag{3.1}$$

where, hereafter,

$$c_k(Y) := \frac{1}{i^k}\,\frac{d^k}{du^k}\ln \mathbb{E}\left(e^{iuY}\right)\Big|_{u=0},$$

represents the k-th cumulant of a r.v. Y. For the VG model, $(c_3(\tau_1), c_4(\tau_1)) = (2\kappa^2, 6\kappa^3)$, while for the NIG model, $(c_3(\tau_1), c_4(\tau_1)) = (3\kappa^2, 15\kappa^3)$ (cf. (see, e.g., Cont & Tankov, 2004, pp. 32 and 117)).

Throughout the rest of the paper, we assume that $\theta = 0$, which is consistent with extensive empirical evidence suggesting that the skewness of high-frequency data is negligible. To assess the latter assertion, we consider the sample central moment of third order, $\hat{\mu}_{3,n}$, divided by $\delta_n = T/n$. According to (3.1), $|\theta|$ will be close to 0 when $\hat{\mu}_{3,n}/\delta_n$ is small[1]. The results are collected in Table 1.

The simplifying assumption $\theta = 0$ allows us to find tractable expressions for the MME of the parameters σ^2 and κ as follows:

$$\tilde{\sigma}_n^2(X) := \frac{1}{\delta_n}\hat{\mu}_{2,n}(X), \qquad \tilde{\kappa}_n(X) := \frac{\delta_n}{3}\left(\frac{\hat{\mu}_{4,n}(X)}{\hat{\mu}_{2,n}^2(X)} - 3\right), \tag{3.2}$$

[1]Note that $c_3(\tau_1) = \mathbb{E}(\tau_1 - 1)^3 = \int_0^\infty u^3 \nu_{\tau_1}(du) > 0$, where ν_{τ_1} is the Lévy measure of τ_1. Thus, $|\mu_3(X_\delta)/\delta| \geq |\theta|^3 c_3(\tau_1) \geq |\theta|^3 \kappa^{3/2}$ by Jensen's inequality.

δ	5 sec	10 sec	30 sec	1 min	5 min	10 min	30 min
INTEL	-1.7×10^{-4}	-1.8×10^{-4}	-1.5×10^{-4}	-1.6×10^{-4}	-1.9×10^{-4}	-1.8×10^{-4}	-2.9×10^{-4}
CVX	-4.8×10^{-6}	-3.4×10^{-6}	9.5×10^{-7}	3.2×10^{-5}	2.3×10^{-5}	4.4×10^{-6}	-3.1×10^{-5}
CSCO	-7.9×10^{-5}	-8.9×10^{-5}	-8.8×10^{-5}	-7.6×10^{-5}	-5.8×10^{-5}	-6.3×10^{-5}	-7.3×10^{-5}
PFE	9.5×10^{-4}	9.2×10^{-4}	9.6×10^{-4}	9.9×10^{-4}	7.9×10^{-4}	7.1×10^{-4}	9.6×10^{-4}

TABLE 1

Computation of the sample central moment of third order, $\hat{\mu}_{3,n}$, divided by $\delta_n = T/n$ for different stocks based on high-frequency data during the year of 2005 ($T = 252$ days).

where $\hat{\mu}_{k,n}(X)$ represents the sample central moment of k^{th} order as defined by

$$\hat{\mu}_{k,n}(X) := \frac{1}{n}\sum_{i=1}^{n}\left(\Delta_i^n X - \overline{\Delta^n X}\right)^k, \quad k \geq 2, \qquad \overline{\Delta^n X} := \frac{1}{n}\sum_{i=1}^{n}\Delta_i^n X = \frac{1}{n}\log\frac{S_T}{S_0}.$$

We can further simplify the above statistics by neglecting the terms of order $O(\delta_n) = O(1/n)$ as follows:

$$\hat{\sigma}_n^2(X) := \frac{1}{T}\,[X,X]_2\,, \qquad \hat{\kappa}_n(X) := \frac{\delta_n}{3}\frac{\frac{1}{n}\sum_{i=1}^n\left(\Delta_i^n X\right)^4}{\left(\frac{1}{n}\sum_{i=1}^n\left(\Delta_i^n X\right)^2\right)^2} = \frac{1}{3}\frac{T^{-1}\widehat{[X,X]}_4}{\left(T^{-1}\widehat{[X,X]}_2\right)^2}, \tag{3.3}$$

where above we have expressed the estimators in terms of the realized variations of order 2 and 4:

$$\widehat{[X,X]}_2 := \sum_{i=1}^{n}\left(\Delta_i^n X\right)^2, \qquad \widehat{[X,X]}_4 = \sum_{i=1}^{n}\left(\Delta_i^n X\right)^4.$$

Note that, in the general case ($\theta \neq 0$), we can see the estimators (3.2)-(3.3) as approximate Method of Moment Estimators up to first order.

We now proceed to show some simple "in-fill" ($n \to \infty$ with fixed T) asymptotic properties of the estimators in (3.2)-(3.3). As above, in the sequel we assume that $\theta = 0$ and neglect $O(\delta_n) = O(1/n)$ terms. In that case, it is easy to see that

$$\mathbb{E}\hat{\sigma}_n^2 = \mathbb{E}\tilde{\sigma}_n^2 = \sigma^2 + O\left(\frac{1}{n}\right), \qquad \mathrm{Var}\left(\hat{\sigma}_n^2\right) = \mathrm{Var}\left(\tilde{\sigma}_n^2\right) = \frac{3\sigma^2\kappa}{T} + O\left(\frac{1}{n}\right). \tag{3.4}$$

From the above formulas, we conclude the (not surprising) fact that $\hat{\sigma}_n^2$ is not a mean-squared consistent estimator for σ^2, at a fixed time horizon T, when the sampling frequency increases. However, the standard error of $\hat{\sigma}_n$ decreases inversely proportional to the time horizon T.

An analysis of the bias and variance of $\hat{\kappa}_n$ and $\tilde{\kappa}_n$ is more complicated due to the non-linearity of the sample kurtosis. However, we can deduce some interesting features of its infill asymptotic behavior using some well known properties of Lévy processes. For reader's convenience, we recall those. First, if $\int_{|x|\geq 1}|x|^k\nu(dx) < \infty$ for $k \geq 2$, then $\mathbb{E}|X_t|^k < \infty$ for any $t > 0$ (see, e.g., Sato (1999)) and, furthermore,

$$\lim_{t\to 0}\frac{1}{t}\mathbb{E}\left(X_t^k\right) = \int x^k\nu(dx) + \upsilon^2\mathbf{1}_{\{k=2\}}, \tag{3.5}$$

where ν denotes the Lévy measure of X and υ^2 is the variance of the Brownian component of X (see, e.g., Lemma 2.1 in Asmussen & Rosiński (2001) or Figueroa-López (2008)). Second, for any $k \geq 2$,

$$\sum_{i=1}^{n} \left(X_{i\delta_n} - X_{(i-1)\delta_n}\right)^k \xrightarrow{\mathbb{P}} [X,X]_T := \sum_{t \leq T} (\Delta X_t)^k + \upsilon^2 T \mathbf{1}_{\{k=2\}} \tag{3.6}$$

as $n \to \infty$, where $\Delta X_t = X_t - X_{t-}$ is the jump size of X at time t and where the above summation is over the countable random set of times t for which $\Delta X_t \neq 0$. It is convenient to express the above limit process in terms of the jump measure M of X, which is defined by

$$M((u,v] \times C) := \#\{t \in (u,v] : \Delta X_t \in C\}.$$

In light of the Lévy-Itô decomposition of X (cf. (Sato, 1999, Section 19)), M is a Poisson random measure on $\mathbb{R}_+ \times \mathbb{R}\backslash\{0\}$ with mean measure $\mathbb{E}M(dt \times dx) = dt\nu(dx)$. Furthermore, the Poisson integral

$$\int_0^T \int f(x) M(dt,dx) := \sum_{t \leq T} f(\Delta X_t)$$

is well-defined if, for instance, $\int_{|x|\leq 1} |f(x)|\nu(dx) < \infty$ (e.g., see Theorem 10.15 in Kallenberg (1997)).

In view of (3.6), we first note the following limit, as $n \to \infty$,

$$\operatorname*{lim^{P}}_{n\to\infty} \hat{\kappa}_n = \operatorname*{lim^{P}}_{n\to\infty} \tilde{\kappa}_n = \frac{1}{3} \frac{\frac{1}{T}\int_0^T \int x^4 M(dt,dx)}{\left(\frac{1}{T}\int_0^T \int x^2 M(dt,dx)\right)^2} := \hat{\kappa}^{(T)}. \tag{3.7}$$

The convergence of the corresponding moments also holds true since $0 \leq \delta_n \hat{\mu}_{4,n}/\hat{\mu}_{2,n}^2 \leq \delta_n n = T < \infty$, which follows from the triangle inequality. Thus,

$$\lim_{n\to\infty} \mathbb{E}\hat{\kappa}_n = \lim_{n\to\infty} \mathbb{E}\tilde{\kappa}_n = \mathbb{E}\hat{\kappa}^{(T)} \quad \text{and} \quad \lim_{n\to\infty} \operatorname{Var}(\hat{\kappa}_n) = \lim_{n\to\infty} \operatorname{Var}(\tilde{\kappa}_n) = \operatorname{Var}\left(\hat{\kappa}^{(T)}\right). \tag{3.8}$$

In order to analyze the limit values in (3.8), let us note that, by the Strong Law of Large Numbers,

$$\hat{\kappa}^{(T)} \xrightarrow{T\to\infty} \frac{\int x^2 \nu(dx)}{3(\int x^2 \nu(dx))^2} = \frac{c_4(X_1)}{3c_2^2(X_1)} = \kappa,$$

which suggests that $\hat{\kappa}_n$ has a small bias and variance when the time horizon T is large. In fact we have the following result, which gives an explicit estimate of the bias and mean-squared error of the statistic $\kappa^{(T)}$. The result is valid for a general pure-jump Lévy process X with finite moments of sufficiently large order and its proof is given in Appendix A:

Proposition 3.1. *Let $c_i := c_i(X_1)$ be the i^{th} cumulant of X_1, $\kappa := c_4/3c_2^2$, and suppose that $\int |x|^i \nu(dx) < \infty$ for any $i \geq 2$. Then, as $T \to \infty$,*

$$\mathbb{E}\,\hat{\kappa}^{(T)} = \kappa + \frac{3c_4^2 - 2c_6c_2}{3c_2^4} T^{-1} + O(T^{-2}), \tag{3.9}$$

$$\mathbb{E}\left(\hat{\kappa}^{(T)} - \kappa\right)^2 = \frac{c_8c_2 - 4c_4c_6 + 4c_4^2c_2}{9c_2^5} T^{-1} + O(T^{-2}). \tag{3.10}$$

Remark 3.2. *One can easily carry on the above analysis for the estimators of σ^2 introduced in (3.2)-(3.3). Concretely, we have*

$$\lim_{n\to\infty}^{\mathrm{P}} \hat{\sigma}_n^2 = \lim_{n\to\infty}^{\mathrm{P}} \tilde{\sigma}_n^2 = \frac{1}{T}\int_0^T \int x^2 M(dt,dx) := \hat{\sigma}_T^2 \overset{T\to\infty}{\longrightarrow} \sigma^2, \tag{3.11}$$

and, due to the well-known mean and variance formulas of Poisson integrals,

$$\mathbb{E}\,\hat{\sigma}_T^2 = \sigma^2, \qquad \mathbb{E}\left(\hat{\sigma}_T^2 - \sigma^2\right)^2 = c_4 T^{-1}. \tag{3.12}$$

4. Basic properties of the MME under microstructure noise

In this part we study how a microstructure noise component affects the estimators introduced in the previous section. In turn, such analyses will help us to develop bias correction techniques in the subsequent section. We adopt the setup introduced at the end of Section 2, under which the observed log-returns are given by

$$\Delta_i^n \widetilde{X} := \widetilde{X}_{i\delta_n} - \widetilde{X}_{(i-1)\delta_n} = \left(X_{i\delta_n} - X_{(i-1)\delta_n}\right) + \left(\varepsilon_{i\delta_n} - \varepsilon_{(i-1)\delta_n}\right) =: \Delta_i^n X + \tilde{\varepsilon}_{i,n}. \tag{4.1}$$

Furthermore, throughout we assume that, for each n, $(\tilde{\varepsilon}_{i,n})_{i\geq 1}$ have identical distribution with zero mean and finite moments of any order. Moreover, the following mild assumptions are imposed for any positive integer $k \geq 1$:

$$(i) \quad \frac{1}{n}\sum_{i=1}^{n}(\tilde{\varepsilon}_{i,n})^k \overset{\mathbb{P}}{\longrightarrow} m_k(\tilde{\varepsilon}), \quad (n\to\infty), \quad \text{for some } m_k(\varepsilon)\in\mathbb{R}; \tag{4.2}$$

$$(ii) \quad \limsup_{n\to\infty} \mathbb{E}(\tilde{\varepsilon}_{i,n})^k < \infty. \tag{4.3}$$

Let us remark that the previous assumptions not only cover the microstructure white-noise case, where $(\varepsilon_t)_t$ are i.i.d., but also block dependent sequences, where $\tilde{\varepsilon}_{i,n}$ and $\tilde{\varepsilon}_{j,n}$ are assumed to be independent, whenever $|i-j|\geq k$, for some fixed positive integer k. Also, note that, under the white noise case, $m_k(\tilde{\varepsilon}) := \mathbb{E}\left((\tilde{\varepsilon}_{1,n})^k\right)$.

Let us first analyze the infill asymptotic behavior of the estimators for σ^2, defined analogously to (3.2)-(3.3), but based on the noisy observations:

$$\tilde{\sigma}_n^2(\widetilde{X}) := \frac{1}{\delta_n n}\sum_{i=1}^{n}(\Delta_i^n\widetilde{X} - \overline{\Delta^n\widetilde{X}})^2, \qquad \hat{\sigma}_n^2(\widetilde{X}) := \frac{1}{T}\widehat{[\widetilde{X},\widetilde{X}]}_2 = \frac{1}{\delta_n n}\sum_{i=1}^{n}(\Delta_i^n\widetilde{X})^2. \tag{4.4}$$

The following simple result is needed in the sequel.

Lemma 4.1. *For each positive integer m and k*

$$\frac{1}{n}\sum_{i=1}^{n}(\Delta_i^n X)^m(\tilde{\varepsilon}_{i,n})^k \overset{\mathbb{P}}{\longrightarrow} 0, \quad \textit{as } n\to\infty.$$

Proof. The second moment of $\sum_{i=1}^n (\Delta_i^n X)^m (\tilde{\varepsilon}_{i,n})^k / n$ can be clearly written as

$$\frac{1}{n}\mathbb{E}(\Delta_1^n X)^{2m}\mathbb{E}(\tilde{\varepsilon}_{1,n})^{2k} + \frac{2}{n^2}\sum_{i\neq j}\mathbb{E}(\Delta_i^n X)^m \mathbb{E}(\Delta_j^n X)^m \mathbb{E}(\tilde{\varepsilon}_{i,n}\tilde{\varepsilon}_{j,n})^k.$$

The first term above is clearly $O(n^{-1}\delta_n)$ due to (3.5) and (4.3). The second term above can be bounded in absolute value by $(\mathbb{E}(\Delta_1^n X)^{2m})^2 \mathbb{E}(\tilde{\varepsilon}_{i,n})^{2k}$, and thus, it is again $O(n^{-1}\delta_n)$ due to (3.5) and (4.3). By Markov's inequality, $\frac{1}{n}\sum_{i=1}^n (\Delta_i^n X)^m (\tilde{\varepsilon}_{i,n})^k$ converges to 0 in probability. □

We are now ready to analyze the asymptotic behavior of the estimators in (4.4). Our first result shows that, when n is large, both $\hat{\sigma}_n^2(\widetilde{X})$ and $\tilde{\sigma}_n^2(\widetilde{X})$ asymptotically behave like $\frac{1}{T}\sum_{s\leq T}(\Delta X_s)^2 + \frac{1}{\delta_n}B$, for some constant B. In that case, for large T, the estimators will asymptotically behave like $\sigma^2 + \frac{1}{\delta_n}B$.

Proposition 4.2. *Both estimators $\hat{\sigma}_n^2(\widetilde{X})$ and $\tilde{\sigma}_n^2(\widetilde{X})$ admit the decomposition*

$$\hat{\sigma}_n^2(\widetilde{X}) = A_n + B_n, \qquad \tilde{\sigma}_n^2(\widetilde{X}) = \widetilde{A}_n + \widetilde{B}_n$$

such that, as $n \to \infty$,

$$\operatorname*{lim^{P}}_{n\to\infty} A_n = \operatorname*{lim^{P}}_{n\to\infty} \widetilde{A}_n = \frac{1}{T}\sum_{s\leq T}(\Delta X_s)^2, \qquad \operatorname*{lim^{P}}_{n\to\infty} \delta_n B_n = m_2(\tilde{\varepsilon}), \qquad \operatorname*{lim^{P}}_{n\to\infty} \delta_n \widetilde{B}_n = m_2(\tilde{\varepsilon}) - (m_1(\tilde{\varepsilon}))^2.$$

Proof. We only give the proof for $\tilde{\sigma}_n^2 := \tilde{\sigma}_n^2(\widetilde{X})$. The proof for $\hat{\sigma}_n^2(\widetilde{X})$ is identical. Let us first note that

$$\begin{aligned}
\tilde{\sigma}_n^2 &= \frac{1}{n\delta_n}\sum_{i=1}^n (\Delta_i^n X - \overline{\Delta^n X} + \tilde{\varepsilon}_{i,n} - \overline{\tilde{\varepsilon}_n})^2 \\
&= \frac{1}{n\delta_n}\sum_{i=1}^n (\Delta_i^n X - \overline{\Delta^n X})^2 + \frac{1}{n\delta_n}\sum_{i=1}^n (\tilde{\varepsilon}_{i,n} - \overline{\tilde{\varepsilon}_n})^2 + \frac{2}{n\delta_n}\sum_{i=1}^n (\Delta_i^n X - \overline{\Delta^n X})(\tilde{\varepsilon}_{i,n} - \overline{\tilde{\varepsilon}_n}) \\
&=: \widetilde{A}_n + \widetilde{B}_{n,1} + \widetilde{B}_{n,2}.
\end{aligned}$$

The term $\widetilde{A}_n$ converges to $\frac{1}{T}\sum_{s\leq T}(\Delta X_s)^2$ as $n \to \infty$ in light of (3.6) and the fact that $\overline{\Delta^n X} = O_P(1/n)$. Clearly, (4.2) implies that

$$\delta_n \widetilde{B}_{n,1} = \frac{1}{n}\sum_{i=1}^n (\tilde{\varepsilon}_{i,n})^2 - (\overline{\tilde{\varepsilon}_n})^2 \xrightarrow{\mathbb{P}} m_2(\tilde{\varepsilon}) - (m_1(\tilde{\varepsilon}))^2.$$

Also, using Lemma 4.1, $\delta_n \widetilde{B}_{n,2} = \frac{2}{n}\sum_{i=1}^n (\Delta_i^n X)(\tilde{\varepsilon}_{i,n}) - 2\overline{\Delta^n X}\,\overline{\tilde{\varepsilon}_n}$ goes to 0 in probability. □

Let us consider the estimators for κ introduced in (3.2)-(3.3), but applied to the noisy process $\widetilde{X}$:

$$\tilde{\kappa}_n(\widetilde{X}) = \frac{\delta_n}{3}\left(\frac{\hat{\mu}_{4,n}(\widetilde{X})}{\hat{\mu}_{2,n}^2(\widetilde{X})} - 3\right), \qquad \hat{\kappa}_n(\widetilde{X}) := \frac{T}{3}\frac{\widehat{[\widetilde{X},\widetilde{X}]}_4}{\widehat{[\widetilde{X},\widetilde{X}]}_2^2}.$$

The following result states that, when n is large, the above estimators behave asymptotically as $\delta_n C$, for some constant C, depending on the ergodic properties of the microstructure noise.

Proposition 4.3. *There exist non-zero constants C and $\widetilde{C}$ such that, as $n \to \infty$,*

$$\frac{1}{\delta_n}\hat{\kappa}_n(\widetilde{X}) \xrightarrow{\mathbb{P}} C, \qquad \frac{1}{\delta_n}\tilde{\kappa}_n(\widetilde{X}) \xrightarrow{\mathbb{P}} \widetilde{C}.$$

Proof. We only give the proof for $\tilde{\kappa}_n := \tilde{\kappa}_n(\widetilde{X})$. The proof for $\hat{\kappa}_n(\widetilde{X})$ is identical. Clearly,

$$\frac{3}{\delta_n}\tilde{\kappa}_n + 3 = \frac{\frac{1}{n}\sum_{i=1}^n \left(\Delta_i^n \widetilde{X} - \overline{\Delta^n \widetilde{X}}\right)^4}{\left(\frac{1}{n}\sum_{i=1}^n \left(\Delta_i^n \widetilde{X} - \overline{\Delta^n \widetilde{X}}\right)^2\right)^2} =: \frac{N_n}{D_n^2} \tag{4.5}$$

Observe that

$$D_n = \frac{1}{n}\sum_{i=1}^n (\Delta_i^n X - \overline{\Delta^n X})^2 + \frac{1}{n}\sum_{i=1}^n (\tilde{\varepsilon}_{i,n} - \bar{\tilde{\varepsilon}}_n)^2 + \frac{2}{n}\sum_{i=1}^n (\Delta_i^n X - \overline{\Delta^n X})(\tilde{\varepsilon}_{i,n} - \bar{\tilde{\varepsilon}}_n).$$

By Lemma 4.1 and (3.6), the first and third terms on the last expression above tend to 0 in probability, while the second term converges to $m_2(\tilde{\varepsilon}) - (m_1(\tilde{\varepsilon}))^2$ by (4.2). Regarding the numerator in (4.5), this can be decomposed as follows:

$$\begin{aligned} N_n = {} & \frac{1}{n}\sum_{i=1}^n (\Delta_i^n X - \overline{\Delta^n X})^4 + \frac{4}{n}\sum_{i=1}^n (\Delta_i^n X - \overline{\Delta^n X})^3(\tilde{\varepsilon}_{i,n} - \bar{\tilde{\varepsilon}}_n) + \frac{6}{n}\sum_{i=1}^n (\Delta_i^n X - \overline{\Delta^n X})^2(\tilde{\varepsilon}_{i,n} - \bar{\tilde{\varepsilon}}_n)^2 \\ & + \frac{4}{n}\sum_{i=1}^n (\Delta_i^n X - \overline{\Delta^n X})(\tilde{\varepsilon}_{i,n} - \bar{\tilde{\varepsilon}}_n)^3 + \frac{1}{n}\sum_{i=1}^n (\tilde{\varepsilon}_{i,n} - \bar{\tilde{\varepsilon}}_n)^4. \end{aligned}$$

Again, by Lemma 4.1 and (3.6), all the terms in the expression above tend to 0 in probability except the last term that converges to $m_4(\tilde{\varepsilon}) - 4m_3(\tilde{\varepsilon})m_1(\tilde{\varepsilon}) + 6m_2(\tilde{\varepsilon})m_1^2(\tilde{\varepsilon}) - 3m_1^4(\tilde{\varepsilon})$. Therefore, as $n \to \infty$,

$$\frac{1}{\delta_n}\tilde{\kappa}_n \xrightarrow{\mathbb{P}} \frac{m_4(\tilde{\varepsilon}) - 4m_1(\varepsilon)m_3(\tilde{\varepsilon}) + 6m_1^2(\tilde{\varepsilon})m_2(\tilde{\varepsilon}) - 3m_1^4(\tilde{\varepsilon})}{3\left(m_2(\tilde{\varepsilon}) - m_1^2(\tilde{\varepsilon})\right)^2} - 1 =: C,$$

□

Remark 4.4. *As a consequence of the proof, it follows that, if $m_1(\tilde{\varepsilon}) = 0$, then*

$$C = \widetilde{C} = \frac{m_4(\tilde{\varepsilon})}{3\left(m_2(\tilde{\varepsilon})\right)^2}.$$

In particular, if the microstructure noise $(\varepsilon_t)_t$ in (2.5) is white-noise, then the constant coincides with the excess kurtosis, $\mathbb{E}\tilde{\varepsilon}^4/3\left(\mathbb{E}\tilde{\varepsilon}^2\right)^2$, of the random variable $\tilde{\varepsilon} := \varepsilon_2 - \varepsilon_1$.

5. Robust Method of Moments Estimators

In this section, we adapt the so-called two-scale bias correction technique of Zhang *et al.* (2005) to develop estimators for σ^2 and κ that are robust against microstructure noises. Roughly, their approach consists of three main ingredients: sparse subsampling, averaging, and bias correction. Let us first introduce some needed notation. Let $\bar{\mathcal{G}}_n := \{t_0, t_1, \dots, t_n\}$ be the complete set of available sampling times as described in (2.3). For a subsample $\mathcal{G} = \{t_{i_1}, \dots, t_{i_m}\}$ with $i_1 \leq \cdots \leq i_m$ and a natural $\ell \in \mathbb{N}$, we define the ℓ^{th}-order realized variation of the process $\widetilde{X}$ over $\mathcal{G}$ as

$$[\widetilde{X}, \widetilde{X}]_\ell^{\mathcal{G}} = \sum_{j=0}^{m-1} \left| \widetilde{X}(t_{i_{j+1}}) - \widetilde{X}(t_{i_j}) \right|^\ell .$$

For simplicity, hereafter we omit the hat in realized power variations. Next, we partition the grid $\bar{\mathcal{G}}_n$ into K mutually exclusive regular sub grids as follows:

$$\mathcal{G}_n^{(i)} := \mathcal{G}_{n,K}^{(i)} := \{t_{i-1}, t_{i-1+K}, t_{i-1+2K}, \dots, t_{i-1+n_iK}\}, \qquad i = 1, \dots, K,$$

with $n_i := n_{i,K} := [(n-i+1)/K]$. As in Zhang *et al.* (2005), the key idea to improve the estimators introduced in (3.3) consists of averaging the relevant realized variations over the different sparse sub grids $\mathcal{G}_n^{(i)}$, instead of using only one realized variation over the complete set $\bar{\mathcal{G}}_n$. Hence, for instance, for estimating σ^2, we shall consider the estimator

$$\hat{\sigma}_n^2 := \hat{\sigma}_{n,K}^2 := \frac{1}{K} \sum_{i=1}^K T^{-1} [\widetilde{X}, \widetilde{X}]_2^{\mathcal{G}_n^{(i)}}, \tag{5.1}$$

which is constructed by averaging estimators of the form $\hat{\sigma}^2(\widetilde{X})$ in (4.4) over sparse sub-grids. The above estimator corresponds to the so-called "second-best estimator" in Zhang *et al.* (2005). This estimator can be improved in two ways. First, by correcting the bias of the estimator and, second, by choosing the number of sub grids, K, in an "optimal" way. We analyze these two approaches in the subsequent two subsections.

At this point it is convenient to recall that we are assuming the subordinated Brownian motion model (2.1) with $\theta = 0$. For simplicity, we also assume that $b = 0$, which won't affect much what follows since we are considering high-frequency type estimators and, thus, the contribution of the drift is negligible in that case. Furthermore, hereafter we assume that the noise process $\{\varepsilon_t\}_{t \geq 0}$ appearing in Eq. (2.5) is white noise; i.e., the variables therein are independently identically distributed with mean 0. In particular, the noises of the increments, $\tilde{\varepsilon}_{i,n} := \varepsilon_{i\delta_n} - \varepsilon_{(i-1)\delta_n}$, follow a stationary Moving Average (MA) process with $\mathbb{E}\left(\tilde{\varepsilon}_{i,n}\right) = 0$ and $\mathbb{E}\left(\tilde{\varepsilon}_{i,n}^2\right) = 2\mathbb{E}\left(\varepsilon_1^2\right)$. For simplicity, in the sequel $\tilde{\varepsilon}$ and ε denote variables with the same distribution as $\tilde{\varepsilon}_{i,n}$ and ε_t, respectively.

5.1. Bias corrected estimators

Let us start by exploring bias correction techniques for the estimator (5.1). Clearly, from (3.1) and the independence of the noise $\tilde{\varepsilon}$ and the process X, we have:

$$\mathbb{E}\left(\hat{\sigma}_{n,K}^2\right) = \mathbb{E}\left(\frac{1}{K} \sum_{i=1}^K T^{-1} [\widetilde{X}, \widetilde{X}]_2^{\mathcal{G}_n^{(i)}}\right) = \sigma^2 + \mathbb{E}\left(\tilde{\varepsilon}^2\right) \frac{1}{KT} \sum_{i=1}^K n_i = \sigma^2 + \mathbb{E}\left(\tilde{\varepsilon}^2\right) \frac{n-K+1}{KT}, \tag{5.2}$$

where we used the identity $\sum_{i=1}^{n} n_i = n - K + 1$. The relation (5.2) shows that the bias of the estimator diverges to infinity when the time span between observation $\delta_n := T/n$ tends to 0. To correct this issue, recall from Proposition 4.2 that

$$\hat{m}_{2,n}(\tilde{\varepsilon}) := \delta_n \hat{\sigma}^2_{n,1} = \frac{1}{n}[\widetilde{X}, \widetilde{X}]_2^{\bar{\mathcal{G}}_n} \xrightarrow{\mathbb{P}} m_2(\tilde{\varepsilon}) = \mathbb{E}\left(\tilde{\varepsilon}^2\right), \tag{5.3}$$

as $n \to \infty$. Hence, a natural "bias-corrected" estimator would be

$$\hat{\hat{\sigma}}^2_n := \hat{\hat{\sigma}}^2_{n,K} := \frac{1}{K}\sum_{i=1}^{K} T^{-1}[\widetilde{X}, \widetilde{X}]_2^{\mathcal{G}_n^{(i)}} - \frac{n-K+1}{KTn}[\widetilde{X}, \widetilde{X}]_2^{\bar{\mathcal{G}}_n}. \tag{5.4}$$

However, from (5.2) with $K = 1$,

$$\mathbb{E}\left(\frac{1}{T}[\widetilde{X}, \widetilde{X}]_2^{\bar{\mathcal{G}}_n}\right) = \sigma^2 + \mathbb{E}\left(\tilde{\varepsilon}^2\right)\frac{n}{T},$$

and, thus,

$$\mathbb{E}\left(\hat{\hat{\sigma}}^2_n\right) = \sigma^2 + \mathbb{E}\left(\tilde{\varepsilon}^2\right)\frac{n-K+1}{KT} - \frac{n-K+1}{Kn}\left(\sigma^2 + \mathbb{E}\left(\tilde{\varepsilon}^2\right)\frac{n}{T}\right) = \frac{(n+1)(K-1)}{nK}\sigma^2,$$

which implies that $\hat{\hat{\sigma}}^2_n$ is not truly unbiased but only asymptotically unbiased when $n \to \infty$ and $K \to \infty$. Nevertheless, the above relationships yield the following unbiased estimator for σ^2:

$$\hat{\hat{\sigma}}^2_{n,K} := \frac{nK}{(n+1)(K-1)}\hat{\hat{\sigma}}_{n,K} := \frac{n}{(n+1)(K-1)}\sum_{i=1}^{K} T^{-1}[\widetilde{X}, \widetilde{X}]_2^{\mathcal{G}_n^{(i)}} - \frac{n-K+1}{T(n+1)(K-1)}[\widetilde{X}, \widetilde{X}]_2^{\bar{\mathcal{G}}_n}. \tag{5.5}$$

The estimator (5.4) corresponds to the small-sample adjusted "First-Best Estimator" of Zhang *et al.* (2005). The estimator (5.4) is said to be a two-scale based estimator in the sense that involves the realized variations of the process in two different "scales" or sampling frequencies.

We next attempt to devise (approximate) bias-corrected estimators for κ. In order to separate the problem of estimating κ and σ^2, let us assume that σ is known. In practice, we will have to replace σ with an "accurate" estimate such as the estimator (5.5). The analog of the estimator (5.1) for κ is given by

$$\hat{\kappa}_n := \hat{\kappa}_{n,K} := \frac{1}{3\sigma^4 K}\sum_{i=1}^{K} T^{-1}[\widetilde{X}, \widetilde{X}]_4^{\mathcal{G}_n^{(i)}} - \frac{TK(n-K+1)}{n^2} \tag{5.6}$$

which is actually an unbiased estimator for κ in the absence of the microstructure noise $\tilde{\varepsilon}$ (see (5.7) below). In the presence of microstructure noise, the bias of (5.6) diverges as the frequency of observations increases. Indeed, from (3.1) and the independence of the noise $\tilde{\varepsilon}$ and X, let us first note that

$$\mathbb{E}\left(X_\delta + \tilde{\varepsilon}\right)^4 = 3\sigma^4\kappa\delta + 6\sigma^2\mathbb{E}\left(\tilde{\varepsilon}^2\right)\delta + \mathbb{E}\left(\tilde{\varepsilon}^4\right) + 3\sigma^4\delta^2.$$

Therefore, for a subsample $\mathcal{G} = \{t_{i_1}, \ldots, t_{i_m}\}$, with $i_1 \leq \cdots \leq i_m$,

$$\mathbb{E}\left(\frac{1}{T}[\widetilde{X}, \widetilde{X}]_4^{\mathcal{G}}\right) = 3\sigma^4\kappa + 6\sigma^2\mathbb{E}\left(\tilde{\varepsilon}^2\right) + \frac{m}{T}\mathbb{E}\left(\tilde{\varepsilon}^4\right) + 3\sigma^4\frac{1}{T}\sum_{j=0}^{m-1}(t_{i_{j+1}} - t_{i_j})^2,$$

and, thus,

$$\mathbb{E}\left(\frac{1}{K}\sum_{i=1}^{K}T^{-1}[\widetilde{X},\widetilde{X}]_4^{\mathcal{G}_n^{(i)}}\right)=3\sigma^4\kappa+6\sigma^2\mathbb{E}\left(\tilde{\varepsilon}^2\right)+\frac{n-K+1}{TK}\mathbb{E}\left(\tilde{\varepsilon}^4\right)+3\sigma^4\frac{TK(n-K+1)}{n^2}. \tag{5.7}$$

It is now clear that, for a fixed K, the bias of the estimator $\hat{\kappa}_n$ diverges to ∞ as $T/n\to 0$.

The formula (5.7) suggests the estimator

$$\hat{\hat{\kappa}}_n := \frac{1}{3\sigma^4 K}\sum_{i=1}^{K}T^{-1}[\widetilde{X},\widetilde{X}]_4^{\mathcal{G}_n^{(i)}}-\frac{2}{\sigma^2}\hat{m}_{2,n}(\tilde{\varepsilon})-\frac{n-K+1}{3\sigma^4 TK}\hat{m}_{4,n}(\tilde{\varepsilon})-\frac{TK(n-K+1)}{n^2}, \tag{5.8}$$

where

$$\hat{m}_{4,n}(\tilde{\varepsilon}) := \frac{1}{n}[\widetilde{X},\widetilde{X}]_4^{\bar{\mathcal{G}}_n}, \tag{5.9}$$

which converges to $\mathbb{E}\left(\tilde{\varepsilon}^4\right)$. However, as with the estimator $\hat{\hat{\sigma}}_n$ above, the above estimator is only asymptotically unbiased for large n and K. The following result provides an unbiased estimator for κ based on the realized variations of the process on two scales. The proof follows from (5.2) and (5.7) and is omitted for the sake of space.

Proposition 5.1. *Let*

$$\begin{aligned}\hat{\hat{\kappa}}_n := {} & \frac{n}{3\sigma^4(n+1)(K-1)}\sum_{i=1}^{K}T^{-1}[\widetilde{X},\widetilde{X}]_4^{\mathcal{G}_n^{(i)}}-\frac{n-K+1}{3\sigma^4T(n+1)(K-1)}[\widetilde{X},\widetilde{X}]_4^{\bar{\mathcal{G}}_n} \\ & -\frac{2}{n\sigma^2}[\widetilde{X},\widetilde{X}]_2^{\bar{\mathcal{G}}_n}-T\frac{nK-K^2-n-1}{n(n+1)}\end{aligned} \tag{5.10}$$

Then, $\hat{\hat{\kappa}}_n$ is an unbiased estimator for κ.

5.2. *Optimal selection of K*

An important issue when using the approach outlined above is the selection of the number of subclasses, K. A natural approach to deal with this issue consists of minimizing the variance of the relevant estimators over all K. This procedure will yield an optimal K^* for the number of subclasses. Let us first illustrate this approach for the estimator $\hat{\sigma}^2_{n,K}$ given in (5.1). The following result gives the variance of $\hat{\sigma}^2_{n,K}$. Its proof is outlined in Appendix A.2.

Theorem 5.2. *The estimator (5.1) is such that*

$$\begin{aligned}\mathrm{Var}\left(\hat{\sigma}^2_{n,K}\right) = {} & \frac{4\sigma^4K}{3n}+\frac{4n\mathbb{E}\varepsilon^4}{K^2T^2}+\frac{4\sigma^4}{3n}+\frac{3\sigma^4\kappa}{T}+\frac{2\sigma^4}{3Kn}+\frac{8\sigma^2\mathbb{E}(\varepsilon^2)}{KT} \\ & +O\left(\frac{K^2}{n^2}\right)+O\left(\frac{K}{Tn}\right)+O\left(\frac{1}{KT^2}\right),\end{aligned} \tag{5.11}$$

where the big-O notation in the last three terms in (5.11) means that

$$\left|O\left(\frac{K^2}{n^2}\right)\right| \leq a\frac{K^2}{n^2}, \quad \left|O\left(\frac{K}{Tn}\right)\right| \leq b\frac{K}{Tn}, \quad \left|O\left(\frac{1}{KT^2}\right)\right| \leq c\frac{1}{KT^2},$$

for some absolute constants a, b, and c which are independent of K, n, and T.

Remark 5.3. *As a consequence of (5.11) and (5.2), for a fixed arbitrary K and a high-frequency/long-horizon sampling setup ($\delta \to 0$, $T \to \infty$), a sufficient asymptotic relationship between T and $\delta := T/n$ for the estimator $\hat{\sigma}^2_{n,K}$ to be mean square consistent is that $\delta T \to \infty$ or, equivalently, $n/T^2 \to 0$. If K is chosen depending on n and T, the feasible values $K := K_{n,T}$ must be such that $K_{n,T}/n \to 0$ and $n/(K^2_{n,T}T^2) \to 0$ as $T \to \infty$ and $\delta = T/n \to 0$.*

Now, we are ready to propose an approximately "optimal" K^*. To that end, let us first recall from (5.2) that the bias of the estimator is

$$\text{Bias}\left(\hat{\sigma}^2_{n,K}\right) = 2\mathbb{E}\varepsilon^2\frac{n-K+1}{TK} = \frac{2n\mathbb{E}\varepsilon^2}{TK} - \frac{2(K-1)\mathbb{E}\varepsilon^2}{TK}. \tag{5.12}$$

Together (5.11)-(5.12) implies that

$$\begin{aligned}\text{MSE}\left(\hat{\sigma}^2_{n,K}\right) = \frac{4\sigma^4K}{3n} + \frac{4\sigma^4}{3n} + \frac{3\sigma^4\kappa}{T} + \frac{2\sigma^4}{3Kn} + \frac{8\sigma^2\mathbb{E}(\varepsilon^2)}{KT} + \frac{4n\mathbb{E}\varepsilon^4}{K^2T^2} + \frac{4n^2\left(\mathbb{E}\varepsilon^2\right)^2}{T^2K^2} \\ + O\left(\frac{K^2}{n^2}\right) + O\left(\frac{K}{Tn}\right) + O\left(\frac{1}{KT^2}\right).\end{aligned} \tag{5.13}$$

Our goal is to minimize the MSE in K when n is large. Note that the only term that is increasing in K is $4\sigma^4K/3n$, while out of the terms decreasing in K, the term $4n^2\left(\mathbb{E}\varepsilon^2\right)^2/T^2K^2$ is the dominant (when n is large). It is then reasonable to consider only these two terms leading to the "approximation":

$$\text{MSE}\left(\hat{\sigma}^2_K\right) \approx \frac{4\sigma^4K}{3n} + \frac{4n^2}{T^2K^2}(\mathbb{E}\varepsilon^2)^2 =: \text{MSE}_1\left(\hat{\sigma}^2_K\right). \tag{5.14}$$

The right-hand side in the above expression attains its minimum at the value:

$$K_1^* = n\left(\frac{6(\mathbb{E}\varepsilon^2)^2}{T^2\sigma^4}\right)^{\frac{1}{3}}. \tag{5.15}$$

Interestingly enough, the value above coincides with the optimal K^* proposed in Zhang *et al.* (2005) (see Eq. (8) therein). Plugging (5.15) in (5.13) and, since $\delta = T/n \to 0$, it follows that

$$MSE\left(\hat{\sigma}^2_{K_1^*}\right) = 2^{\frac{4}{3}}3^{\frac{1}{3}}\left(\mathbb{E}\varepsilon^2\right)^{\frac{2}{3}}\sigma^{\frac{8}{3}}T^{-\frac{2}{3}} + 3\kappa\sigma^4T^{-1} + o(T^{-1}). \tag{5.16}$$

In particular, the above expression shows that, in the presence of a microstructure noise component, the rate of convergence reduces from $O(T^{-1})$ to only $O(T^{-2/3})$ and, furthermore, that the convergence is worst when σ, $\mathbb{E}\varepsilon^2$, and κ are larger.

The following result gives an estimate of the variance of the unbiased estimator (5.5). Its proof is given in Appendix A.2.

Proposition 5.4. *The estimator (5.5) is such that*

$$\operatorname{Var}\left(\hat{\hat{\sigma}}^2_{n,K}\right) = \frac{4\sigma^4 K}{3n} + \frac{4n\left(\mathbb{E}\varepsilon^4 + (\mathbb{E}\varepsilon^2)^2\right)}{T^2K^2} + O\left(\frac{1}{n}\right) + O\left(\frac{n}{K^3T^2}\right) + O\left(\frac{1}{TK}\right). \tag{5.17}$$

As before, the previous result suggests to fix K so that to minimize the first two leading terms in (5.17). Such a minimum is given by

$$K_2^* = n^{\frac{2}{3}}\left(\frac{6\left(\mathbb{E}\varepsilon^4 + (\mathbb{E}\varepsilon^2)^2\right)}{T^2\sigma^4}\right)^{\frac{1}{3}}, \tag{5.18}$$

which is similar[2] (but not identical) to the analog optimal K^* proposed in Zhang *et al.* (2005) (see Eq. (58) & (63) therein). After plugging K_1^* in (5.17), the resultant estimator attains the MSE:

$$MSE\left(\hat{\hat{\sigma}}^2_{K_2^*}\right) = 2^{\frac{4}{3}}3^{\frac{1}{3}}\left(\mathbb{E}\varepsilon^4 + (\mathbb{E}\varepsilon^2)^2\right)^{\frac{1}{3}}\sigma^{\frac{8}{3}}n^{-\frac{1}{3}}T^{-\frac{2}{3}} + o(T^{-1}). \tag{5.19}$$

Interestingly enough, since $T/n \to 0$, the estimator $\hat{\hat{\sigma}}^2_{K_1^*}$ attains the order $o(T^{-1})$, which was not achievable by the estimators $\hat{\sigma}^2_K$, even in the absence of microstructure noise, nor by the standard estimators introduced in Section 3 (see (3.4)).

Now, we proceed to study the optimal selection problem of K for the estimator (5.6) for κ. As with $\hat{\hat{\sigma}}^2_{n,K}$, we first need to analyze the variance of the estimator.

Theorem 5.5. *The estimator (5.6) is such that*

$$\operatorname{Var}\left(\hat{\kappa}_{n,K}\right) = \frac{64}{5}\frac{T^2K^3}{n^3} + O\left(\frac{T^2K^2}{n^3}\right). \tag{5.20}$$

Now, we propose a method to choose a value of K that approximately minimizes the MSE of the estimator $\hat{\kappa}_{n,K}$. Let us first recall from (5.7) that the bias of the estimator $\hat{\kappa}_{n,K}$ is

$$\operatorname{Bias}\left(\hat{\kappa}_{n,K}\right) = \mathbb{E}\left(\hat{\kappa}_{n,K}\right) - \kappa = \mathbb{E}\left(\tilde{\varepsilon}^4\right)\frac{n-K+1}{TK\sigma^4} + 2\frac{\mathbb{E}\left(\tilde{\varepsilon}^2\right)}{\sigma^2} = \mathbb{E}\left(\tilde{\varepsilon}^4\right)\frac{n}{TK\sigma^4} + \text{l.o.t.} \tag{5.21}$$

where l.o.t. mean "lower order terms". Together, (5.20)-(5.21) implies that

$$\operatorname{MSE}\left(\hat{\kappa}_{n,K}\right) = \frac{64}{5}\frac{T^2K^3}{n^3} + \frac{n^2\left(\mathbb{E}\tilde{\varepsilon}^4\right)^2}{T^2K^2\sigma^8} + \text{l.o.t.} \tag{5.22}$$

As with the estimator $\hat{\sigma}_{n,K}$, it is then reasonable to select K so that the leading terms of the MSE are minimized. The aforementioned minimum is reached at

$$K_3^* = n\left(\frac{5(\mathbb{E}\tilde{\varepsilon}^4)^2}{96T^4\sigma^8}\right)^{\frac{1}{5}}. \tag{5.23}$$

[2] The optimal value of K proposed in Zhang *et al.* (2005) lacks the term $\mathbb{E}\varepsilon^4$ in the numerator.

Plugging (5.23) in (5.22), it follows that

$$MSE\left(\hat{\kappa}_{K_3^*}\right) = (4)5^{\frac{3}{5}}3^{-\frac{3}{5}}\left(\mathbb{E}\tilde{\varepsilon}^4\right)^{\frac{6}{5}}\sigma^{-\frac{24}{5}}T^{-\frac{2}{5}} + o\left(T^{-\frac{2}{5}}\right),$$

whose rate of convergence to 0 is slower than the rate of $O\left(T^{-2/3}\right)$ attained by the estimator $\hat{\sigma}^2_{K^*}$.

Finally, we consider the unbiased estimator for κ introduced in Proposition 5.1. Below, l.o.t. refers to lower order terms.

Theorem 5.6. *The estimator (5.10) is such that*

$$\operatorname{Var}\left(\hat{\hat{\kappa}}_{n,K}\right) = \frac{64}{5}\frac{T^2K^3}{n^3} + \frac{2n}{9\sigma^8T^2K^2}e(\varepsilon) + \text{l.o.t.}, \tag{5.24}$$

where $e(\varepsilon) = \operatorname{Var}\left((\varepsilon_2-\varepsilon_1)^4\right)$.

The two terms on the right-hand side of (5.24) reach their minimum value at

$$K_4^* = n^{\frac{4}{5}}\left(\frac{5e(\varepsilon)}{(27)(16)T^4\sigma^8}\right)^{\frac{1}{5}}. \tag{5.25}$$

After plugging K_4^* in (5.24), we obtain that

$$MSE\left(\hat{\kappa}_{K_4^*}\right) = 2^{\frac{28}{5}}5^{-\frac{2}{5}}3^{-\frac{9}{5}}e(\varepsilon)^{\frac{3}{5}}\sigma^{-\frac{24}{5}}n^{-\frac{3}{5}}T^{-\frac{2}{5}} + o\left(T^{-1}\right),$$

which again, since $T/n \to 0$, implies that $MSE\left(\hat{\kappa}_{K_4^*}\right) = o(T^{-1})$. The aforementioned result should be compared to (3.10), which essentially says that the estimator $\hat{\kappa}_{K_4^*}$ has better efficiency than the continuous-time based estimator $\hat{\kappa}^{(T)}$, obtained by making $n\to\infty$ in the estimators $\hat{\kappa}_n$ and $\tilde{\kappa}_n$ (see (3.7)). It is worth pointing out here that one can devise a consistent estimator for $e(\varepsilon)$ in light of the relationships

$$\frac{1}{n}[\widetilde{X},\widetilde{X}]_4^{\bar{\mathcal{G}}_n} \xrightarrow{\mathbb{P}} \mathbb{E}\left(\varepsilon_2-\varepsilon_1\right)^4, \quad \frac{1}{n}[\widetilde{X},\widetilde{X}]_8^{\bar{\mathcal{G}}_n} \xrightarrow{\mathbb{P}} \mathbb{E}\left(\varepsilon_2-\varepsilon_1\right)^8. \tag{5.26}$$

6. Numerical Performance and Empirical Evidence

In the sequel, we consider a Variance Gamma (VG) model with white Gaussian microstructure noise. The variance of the noise ε_t is denoted by ϱ^2 so that the noise of the i^{th} increment, $\tilde{\varepsilon}_{i,n}$, is $\mathcal{N}(0,2\varrho^2)$. Other parameters are set as follows:

$$\sigma = 0.02, \quad \kappa = 0.3, \quad \varrho = 0.005.$$

where the time unit is a day. In particular, the above value of σ corresponds to an annualized volatility of $0.02\sqrt{252} = 0.31$.

In this section, we propose an iterative method to implement the estimators described in the previous section, with the corresponding optimal choices of K^*. The main issue arises from the fact that in order to accurately estimate σ, we need to choose K as in (5.18) (or (5.15)), which precisely depends on what we want to estimate: σ. So, we propose to start with an initial reasonable guess for σ^2 to find K^*, which in turn is then used to improve the initial guess of σ, and so forth. The superior finite-sample and empirical performance of the resulting estimators are illustrated by simulation and a real high-frequency data application.

6.1. Estimators for σ

We compare the finite sample performance of the following estimators:

1. The estimator $\hat{\sigma}^2_{n,K}$ given in (5.1) with K determined by a suitable estimate of the optimal value K_1^* given in (5.15). As shown in Proposition 4.2, a consistent estimator for $m_2(\tilde{\varepsilon}) = \mathbb{E}\tilde{\varepsilon}^2 = 2\mathbb{E}\varepsilon^2 = 2\varrho^2$ is provided by $\hat{m}_{2,n}(\tilde{\varepsilon}) := \delta_n \tilde{\sigma}^2_n = [\widetilde{X}, \widetilde{X}]_2^{\bar{\mathcal{G}}_n}/n$, which suggests the following consistent estimate for $\mathbb{E}\varepsilon^2$:
$$\hat{\varrho}^2 := \widehat{\mathbb{E}\varepsilon^2} := \frac{1}{2n}[\widetilde{X}, \widetilde{X}]_2^{\bar{\mathcal{G}}_n}. \tag{6.1}$$
The only missing ingredient for estimating (5.15) is an initial preliminary estimate of σ^2, which we then proceed to improve via $\hat{\sigma}^2_{n,K^*}$. Concretely, we propose the following procedure. First, we evaluate the estimate:
$$\hat{K}_1^* := n\left(\frac{6\left(\widehat{\mathbb{E}\varepsilon^2}\right)^2}{T^2\sigma_0^4}\right)^{1/3}, \tag{6.2}$$
where σ_0 is an initial "reasonable" value for the volatility. Second, we estimate σ via $\hat{\sigma}_1' := \hat{\sigma}_{n,\hat{K}_1^*}$. Next, we use $\hat{\sigma}_1'$ to improve our estimate of K^* by setting
$$\hat{\hat{K}}_1^* := n\left(\frac{6\left(\widehat{\mathbb{E}\varepsilon^2}\right)^2}{T^2\left(\hat{\sigma}_1'\right)^4}\right)^{1/3}. \tag{6.3}$$
Finally, we set $\hat{\sigma}_1'' := \hat{\sigma}_{n,\hat{\hat{K}}_1^*}$
2. We consider the bias-corrected estimator $\hat{\hat{\sigma}}^2_{n,K}$ introduced in (5.5), with a value of K given by (6.2). We denote this estimator $\hat{\sigma}_2'$. We also analyze an iterative procedure similar to that in item 1, but using $\hat{\sigma}_2'$. Concretely, we set
$$\hat{\sigma}_2'' = \hat{\hat{\sigma}}^2_{n,\hat{\hat{K}}_1^*},$$
where $\hat{\hat{K}}_1^*$ is defined analogously to (6.3) but replacing $\hat{\sigma}_1'$ with $\hat{\sigma}_2'$.
3. Finally, we also consider the estimator $\hat{\hat{\sigma}}^2_{n,K}$ introduced in (5.5) but using the optimal value K_2^* in (5.18). Concretely, we set $\hat{\sigma}_3' = \hat{\hat{\sigma}}^2_{n,\hat{K}_2^*}$ with
$$\hat{K}_2^* := n^{2/3}\left(\frac{6\left[\widehat{\mathbb{E}\varepsilon^4} + \left(\widehat{\mathbb{E}\varepsilon^2}\right)^2\right]}{T^2\sigma_0^4}\right)^{1/3},$$
where σ_0 is an initial reasonable value for σ and $\widehat{\mathbb{E}\varepsilon^4}$ is a consistent estimator for $\mathbb{E}\varepsilon^4$. Next, we improve

the estimate of $\hat{\sigma}_3'$ by setting

$$\hat{\sigma}_3'' := \hat{\bar{\sigma}}^2_{n,\hat{\hat{K}}_2^*}, \quad \text{with} \quad \hat{\hat{K}}_2^* := n^{2/3}\left(\frac{6\left[\widehat{\mathbb{E}\varepsilon^4}+\left(\widehat{\mathbb{E}\varepsilon^2}\right)^2\right]}{T^2(\hat{\sigma}_3')^4}\right)^{1/3}. \tag{6.4}$$

To estimate $\mathbb{E}\varepsilon^4$, we use (5.9). Concretely, as shown in the proof of Proposition 4.3, we have

$$\hat{m}_{4,n}(\tilde{\varepsilon}) := \frac{1}{n}[\widetilde{X},\widetilde{X}]_4^{\bar{\mathcal{G}}_n} \to m_4(\tilde{\varepsilon}) := \mathbb{E}\left(\tilde{\varepsilon}^4\right) = 2\mathbb{E}\varepsilon^4 + 6\left(\mathbb{E}\varepsilon^2\right)^2. \tag{6.5}$$

Therefore, a consistent estimate for $\mathbb{E}\varepsilon^4$ is given by

$$\widehat{\mathbb{E}\varepsilon^4} := \frac{1}{2n}[\widetilde{X},\widetilde{X}]_4^{\bar{\mathcal{G}}_n} - \frac{1}{3}\left(\widehat{\mathbb{E}\varepsilon^2}\right)^2.$$

The sample mean, standard deviation, and mean-squared error (MSE) based on 1000 simulations are presented in the Table 2. Here, we take $T = 252$ days and $\sigma_0 \approx 0.063$, which corresponds to an annualized volatility of 1. As expected, the estimator $\hat{\sigma}_1'$ exhibits a noticeable bias and that this bias is corrected by $\hat{\sigma}_2'$. However, $\hat{\sigma}_3''$ is much more superior to other considered estimators, which is consistent with the asymptotic results for the mean-squared errors described in Eqs. (5.16) and (5.19).

	$\hat{\sigma}_1'$	$\hat{\sigma}_1''$	$\hat{\sigma}_2'$	$\hat{\sigma}_2''$	$\hat{\sigma}_3'$	$\hat{\sigma}_3''$
			$\delta_n = 5$ min			
Mean	0.02274333	0.02066226	0.01998258	0.01988843	0.01999695	0.01999614
Std Dev	0.0006854182	0.0011434344	0.0007945224	0.0012479476	0.0008839566	0.0007044640
MSE	7.995654e-06	1.746024e-06	6.315694e-07	1.569822e-06	7.813885e-07	4.962843e-07
			$\delta_n = 1$ min			
Mean	0.02288498	0.02066931	0.01995456	0.01984824	0.01997237	0.02000242
Std Dev	0.0006482329	0.0010605652	0.0007468549	0.0011609025	0.0007887707	0.0006469303
MSE	8.743311e-06	1.572774e-06	5.598574e-07	1.370725e-06	6.229225e-07	4.185247e-07
			$\delta_n = 30$ sec			
Mean	0.02293765	0.02075251	0.01998865	0.01993685	0.02000009	0.02001709
Std Dev	0.0006537998	0.0010611910	0.0007515176	0.0011497640	0.0007185258	0.0006364266
MSE	9.057229e-06	1.692391e-06	5.649076e-07	1.325945e-06	5.162794e-07	4.053310e-07
			$\delta_n = 10$ sec			
Mean	0.02291554	0.02068342	0.01994569	0.01985480	0.01998979	0.01998727
Std Dev	0.0006660059	0.0010835667	0.0007655451	0.0011754339	0.0006954765	0.0006270494
MSE	8.943940e-06	1.641179e-06	5.890091e-07	1.402727e-06	4.837918e-07	3.933530e-07
			$\delta_n = 1$ sec			
Mean	0.02296041	0.02076158	0.01998938	0.01994110	0.02000240	0.02000628
Std Dev	0.0006346972	0.0010546469	0.0007285086	0.0011415267	0.0006393828	0.0005973219
MSE	9.166839e-06	1.692287e-06	5.308377e-07	1.306553e-06	4.088161e-07	3.568328e-07

TABLE 2

Sample means, standard deviations, and mean-squared errors for different estimator of $\sigma = 0.02$ based on 1000 simulations.

6.2. Estimators for κ

We compare the finite sample performance of the following three estimators, which are respectively denoted by $\hat{\kappa}_1, \hat{\kappa}_2, \hat{\kappa}_3$.

1. The estimator $\hat{\kappa}_{n,K}$ given in (5.6) with σ replaced with the estimate $\hat{\sigma}_3''$ in Eq. (6.4) and K determined by a suitable estimate of the optimal value K_3^* given in (5.23). To estimate $\mathbb{E}\tilde{\varepsilon}^4$, we used the statistic in the first limit in Eq. (5.26).
2. The unbiased estimator $\hat{\hat{\kappa}}_n$ defined in (5.10) with the same value of K as the previous item. As before, we replace σ by the estimator $\hat{\sigma}_3''$.
3. Again, the unbiased estimator $\hat{\hat{\kappa}}_n$ in (5.10) replacing σ with $\hat{\sigma}_3''$, but now the value of K is given by (5.25). We replace σ therein with $\hat{\sigma}_3''$, while to estimate $e(\varepsilon) = \mathrm{Var}\left((\varepsilon_2 - \varepsilon_1)^4\right)$, we exploit the limits in (5.26).

The sample mean, standard deviation, and mean-squared error (MSE) based on 1000 simulations are presented in Table 3. Here, we take $T = 252$ days and $\sigma_0 = 0.063$. As expected, the estimator $\hat{\sigma}_3$ has much better performance than any other estimator therein.

	$\hat{\kappa}_1$	$\hat{\kappa}_2$	$\hat{\kappa}_3$
$\delta_n = 5$ min			
Mean	0.57771957	0.29982420	0.29967835
Std Dev	0.1783289311	0.1832631941	0.0979104650
MSE	1.089294e-01	3.358543e-02	9.586563e-03
$\delta_n = 1$ min			
Mean	0.57428966	0.29189326	0.29686684
Std Dev	0.1571320926	0.1599275870	0.0758019358
MSE	9.992531e-02	2.564255e-02	5.755750e-03
$\delta_n = 30$ sec			
Mean	0.58111784	0.29929056	0.29677713
Std Dev	0.161799873	0.163678990	0.069347518
MSE	1.052064e-01	2.679132e-02	4.819465e-03
$\delta_n = 10$ sec			
Mean	0.57194184	0.28868504	0.29216507
Std Dev	0.1581111167	0.1605442985	0.0609263392
MSE	9.895149e-02	2.590250e-02	3.773405e-03
$\delta_n = 1$ sec			
Mean	0.57371817	0.29046728	0.29455234
Std Dev	0.162874998	0.165066890	0.066836990
MSE	1.014499e-01	2.733795e-02	4.496860e-03

TABLE 3

Sample means, standard deviations, and mean-squared errors for different estimator of $\kappa = 0.3$ *based on* 1000 *simulations.*

6.3. Empirical study

We now proceed to analyze the performance of the proposed estimators when applied to real data. As it was explained above and was theoretically verified by Propositions 4.2-4.3, traditional estimators are not stable as the sampling frequency increases. Indeed, $\hat{\sigma}_n$ and $\tilde{\sigma}_n$ both diverge to ∞ while $\hat{\kappa}_n$ and $\tilde{\kappa}_n$ converge to 0, as $n \to \infty$. The objective is to verify that the proposed estimators do not exhibit the aforementioned behaviors at very high-frequencies.

We consider high-frequency stock data for four stocks during 2005, which were obtained from the NYSE TAQ database of Wharton's WRDS system. We then compute the estimator $\hat{\varrho}$ defined in (6.1), the estimator $\hat{\sigma}_{n,K}$ defined in (5.1) with $K = 1$, the estimator $\hat{\hat{\sigma}}_{n,K}$ defined in (5.5) with $K = \hat{\hat{K}}_1^*$ as given in (6.4), the estimator $\hat{\kappa}_{n,K}$ defined in (5.6) with $K = 1$, and finally the estimator $\hat{\hat{\kappa}}_{n,K}$ defined in (5.10) with $K = \hat{K}_4^*$ as given in (5.25). In the case of $\hat{\kappa}_{n,1}$, we used $\sigma = \hat{\sigma}_{n,1}$. Both $\hat{\sigma}_{n,1}$ and $\hat{\kappa}_{n,1}$ represent the estimators without any technique to alleviate the effect of the microstructure noise. As one can see in Tables 4-7, the estimators $\hat{\hat{\sigma}}$ and $\hat{\hat{\kappa}}$ do not exhibit the drawbacks of the estimators $\hat{\sigma}$ and $\hat{\kappa}$ at high frequencies. As a consequence of the empirical results therein, for instance, Intel's stock exhibits an annualized volatility σ of about $0.014 * \sqrt{252} = 0.22$ per year, while its excess kurtosis increases with $1/\delta$ at a rate of about 0.5. In comparison, even though the volatility of Pfizer's stock is just slightly larger (about $0.015 * \sqrt{252} = 0.23$), its excess kurtosis increases at a rate of about 2.3 with $1/\delta$, showing much more riskiness due to the much heavier tails of its return's distribution. This example illustrates the importance of considering a parameter which measures the tail heaviness of the return distribution and not only its variance.

	$\hat{\varrho}$	$\hat{\sigma}_{n,1}$	$\hat{\hat{\sigma}}_{n,\hat{K}_1^*}$	$\hat{\kappa}_{n,1}$	$\hat{\hat{\kappa}}_{n,\hat{K}_4^*}$
20 min	0.002198811	0.013732969	0.013115165	0.772846688	0.645084939
10 min	0.001584536	0.013995671	0.013112833	0.589344904	0.727208959
5 min	0.001152404	0.014394983	0.013253727	0.495378704	0.768302688
1 min	0.0005581856	0.0155908617	0.0136519981	0.3499494734	0.7293149570
30 sec	0.0004113675	0.0162494093	0.0139405766	0.2817929514	0.6875741045
20 sec	0.0003483541	0.0168528945	0.0141596310	0.2566280373	0.6575495762
10 sec	0.0002712869	0.0185608431	0.0145174963	0.1831341414	0.5921934015
5 sec	0.0002174315	0.0210381061	0.0147818871	0.1084570206	0.4987667343

TABLE 4
Estimation of the parameters σ and κ of a subordinated Brownian motion with microstructure noise for INTC (Intel) stock.

Appendix A: Proofs

A.1. Proof of Proposition 3.1.

We shall need the following result that can be shown easily from the moment generating function for Poisson integrals:

	$\hat{\varrho}$	$\hat{\sigma}_{n,1}$	$\hat{\hat{\sigma}}_{n,\hat{K}_1^*}$	$\hat{\kappa}_{n,1}$	$\hat{\hat{\kappa}}_{n,\hat{K}_4^*}$
20 min	0.002310884	0.014432934	0.014279133	3.552809339	3.665645436
10 min	0.001678615	0.014826633	0.013921679	3.330420039	4.192632331
5 min	0.001223294	0.015280492	0.013758805	3.395593192	4.458814370
1 min	0.000581559	0.016243711	0.014289601	2.885849749	3.074717720
30 sec	0.0004379718	0.0173003060	0.0147847384	2.1009477905	2.5399891978
20 sec	0.0003733763	0.0180634325	0.0149589310	1.8189209947	2.3582752416
10 sec	0.0003021168	0.0206701623	0.0150440707	1.0395706194	2.3194219287
5 sec	0.0002547010	0.0246442060	0.0151395852	0.5255478783	2.3750789809

Table 5

Estimation of the parameters σ and κ of a subordinated Brownian motion with microstructure noise for PFE (Pfeizer) stock.

	$\hat{\varrho}$	$\hat{\sigma}_{n,1}$	$\hat{\hat{\sigma}}_{n,\hat{K}_1^*}$	$\hat{\kappa}_{n,1}$	$\hat{\hat{\kappa}}_{n,\hat{K}_4^*}$
20 min	0.002001343	0.012499653	0.011799576	0.239780179	0.652064235
10 min	0.001444737	0.012760874	0.011941252	0.219621789	0.639680272
5 min	0.001065309	0.013307053	0.012056169	0.180598527	0.533660824
1 min	0.0005279981	0.0147476856	0.0126650388	0.1255556037	0.2780196460
30 sec	0.0004058086	0.0160298273	0.0129515252	0.0912895956	0.2414279693
20 sec	0.0003567437	0.0172587727	0.0131456378	0.0681058924	0.2208446641
10 sec	0.0002976521	0.0203646952	0.0132914281	0.0347817660	0.1993772187
5 sec	0.0002546159	0.0246359741	0.0133782455	0.0156713991	0.1890478414

Table 6

Estimation of the parameters σ and κ of a subordinated Brownian motion with microstructure noise for CSCO (Cisco) index.

Lemma A.1. *Suppose that M is a Poisson random measure with mean measure m defined on an open domain of $\mathbb{R}^d$ and let $\bar{M}(f) = \int f(z)(M-m)(dz)$ denote the integral of f with respect the compensated random measure $\bar{M} = M - m$. If $m(|f|^k) := \int |f(z)|^k m(dz) < \infty$ for $k = 1, \ldots, 5$, then the following formulas hold true:*

$$\mathbb{E}\left(\bar{M}(f)\right)^k = \begin{cases} 0, & k=1, \\ m(f^k), & k=2,3 \\ 3m(f^2)^2 + m(f^4), & k=4, \\ 10m(f^2)m(f^3) + m(f^5), & k=5. \end{cases}$$

$$\mathbb{E}\left\{\bar{M}(g)\bar{M}(f)^k\right\} = \begin{cases} m(gf^k), & k=1,2 \\ m(gf^3) + 3m(f^2)m(gf), & k=3. \end{cases}$$

Lemma A.2. *Let M be the Poisson jump measure of a Lévy process X with Lévy measure ν and let $\bar{M}(dx,dt) := M(dx,dt) - \nu(dx)dt$ be the corresponding compensated measure. Also, suppose that f is such that $\int |f(x)|^k \nu(dx) < \infty$ for some $k \geq 2$. Then, for any $T \geq 1$, there exists a constant $A_k(f)$ such that*

$$\mathbb{E}\left|\frac{1}{T}\int_0^T \int f(x)\bar{M}(dx,dt)\right|^k \leq A_k(f) T^{-k/2}.$$

Proof. If T is a positive integer, the result is a direct consequence of the following inequality (see Lemma 5.3.1 in Bickel & Doksum (2001)):

$$\mathbb{E}|\bar{Z}_n - \mu_Z|^k \leq C_k \mathbb{E}|Z_1|^k n^{-k/2}, \tag{A.1}$$

	$\hat{\varrho}$	$\hat{\sigma}_{n,1}$	$\hat{\bar{\sigma}}_{n,\hat{K}_1^*}$	$\hat{\kappa}_{n,1}$	$\hat{\bar{\kappa}}_{n,\hat{K}_4^*}$
20 min	0.002654995	0.016582127	0.015223242	0.204160629	0.254309823
10 min	0.001923067	0.016985799	0.015798588	0.256973339	0.204201473
5 min	0.001390687	0.017371437	0.016239610	0.162056277	0.229330646
1 min	0.000666907	0.018627596	0.016883173	0.088320021	0.232902894
30 sec	0.0004982082	0.0196797006	0.0172213711	0.0597469161	0.1919979467
20 sec	0.0004287861	0.0207440869	0.0173514240	0.0448251291	0.1764936472
10 sec	0.0003347169	0.0229005848	0.0175005599	0.0306273072	0.1690680474
5 sec	0.0002653463	0.0256742175	0.0177124265	0.0193756187	0.1521620542

TABLE 7

Estimation of the parameters σ and κ of a subordinated Brownian motion with microstructure noise for AMAT (Applied Materials) stock.

where $\bar{Z}_n = \frac{1}{n}\sum_{i=1}^n Z_i$, $\mu_Z = \mathbb{E}Z_1$, and $\{Z_i\}_i$ are i.i.d. such that $\mathbb{E}|Z_1|^k < \infty$. For general T, let $[T]$ the integer part of T. Then, we can write

$$\frac{1}{T}\int_0^T \int f(x)\bar{M}(dx,dt) = \frac{1}{T}\int_0^{[T]} \int f(x)\bar{M}(dx,dt) + \frac{1}{T}\int_{[T]}^T \int f(x)\bar{M}(dx,dt).$$

Hence,

$$\mathbb{E}\left|\frac{1}{T}\int_0^T \int f(x)\bar{M}(dx,dt)\right|^k \leq 2^k\mathbb{E}\left|\frac{1}{[T]}\int_0^{[T]} \int \frac{[T]}{T} f(x)\bar{M}(dx,dt)\right|^k + 2^k T^{-k}\mathbb{E}\left|\int_0^{T-[T]} \int f(x)\bar{M}(dx,dt)\right|^k$$

For the first term on the right-hand side above, we apply inequality (A.1). For the second term, note that by Burkholder-Davis-Gundy inequality (see Protter (2004)),

$$\mathbb{E}\left|\int_0^{T-[T]} \int f(x)\bar{M}(dx,dt)\right|^k \leq \mathbb{E}\left|\sup_{t\leq 1}\int_0^t \int f(x)\bar{M}(dx,dt)\right|^k \leq B_k^k\mathbb{E}\left|\int_0^1 \int f^2(x)M(dx,dt)\right|^{k/2}.$$

This suffices to obtain the inequality of the lemma. □

Proof of Proposition 3.1. Using the identity

$$\frac{1}{(1+x)^2} = \sum_{i=0}^{k-1}(-1)^i(i+1)x^i + \frac{(-1)^k x^k}{(1+x)^2}(k+1+kx), \tag{A.2}$$

and the notation

$$\hat{\mu}_k^{(T)} := \frac{1}{T}\int_0^T \int x^k M(dt,dx), \quad \hat{D}_T := \frac{\hat{\mu}_2^{(T)}}{c_2(X_1)} - 1,$$

we have the following decomposition:

$$\begin{aligned}\mathbb{E}\hat{\kappa}^{(T)} = \frac{1}{3c_2^2(X_1)}\mathbb{E}\left\{\frac{\hat{\mu}_4^{(T)}}{(1+\hat{D}_T)^2}\right\} &= \frac{1}{3c_2^2(X_1)}\mathbb{E}\left\{\hat{\mu}_4^{(T)}\left(1 - 2\hat{D}_T + 3\hat{D}_T^2 - 4\hat{D}_T^3 + 5\hat{D}_T^4 - 6\hat{D}_T^5\right)\right\}\\ &\quad + \frac{1}{3}\mathbb{E}\left\{\hat{\mu}_4^{(T)}\left(\hat{\mu}_2^{(T)}\right)^{-2}\left(7+6\hat{D}_T\right)\hat{D}_T^6\right\}\\ &=: L_T + R_T.\end{aligned}$$

Let us first analyze the residual term R_T using the following easy consequence of the triangle inequality:

$$(\hat{\mu}_4^{(T)})^{1/2} = \frac{1}{T^{1/2}}\left(\sum_{s\le T}(\Delta X_s)^4\right)^{1/2} \le \frac{1}{T^{1/2}}\sum_{s\le T}(\Delta X_s)^2 = T^{1/2}\mu_2^{(T)}. \tag{A.3}$$

Thus, since $7+6\hat{D}_T = 1+6(1+\hat{D}_T) = 1+6\hat{\mu}_2^{(T)}/c_2(X_1) > 0$, we have that

$$\begin{aligned}0 \le R_T &\le \frac{7T}{3}\mathbb{E}\left(\hat{D}_T^6\right) + \frac{6T}{3}\mathbb{E}\left(\hat{D}_T^7\right)\\ &= \frac{7T}{3c_2^6(X_1)}\mathbb{E}\left(\hat{\mu}_2^{(T)} - c_2(X_1)\right)^6 + \frac{6T}{3c_2^7(X_1)}\mathbb{E}\left(\hat{\mu}_2^{(T)} - c_2(X_1)\right)^7.\end{aligned}$$

Using that $\mathbb{E}\hat{\mu}_2^{(T)} = c_2(X_1)$ and Lemma A.2, $R_T = O(T^{-2})$. Similarly, using Lemma A.1, the first four terms of L_T (i.e. those multiplying $\hat{D}_T^i$ up to $i=3$) can be seen to be

$$\frac{c_4(X_1)}{3c_2^2(X_1)} - \frac{2c_6(X_1)}{3c_2^3(X_1)}T^{-1} + \frac{c_4^2(X_1)}{c_2^4(X_1)}T^{-1} + O(T^{-2}).$$

The last two term of L_T can be seen to be $O(T^{-2})$ from Lemma A.2 and Cauchy inequality. Indeed,

$$\begin{aligned}\left|\mathbb{E}\hat{\mu}_4^{(T)}\hat{D}_T^4\right| &\le c_4(X_1)\left|\mathbb{E}\hat{D}_T^4\right| + \frac{1}{c_2^4(X_1)}\left|\mathbb{E}\left(\hat{\mu}_4^{(T)} - c_4(X_1)\right)\left(\hat{\mu}_2^{(T)} - c_2(X_1)\right)^4\right|\\ &\le Kc_2T^{-2} + \left(\mathbb{E}\left(\hat{\mu}_4^{(T)} - c_4(X_1)\right)^2\mathbb{E}\left(\hat{\mu}_2^{(T)} - c_2(X_1)\right)^8\right)^{1/2},\end{aligned}$$

which is $O(T^{-2})$ in light of Lemma A.2. We finally obtain that

$$\mathbb{E}\hat{\kappa}_n \overset{n\to\infty}{\longrightarrow} \mathbb{E}\hat{\kappa}^{(T)} = \frac{c_4(X_1)}{3c_2^2(X_1)} - \frac{2c_6(X_1)}{3c_2^3(X_1)}T^{-1} + \frac{c_4^2(X_1)}{c_2^4(X_1)}T^{-1} + O(T^{-2}).$$

In order to show the bound for the variance, we use again (A.2) to get

$$\hat{\kappa}^{(T)} = \frac{\hat{\mu}_4^{(T)}}{3c_2^2(X_1)}\left(1 - 2\hat{D}_T + 3\hat{D}_T^2 - 4\hat{D}_T^3\right) + \frac{1}{3}\frac{\hat{\mu}_4^{(T)}}{\left(\hat{\mu}_2^{(T)}\right)^2}\left(5 + 4\hat{D}_T\right)\hat{D}_T^4.$$

Then,

$$\begin{aligned}\hat{\kappa}^{(T)} - \frac{c_4(X_1)}{3c_2^2(X_1)} &= \frac{1}{3c_2^2(X_1)}\left(\hat{\mu}_4^{(T)} - c_4(X_1)\right) - \frac{2\hat{\mu}_4^{(T)}}{3c_2^2(X_1)}\hat{D}_T + \frac{\hat{\mu}_4^{(T)}}{c_2^2(X_1)}\hat{D}_T^2\\ &\quad - \frac{4\hat{\mu}_4^{(T)}}{3c_2^2(X_1)}\hat{D}_T^3 + \frac{1}{3}\frac{\hat{\mu}_4^{(T)}}{\left(\hat{\mu}_2^{(T)}\right)^2}\left(5 + 4\hat{D}_T\right)\hat{D}_T^4.\end{aligned}$$

After expanding the squares, taking expectations both sides, and using Cauchy's inequality together with Lemmas A.1 and A.2, one can check that all the terms are at least $O(T^{-2})$ except possibly the following terms:

$$\frac{1}{9c_2^4(X_1)}\mathbb{E}\left\{\left(\hat{\mu}_4^{(T)}-c_4(X_1)\right)^2\right\}-\frac{4}{9c_2^4(X_1)}\mathbb{E}\left\{\left(\hat{\mu}_4^{(T)}-c_4(X_1)\right)\hat{\mu}_4^{(T)}\hat{D}_T\right\}$$
$$+\frac{4}{9c_2^4(X_1)}\mathbb{E}\left\{(\hat{\mu}_4^{(T)})^2\hat{D}_T^2\right\}.$$

Subtracting $c_4(X_1)$ from $\hat{\mu}_4^{(T)}$ in the second and third terms above, and using again Lemmas A.1 and A.2, we can check that the above expression indeed coincides with the expression in (3.10). $\square$

A.2. Proofs of Section 5.

Proof of Theorem 5.2. Clearly,

$$\mathrm{Var}\left(\hat{\sigma}^2_{n,K}\right)=\frac{2}{K^2T^2}\sum_{1\le i<j\le K}\mathrm{Cov}\left([\widetilde{X},\widetilde{X}]_2^{\mathcal{G}_n^{(i)}},[\widetilde{X},\widetilde{X}]_2^{\mathcal{G}_n^{(j)}}\right)+\frac{1}{K^2T^2}\sum_{i=1}^{K}\mathrm{Var}\left([\widetilde{X},\widetilde{X}]_2^{\mathcal{G}_n^{(i)}}\right). \qquad \text{(A.4)}$$

Each covariance in the first term on the right hand side above is given by

$$\begin{aligned}
A_{i,j}&:=\mathrm{Cov}\left([\widetilde{X},\widetilde{X}]_2^{\mathcal{G}_n^{(i)}},[\widetilde{X},\widetilde{X}]_2^{\mathcal{G}_n^{(j)}}\right)\\
&=\sum_{q=0}^{n_i-1}\sum_{r=0}^{n_j-1}\mathrm{Cov}\left(\left|\widetilde{X}(t_{i-1+(q+1)K})-\widetilde{X}(t_{i-1+qK})\right|^2,\left|\widetilde{X}(t_{j-1+(r+1)K})-\widetilde{X}(t_{j-1+rK})\right|^2\right)\\
&=n_i\mathrm{Cov}\left(\left|\widetilde{X}(t_{i-1+K})-\widetilde{X}(t_{i-1})\right|^2,\left|\widetilde{X}(t_{j-1+K})-\widetilde{X}(t_{j-1})\right|^2\right)\\
&\quad+(n_j-1)\mathrm{Cov}\left(\left|\widetilde{X}(t_{i-1+2K})-\widetilde{X}(t_{i-1+K})\right|^2,\left|\widetilde{X}(t_{j-1+K})-\widetilde{X}(t_{j-1})\right|^2\right)\\
&=n_iC\left((K+i-j)\delta_n\right)+(n_j-1)\,C\left((j-i)\delta_n\right),
\end{aligned}$$

where, for any $u<t<t+\delta<v$,

$$C(\delta):=\mathrm{Cov}\left(\left|\widetilde{X}(t+\delta)-\widetilde{X}(u)\right|^2,\left|\widetilde{X}(v)-\widetilde{X}(t)\right|^2\right),$$

which can be proved to depend only on $\delta>0$. More specifically, note that $C(\delta)=\mathrm{Cov}\left(|S+U|^2,|S+V|^2\right)$, where $S:=X(t+\delta)-X(t)$, $U:=X(t)-X(u)+\varepsilon_{t+\delta}-\varepsilon_u$, and $V:=X(v)-X(t+\delta)+\varepsilon_v-\varepsilon_t$. Next, using that independence of S, U, and V,

$$\begin{aligned}
C(\delta)&=\mathrm{Var}\left(S^2\right)+2\mathrm{Cov}\left(S^2,SV\right)+2\mathrm{Cov}\left(SU,S^2\right)+4\mathrm{Cov}\left(SU,SV\right)\\
&=\mathrm{Var}\left(S^2\right)+2\mathbb{E}(V)\mathrm{Cov}\left(S^2,S\right)+2\mathbb{E}(U)\mathrm{Cov}\left(S,S^2\right)+4\mathbb{E}(U)\mathbb{E}(V)\mathrm{Var}\left(S\right).
\end{aligned}$$

Finally, using that $\mathbb{E}U = \mathbb{E}V = 0$ as well as the moment formulas in (3.1), $C(\delta) = \mathrm{Var}\left(S^2\right) = 2\sigma^4\delta^2 + 3\sigma^4\kappa\delta$. Using the previous formula together with the fact that $|n_i - n/K| < 2$, the first term in (A.4), which we denote A, can be computed as follows:

$$\begin{aligned} A &= \frac{n}{K}\frac{2}{K^2T^2}\sum_{1\leq i<j\leq K}\left(2\sigma^4(j-i)^2\delta_n^2 + 3\sigma^4\kappa(j-i)\delta_n\right) \\ &\quad + \frac{n}{K}\frac{2}{K^2T^2}\sum_{1\leq i<j\leq K}\left(2\sigma^4(K+i-j)^2\delta_n^2 + 3\sigma^4\kappa(K+i-j)\delta_n\right) + \mathcal{R} \\ &= \frac{n}{K}\frac{K-1}{KT^2}\left(3\sigma^4\kappa K\delta_n + \frac{2}{3}\sigma^4K(2K-1)\delta_n^2\right) + \mathcal{R}, \end{aligned}$$

where $\mathcal{R}$ is such that

$$|\mathcal{R}| \leq \frac{4(K-1)}{KT^2}\left(3\sigma^4\kappa K\delta_n + \frac{2}{3}\sigma^4K(2K-1)\delta_n^2\right). \tag{A.5}$$

Now, we consider the second term in (A.4), which we denote B. Each variance term of B can be written as

$$\begin{aligned} B_i &:= \mathrm{Var}\left([\widetilde{X},\widetilde{X}]_2^{\mathcal{G}_n^{(i)}}\right) \\ &= \sum_{q=0}^{n_i-1}\mathrm{Var}\left(\left|\widetilde{X}(t_{i-1+(q+1)K}) - \widetilde{X}(t_{i-1+qK})\right|^2\right) \\ &\quad + 2\sum_{q=0}^{n_i-2}\mathrm{Cov}\left(\left|\widetilde{X}(t_{i-1+(q+1)K}) - \widetilde{X}(t_{i-1+qK})\right|^2, \left|\widetilde{X}(t_{i-1+(q+2)K}) - \widetilde{X}(t_{i-1+(q+1)K})\right|^2\right). \end{aligned}$$

Next, using the relationships

$$\begin{aligned} \mathrm{Var}\left(\left|\widetilde{X}(t+\delta) - \widetilde{X}(t)\right|^2\right) &= 2\sigma^4\delta^2 + 3\sigma^4\kappa\delta + 8\sigma^2\mathbb{E}\left(\varepsilon^2\right)\delta + 2\mathbb{E}\left(\varepsilon^2\right)^2 + 2\mathbb{E}\left(\varepsilon^4\right) \\ \mathrm{Cov}\left(\left|\widetilde{X}(t+\delta) - \widetilde{X}(t)\right|^2, \left|\widetilde{X}(v) - \widetilde{X}(t+\delta)\right|^2\right) &= \mathbb{E}\left(\varepsilon^4\right) - \mathbb{E}\left(\varepsilon^2\right)^2, \end{aligned}$$

valid for any $t < t+\delta < v$, we get

$$\begin{aligned} B_i &= n_i\left(2\sigma^4\left(K\delta_n\right)^2 + 3\sigma^4\kappa\left(K\delta_n\right) + 8\sigma^2\mathbb{E}\left(\varepsilon^2\right)\left(K\delta_n\right) + 2\mathbb{E}\left(\varepsilon^2\right)^2 + 2\mathbb{E}\left(\varepsilon^4\right)\right) + 2(n_i-1)\left(\mathbb{E}\left(\varepsilon^4\right) - \mathbb{E}\left(\varepsilon^2\right)^2\right) \\ &= n_i\left(2\sigma^4\left(K\delta_n\right)^2 + 3\sigma^4\kappa\left(K\delta_n\right) + 8\sigma^2\mathbb{E}\left(\varepsilon^2\right)\left(K\delta_n\right)\right) + 2(2n_i-1)\mathbb{E}\left(\varepsilon^4\right) + 2\mathbb{E}\left(\varepsilon^2\right)^2. \end{aligned} \tag{A.6}$$

Therefore, using that $\sum_{i=1}^K n_i = n - K + 1$,

$$B = \frac{n-K+1}{K^2T^2}\left(2\sigma^4\left(K\delta_n\right)^2 + 3\sigma^4\kappa\left(K\delta_n\right) + 8\sigma^2\mathbb{E}\left(\varepsilon^2\right)\left(K\delta_n\right) + 4\mathbb{E}\left(\varepsilon^4\right)\right) - \frac{2}{KT^2}\left(\mathbb{E}\left(\varepsilon^4\right) - \mathbb{E}\left(\varepsilon^2\right)^2\right).$$

Putting together A and B above,

$$\begin{aligned}\operatorname{Var}\left(\hat{\sigma}^2_{n,K}\right) &= \frac{n}{K}\frac{K-1}{KT^2}\left(3\sigma^4\kappa K\delta_n + \frac{2}{3}\sigma^4 K(2K-1)\delta_n^2\right) + \mathcal{R} \\ &+ \frac{n-K+1}{K^2T^2}\left(2\sigma^4\left(K\delta_n\right)^2 + 3\sigma^4\kappa\left(K\delta_n\right) + 8\sigma^2\mathbb{E}\left(\varepsilon^2\right)\left(K\delta_n\right) + 4\mathbb{E}\left(\varepsilon^4\right)\right) \\ &- \frac{2}{KT^2}\left(\mathbb{E}\left(\varepsilon^4\right) - \mathbb{E}\left(\varepsilon^2\right)^2\right).\end{aligned} \tag{A.7}$$

Recalling that $\delta_n = T/n$ and using (A.5), we get the expression (5.11). □

Proof of Proposition 5.4. Let $a_{n.K} := \frac{nK}{(n+1)(K-1)}$ and $b_{n,K} := \frac{n-K+1}{T(n+1)(K-1)}$. Clearly,

$$\operatorname{Var}\left(\hat{\bar{\sigma}}^2_{n,K}\right) = a^2_{n,K}\operatorname{Var}\left(\hat{\sigma}^2_{n,K}\right) + b^2_{n,K}\operatorname{Var}\left([\widetilde{X},\widetilde{X}]_2^{\bar{\mathcal{G}}_n}\right) - 2a_{n,K}b_{n,K}\operatorname{Cov}\left(\hat{\sigma}^2_{n,K}, [\widetilde{X},\widetilde{X}]_2^{\bar{\mathcal{G}}_n}\right)$$

From the expressions in Eqs. (A.6)-(A.7), we have

$$\begin{aligned}\operatorname{Var}\left([\widetilde{X},\widetilde{X}]_2^{\bar{\mathcal{G}}_n}\right) &= n\left(2\sigma^4\delta_n^2 + 3\sigma^4\kappa\delta_n + 8\sigma^2\mathbb{E}\left(\varepsilon^2\right)\delta_n\right) + 2(2n-1)\mathbb{E}\left(\varepsilon^4\right) + 2\mathbb{E}\left(\varepsilon^2\right)^2 \\ \operatorname{Var}\left(\hat{\sigma}^2_{n,K}\right) &= \frac{n}{K}\frac{K-1}{KT^2}\left(3\sigma^4\kappa K\delta_n + \frac{2}{3}\sigma^4 K(2K-1)\delta_n^2\right) + \mathcal{R} \\ &+ \frac{n-K+1}{K^2T^2}\left(2\sigma^4\left(K\delta_n\right)^2 + 3\sigma^4\kappa\left(K\delta_n\right) + 8\sigma^2\mathbb{E}\left(\varepsilon^2\right)\left(K\delta_n\right) + 4\mathbb{E}\left(\varepsilon^4\right)\right) \\ &- \frac{2}{KT^2}\left(\mathbb{E}\left(\varepsilon^4\right) - \mathbb{E}\left(\varepsilon^2\right)^2\right).\end{aligned}$$

To compute the last covariance, let us first note that

$$\operatorname{Cov}\left(\hat{\sigma}^2_{n,K}, [\widetilde{X},\widetilde{X}]_2^{\bar{\mathcal{G}}_n}\right) = \frac{1}{KT}\sum_{i=1}^{K}\operatorname{Cov}\left([\widetilde{X},\widetilde{X}]_2^{\mathcal{G}_n^{(i)}}, [\widetilde{X},\widetilde{X}]_2^{\bar{\mathcal{G}}_n}\right) =: \frac{1}{KT}\sum_{i=1}^{K}B_i. \tag{A.8}$$

Each covariance term on the right hand side above is given by

$$\begin{aligned}B_i &= \sum_{q=0}^{n_i-1}\sum_{r=0}^{n-1}\operatorname{Cov}\left(\left|\widetilde{X}(t_{i-1+(q+1)K}) - \widetilde{X}(t_{i-1+qK})\right|^2, \left|\widetilde{X}(t_{r+1}) - \widetilde{X}(t_r)\right|^2\right) \\ &= (n_i - e_i)\sum_{r=0}^{n-1}\operatorname{Cov}\left(\left|\widetilde{X}(t_{i-1+2K}) - \widetilde{X}(t_{i-1+K})\right|^2, \left|\widetilde{X}(t_{r+1}) - \widetilde{X}(t_r)\right|^2\right) \\ &+ e_i\sum_{r=0}^{n-1}\operatorname{Cov}\left(\left|\widetilde{X}(t_K) - \widetilde{X}(t_0)\right|^2, \left|\widetilde{X}(t_{r+1}) - \widetilde{X}(t_r)\right|^2\right),\end{aligned}$$

where above e_i denote the number of subintervals in $\left\{[t_{i-1+qK}, t_{i-1+(q+1)K}]\right\}_{i=1}^{n_i}$ which intersect the end points 0 and T. Note that $\sum_{i=1}^{K} e_i = 2$. Now, we use the following formulas, which can be directly computed

$$\begin{aligned}\operatorname{Cov}\left(|\widetilde{X}(v) - \widetilde{X}(u)|^2, |\widetilde{X}(v') - \widetilde{X}(u')|^2\right) &= 2\sigma^4(v'-u')^2 + 3\kappa\sigma^4(v'-u'), \qquad u < u' < v' < v \\ \operatorname{Cov}\left(|\widetilde{X}(t) - \widetilde{X}(s)|^2, |\widetilde{X}(u) - \widetilde{X}(t)|^2\right) &= \mathbb{E}\varepsilon^4 - (\mathbb{E}\varepsilon^2)^2, \qquad s < t < u.\end{aligned}$$

We then get $B_i = n_i \left\{ K\left(2\sigma^4\delta_n^2 + 3\kappa\sigma^4\delta_n\right) + 2(\mathbb{E}\varepsilon^4 - (\mathbb{E}\varepsilon^2)^2)\right\} - e_i(\mathbb{E}\varepsilon^4 - (\mathbb{E}\varepsilon^2)^2)$. Next, using that $\sum_{i=1}^K n_i = n - K + 1$ and $\sum_{i=1}^K e_i = 2$,

$$\mathrm{Cov}\left(\hat{\sigma}^2_{n,K}, [\widetilde{X},\widetilde{X}]_2^{\bar{\mathcal{G}}_n}\right) = \frac{n-K+1}{T}\left(2\sigma^4\delta_n^2 + 3\kappa\sigma^4\delta_n\right) + 2\frac{n-K}{TK}(\mathbb{E}\varepsilon^4 - (\mathbb{E}\varepsilon^2)^2).$$

Putting together the previous relationships,

$$\mathrm{Var}\left(\hat{\hat{\sigma}}^2_{n,K}\right) = \frac{4\sigma^4 K}{3n} + \frac{4n\left(\mathbb{E}\left(\varepsilon^4\right) + (\mathbb{E}\varepsilon^2)^2\right)}{T^2K^2} + O\left(\frac{1}{n}\right) + O\left(\frac{n}{K^3T^2}\right) + O\left(\frac{1}{TK}\right).$$

□

Proof of Theorem 5.5. The proof is similar to that of Theorem 5.2. Again, first note that

$$\mathrm{Var}\left(\hat{\kappa}_{n,K}\right) = \frac{2}{9\sigma^8K^2T^2}\sum_{1\le i<j\le K}\mathrm{Cov}\left([\widetilde{X},\widetilde{X}]_4^{\mathcal{G}_n^{(i)}}, [\widetilde{X},\widetilde{X}]_4^{\mathcal{G}_n^{(j)}}\right) + \frac{1}{9\sigma^8K^2T^2}\sum_{i=1}^K \mathrm{Var}\left([\widetilde{X},\widetilde{X}]_4^{\mathcal{G}_n^{(i)}}\right) \tag{A.9}$$

$$=: A + B. \tag{A.10}$$

Each covariance in the first term on the right hand side above is given by

$$\begin{aligned}
A_{i,j} &:= \mathrm{Cov}\left([\widetilde{X},\widetilde{X}]_4^{\mathcal{G}_n^{(i)}}, [\widetilde{X},\widetilde{X}]_4^{\mathcal{G}_n^{(j)}}\right)\\
&= \sum_{q=0}^{n_i-1}\sum_{r=0}^{n_j-1}\mathrm{Cov}\left(\left|\widetilde{X}(t_{i-1+(q+1)K}) - \widetilde{X}(t_{i-1+qK})\right|^4, \left|\widetilde{X}(t_{j-1+(r+1)K}) - \widetilde{X}(t_{j-1+rK})\right|^4\right)\\
&= n_i\mathrm{Cov}\left(\left|\widetilde{X}(t_{i-1+K}) - \widetilde{X}(t_{i-1})\right|^4, \left|\widetilde{X}(t_{j-1+K}) - \widetilde{X}(t_{j-1})\right|^4\right)\\
&\quad + (n_j-1)\mathrm{Cov}\left(\left|\widetilde{X}(t_{i-1+2K}) - \widetilde{X}(t_{i-1+K})\right|^4, \left|\widetilde{X}(t_{j-1+K}) - \widetilde{X}(t_{j-1})\right|^4\right)\\
&= n_iC\left((j-i)\delta_n, (K+i-j)\delta_n, (j-i)\delta_n\right) + (n_j-1)\,C\left((K+i-j)\delta_n, (j-i)\delta_n, (K+i-j)\delta_n\right),
\end{aligned}$$

where, for any $t, s_1, s_2, s_3 > 0$,

$$C(s_1,s_2,s_3) := \mathrm{Cov}\left(\left|\widetilde{X}_{t+s_1+s_2} - \widetilde{X}_t\right|^4, \left|\widetilde{X}_{t+s_1+s_2+s_3} - \widetilde{X}_{t+s_1}\right|^4\right), \tag{A.11}$$

which again can be proved to be independent of t. Concretely, with the notation $S := X_{t+s_1+s_2} - X_{t+s_1}$, $U := X_{t+s_1} - X_t + \varepsilon_{t+s_1+s_2} - \varepsilon_t$, and $V := X_{t+s_1+s_2+s_3} - X_{t+s_1+s_2} + \varepsilon_{t+s_1+s_2+s_3} - \varepsilon_{t+s_1}$

$$\begin{aligned}
C(s_1,s_2,s_3) &= \mathrm{Cov}\left(|S+U|^4, |S+V|^4\right)\\
&= \mathrm{Var}\left(S^4\right) + 6\left[\mathbb{E}(U^2) + \mathbb{E}(V^2)\right]\mathrm{Cov}\left(S^4, S^2\right)\\
&\quad + 36\mathbb{E}(U^2)\mathbb{E}(V^2)\mathrm{Var}\left(S^2\right) + 16\mathbb{E}(U^3)\mathbb{E}(V^3)\mathrm{Var}\left(S\right)
\end{aligned}$$

where above we used the independence of S, U, and V as well as the fact that $\mathbb{E}U = \mathbb{E}V = \mathbb{E}S^k = 0$ for any odd positive integer k. Upon computation of the relevant moments of U and V, we get

$$\begin{aligned} C(s_1, s_2, s_3) &= \operatorname{Var}\left(X_{s_2}^4\right) + 6\left[\sigma^2(s_1+s_3) + 4\mathbb{E}\varepsilon^2\right]\operatorname{Cov}\left(X_{s_2}^4, X_{s_2}^2\right) \\ &\quad + 6^2\left(\sigma^2 s_1 + 2\mathbb{E}\varepsilon\right)\left(\sigma^2 s_3 + 2\mathbb{E}\varepsilon\right)\operatorname{Var}\left(X_{s_2}^2\right) + 4^2\left(2\mathbb{E}\varepsilon^3\right)^2\operatorname{Var}\left(X_{s_2}\right). \end{aligned} \tag{A.12}$$

Note that

$$\mathbb{E}X_s^k = \mathbb{E}\left((\sigma W_{\tau_s})^k\right) = \sigma^k\mathbb{E}\left(W_1^k\right)\mathbb{E}\left(\tau_s^{k/2}\right) = \sigma^k\mathbb{E}\left(W_1^k\right)\left(s^{k/2} + \sum_{i=1}^{k/2-1} a_{k,i}s^i\right),$$

for some constant $a_{k,i}$'s. we now proceed to analyze each term separately:

- The contribution to A due to $\operatorname{Var}\left(X_{s_2}^4\right)$:

$$A^{(1)} := \frac{n}{K}\frac{2}{9\sigma^8K^2T^2}\sum_{1\le i<j\le K}\operatorname{Var}\left(X^4_{(K+i-j)\delta_n}\right) + \frac{n}{K}\frac{2}{9\sigma^8K^2T^2}\sum_{1\le i<j\le K}\operatorname{Var}\left(X^4_{(j-i)\delta_n}\right).$$

  Using that $\operatorname{Var}\left(X_t^4\right)$ is a polynomial of degree 4 in t with the highest-degree term being $96\sigma^8t^4$,

$$\begin{aligned} A^{(1)} &= \frac{n}{K}\frac{192\delta_n^4}{9K^2T^2}\left(\sum_{1\le i<j\le K}(K+i-j)^4 + \sum_{1\le i<j\le K}(j-i)^4 + O\left(K^5\right)\right) \\ &= \frac{n}{K}\frac{192\delta_n^4}{9K^2T^2}\left(\sum_{i=1}^{K-1}\sum_{\ell=1}^{K-i}(K-\ell)^4 + \sum_{i=1}^{K-1}\sum_{\ell=1}^{K-i}\ell^4 + O\left(K^5\right)\right) \\ &= \frac{n}{K}\frac{192\delta_n^4}{9K^2T^2}K\sum_{\ell=1}^{K-1}\ell^4 + O\left(\frac{T^2K^2}{n^3}\right) \\ &= \frac{192}{5(9)}\frac{T^2K^3}{n^3} + O\left(\frac{T^2K^2}{n^3}\right). \end{aligned}$$

- Let us analyze the contribution to A due to $\operatorname{Var}\left(X_{s_2}^2\right)$. The leading term is given by:

$$\begin{aligned} A^{(2)} &:= 6^2\frac{n}{K}\frac{2}{9\sigma^8K^2T^2}\sum_{1\le i<j\le K}\left(\sigma^2(j-i)\delta_n\right)^2\operatorname{Var}\left(X^2_{(K+i-j)\delta_n}\right) \\ &\quad + 6^2\frac{n}{K}\frac{2}{9\sigma^8K^2T^2}\sum_{1\le i<j\le K}\left(\sigma^2(K+i-j)\delta_n\right)^2\operatorname{Var}\left(X^2_{(j-i)\delta_n}\right) \\ &= 6^2\frac{n}{K}\frac{2}{9\sigma^8K^2T^2}\sum_{1\le i<j\le K}\left(\sigma^2(j-i)\delta_n\right)^2\left(3\sigma^4\kappa(K+i-j)\delta_n + 2\sigma^4(K+i-j)^2\delta_n^2\right) \\ &\quad + 6^2\frac{n}{K}\frac{2}{9\sigma^8K^2T^2}\sum_{1\le i<j\le K}\left(\sigma^2(K+i-j)\delta_n\right)^2\left(3\sigma^4\kappa(j-i)\delta_n + 2\sigma^4(j-i)^2\delta_n^2\right) \\ &= \frac{(6)(4)(13)}{5(9)}\frac{T^2K^3}{n^3} + O\left(\frac{T^2K^2}{n^3}\right) \end{aligned}$$

- The contribution to A due to $\mathrm{Cov}\left(X_{s_2}^4, X_{s_2}^2\right)$ is given by:

$$\begin{aligned}
A^{(3)} &:= 6\frac{n}{K}\frac{2}{9\sigma^8K^2T^2}\sum_{1\le i<j\le K}\left(2\sigma^2(j-i)\delta_n\right)\mathrm{Cov}\left(X^2_{(K+i-j)\delta_n}, X^4_{(K+i-j)\delta_n}\right)\\
&\quad + 6\frac{n}{K}\frac{2}{9\sigma^8K^2T^2}\sum_{1\le i<j\le K}\left(2\sigma^2(K+i-j)\delta_n\right)\mathrm{Cov}\left(X^2_{(j-i)\delta_n}, X^4_{(j-i)\delta_n}\right)\\
&= \frac{12^2}{9}\sigma^8\frac{n}{K}\frac{2\delta_n^4}{\sigma^8K^2T^2}\left(\sum_{1\le i<j\le K}(j-i)(K+i-j)^3+\sum_{1\le i<j\le K}(K+i-j)(j-i)^3+O\left(K^5\right)\right)\\
&= \frac{12^2(2)}{(5)(4)(9)}\frac{T^2K^3}{n^3}+O\left(\frac{T^2K^2}{n^3}\right)
\end{aligned}$$

  where above we used that $\mathrm{Cov}\left(X_s^2, X_s^4\right) = \mathbb{E}X_s^6 - \mathbb{E}(X_s^2)\mathbb{E}(X_s^4) = 12\sigma^6s^3 + \text{l.o.t.}.$
- Finally, the contribution to A due to $\mathrm{Var}\left(X_{s_2}\right)$ will generate a term of smaller order than T^2K^3/n^3. Indeed,

$$\begin{aligned}
A^{(4)} &:= 4^2\left(2\mathbb{E}\varepsilon^3\right)^2\frac{n}{K}\frac{2}{9\sigma^8K^2T^2}\sum_{1\le i<j\le K}\left(\mathrm{Var}\left(X_{(K+i-j)\delta_n}\right)+\mathrm{Var}\left(X_{(j-i)\delta_n}\right)\right)\\
&:= 4^2\left(2\mathbb{E}\varepsilon^3\right)^2\frac{n}{K}\frac{2}{9\sigma^8K^2T^2}\sigma^2\delta_n\sum_{1\le i<j\le K}\left((K+i-j)+(j-i)\right)\\
&= \frac{4^2}{9}\left(2\mathbb{E}\varepsilon^3\right)^2\frac{1}{\sigma^6T}.
\end{aligned}$$

Putting together the above relationships,

$$\begin{aligned}
A &= \frac{192}{5(9)}\frac{T^2K^3}{n^3}+\frac{(6)(4)(13)}{5(9)}\frac{T^2K^3}{n^3}+\frac{6^2(2)}{5(9)}\frac{T^2K^3}{n^3}++O\left(\frac{T^2K^2}{n^3}\right)\\
&= \frac{576}{5(9)}\frac{T^2K^3}{n^3}+O\left(\frac{T^2K^2}{n^3}\right).
\end{aligned}$$

Now, we consider the second term in (A.14), which we denote B. Each variance term, $B_i := \mathrm{Var}\big([\widetilde{X},\widetilde{X}]_4^{\mathcal{G}_n^{(i)}}\big)$, of B can be written as

$$\begin{aligned}
B_i &= \sum_{q=0}^{n_i-1}\mathrm{Var}\left(\left|\widetilde{X}(t_{i-1+(q+1)K})-\widetilde{X}(t_{i-1+qK})\right|^4\right)\\
&\quad + 2\sum_{q=0}^{n_i-2}\mathrm{Cov}\left(\left|\widetilde{X}(t_{i-1+(q+1)K})-\widetilde{X}(t_{i-1+qK})\right|^4, \left|\widetilde{X}(t_{i-1+(q+2)K})-\widetilde{X}(t_{i-1+(q+1)K})\right|^4\right).
\end{aligned}$$

Next, using arguments similar to those following (A.11),

$$\mathrm{Var}\left(\left|\widetilde{X}_{t+s}-\widetilde{X}_t\right|^4\right) = \mathrm{Var}\left(\left|X_{t+s}-X_t\right|^4\right)+\text{l.o.t.} = 96\sigma^8s^4+\text{l.o.t.}, \tag{A.13}$$

$$\mathrm{Cov}\left(\left|\widetilde{X}(t+s_1)-\widetilde{X}(t)\right|^4, \left|\widetilde{X}(t+s_1+s_2)-\widetilde{X}(t+s_1)\right|^4\right) = -36\mathbb{E}X_{s_1}^2\mathbb{E}X_{s_2}^2\mathbb{E}\varepsilon^2+\text{l.o.t.} = -36\sigma^4\mathbb{E}\varepsilon^2s_1s_2+\text{l.o.t.}$$

valid for any $t, s_1, s_2 > 0$ and where l.o.t. mean lower order terms. Therefore,

$$B_i = n_i \left(96\sigma^8 \left(K\delta_n \right)^4 \right) + O\left((K\delta_n)^3 \right).$$

Therefore,

$$B = \frac{96}{9} \frac{K^2T^2}{n^3} + O\left(\frac{KT}{n^2} \right).$$

which shows that $B = O(T^2K^2/n^3)$. Finally,

$$\mathrm{Var}\left(\hat{\kappa}_{n,K} \right) = \frac{576}{5(9)} \frac{T^2K^3}{n^3} + \frac{96}{9} \frac{K^2T^2}{n^3} + O\left(\frac{KT}{n^2} \right),$$

which implies the result.

$\square$

Proof of Theorem 5.6. Let $a_{n.K} := \frac{nK}{(n+1)(K-1)}$, $b_{n,K} := \frac{n-K+1}{3\sigma^4 T(n+1)(K-1)}$, and $c_{n,K} := \frac{2}{n\sigma^2}$. Clearly,

$$\begin{aligned}
\mathrm{Var}\left(\hat{\bar{\kappa}}_n \right) &= a_{n,K}^2 \mathrm{Var}\left(\hat{\kappa}_{n,K} \right) + b_{n,K}^2 \mathrm{Var}\left([\widetilde{X}, \widetilde{X}]_4^{\bar{\mathcal{G}}_n} \right) \\
&\quad + c_{n,K}^2 \mathrm{Var}\left([\widetilde{X}, \widetilde{X}]_2^{\bar{\mathcal{G}}_n} \right) - 2a_{n,K}b_{n,K} \mathrm{Cov}\left(\hat{\kappa}_{n,K}, [\widetilde{X}, \widetilde{X}]_4^{\bar{\mathcal{G}}_n} \right) \\
&\quad - 2a_{n,K}c_{n,K} \mathrm{Cov}\left(\hat{\kappa}_{n,K}, [\widetilde{X}, \widetilde{X}]_2^{\bar{\mathcal{G}}_n} \right) + 2b_{n,K}c_{n,K} \mathrm{Cov}\left([\widetilde{X}, \widetilde{X}]_4^{\bar{\mathcal{G}}_n}, [\widetilde{X}, \widetilde{X}]_2^{\bar{\mathcal{G}}_n} \right)
\end{aligned}$$

As in the case of the variance of $\hat{\bar{\sigma}}_{n,K}$, we are looking for the terms of the highest power of K and the terms with the highest power of n (and least negative power of K). For $\mathrm{Var}\left(\hat{\kappa}_{n,K} \right)$, the highest power of K is given by (5.20). To find the highest power of n, we recall from the proof of Theorem 5.5 that the variance can be decomposed into two terms, called A and B therein. The term with the highest power n in A is due to the term $4^2(2\mathbb{E}\varepsilon^3)^2\mathrm{Var}(X_{s_2})$ in (A.12) and is of order n^0. In order to determine the term with the highest power of n in B, note that this will be due to the constant terms of the variance and covariance in Eqs. (A.13). These are given by

$$\mathrm{Var}\left(\left| \widetilde{X}_{t+s} - \widetilde{X}_t \right|^4 \right) = \mathrm{Var}\left((\varepsilon_2 - \varepsilon_1)^4 \right) + \text{h.o.t.}, \tag{A.14}$$

$$\mathrm{Cov}\left(\left| \widetilde{X}(t+s_1) - \widetilde{X}(t) \right|^4, \left| \widetilde{X}(t+s_1+s_2) - \widetilde{X}(t+s_1) \right|^4 \right) = \mathrm{Cov}\left(\left| \varepsilon_2 - \varepsilon_1 \right|^4, \left| \varepsilon_3 - \varepsilon_2 \right|^4 \right) + \text{h.o.t.}$$

where h.o.t. means higher order term (as powers of s, s_1, and s_2). These terms contribute to B as follows:

$$\begin{aligned}
B &:= \frac{1}{9\sigma^8 K^2 T^2} \sum_{i=1}^{K} \mathrm{Var}\left([\widetilde{X}, \widetilde{X}]_4^{\mathcal{G}_n^{(i)}} \right) \\
&= \frac{1}{9\sigma^8 K^2 T^2} \sum_{i=1}^{K} \left(n_i \mathrm{Var}\left((\varepsilon_2 - \varepsilon_1)^4 \right) + 2(n_i - 1) \mathrm{Cov}\left(\left| \varepsilon_2 - \varepsilon_1 \right|^4, \left| \varepsilon_3 - \varepsilon_2 \right|^4 \right) \right) + \text{o.t.} \\
&= \frac{n}{9\sigma^8 K^2 T^2} d(\varepsilon) + \text{o.t.},
\end{aligned}$$

where $d(\varepsilon) := \mathrm{Var}\left((\varepsilon_2 - \varepsilon_1)^4\right) + 2\mathrm{Cov}\left(|\varepsilon_2 - \varepsilon_1|^4, |\varepsilon_3 - \varepsilon_2|^4\right)$.

Now we consider $b_{n,K}^2 \mathrm{Var}\left([\widetilde{X}, \widetilde{X}]_4^{\bar{\mathcal{G}}_n}\right)$. As done with B, the term with the highest degree in n is $\frac{n}{9\sigma^8 T^2 K^2} d(\varepsilon)$. Clearly, all the terms in $c_{n,K}^2 \mathrm{Var}\left([\widetilde{X}, \widetilde{X}]_2^{\bar{\mathcal{G}}_n}\right)$ are of lower order than $n/(T^2K^2)$.

To compute $\mathrm{Cov}\left(\hat{\kappa}_{n,K}^2, [\widetilde{X}, \widetilde{X}]_4^{\bar{\mathcal{G}}_n}\right)$, let us first note that

$$\mathrm{Cov}\left(\hat{\kappa}_{n,K}, [\widetilde{X}, \widetilde{X}]_4^{\bar{\mathcal{G}}_n}\right) = \frac{1}{3\sigma^4 KT} \sum_{i=1}^{K} \mathrm{Cov}\left([\widetilde{X}, \widetilde{X}]_4^{\mathcal{G}_n^{(i)}}, [\widetilde{X}, \widetilde{X}]_4^{\bar{\mathcal{G}}_n}\right) =: \frac{1}{3\sigma^4 KT} \sum_{i=1}^{K} B_i. \tag{A.15}$$

Each covariance term on the right hand side above is given by

$$\begin{aligned} B_i &= \sum_{q=0}^{n_i-1} \sum_{r=0}^{n-1} \mathrm{Cov}\left(\left|\widetilde{X}(t_{i-1+(q+1)K}) - \widetilde{X}(t_{i-1+qK})\right|^4, \left|\widetilde{X}(t_{r+1}) - \widetilde{X}(t_r)\right|^4\right) \\ &= (n_i - e_i) \sum_{r=0}^{n-1} \mathrm{Cov}\left(\left|\widetilde{X}(t_{i-1+2K}) - \widetilde{X}(t_{i-1+K})\right|^4, \left|\widetilde{X}(t_{r+1}) - \widetilde{X}(t_r)\right|^4\right) \\ &\quad + e_i \sum_{r=0}^{n-1} \mathrm{Cov}\left(\left|\widetilde{X}(t_K) - \widetilde{X}(t_0)\right|^4, \left|\widetilde{X}(t_{r+1}) - \widetilde{X}(t_r)\right|^4\right), \end{aligned}$$

where above e_i denote the number of subintervals in $\left\{[t_{i-1+qK}, t_{i-1+(q+1)K}]\right\}_{i=1}^{n_i}$ which intersect the end points 0 and T. Note that $\sum_{i=1}^{K} e_i = 2$. Now, it turns out that

$$\begin{aligned} &\mathrm{Cov}\left(|\widetilde{X}(v) - \widetilde{X}(u)|^4, |\widetilde{X}(v') - \widetilde{X}(u')|^4\right) \asymp n^{-1}, \qquad u < u' < v' < v \\ &\mathrm{Cov}\left(|\widetilde{X}(t) - \widetilde{X}(s)|^4, |\widetilde{X}(u) - \widetilde{X}(t)|^4\right) = \mathrm{Cov}\left(|\varepsilon_2 - \varepsilon_1|^4, |\varepsilon_3 - \varepsilon_2|^4\right) =: g(\varepsilon), \qquad s < t < u, \end{aligned} \tag{A.16}$$

where here $a_n \asymp b_n$ means $\lim_{n\to\infty} a_n/b_n \in \mathbb{R}\backslash\{0\}$. We then get $B_i = 2n_i g(\varepsilon) - e_i g(\varepsilon) + o.t.$. Next, using that $\sum_{i=1}^{K} n_i = n - K + 1$ and $\sum_{i=1}^{K} e_i = 2$,

$$\mathrm{Cov}\left(\hat{\kappa}_{n,K}, [\widetilde{X}, \widetilde{X}]_4^{\bar{\mathcal{G}}_n}\right) = \frac{2}{3\sigma^4} \frac{n-K}{TK} g(\varepsilon) + \text{o.t.}$$

Therefore, the contribution here is $-\frac{4n}{9\sigma^8 T^2 K^2} g(\varepsilon)$ Given that $c_{n,K}$ is of order n^{-1}, it is not hard to see that the term $-2a_{n,K} c_{n,K} \mathrm{Cov}\left(\hat{\kappa}_{n,K}, [\widetilde{X}, \widetilde{X}]_2^{\bar{\mathcal{G}}_n}\right)$ is of an order smaller than n. Finally, consider the term corresponding to $D_n := \mathrm{Cov}\left([\widetilde{X}, \widetilde{X}]_4^{\bar{\mathcal{G}}_n}, [\widetilde{X}, \widetilde{X}]_2^{\bar{\mathcal{G}}_n}\right)$. Note that

$$\begin{aligned} D_n &= \sum_{q=0}^{n-1} \sum_{r=0}^{n-1} \mathrm{Cov}\left(\left|\widetilde{X}_{t_{q+1}} - \widetilde{X}_{t_q}\right|^4, \left|\widetilde{X}_{t_{r+1}} - \widetilde{X}_{t_r}\right|^2\right) \\ &= n\left(\mathrm{Cov}\left(\left|\widetilde{X}_{t_1} - \widetilde{X}_{t_0}\right|^4, \left|\widetilde{X}_{t_1} - \widetilde{X}_{t_0}\right|^2\right) + 2\mathrm{Cov}\left(\left|\widetilde{X}_{t_1} - \widetilde{X}_{t_0}\right|^4, \left|\widetilde{X}_{t_2} - \widetilde{X}_{t_1}\right|^2\right)\right) \\ &\quad - \mathrm{Cov}\left(\left|\widetilde{X}_{t_1} - \widetilde{X}_{t_0}\right|^4, \left|\widetilde{X}_{t_2} - \widetilde{X}_{t_1}\right|^2\right) \end{aligned}$$

Using (A.16), it is clear that $D_n \asymp 1$. Hence,

$$2b_{n,K}c_{n,K}\mathrm{Cov}\left([\widetilde{X},\widetilde{X}]_4^{\bar{\mathcal{G}}_n},[\widetilde{X},\widetilde{X}]_2^{\bar{\mathcal{G}}_n}\right) \asymp \frac{2}{3\sigma^4 TK}.$$

Finally, we obtain that

$$\mathrm{Var}\left(\hat{\hat{\kappa}}_{n,K}\right) = \frac{64}{5}\frac{T^2K^3}{n^3} + \frac{n}{9\sigma^8K^2T^2}d(\varepsilon) + \frac{n}{9\sigma^8T^2K^2}d(\varepsilon) - \frac{4n}{9\sigma^8T^2K^2}g(\varepsilon) + \text{l.o.t.}$$

which implies the result. □

References

Aït-Sahalia, Y., Mykland, P. & Zhang, L. (2005). How often to sample a continuous-time process in the presence of market microstructure noise. *The review of financial studies* **18**(2).

Asmussen, S. & Rosiński, J. (2001). Approximations of small jumps of Levy processes with a view towards simulation. *Journal of Applied Probability* **38**, 482–493.

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

Barndorff-Nielsen, O. (1998). Processes of normal inverse Gaussian type. *Finance and Stochastics* **2**, 41–68.

Behr, A. & Pötter, U. (2009). Alternatives to the normal model of stock returns: Gaussian mixture, generalised logF and generalised hyperbolic models. *Annals of Finance* **5**, 49–68.

Bickel, P. & Doksum, K. (2001). *Mathematical Statistics. Basic ideas and selected topics. Vol. I.* Prentice Hall.

Campbell, J., Lo, A. & MacKinlay, A. (1997). *The econometrics of Financial Markets.* Princeton.

Carr, P., Madan, D. & Chang, E. (1998). The variance Gamma process and option pricing. *European Finance Review* **2**, 79–105.

Cont, R. & Tankov, P. (2004). *Financial modelling with Jump Processes.* Chapman & Hall.

Figueroa-López, J. (2008). Small-time moment asymptotics for Lévy processes. *Statistics and Probability Letters* **78**, 3355–3365.

Figueroa-López, J., Lancette, S., Lee, K. & Mi, Y. (2011). Estimation of NIG and VG models for high frequency financial data. *Handbook of Modeling High-Frequency Data in Finance, F. Viens, M.C. Mariani, and I. Florescu (eds.). J. Wiley.* .

Hansen, P. & Lunde, A. (2006). Realized variance and market microstructure noise. *J. Bus. Econom. Statist.* **24**, 127–218.

KALLENBERG, O. (1997). *Foundations of Modern Probability.* Springer-Verlag, Berlin, New York, Heidelberg.

MYKLAND, P. & ZHANG, L. (2012). The econometrics of high-frequency data. *In Statistical Methods for Stochastic Differential Equations, M. Kessler, A. Lindner, and M. Srensen, eds.* , 109–190.

OOMEN, R. (2006). Properties of realized variance under alternative sampling schemes. *Journal of Business and Economic Statistics* **24**, 219–237.

PROTTER, P. (2004). *Stochastic Integration and Differentil Equations.* Springer. 2nd Edition.

RAMEZANI, C. & ZENG, Y. (2007). Maximum likelihood estimation of the double exponential jump-diffusion process. *Annals of Finance* **3**, 487–507.

ROLL, R. (1984). A simple implicit measure of the effective bid-ask spread in an efficient market. *Journal of Finance* **39**, 1127–1139.

SATO, K. (1999). *Lévy Processes and Infinitely Divisible Distributions.* Cambridge University Press.

SENETA, E. (2004). Fitting the variance-gamma model to financial data. *Journal of Applied Probability* **41A**, 177–187.

ZENG, Y. (2003). A partially observed model for micromovement of asset prices with Bayes estimation via filtering. *Mathematical Finance* **13**(3), 411–444.

ZHANG, L., MYKLAND, P. & AIT-SAHALIA, Y. (2005). A tale of two time scales: Determining integrated volatility with noisy high-frequency data. *J. Amer. Statist. Assoc.* **100**, 1294–1411.