博士論文（要約）

論文題目　Estimation and statistical testing of the correlation between latent processes
（潜在確率過程間の相関の推定及び統計的検定）

氏　　名　木村　晃敏

Estimation and statistical testing of the correlation between latent processes *

Akitoshi Kimura †

1 Histrical Review

The volatility of a financial asset price is an important risk measure. The realized volatility is a well-known consistent estimator of the integrated volatility. Its convergence rate and asymptotic mixed normality has been established and it is known that high frequency sampling of the price process gives accurate estimation. The covariance and correlation of two financial assets are also significant risk measures especially in the portfolio risk management. The realized covariance and realized correlation have similar asymptotic properties in high frequency sampling.

Epps [16] pointed out that the sample correlation between the returns of two different stocks decreases as the sampling frequency of data increases. It is considered that non-synchronicity of trading and market microstructure cause this phenomenon.

Since the trade timing of two financial assets is rarely synchronous in financial market,some synchronization is necessary to use the realized covariance. However, such estimator has serious bias when the interval size of synchronized sampling is small relative to the frequency of original trade timing (Hayashi and Yoshida [23]). Avoiding such synchronization, non-synchronous covariance estimation schemes have been developed: Fourier analytic approach (Malliavin and Mancino [33], Malliavin et al. [34]) and the cumulative covariance estimator (Hayashi and Yoshida [23, 21, 22], Mykland [41]).

The market microstructure is often modeled as the noise added to the latent price process. This modeling was successful and denoising techniques have also been developed: sub-sampling (Zhang et al. [49], Zhang [50]), pre-averaging (Podolskij and Vetter [44], Jacod et al. [25]), and others (Zhou [51]). There are many studies that treat both non-synchoronicity

*This thesis is based on the two of author's papers: Kimura and Yoshida [27] is published and Kimura [26] is submitted.

†Graduate School of Mathematical Sciences, University of Tokyo, 3-8-1 Komaba, Meguro-ku, Tokyo 153-8914, Japan, CREST, Japan Science and Technology Agency

and market microstructure noise: Malliavin and Mancino [35], Mancino and Sanfelici [36], Park and Linton [43], Voev and Lunde [48], Griffin and Oomen [19], Christensen et al. [12, 13], Koike [28, 30, 29], Aït-Sahalia et al. [4], Barndorff-Nielsen et al. [6], Bibinger [8, 9].

The followings are considered as the cause of market microstructure: bid-ask spread (Roll [45]), discretization error (Gottlieb and Kalay [18]), and asymmetric information (Glosten and Milgrom [17]). On the other hand, the relationship between the additive noise modeling and the market microstructure is not so clearly explained.

As the measurement and storage technology develop, the sampling time of trade becomes increasingly precise, and the ultra high frequency data of orders in the financial market is available today. In ultra high frequency sampling, the market microstructure is modeled as dynamics of the limit order book (LOB) rather than the noise, recently. Many studies adopt approaches modeling LOB with Poisson processes (Cont et al. [15], Abergel and Jedidi [2], Muni Toke [47, 38], Smith et al. [46], Muni Toke and Yoshida [40]), Hawkes processes (Hewlett [24], Large [32], Bowsher [10], Bacry et al. [5], Abergel and Jedidi [3], Muni Toke and Pomponio [39], Muni Toke [37], Clinet and Yoshida [14], Ogihara and Yoshida [42]), and doubly stochastic Poisson processes (Abergel et al. [1], Guilbaud and Pham [20], Chertok et al. [11], Korolev et al. [31]). These approaches can also treat the non-synchronicity of trading by nature.

In this stream, this thesis treats the model mentioned later that includes doubly stochastic Poisson processes. We present a simple estimator of the correlation between two latent processes indirectly observed high frequently and study its asymptotic behavior.

2 Model

Now, we consider a stochastic basis $\mathcal{B} = (\Omega, \mathcal{F}, \mathbb{F}, P)$, $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$. On $\mathcal{B}$, let $\mathbb{X} = (X^1, X^2)$ be an $\mathbb{R}^2$-valued Itô process given by

$$\mathbb{X}_t = \mathbb{X}_0 + \int_0^t \mathbb{X}_s^0\, ds + \int_0^t \mathbb{X}_s^1\, dw_s \quad (t \in [0,T]), \tag{2.1}$$

where w is an r-dimensional $\mathbb{F}$-Wiener process, $\mathbb{X}_0$ is an $\mathcal{F}_0$-measurable random variable, $\mathbb{X}^0$ is a two-dimensional $\mathbb{F}$-adapted process, and $\mathbb{X}^1$ is an $\mathbb{R}^2 \otimes \mathbb{R}^r$-valued $\mathbb{F}$-adapted process. satisfies condition [A] mentioned later. Let a_n be a positive number depending on n such that $a_n \to \infty$ as $n \to \infty$. On $\mathcal{B}$, consider a two-dimensional measurable process $\mathbb{Y}^n = (Y^{n,1}, Y^{n,2})$ having a decomposition

$$\mathbb{Y}_t^n = \mathbb{Y}_0^n + \int_0^t a_n \mathbb{X}_s\, ds + \mathbb{M}_t^n \quad (t \in [0,T]),$$

where $\mathbb{M}^n$ is a two-dimensional measurable process with $\mathbb{M}^n_0 = 0$. satisfies condition [B$^\flat$], [B], [B'], [B$^\sharp$], or [B$^{\sharp\sharp}$] mentioned later.

Example 1. *Suppose that each X^α is $\mathbb{R}_+ = [0,\infty)$-valued and $Y^{n,\alpha}$ is a counting process with intensity process $a_n X^\alpha$. This model describes the high frequency counting data of the orders or transactions in the active market, for example.*

Now, let $I_k = [t_{k-1}, t_k)$ for a sampling design $\Pi = (t_k)_{k=0,\dots,b_n}$ with $0 = t_0 < t_1 < \cdots < t_{b_n} = T$ and $h_k = t_k - t_{k-1}$. We assume $b_n \to \infty$ as $n \to \infty$. $\mathbb{X}$ is not observed, but $\mathbb{Y}^n$ is high frequently observed in Π.

The first aim of this thesis is to construct the estimator from the observed data of $\mathbb{Y}^n$ for the correlation of $\mathbb{X}$ observed indirectly. Then, we investigate the asymptotic behavior of the correlation estimator.

3 Assumptions

For simplicity, we assume $t_k = k(T/b_n)$, $h_k = T/b_n =: \delta_n$. We write $\Delta_k V = V_{t_k} - V_{t_{k-1}}$ for a process V. Consider the following conditions.

[A] Process $\mathbb{X}$ admits the representation (2.1) for an $\mathbb{R}^2$-valued $\mathcal{F}_0$-measurable random variable $\mathbb{X}_0$ and coefficients $\mathbb{X}^\kappa$ ($\kappa = 0, 1$) such that $\mathbb{X}^0$ is a cádlág $\mathbb{F}$-adapted process and that $\mathbb{X}^1$ has a representation

$$\mathbb{X}^1_t = \mathbb{X}^1_0 + \int_0^t \mathbb{X}^{10}_s \, ds + \int_0^t \sum_{\kappa'=1}^{r'} \mathbb{X}^{1\kappa'}_s \, d\tilde{w}^{\kappa'}_s \quad (t \in [0,T]),$$

where $\mathbb{X}^1_0$ is an $\mathbb{R}^2 \otimes \mathbb{R}^r$-valued $\mathcal{F}_0$-measurable random variable, $\tilde{w} = (\tilde{w}^1, ..., \tilde{w}^{r'})$ is an r'-dimensional $\mathbb{F}$-Wiener process (not necessary independent of w), and $\mathbb{X}^{1\kappa'}_s$ ($\kappa' = 0, 1, ..., r'$) are $\mathbb{R}^2 \otimes \mathbb{R}^r$-valued cádlág $\mathbb{F}$-adapted processes.

[B$^\flat$]
(i) $\lim_{n\to\infty} b_n^2/a_n = 0$.
(ii) $\sum_{k=1}^{b_n} |\Delta_k \mathbb{M}^n|^2 = O_p(a_n)$ as $n \to \infty$.

[B] $\mathbb{M}^n = (M^{n,\alpha})_{\alpha=1,2}$ is a two-dimensional $\mathbb{F}$-local martingale with $\mathbb{M}^n_0 = 0$ and such that

(i) $\lim_{n\to\infty} b_n^{5/2}/a_n = 0$.

(ii) $\sum_{k=1}^{b_n} |\Delta_k \mathbb{M}^n|^2 = O_p(a_n)$ as $n \to \infty$, $\sup_{t\in[0,1]} |\Delta \mathbb{M}^n| \le c a_n^{1/2}$ for a constant c independent of n.

(iii) The absolutely continuous (w.r.t. the Lebesgue measure a.s.) mapping $[0,T] \ni t \mapsto \langle \mathbb{M}^n, w\rangle_t \in \mathbb{R}^2 \otimes \mathbb{R}^r$ satisfies $\sup_{t\in[0,T]} |d\langle \mathbb{M}^n, w\rangle_t/dt| = O_p(b_n)$ as $n \to \infty$.

Here, $\langle \mathbb{M}^n, w\rangle_t$ is the $2\times r$ matrix of angle brackets $\langle M^{n,\alpha}, w^k\rangle_t$ for $\mathbb{M}^n = (M^{n,\alpha})_{\alpha=1,2}$ and $w = (w^k)_{k=1,\dots,r}$.

[B′] $\mathbb{M}^n$ is a two-dimensional $\mathbb{F}$-local martingale with $\mathbb{M}^n_0 = 0$, satisfies [B] (i), (iii), and

(ii′) $E[\sum_{k=1}^{b_n} |\Delta_k \mathbb{M}^n|^2] = O(a_n)$ as $n\to\infty$.

[B♯] (i) $\lim_{n\to\infty} b_n^3/a_n = 0$.

(ii) $\sum_{k=1}^{b_n} |\Delta_k \mathbb{M}^n|^2 = O_p(a_n)$ as $n\to\infty$.

[B♯♯] $\mathbb{M}^n$ is a two-dimensional $\mathbb{F}$-local martingale with $\mathbb{M}^n_0 = 0$, satisfies [B] (ii), (iii), and

(i♯♯) $\lim_{n\to\infty} b_n^3/a_n = 0$.

[C] $E[\sum_{k=1}^{b_n} |\Delta_k \mathbb{M}^n|^4] = O(a_n^2 b_n^{-1})$ as $n\to\infty$

4 Results

Before the statement of the results, we introduce some notations to simplify the description. Let $A^{\otimes(i,j)}$ is an (i,j)-element of $A^{\otimes}$, $A^{\otimes} = A\otimes A = AA^{\star}$, and * denotes transpose. Let $x\tilde{\otimes}y = ((x_iy_j + x_jy_i)/2) \in \mathbb{R}^{\mathsf{r}}\otimes\mathbb{R}^{\mathsf{r}}$ for $x=(x_i)$, $y=(y_i)\in\mathbb{R}^{\mathsf{r}}$, and let $x^{\tilde{\otimes}(\alpha,\beta)} = x^{\alpha}_{\cdot}\tilde{\otimes}x^{\beta}_{\cdot}$ for $x = (x^{\alpha}_i)\in\mathbb{R}^2\otimes\mathbb{R}^{\mathsf{r}}$. We write $x\cdot y = \sum_{i=1}^{\mathsf{r}} x_iy_i$ for $x=(x_i)$, $y=(y_i)\in\mathbb{R}^{\mathsf{r}}$, and $x\cdot y = \sum_{i,j=1}^{\mathsf{r}} x_{i,j}y_{i,j}$ for $x=(x_{i,j})$, $y=(y_{i,j})\in\mathbb{R}^{\mathsf{r}}\otimes\mathbb{R}^{\mathsf{r}}$. Write

$$\tilde{V}_k := \frac{\Delta_k V}{a_n h_k} = \frac{\Delta_k V}{a_n\delta_n}.$$

for a stochastic process V.

4.1 Correlation estimator

For $\alpha,\beta = 1,2$,

$$\begin{aligned} S_n^{\alpha\beta} &= \sum_{k=2}^{b_n}\Big(\frac{Y^{n,\alpha}_{t_k} - Y^{n,\alpha}_{t_{k-1}}}{a_nh_k} - \frac{Y^{n,\alpha}_{t_{k-1}} - Y^{n,\alpha}_{t_{k-2}}}{a_nh_{k-1}}\Big)\Big(\frac{Y^{n,\beta}_{t_k} - Y^{n,\beta}_{t_{k-1}}}{a_nh_k} - \frac{Y^{n,\beta}_{t_{k-1}} - Y^{n,\beta}_{t_{k-2}}}{a_nh_{k-1}}\Big) \\ &= \sum_{k=2}^{b_n}(\tilde{\mathbb{Y}}^n_k - \tilde{\mathbb{Y}}^n_{k-1})^{\otimes(\alpha,\beta)}. \end{aligned}$$

Here, $S_n = (S_n^{12}, S_n^{11}, S_n^{22})^{\star}$ is the (co)variance estimator. S_n depends on the scaling parameter a_n which is not derived from the data of $\mathbb{Y}$. Therefore, treat the correlation estimator $C_n^{12} = S_n^{12}/\sqrt{S_n^{11}S_n^{22}}$ which does not depend on the scaling parameter a_n.

Let

$$U^{\alpha\beta} = \frac{2}{3}\langle X^{\alpha,c}, X^{\beta,c}\rangle_T \;=\; \frac{2}{3}\int_0^T X_t^{\alpha 1}\cdot X_t^{\beta 1}dt \quad (\alpha,\beta=1,2)$$

and $U=(U^{12},U^{11},U^{22})^\star$, where $X^{\alpha,c}$ is the continuous part of X^α and $X_t^{\alpha 1}$ is the α-th row of $\mathbb{X}_t^1$. Let

$$\begin{aligned}\gamma^{(\alpha_1,\beta_1),(\alpha_2,\beta_2)} &= \int_0^T \sum_{i,j=1}^r \frac{X_{i,s}^{\alpha_1 1}X_{j,s}^{\beta_1 1}+X_{j,s}^{\alpha_1 1}X_{i,s}^{\beta_1 1}}{2}\frac{X_{i,s}^{\alpha_2 1}X_{j,s}^{\beta_2 1}+X_{j,s}^{\alpha_2 1}X_{i,s}^{\beta_2 1}}{2}\, ds\\ &= \int_0^T (\mathbb{X}_s^1)^{\tilde{\otimes}(\alpha_1,\beta_1)}\cdot(\mathbb{X}_s^1)^{\tilde{\otimes}(\alpha_2,\beta_2)}\; ds.\end{aligned}$$

and $\Gamma=(\gamma^{pq})_{p,q=(1,2),(1,1),(2,2)}$, where $X_{i,s}^{\alpha 1}$ is the (α,i)-element of $\mathbb{X}_s^1$.

Theorem 4.1. *(a) $S_n\to^p U$ as $n\to\infty$ under $[A]$ and $[B^\flat]$, where p denotes the convergence in probability.*

(b) Under $[A]$ and any one of $[B]$, $[B']$ and $[B^\sharp]$,

$$\Big(\frac{T}{b_n}\Big)^{-1/2}\Big(S_n-U\Big)\to^{d_s}\Gamma^{\frac{1}{2}}\zeta$$

as $n\to\infty$, where ζ is an $\mathbb{R}^3$-valued standard normal variable independent of $\mathcal{F}$ and d_s denotes the $\mathcal{F}$-stable convergence.

Let $R=U^{12}/\sqrt{U^{11}U^{22}}$. For the correlation estimator, we have

Theorem 4.2. *Suppose that $U^{11}U^{22}\neq 0$ a.s. Then*

(a) $C_n^{12}\to^p R$ as $n\to\infty$ under $[A]$ and $[B^\flat]$.

(b) Under $[A]$ and any one of $[B]$, $[B']$ and $[B^\sharp]$,

$$\Big(\frac{T}{b_n}\Big)^{-1/2}(C_n^{12}-R)\to^{d_s}\Xi^{1/2}\zeta$$

as $n\to\infty$, where ζ is an $\mathbb{R}^3$-valued standard normal variable independent of $\mathcal{F}$ and

$$\Xi^{1/2}:=\Big(\frac{1}{\sqrt{U^{11}U^{22}}},\frac{-U^{12}}{2\sqrt{(U^{11})^3U^{22}}},\frac{-U^{12}}{2\sqrt{U^{11}(U^{22})^3}}\Big)\Gamma^{1/2}.$$

$\Xi=(\Xi^{1/2})(\Xi^{1/2})^\star$ is the asymptotic variance of the correlation estimator.

4.2 Asymptotic variance estimators

Thanks to the stable convergence of C_n^{12}, once a consistent estimator of the asymptotic variance of C_n^{12} is obtained, it enables interval estimation and detection the correlation based on the hypothesis testing. We propose two types of asymptotic variance estimators based on Barndorff-Nielsen and Shephard [7], and Hayashi and Yoshida [22] respectively. We prove their consistency in the same setting for the stable convergence. Furthermore, we prove the faster convergence rate of the former one in some situation.

We need the estimator of Γ to obtain the asymptotic variance estimator. Here, we give two types of the Γ estimator.

4.2.1 Γ estimator

Now, we define

$$\begin{aligned}\hat{\gamma}_{n,1}^{pq} = \frac{9}{8}\Big\{ &\sum_{k=2}^{b_n}(\tilde{\mathbb{Y}}_k^n - \tilde{\mathbb{Y}}_{k-1}^n)^{\otimes p}(\tilde{\mathbb{Y}}_k^n - \tilde{\mathbb{Y}}_{k-1}^n)^{\otimes q} \\ &- \frac{1}{2}\sum_{k=2}^{b_n-2}\Big((\tilde{\mathbb{Y}}_k^n - \tilde{\mathbb{Y}}_{k-1}^n)^{\otimes p}(\tilde{\mathbb{Y}}_{k+2}^n - \tilde{\mathbb{Y}}_{k+1}^n)^{\otimes q} \\ &+ (\tilde{\mathbb{Y}}_{k+2}^n - \tilde{\mathbb{Y}}_{k+1}^n)^{\otimes p}(\tilde{\mathbb{Y}}_k^n - \tilde{\mathbb{Y}}_{k-1}^n)^{\otimes q}\Big)\Big\}\Big(\frac{T}{b_n}\Big)^{-1}\end{aligned}$$

based on the idea in Barndorff-Nielsen and Shephard [7], and let

$$\begin{aligned}\hat{\gamma}_{n,2}^{pq} = \frac{9}{8}\sum_{k=2}^{b_n-2}\frac{1}{2}&\Big\{(\tilde{\mathbb{Y}}_{k+2}^n - \tilde{\mathbb{Y}}_{k+1}^n)^{\otimes p} - (\tilde{\mathbb{Y}}_k^n - \tilde{\mathbb{Y}}_{k-1}^n)^{\otimes p}\Big\} \\ &\times\Big\{(\tilde{\mathbb{Y}}_{k+2}^n - \tilde{\mathbb{Y}}_{k+1}^n)^{\otimes q} - (\tilde{\mathbb{Y}}_k^n - \tilde{\mathbb{Y}}_{k-1}^n)^{\otimes q}\Big\}\Big(\frac{T}{b_n}\Big)^{-1}.\end{aligned}$$

4.2.2 Kernel based Γ estimator

In reference to Hayashi and Yoshida [22], we define the kernel based estimator.

$$\partial_h\{X^\alpha, X^\beta\}_k := \sum_{l=(k-n(h)+1)_+}^{k}(\Delta_l X^\alpha \Delta_l X^\beta)h^{-1},$$

where $h = h_n$ is a parameter satisfying $h_n \to 0$ and $b_n^{-1/2}h_n^{-1} \to 0$, and $n(h) := \max_m\{t_m \leq h\}$.

$$\begin{aligned}\hat{\gamma}_{n,h}^{pq} =& \frac{9}{8}\sum_{k=2}^{b_n}\Big\{\partial_h\{\tilde{Y}^{n,\alpha_1}, \tilde{Y}^{n,\alpha_2}\}_k\ \partial_h\{\tilde{Y}^{n,\beta_1}, \tilde{Y}^{n,\beta_2}\}_k \\ &+ \partial_h\{\tilde{Y}^{n,\alpha_1}, \tilde{Y}^{n,\beta_2}\}_k\ \partial_h\{\tilde{Y}^{n,\beta_1}, \tilde{Y}^{n,\alpha_2}\}_k\Big\}\Big(\frac{T}{b_n}\Big).\end{aligned}$$

4.2.3 Ξ estimator

Here, we define

$$\hat{\Xi}_{n,*}^{1/2} = \Big(\frac{1}{\sqrt{S_n^{11}S_n^{22}}}, \frac{-S_n^{12}}{2\sqrt{(S_n^{11})^3 S_n^{22}}}, \frac{-S_n^{12}}{2\sqrt{S_n^{11}(S_n^{22})^3}}\Big)\hat{\Gamma}_{n,*}^{1/2}$$

and $\hat{\Xi}_{n,*} = (\hat{\Xi}_{n,*}^{1/2})^{\otimes}$, for $* = 1, 2, h$.

Theorem 4.3. *(a) $\hat{\gamma}_{n,i}^{pq} = \gamma^{pq} + o_p(1)$ under [A] and any one of [B], [B′] and [B$^{\sharp\sharp}$], $(i = 1, 2)$.*

(b) $\hat{\gamma}_{n,i}^{pq} = \gamma^{pq} + O_p(b_n^{-1/2})$ under [A] and [B$^{\sharp\sharp}$], $(i = 1, 2)$.

Theorem 4.4. *$\hat{\gamma}_{n,h}^{pq} \to^p \gamma^{pq}$ under [A] and any one of [B], [B′] and [B$^{\sharp\sharp}$].*

Corollary 4.5. *(1a) $\hat{\Xi}_{n,i} = \Xi + o_p(1)$ under [A] and any one of [B], [B′] and [B$^{\sharp\sharp}$], $(i = 1, 2)$.*

(1b) $\hat{\Xi}_{n,i} = \Xi + O_p(b_n^{-1/2})$ under [A] and [B$^{\sharp\sharp}$], $(i = 1, 2)$.

(2) $\hat{\Xi}_{n,h} = \Xi + o_p(1)$ under [A] and any one of [B], [B′] and [B$^{\sharp\sharp}$].

5 Simulation

For simulation, we treat the doubly stochastic Poisson processes whose intensity processes are correlated. We compare the proposed asymptotic variance estimators, using the mean squared error and the coverage probability of confidence interval.

When we apply our estimation method to real data, b_n is the only tuning parameter. Too big b_n does not satisfy the assumptions and too small b_n gives bad convergence of the estimators. Then, in order to find a suitable b_n for estimation empirically, we plot the coverage probability of confidence interval against the sampling interval (T/b_n). We also look for an appropriate b_n for statistical hypothesis testing by the plot of the power of the test versus the sampling interval. We simulate it in various parameter settings and examine the response to the good b_ns. As a result, it is suggested that they are mainly affected by the smaller number of times of occurrence of two event series and the correlation parameter.

6 Real data analysis

We treat the order data of the stocks compose TOPIX Core30. TOPIX Core 30 is the stock market index consists of the 30 stocks that are listed with first section of the Tokyo Stock Exchange and have the highest market capitalization and liquidity. We see the plot of the correlation and its acceptance region for the order data against the sampling interval, and we determine a good sampling interval with the observation of the simulated graph in mind. Furthermore, we apply the graphical modeling to the data and observe the relationships among the stocks.

7 Conclusion

In this thesis, we regard intensity processes of point processes as latent processes and construct the correlation estimator. Then, we prove its consistency and asymptotic mixed normality and specify the conditions for the asymptotic property. We also construct the asymptotic variance estimators and prove its consistency under the same conditions for the asymptotic mixed normality of the correlation estimator.

With the simulation of doubly stochastic Poisson processes, we confirm the asymptotic behaviors of the proposed estimators and discuss the suitable sampling interval. Then, we analyze the real data (the order data of stocks) based on the simulation result.

Acknowledgment

My heartfelt appreciation goes to Prof. N. Yoshida whose enormous support and insightful comments were invaluable during the course of my study. I would also like to express my gratitude to my family for their moral support and warm encouragements. This work is partly supported by CREST, JST.

References

[1] Abergel, F., Huré, C., Pham, H.: Algorithmic trading in a microstructural limit order book model. ArXiv e-prints (2017)

[2] Abergel, F., Jedidi, A.: A mathematical approach to order book modeling. International Journal of Theoretical and Applied Finance **16**(05), 1350,025 (2013)

[3] Abergel, F., Jedidi, A.: Long-time behavior of a Hawkes process–based limit order book. SIAM Journal on Financial Mathematics **6**(1), 1026–1043 (2015)

[4] Aït-Sahalia, Y., Fan, J., Xiu, D.: High-frequency covariance estimates with noisy and asynchronous financial data. Journal of the American Statistical Association **105**(492), 1504–1517 (2010)

[5] Bacry, E., Delattre, S., Hoffmann, M., Muzy, J.F.: Modelling microstructure noise with mutually exciting point processes. Quantitative Finance **13**(1), 65–77 (2013)

[6] Barndorff-Nielsen, O.E., Hansen, P.R., Lunde, A., Shephard, N.: Multivariate realised kernels: consistent positive semi-definite estimators of the covariation of equity prices with noise and non-synchronous trading. Journal of Econometrics **162**(2), 149–169 (2011)

[7] Barndorff-Nielsen, O.E., Shephard, N.: Econometric analysis of realized covariation: High frequency based covariance, regression, and correlation in financial economics. Econometrica **72**(3), 885–925 (2004). URL http://www.jstor.org/stable/3598838

[8] Bibinger, M.: Efficient covariance estimation for asynchronous noisy high-frequency data. Scandinavian Journal of Statistics **38**(1), 23–45 (2011)

[9] Bibinger, M.: An estimator for the quadratic covariation of asynchronously observed Itô processes with noise: asymptotic distribution theory. Stochastic Processes and their Applications **122**(6), 2411–2453 (2012)

[10] Bowsher, C.G.: Modelling security market events in continuous time: Intensity based, multivariate point process models. Journal of Econometrics **141**(2), 876–912 (2007)

[11] Chertok, A.V., Korolev, V.Y., Korchagin, A.Y.: Modeling high-frequency non-homogeneous order flows by compound Cox processes*. Journal of Mathematical Sciences **214**(1), 44–68 (2016). DOI 10.1007/s10958-016-2757-6. URL https://doi.org/10.1007/s10958-016-2757-6

[12] Christensen, K., Kinnebrock, S., Podolskij, M.: Pre-averaging estimators of the ex-post covariance matrix in noisy diffusion models with non-synchronous data. Journal of Econometrics **159**(1), 116–133 (2010)

[13] Christensen, K., Podolskij, M., Vetter, M.: On covariation estimation for multivariate continuous Itô semimartingales with noise in non-synchronous observation schemes. Journal of Multivariate Analysis **120**, 59–84 (2013)

[14] Clinet, S., Yoshida, N.: Statistical inference for ergodic point processes and application to limit order book. Stochastic Processes and their Applications **127**(6), 1800 – 1839 (2017). DOI http://dx.doi.org/10.1016/j.spa.2016.09.014. URL http://www.sciencedirect.com/science/article/pii/S0304414916301740

[15] Cont, R., Stoikov, S., Talreja, R.: A stochastic model for order book dynamics. Operations research **58**(3), 549–563 (2010)

[16] Epps, T.W.: Comovements in stock prices in the very short run. Journal of the American Statistical Association **74**(366a), 291–298 (1979)

[17] Glosten, L.R., Milgrom, P.R.: Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. Journal of Financial Economics **14**(1), 71 – 100 (1985). DOI http://dx.doi.org/10.1016/0304-405X(85)90044-3. URL http://www.sciencedirect.com/science/article/pii/0304405X85900443

[18] Gottlieb, G., Kalay, A.: Implications of the discreteness of observed stock prices. The Journal of Finance **40**(1), 135–153 (1985). URL http://www.jstor.org/stable/2328052

[19] Griffin, J.E., Oomen, R.C.: Covariance measurement in the presence of non-synchronous trading and market microstructure noise. Journal of Econometrics **160**(1), 58–68 (2011)

[20] Guilbaud, F., Pham, H.: Optimal High Frequency Trading with limit and market orders. ArXiv e-prints (2011)

[21] Hayashi, T., Yoshida, N.: Asymptotic normality of a covariance estimator for nonsynchronously observed diffusion processes. Annals of the Institute of Statistical Mathematics **60**(2), 367–406 (2008)

[22] Hayashi, T., Yoshida, N.: Nonsynchronous covariation process and limit theorems. Stochastic Processes and their Applications **121**(10), 2416–2454 (2011)

[23] Hayashi, T., Yoshida, N., et al.: On covariance estimation of non-synchronously observed diffusion processes. Bernoulli **11**(2), 359–379 (2005)

[24] Hewlett, P.: Clustering of order arrivals, price impact and trade path optimisation. In: Workshop on Financial Modeling with Jump processes, Ecole Polytechnique, pp. 6–8 (2006)

[25] Jacod, J., Li, Y., Mykland, P.A., Podolskij, M., Vetter, M.: Microstructure noise in the continuous case: the pre-averaging approach. Stochastic processes and their applications **119**(7), 2249–2276 (2009)

[26] Kimura, A.: Confidence interval for correlation estimator between latent processes. ArXiv e-prints (2017)

[27] Kimura, A., Yoshida, N.: Estimation of Correlation Between Latent Processes, pp. 131–146. Springer International Publishing, Cham (2016). DOI 10.1007/978-3-319-45875-5_6. URL https://doi.org/10.1007/978-3-319-45875-5_6

[28] Koike, Y.: Estimation of integrated covariances in the simultaneous presence of nonsynchronicity, microstructure noise and jumps. Econometric Theory pp. 1–79 (2013)

[29] Koike, Y.: An estimator for the cumulative co-volatility of asynchronously observed semimartingales with jumps. Scandinavian Journal of Statistics **41**(2), 460–481 (2014)

[30] Koike, Y.: Limit theorems for the pre-averaged Hayashi–Yoshida estimator with random sampling. Stochastic Processes and their Applications **124**(8), 2699–2753 (2014)

[31] Korolev, V.Y., Chertok, A.V., Korchagin, A.Y., Zeifman, A.I.: Modeling high-frequency order flow imbalance by functional limit theorems for two-sided risk processes. ArXiv e-prints (2014)

[32] Large, J.: Measuring the resiliency of an electronic limit order book. Journal of Financial Markets **10**(1), 1–25 (2007)

[33] Malliavin, P., Mancino, M.E.: Fourier series method for measurement of multivariate volatilities. Finance and Stochastics **6**(1), 49–61 (2002)

[34] Malliavin, P., Mancino, M.E., Recchioni, M.C.: A non-parametric calibration of the HJM geometry: an application of Itô calculus to financial statistics. Japanese Journal of Mathematics **2**(1), 55–77 (2007). DOI 10.1007/s11537-007-0666-7. URL https://doi.org/10.1007/s11537-007-0666-7

[35] Malliavin, P., Mancino, M.E., et al.: A Fourier transform method for nonparametric estimation of multivariate volatility. The Annals of Statistics **37**(4), 1983–2010 (2009)

[36] Mancino, M.E., Sanfelici, S.: Estimating covariance via Fourier method in the presence of asynchronous trading and microstructure noise. Journal of financial econometrics **9**(2), 367–408 (2011)

[37] Muni Toke, I.: “Market making” behaviour in an order book model and its impact on the bid-ask spread. ArXiv e-prints (2010)

[38] Muni Toke, I.: Stationary distribution of the volume at the best quote in a Poisson order book model. ArXiv e-prints (2015)

[39] Muni Toke, I., Pomponio, F.: Modelling trades-through in a limited order book using Hawkes processes. Economics discussion paper (2011-32) (2011)

[40] Muni Toke, I., Yoshida, N.: Modelling intensities of order flows in a limit order book. ArXiv e-prints (2016)

[41] Mykland, P.A.: A Gaussian calculus for inference from high frequency data. Annals of finance **8**(2-3), 235–258 (2012)

[42] Ogihara, T., Yoshida, N.: Quasi likelihood analysis for point process regression models. preprint (2015)

[43] Park, S., Linton, O.B., et al.: Estimating the quadratic covariation matrix for an asynchronously observed continuous time signal masked by additive noise (2012)

[44] Podolskij, M., Vetter, M., Sommer, M.: Estimation of volatility functionals in the simultaneous presence of microstructure noise and jumps. Available at SSRN 950344 (2007)

[45] Roll, R.: A simple implicit measure of the effective bid-ask spread in an efficient market. The Journal of Finance **39**(4), 1127–1139 (1984). URL http://www.jstor.org/stable/2327617

[46] Smith, E., Farmer, J.D., Gillemot, L.s., Krishnamurthy, S., et al.: Statistical theory of the continuous double auction. Quantitative finance **3**(6), 481–514 (2003)

[47] Toke, I.M.: The order book as a queueing system: average depth and influence of the size of limit orders. Quantitative Finance **15**(5), 795–808 (2015). DOI 10.1080/14697688.2014.963654. URL http://dx.doi.org/10.1080/14697688.2014.963654

[48] Voev, V., Lunde, A.: Integrated covariance estimation using high-frequency data in the presence of noise. Journal of Financial Econometrics **5**(1), 68–104 (2007)

[49] Zhang, L., Mykland, P.A., Aït-Sahalia, Y.: A tale of two time scales. Journal of the American Statistical Association (2012)

[50] Zhang, L., et al.: Efficient estimation of stochastic volatility using noisy observations: A multi-scale approach. Bernoulli **12**(6), 1019–1043 (2006)

[51] Zhou, B.: High-frequency data and volatility in foreign-exchange rates. Journal of Business & Economic Statistics **14**(1), 45–52 (1996)