

# SECOND-ORDER ASYMPTOTIC EXPANSION FOR THE COVARIANCE ESTIMATOR OF TWO ASYNCHRONOUSLY OBSERVED DIFFUSION PROCESSES

ARNAK DALALYAN\* AND NAKAHIRO YOSHIDA\*\*

ABSTRACT. In this paper, we study the asymptotic properties of the Hayashi-Yoshida estimator, hereafter HY-estimator, of two diffusion processes when observations are subject to non-synchronicity. Our setup includes random sampling schemes, provided that the observation times are independent of the underlying diffusions. We first derive second-order asymptotic expansions for the distribution of the Hayashi-Yoshida estimator in the case when observed diffusions have no drift. We then focus on the drifted case and carry out a stochastic decomposition of the HY-estimator itself. This decomposition, in conjunction with the evaluation of the Malliavin covariance, leads to a second-order asymptotic expansion of the distribution of the HY-estimator. This result lies in continuity of the consistency and the asymptotic normality results proved by Hayashi and Yoshida [12, 13]. We compute the constants involved in the obtained expansions for the particular case where the sampling scheme is generated by two independent Poisson processes.

## 1. INTRODUCTION

Let  $\mathbf{X} = (X_1, X_2)$  be a two dimensional diffusion process given by

$$d\mathbf{X}_t = \boldsymbol{\beta}_t dt + \text{diag}(\boldsymbol{\sigma}_t) d\mathbf{B}_t \quad (1)$$

where  $\mathbf{B} = ((B_{1,t}, B_{2,t})^\mathbf{T}, t \geq 0)$  is a two dimensional Gaussian process with independent increments, zero mean and covariance matrix

$$\mathbf{E}[\mathbf{B}_t \cdot \mathbf{B}_t^\mathbf{T}] = \begin{pmatrix} t & \int_0^t \rho_s ds \\ \int_0^t \rho_s ds & t \end{pmatrix}, \quad \forall t \geq 0.$$

In (1),  $\boldsymbol{\beta} = (\beta_1, \beta_2)^\mathbf{T}$  and  $\boldsymbol{\sigma} = (\sigma_1, \sigma_2)^\mathbf{T}$  are some progressively measurable processes,  $\text{diag}(\boldsymbol{\sigma})$  stands for the diagonal matrix having  $\sigma_i$  as  $i^{\text{th}}$  diagonal entry,  $i = 1, 2$ . In what follows, we restrict our attention to the case when  $\sigma_1, \sigma_2$  and  $\rho$  are deterministic functions; the functions  $\sigma_i, i = 1, 2$  take positive values while  $\rho$  takes values in the interval  $[-1, 1]$ . Note that the marginal processes  $B_1$  and  $B_2$  are Brownian motions (BM). Moreover, we can define a process  $B_t^*$  such that  $(B_{1,t}, B_t^*)_{t \geq 0}$  is a two-dimensional BM and  $dB_{2,t} = \rho_t dB_{1,t} + \sqrt{1 - \rho_t^2} dB_t^*$  for every  $t \geq 0$ .

We will assume that the processes  $X_1$  and  $X_2$  are observed respectively at time instants  $0 = S^0 < S^1 < \dots < S^{N_1} = T$  and  $0 = T^0 < \dots < T^{N_2} = T$ . Let us denote  $I^i = (S^{i-1}, S^i]$  and  $J^j = (T^{j-1}, T^j]$ . The families  $\Pi^1 = \{I^i, i = 1, \dots, N_1\}$  and  $\Pi^2 = \{J^j, j = 1, \dots, N_2\}$  are partitions of the interval  $[0, T]$ . We will also use the notation  $\Delta_i X_1 = X_{1,S^i} - X_{1,S^{i-1}}$  and  $\Delta_j X_2 = X_{2,T^j} - X_{2,T^{j-1}}$ .

---

*Date:* April 4, 2008.

\*: University Pierre et Marie Curie Paris 6 and \*\*: University of Tokyo.

In this paper, we are concerned with the problem of estimating the parameter

$$\theta = \int_0^T \rho_t \sigma_{1,t} \sigma_{2,t} dt = \langle X_1, X_2 \rangle_T$$

based on the observations  $(X_{1,S^i}, X_{2,T^j}, i = 0, \dots, N_1, j = 0, \dots, N_2)$ . The parameter  $\theta$  represents the covariance between martingale parts of  $X_1$  and  $X_2$ . Therefore, it can be used to evaluate the correlation between the two BMs  $B_1$  and  $B_2$ .

If the two sequences of data are synchronously observed, the sum of cross products  $\sum_{i=1}^{N_1} \Delta_i X_1 \cdot \Delta_i X_2$  is a natural estimator of  $\theta$  because it converges in probability to  $\theta$  when the maximum lag of the time points tends to 0 in probability, as it is well known in the stochastic analysis. In the field of statistical inference for stochastic processes, this fact has been applied to estimating volatility and covariation between semimartingales. The asymptotic distributions are well investigated. There is a long list of studies, e.g., Dacunha-Castelle and Florens-Zmirou [7], Florens-Zmirou [8], Prakasa Rao [23, 24], Yoshida [30], Genon-Catalot and Jacod [9], Kessler [18], and Mykland and Zhang [21].

Recently, there appears extensive literature on the covariance estimation in time series analysis and econometrics as well; see e.g. Andersen and Bollerslev [1], Comte and Renault [6], Andersen *et al.* [2, 3], Barndorff-Nielsen and Shephard [4]. Contrarily, estimation under a “non-synchronous” sampling scheme has rarely been treated theoretically. One approach for tackling the non-synchronicity consists in generating by an interpolation method a set of equally spaced data from the original nonsynchronous data and in using the realized covariance estimator for the interpolated data. However, it is known that such a synchronization technique causes estimation bias, which is often referred to as the *Epps effect*. See Hayashi and Yoshida [12] for details. Besides, the synchronization method involves a tuning parameter the choice of which is a delicate issue. Another estimator of covariance based on harmonic analysis has been proposed by Malliavin and Mancino [19]. Their estimator as well involves an implicit interpolation and a tuning parameter corresponding to the number of Fourier coefficients used for approximating the integrated volatility.

An estimator of  $\theta$ , which is unbiased when the drift  $\beta$  is identically zero, has been proposed by Hayashi and Yoshida. Henceforth called HY-estimator, it is defined as follows

$$\hat{\theta} = \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} \Delta_i X_1 \cdot \Delta_j X_2 \cdot \mathbf{1}(I^i \cap J^j \neq \emptyset). \quad (2)$$

Under mild assumptions,  $\hat{\theta}$  is consistent as the maximum lag of the observation times tends to 0 in probability; see Hayashi and Yoshida [12]. Kusuoka and Hayashi [11] proved consistency in a more general sampling scheme. Asymptotic normality of the estimator was proved in Hayashi and Yoshida [13] under the sampling independent of the stochastic processes. For related literature, see e.g. Hoshikawa, Kanatani, Nagai and Nishiyama [15], Griffin and Oomen [10], and Voev and Lunde [29]. It is reported as an empirical fact that the kernel choosing the overlapping intervals as (2) gives accurate estimates among quadratic type estimators with various kernels. We will consider this property in Section 2. In Hayashi and Yoshida [14], the authors treated a sampling scheme possibly depending on the stochastic processes and presented a stochastic analytic proof of the asymptotic normality by regarding the sampling scheme as a certain point process. An

estimator for the variance of the HY-estimator in the case when the observed process has no drift has been recently constructed by Mykland [20].

In Section 3, we will give a representation of the cumulant of  $\hat{\theta}$  as a functional of the configuration of the observation times, and give asymptotic estimates for them. Our interest is in the second-order expansion of the characteristic function of the estimator while asymptotic normality is also proved as an application of those estimates. Section 4 gives asymptotic expansion of the distribution of the HY-estimator. This applies to the Poisson sampling scheme in Section 5.

When there are (possibly random) drift terms in the stochastic differential equation of  $\mathbf{X}_t$ , certain extra terms appear in the asymptotic expansion. In order to identify these terms, we derive in Section 6 a stochastic decomposition of the estimator and investigate the asymptotic behavior of the variables appearing in the second-order. To this end, the limit theory for semimartingales is necessary because the asymptotics is non-Gaussian differently from the usual theory of the Edgeworth expansion. The result providing the asymptotic expansion of the distribution of HY-estimator is carried out using a perturbation method presented in Section 7. We also apply the Malliavin calculus to ensure that the regularity of the distribution of the principal part, a quadratic form of Gaussian random variables, remains under the perturbation. It should be noted that the higher-order term can in general destroy the regularity of the lower-order term; see an example in [31]. Finally the perturbation method merges the obtained convergences with the asymptotic expansion for the principal part in Section 7 to give asymptotic expansion, which is a main result.

## 2. ELEMENTARY PROPERTIES OF $\hat{\theta}$

As noticed by Mykland [20], the estimator  $\hat{\theta}$  is the Maximum Likelihood Estimator (MLE) of  $\theta$ . Let us present here some computations that not only show that  $\hat{\theta}$  is the MLE of  $\theta$ , but also give some interesting insight concerning the efficiency properties of the HY-estimator  $\hat{\theta}$ .

Let us deal with a slightly more general setup. Assume that  $\boldsymbol{\xi} \in \mathbb{R}^N$  is a random vector having centered Gaussian distribution with unknown covariance matrix  $\Sigma$ . The entries of the matrix  $\Sigma$  are  $\sigma_{\ell,\ell'} = E[\xi_\ell \xi_{\ell'}]$  for  $\ell, \ell' = 1, \dots, N$ . We want to estimate a linear combination

$$\theta = \sum_{\ell,\ell'=1}^N a_{\ell,\ell'} \sigma_{\ell,\ell'},$$

where  $a_{\ell,\ell'} \in \mathbb{R}$ ,  $\ell, \ell' = 1, \dots, N$  are some known numbers verifying  $a_{\ell,\ell'} = a_{\ell',\ell}$ .

In order that this model lies in the setting of distributions belonging to the exponential family, it is convenient to consider the parametrization by the entries of the inverse of the covariance matrix  $V = \Sigma^{-1}$ . Set  $p = (N^2 + N)/2$  and write

$$V = \begin{pmatrix} v_1 & v_2 & \dots & v_N \\ v_2 & v_{N+1} & \dots & v_{2N-1} \\ \vdots & \vdots & \ddots & \vdots \\ v_N & v_{2N-1} & \dots & v_p \end{pmatrix}.$$

The log-likelihood function can now be written as follows:

$$\ell(V) = \frac{1}{2} \log |V| - \frac{1}{2} \sum_{k=1}^p v_k \mathsf{T}_k(\boldsymbol{\xi}), \quad (3)$$

where  $|V|$  denotes the determinant of the matrix  $V$  and  $\mathsf{T}(\boldsymbol{\xi}) = (\mathsf{T}_1(\boldsymbol{\xi}), \mathsf{T}_2(\boldsymbol{\xi}), \dots)$  is defined by

$$\mathsf{T}_1(\boldsymbol{\xi}) = \xi_1^2, \quad \mathsf{T}_2(\boldsymbol{\xi}) = 2\xi_1\xi_2, \quad \mathsf{T}_3(\boldsymbol{\xi}) = 2\xi_1\xi_3, \quad \dots, \quad \mathsf{T}_p(\boldsymbol{\xi}) = \xi_N^2.$$

It follows from (3) that the distribution  $P_V$  of the Gaussian vector  $\boldsymbol{\xi} \sim \mathcal{N}_N(0, V^{-1})$  belongs to the (simple) exponential family. This implies that the statistic  $\mathsf{T}(\boldsymbol{\xi})$  is the MLE of the parameter

$$\tau = E[\mathsf{T}(\boldsymbol{\xi})] = (\sigma_{11}, 2\sigma_{12}, \dots, \sigma_{NN})^{\mathbf{T}}.$$

Hence, the MLE of  $\theta = \sum_{\ell, \ell'} a_{\ell, \ell'} \sigma_{\ell, \ell'}$  is  $\hat{\theta} = \sum_{\ell, \ell'} a_{\ell, \ell'} \xi_{\ell} \xi_{\ell'}$ . It is easily seen that this estimator is unbiased. Since  $\mathsf{T}(\boldsymbol{\xi})$  is a complete sufficient statistic, the MLE  $\hat{\theta} = \sum_{\ell, \ell'} a_{\ell, \ell'} \xi_{\ell} \xi_{\ell'}$  is the best unbiased estimator of  $\theta$  in the sense that any other unbiased estimator will have a variance at least as large as that of  $\hat{\theta}$ .

We can now return to our model. The vector

$$\boldsymbol{\xi} = (\Delta_1 X_1, \dots, \Delta_{N_1} X_1, \Delta_1 X_2, \dots, \Delta_{N_2} X_2)^{\mathbf{T}}$$

follows an  $N = N_1 + N_2$  dimensional centered Gaussian distribution. The parameter  $\theta = \text{Cov}(X_{1,T}, X_{2,T})$  can be represented in the form  $\sum_{\ell, \ell'} a_{\ell, \ell'} \sigma_{\ell, \ell'}$  with

$$a_{\ell, \ell'} = \frac{1}{2} \mathbf{1}(\ell \leq N_1, \ell' > N_1, I^{\ell} \cap J^{\ell' - N_1} \neq \emptyset)$$

for every  $\ell \leq \ell'$  and  $a_{\ell, \ell'} = a_{\ell', \ell}$  for  $\ell > \ell'$ . Therefore, the arguments presented above yield the following result.

**Proposition 1.** *The estimator  $\hat{\theta}$  defined by (2) is the MLE of  $\theta$ . Moreover, it is the estimator having the smallest quadratic risk among all unbiased estimators of  $\theta$ .*

### 3. GAUSSIAN ANALYSIS AND EXPANSION OF THE CHARACTERISTIC FUNCTION

The goal of this section is to derive an asymptotic expansion for the distribution of the estimator  $\hat{\theta}$ . It will be useful for our purposes to consider the more general setup defined via Gaussian vector  $\boldsymbol{\xi}$  and the matrix  $A = (a_{\ell, \ell'})_{\ell, \ell'=1}^N$ , see Section 2.

**3.1. General Gaussian setup.** In order to determine the asymptotic expansion of the distribution of  $\hat{\theta}$ , we start with expanding its characteristic function. Recall that

$$\hat{\theta} = \boldsymbol{\xi}^{\mathbf{T}} A \boldsymbol{\xi} \quad \text{and} \quad \boldsymbol{\xi} \sim \mathcal{N}_N(0, \Sigma).$$

Since  $A$  is a symmetric matrix, the  $N$ -by- $N$  matrix  $\Sigma^{1/2} A \Sigma^{1/2}$  is symmetric and therefore diagonalizable. Let  $\Lambda$  and  $U$  be respectively the  $N$ -by- $N$  diagonal and orthogonal matrices such that  $\Sigma^{1/2} A \Sigma^{1/2} = U^{\mathbf{T}} \Lambda U$ . Let  $\boldsymbol{\zeta}$  be a Gaussian  $\mathcal{N}_N(0, I_N)$  vector such that  $\boldsymbol{\xi} = \Sigma^{1/2} \cdot U^{\mathbf{T}} \boldsymbol{\zeta}$ . Such a vector exists always and it is unique if  $\Sigma$  is invertible. In this notation, we have

$$\hat{\theta} = \boldsymbol{\zeta}^{\mathbf{T}} \Lambda \boldsymbol{\zeta} = \sum_{\ell=1}^N \lambda_{\ell} \zeta_{\ell}^2,$$

where  $\lambda_1, \dots, \lambda_N$  are the eigenvalues of the matrix  $\Sigma^{1/2} A \Sigma^{1/2}$  and  $\zeta_1, \dots, \zeta_N$  are independent Gaussian random variables. This implies that  $\zeta_\ell^2$ s are independent and distributed according to the  $\chi_1^2$  distribution. Hence  $\mathbf{E}[e^{iu\zeta_\ell^2}] = (1 - 2iu)^{-1/2}$  and

$$\varphi_{\hat{\theta}}(u) := \mathbf{E}[e^{iu\hat{\theta}}] = \prod_{\ell=1}^N (1 - 2i\lambda_\ell u)^{-1/2}.$$

By taking the logarithm and using its Taylor series we get

$$\log \varphi_{\hat{\theta}}(u) = -\frac{1}{2} \sum_{\ell=1}^N \log(1 - 2i\lambda_\ell u) = \frac{1}{2} \sum_{\ell=1}^N \sum_{k=1}^{\infty} \frac{(2i\lambda_\ell u)^k}{k},$$

as soon as  $|u| < 1/(2 \max_\ell |\lambda_\ell|)$ . Since all the series in the above formula are absolutely convergent, we can change the order of summation. This yields

$$\log \varphi_{\hat{\theta}}(u) = \sum_{k=1}^{\infty} \frac{(2iu)^k}{2k} \mu_k, \quad |u| < 1/(2\|\boldsymbol{\lambda}\|_\infty), \quad (4)$$

with  $\|\boldsymbol{\lambda}\|_\infty = \max_\ell |\lambda_\ell|$  and

$$\mu_k = \sum_{\ell=1}^N \lambda_\ell^k = \text{Tr}[(\Sigma^{1/2} A \Sigma^{1/2})^k] = \text{Tr}[(\Sigma \cdot A)^k],$$

where the last equality follows from the property  $\text{Tr}(M_1 \cdot M_2) = \text{Tr}(M_2 \cdot M_1)$  provided that both products are well defined. Separating the first two terms in the RHS of (4), we arrive at

$$\log \varphi_{\hat{\theta}}(u) = i\theta u - u^2 \mu_2 + \sum_{k=3}^{\infty} \frac{(2iu)^k}{2k} \mu_k, \quad |u| < 1/(2\|\boldsymbol{\lambda}\|_\infty). \quad (5)$$

Let us define  $\bar{\alpha} = \|\boldsymbol{\lambda}\|_\infty / \|\boldsymbol{\lambda}\|_2$ . Using simple inequalities, one checks that  $|\mu_k| \leq \bar{\alpha}^{k-2} \mu_2^{k/2}$  for every  $k \geq 3$ . Therefore,

$$\left| \sum_{k=3}^{\infty} \frac{(2iu)^k \mu_k}{2k} \right| \leq 2\mu_2 |u|^2 \sum_{k \geq 0} \frac{(2|u| \bar{\alpha} \sqrt{\mu_2})^{k+1}}{k+1} = -2\mu_2 |u|^2 \log(1 - 2|u| \bar{\alpha} \sqrt{\mu_2}),$$

for every  $u$  satisfying  $|u| < (2\bar{\alpha}\sqrt{\mu_2})^{-1}$ . This leads to the inequality

$$|\log \varphi_{\hat{\theta}-\theta}(v/\sqrt{2\mu_2}) + \frac{v^2}{2}| \leq -v^2 \log(1 - \sqrt{2}|v|\bar{\alpha}), \quad (6)$$

for every  $|v| < (\sqrt{2}\bar{\alpha})^{-1}$ .

As the first application of our approach, we obtain a central limit theorem for  $\hat{\theta}_n$ .

**Proposition 2.** *Suppose that the matrices  $A = A_n$  and  $\Sigma = \Sigma_n$  as well as the number  $N = N_n$  depend on  $n \in \mathbb{N}$ . If  $\lambda_{1,n}, \dots, \lambda_{N,n}$ , the eigenvalues of  $\Sigma_n^{1/2} A_n \Sigma_n^{1/2}$ , satisfy  $\lim_{n \rightarrow \infty} \|\boldsymbol{\lambda}_n\|_\infty^2 / \mu_{2,n} = 0$ , then*

$$\frac{\hat{\theta}_n - \theta_n}{\sqrt{2\mu_{2,n}}} \xrightarrow[n \rightarrow \infty]{D} \mathcal{N}(0, 1),$$

where  $\hat{\theta}_n = \boldsymbol{\xi}^T A_n \boldsymbol{\xi}$ ,  $\theta_n = \mathbf{E}[\hat{\theta}_n] = \text{Tr}[\Sigma_n A_n]$ ,  $\mu_{2,n} = \text{Tr}[(\Sigma_n A_n)^2]$  and  $\xrightarrow{D}$  stands for the convergence in distribution.

*Proof.* Set  $\mu_{k,n} = \text{Tr}[(\Sigma_n A_n)^k] = \sum_{\ell} \lambda_{\ell,n}^k$  and  $\eta_n = (\hat{\theta}_n - \theta_n)/\sqrt{2\mu_{2,n}}$ . The inequality (6) and the condition  $\lim_{n \rightarrow \infty} \|\boldsymbol{\lambda}_n\|_{\infty}^2/\mu_{2,n} = 0$  imply that the characteristic function of  $\eta_n$  converges pointwise to the characteristic function of a standard Gaussian distribution. This completes the proof of the proposition.  $\square$

This result states that the distribution of the estimator  $\hat{\theta}_n$  is well approximated by a Gaussian distribution. In order to give a more precise sense to this approximation and to obtain more accurate approximations, we focus our attention on a second-order asymptotic expansion of the distribution of  $\hat{\theta}_n$ . To this end, we prove first that the tails of this distribution are sufficiently small.

**Lemma 1.** *If for some  $p \in \mathbb{N}$  the inequality  $\|\boldsymbol{\lambda}\|_{\infty}^2 \leq \mu_2/(2p)$  holds, then for every  $j \in \mathbb{N}$*

$$\left| \frac{d^j}{du^j} \mathbf{E}[e^{iu(\hat{\theta}-\theta)}] \right| \leq j!(2N\|\boldsymbol{\lambda}\|_{\infty} + |\theta|)^j (p/2)^{p/4} (1 + \mu_2 u^2)^{-p/4}, \quad \forall u \in \mathbb{R}.$$

*Proof.* Using the fact that  $\zeta_{\ell}^2$  is distributed according to the  $\chi_1^2$  distribution, one easily checks that

$$|\varphi_{\hat{\theta}}(u)| = \left| \prod_{\ell=1}^N \frac{1}{(1 - 2iu\lambda_{\ell})^{1/2}} \right| \leq \prod_{\ell=1}^N (1 + 4u^2\lambda_{\ell}^2)^{-1/4}.$$

In view of the assumptions of the lemma, for every  $i = 1, \dots, p$ , there exists an integer  $\ell_i$  verifying  $\mu_2^{-1} \sum_{\ell=1}^{\ell_i} \lambda_{\ell}^2 < i/p$  and  $\mu_2^{-1} \sum_{\ell=1}^{\ell_i+1} \lambda_{\ell}^2 \geq i/p$ . For this sequence  $\ell_i$ , we get  $\mu_2^{-1} \sum_{\ell=\ell_i+1}^{\ell_i+1} \lambda_{\ell}^2 \geq (i+1)/p - 1/(2p) - i/p = 1/(2p)$  and therefore

$$\prod_{\ell=1}^N (1 + 4u^2\lambda_{\ell}^2)^{-1/4} \leq \prod_{i=1}^p \left( 1 + 4u^2 \sum_{\ell=\ell_i+1}^{\ell_i+1} \lambda_{\ell}^2 \right)^{-1/4} \leq (p/2)^{p/4} (1 + \mu_2 u^2)^{-p/4}. \quad (7)$$

This gives the desired estimate in the case where  $j = 0$ .

For  $j > 0$ , the explicit form of  $\varphi_{\hat{\theta}}$  allows one to check that

$$\varphi_{\hat{\theta}}^{(j)}(u) = \sum_{j_1 + \dots + j_N = j} \frac{j!}{j_1! \dots j_N!} \prod_{\ell=1}^N \frac{d^{j_{\ell}}}{du^{j_{\ell}}} (1 - 2iu\lambda_{\ell})^{-1/2}.$$

Simple computations yield

$$\begin{aligned} \left| \frac{d^{j_{\ell}}}{du^{j_{\ell}}} (1 - 2iu\lambda_{\ell})^{-1/2} \right| &\leq \left| \frac{j_{\ell}! (2i\lambda_{\ell})^{j_{\ell}}}{(1 - 2iu\lambda_{\ell})^{j_{\ell}+1/2}} \right| = \frac{j_{\ell}! (2\lambda_{\ell})^{j_{\ell}}}{(1 + 4u^2\lambda_{\ell}^2)^{(2j_{\ell}+1)/4}} \\ &\leq \frac{j_{\ell}! \|\boldsymbol{\lambda}\|_{\infty}^{j_{\ell}}}{(1 + 4u^2\lambda_{\ell}^2)^{1/4}}. \end{aligned}$$

Therefore,

$$\left| \frac{d^j}{du^j} \varphi_{\hat{\theta}}(u) \right| \leq j!(2N\|\boldsymbol{\lambda}\|_{\infty})^j \prod_{\ell=1}^N (1 + 4u^2\lambda_{\ell}^2)^{-1/4}$$

and the desired inequality for  $\theta = 0$  follows from (7). For  $\theta \neq 0$ , it suffices to use the relation  $|\varphi_{\hat{\theta}-\theta}^{(j)}(u)| \leq \sum_{k=0}^j C_j^k |\theta|^k |\varphi_{\hat{\theta}}^{(j-k)}(u)|$  and the obtained estimate for  $|\varphi_{\hat{\theta}}^{(j-k)}(u)|$ .  $\square$

**Remark 1.** We will use the result of Lemma 1 in the asymptotic setup described in Proposition 2, essentially for bounding the tails of the derivatives of the characteristic function  $\varphi_{\hat{\theta}-\theta}(u)$  of  $\hat{\theta} - \theta$ , when  $u$  is in absolute value larger than  $N^{q_0}/\sqrt{\mu_2}$  for some  $q_0 > 0$ . As we see later, in the asymptotic setup, the ratio  $\|\boldsymbol{\lambda}\|_\infty^2/\mu_2$  tends to zero under mild assumptions on the sampling schemes. This will allow us to take the parameter  $p$  of Lemma 1 large enough to guarantee nice decay properties for the tails of the derivatives of  $\varphi_{\hat{\theta}-\theta}$ .

**3.2. Computation of  $\mu_k$ .** We showed in the previous subsection that the asymptotic expansion of the characteristic function of  $\hat{\theta}$  involves the traces of integer powers of the matrix  $\Sigma \cdot A$ . In our setup, both matrices  $A$  and  $\Sigma$  have special forms. In particular, they contain a great number of zeros and, therefore, the expression of  $\mu_k$  takes a simplified form.

Prior to presenting the formula for  $\mu_k$ , we need a definition. Let  $k > 0$  be an integer.

**Definition 1.** We call chain of length  $k$ , any vector  $(\mathbf{i}, \mathbf{j}) \in \{1, \dots, N_1\}^k \times \{1, \dots, N_2\}^k$  such that  $I^{i_p} \cap J^{j_p} \neq \emptyset$  and  $J^{j_p} \cap I^{i_{p+1}} \neq \emptyset$  for all  $p \in \{1, \dots, k\}$  with the convention  $i_{k+1} = i_1$ . The set of all chains of length  $k$  will be denoted by  $\mathcal{C}_k$ .

In the definition of  $\mathcal{C}_k$ ,  $i_p$  (resp.  $j_p$ ) stands for the  $p$ th coordinate of  $\mathbf{i}$  (resp.  $\mathbf{j}$ ).

**Proposition 3.** For any Borel set  $S \subset \mathbb{R}$  let us denote

$$v(S) = \int_S \rho_t \sigma_{1,t} \sigma_{2,t} dt, \quad v_1(S) = \int_S \sigma_{1,t}^2 dt, \quad v_2(S) = \int_S \sigma_{2,t}^2 dt.$$

The coefficients  $\mu_2$  and  $\mu_3$  can be computed by the formulae

$$\begin{aligned} \mu_2 &= \frac{1}{2} \sum_{(\mathbf{i}, \mathbf{j}) \in \mathcal{C}_2} \prod_{p=1}^2 v(I^{i_p} \cap J^{j_p}) + \frac{1}{2} \sum_{(\mathbf{i}, \mathbf{j}) \in \mathcal{C}_1} v_1(I^i) v_2(J^j), \\ \mu_3 &= \frac{1}{4} \sum_{(\mathbf{i}, \mathbf{j}) \in \mathcal{C}_3} \prod_{p=1}^3 v(I^{i_p} \cap J^{j_p}) + \frac{3}{4} \sum_{(\mathbf{i}, \mathbf{j}) \in \mathcal{C}_2} v_1(I^{i_1}) v_2(J^{j_1}) v(I^{i_2} \cap J^{j_2}). \end{aligned}$$

*Proof.* We give only the proof of the second formula. The proof of the first formula is analogous but simpler, therefore it is omitted. Since  $\mu_3 = \text{Tr}[(\Sigma \cdot A)^3]$ , we have

$$\mu_3 = \sum_{\ell_1, \dots, \ell_6=1}^N \sigma_{\ell_1 \ell_2} a_{\ell_2 \ell_3} \sigma_{\ell_3 \ell_4} a_{\ell_4 \ell_5} \sigma_{\ell_5 \ell_6} a_{\ell_6 \ell_1}. \quad (8)$$

In our setup, the entries of the matrix  $A$  are

$$\begin{aligned} a_{\ell, \ell'} &= \frac{1}{2} \cdot \mathbf{1}(\ell \leq N_1, \ell' > N_1, I^\ell \cap J^{\ell'-N_1} \neq \emptyset) \\ &\quad + \frac{1}{2} \cdot \mathbf{1}(\ell > N_1, \ell' \leq N_1, I^{\ell'} \cap J^{\ell-N_1} \neq \emptyset), \end{aligned} \quad (9)$$

and those of  $\Sigma$  are

$$\sigma_{\ell,\ell'} = \begin{cases} v(I^\ell \cap J^{\ell'-N_1}), & \text{if } \ell \leq N_1, \ell' > N_1, \\ v(I^{\ell'} \cap J^{\ell-N_1}), & \text{if } \ell' \leq N_1, \ell > N_1, \\ v_1(I^\ell), & \text{if } \ell = \ell' \leq N_1, \\ v_2(J^{\ell-N_1}), & \text{if } \ell = \ell' > N_1, \\ 0, & \text{otherwise.} \end{cases} \quad (10)$$

To compute the sum in the left hand side of (8), we consider different cases separately.

Case A:  $\ell_1 \leq N_1$  Our aim now is to compute

$$\mu_{3,A} = \sum_{\ell_1 \leq N_1} \sum_{\ell_2, \dots, \ell_6=1}^N \sigma_{\ell_1 \ell_2} a_{\ell_2 \ell_3} \sigma_{\ell_3 \ell_4} a_{\ell_4 \ell_5} \sigma_{\ell_5 \ell_6} a_{\ell_6 \ell_1}.$$

This can be done by considering the following four subcases:

$$\begin{array}{ll} \text{Case A.1 } \ell_1 \neq \ell_2 \text{ and } \ell_3 \neq \ell_4, & \text{Case A.2 } \ell_1 = \ell_2 \text{ and } \ell_3 = \ell_4, \\ \text{Case A.3 } \ell_1 \neq \ell_2 \text{ and } \ell_3 = \ell_4, & \text{Case A.4 } \ell_1 = \ell_2 \text{ and } \ell_3 \neq \ell_4. \end{array}$$

In the case A.1, in order that the corresponding term in (8) be nonzero, the indices  $\ell_i, i \leq 6$ , should satisfy  $\ell_1 \leq N_1, \ell_2 > N_1, \ell_3 \leq N_1, \ell_4 > N_1, \ell_5 \leq N_1$  and  $\ell_6 > N_1$ . Moreover, if we set  $\mathbf{i} = (\ell_1, \ell_3, \ell_5)$  and  $\mathbf{j} = (\ell_2, \ell_4, \ell_6)$ , then  $(\mathbf{i}, \mathbf{j})$  should belong to  $\mathcal{C}_3$ . Therefore,  $\sigma_{i_p j_p} = v(I^{i_p} \cap J^{j_p})$  for  $p = 1, 2, 3$  and

$$\sigma_{\ell_1 \ell_2} a_{\ell_2 \ell_3} \sigma_{\ell_3 \ell_4} a_{\ell_4 \ell_5} \sigma_{\ell_5 \ell_6} a_{\ell_6 \ell_1} = \frac{1}{8} \mathbf{1}((\mathbf{i}, \mathbf{j}) \in \mathcal{C}_3) \prod_{p=1}^3 v(I^{i_p} \cap J^{j_p}). \quad (11)$$

In the case A.2, in order to get nonzero term in (8), the indices  $\ell_i, i \leq 6$ , should satisfy  $\ell_1 = \ell_2 \leq N_1, \ell_3 = \ell_4 > N_1, \ell_5 \leq N_1$  and  $\ell_6 > N_1$ . Moreover, if we set  $\mathbf{i} = (\ell_1, \ell_5)$  and  $\mathbf{j} = (\ell_3, \ell_6)$ , then  $(\mathbf{i}, \mathbf{j})$  should belong to  $\mathcal{C}_2$ . Therefore,

$$\begin{aligned} \sigma_{\ell_1 \ell_2} a_{\ell_2 \ell_3} \sigma_{\ell_3 \ell_4} a_{\ell_4 \ell_5} \sigma_{\ell_5 \ell_6} a_{\ell_6 \ell_1} &= \sigma_{i_1 i_1} a_{i_1 j_1} \sigma_{j_1 j_1} a_{j_1 i_2} \sigma_{i_2 j_2} a_{j_2 i_1} \\ &= \frac{1}{8} \mathbf{1}((\mathbf{i}, \mathbf{j}) \in \mathcal{C}_2) v_1(I^{i_1}) v_2(J^{j_1}) v(I^{i_2} \cap J^{j_2}). \end{aligned} \quad (12)$$

In the cases A.3 and A.4, it is easily seen that the corresponding summand in the right hand side of (8) is different from 0 only if  $\ell_5 = \ell_6$ . Using the symmetry of  $a_{\ell\ell'}$ s and  $\sigma_{\ell\ell'}$ s, we infer that the results in these cases are equal and equal to the result of the case A.2.

Case B:  $\ell_1 > N_1$  We want to evaluate the term

$$\mu_{3,B} = \sum_{\ell_1 > N_1} \sum_{\ell_2, \dots, \ell_6=1}^N \sigma_{\ell_1 \ell_2} a_{\ell_2 \ell_3} \sigma_{\ell_3 \ell_4} a_{\ell_4 \ell_5} \sigma_{\ell_5 \ell_6} a_{\ell_6 \ell_1}.$$

In view of the symmetry of matrices  $A$  and  $\Sigma$ , we can rewrite  $\mu_{3,B}$  in the form

$$\mu_{3,B} = \sum_{\ell_1 > N_1} \sum_{\ell_2, \dots, \ell_6=1}^N \sigma_{\ell_6 \ell_5} a_{\ell_5 \ell_4} \sigma_{\ell_4 \ell_3} a_{\ell_3 \ell_2} \sigma_{\ell_2 \ell_1} a_{\ell_1 \ell_6}.$$



Since  $a_{\ell_1 \ell_6} \neq 0$  and  $\ell_1 > N_1$  entails  $\ell_6 \leq N_1$ , and  $a_{\ell_1 \ell_6} \neq 0$  and  $\ell_6 \leq N_1$  entails  $\ell_1 > N_1$ , we get

$$\mu_{3,B} = \sum_{\ell_6 \leq N_1} \sum_{\ell_1, \ell_2, \dots, \ell_5=1}^N \sigma_{\ell_6 \ell_5} a_{\ell_5 \ell_4} \sigma_{\ell_4 \ell_3} a_{\ell_3 \ell_2} \sigma_{\ell_2 \ell_1} a_{\ell_1 \ell_6}.$$

By reordering the indices we get  $\mu_{3,B} = \mu_{3,A}$  and the assertion of the proposition follows.  $\square$

**Corollary 1.** *The terms  $\mu_2$  and  $\mu_3$  may alternatively be written in the form*

$$2\mu_2 = \sum_{I,J} v_1(I)v_2(J)K_{IJ} + \sum_{I \in \Pi^1} v(I)^2 + \sum_{J \in \Pi^2} v(J)^2 - \sum_{I,J} v(I \cap J)^2, \quad (13)$$

$$\begin{aligned} 4\mu_3 &= \sum_{I \in \Pi^1} v(I)^3 + \sum_{J \in \Pi^2} v(J)^3 + 2 \sum_{I,J} v(I \cap J)^3 \\ &\quad - 3 \sum_{I,J} [v(I \cap J)^2(v(I) + v(J)) - v(I \cap J)v(I)v(J)] \\ &\quad + 3 \sum_{I,J} v_1(I)v_2(J)v(I \cup J)K_{IJ}, \end{aligned} \quad (14)$$

where  $K_{IJ} = \mathbf{1}(I \cap J \neq \emptyset)$  and  $\sum_{I,J} = \sum_{I \in \Pi^1} \sum_{J \in \Pi^2}$ .

*Proof.* Let us prove the second equality. Let us denote by  $T_1$  and  $T_2$  respectively the first and the second sums in the expression of  $\mu_3$  given in Proposition 3. In this notation,  $4\mu_3 = T_1 + 3T_2$ .

On the one hand,  $(i, j) \in \mathcal{C}_2$  implies that both  $I^{i_1}$  and  $I^{i_2}$  have non-empty intersections with each of  $J^{j_1}$  and  $J^{j_2}$ . This obviously implies that  $i_1 = i_2$  or  $j_1 = j_2$ . Therefore,

$$\begin{aligned} T_2 &= \sum_{(i,j) \in \mathcal{C}_2} v_1(I^{i_1})v_2(J^{j_1})v(I^{i_2} \cap J^{j_2}) \\ &= \sum_{I,J,J'} v_1(I)v_2(J)v(I \cap J')K_{IJ} + \sum_{I,I',J} v_1(I)v_2(J)v(I' \cap J)K_{IJ} \\ &\quad - \sum_{I,J} v_1(I)v_2(J)v(I \cap J), \end{aligned}$$

the last term resulting from the fact that the terms with  $i_1 = i_2$  and  $j_1 = j_2$  are present both in the first and in the second sums of the right hand side. Since the set of intervals  $\Pi^2 = \{J^j\}$  forms a partition of  $[0, T]$ , we have  $\sum_{J'} v(I \cap J') = v(I)$ . Similarly,  $\sum_{I'} v(I' \cap J) = v(J)$ . Therefore

$$T_2 = \sum_{I,J} v_1(I)v_2(J)[(v(I) + v(J))K_{IJ} - v(I \cap J)] = v(I \cup J)K_{IJ}. \quad (15)$$

To compute the term  $T_1$ , we decompose the sum  $\sum_{(i,j) \in \mathcal{C}_3}$  into the sum of three terms

$$T_{1q} = \sum_{\substack{(i,j) \in \mathcal{C}_3 \\ \#\{j_1, j_2, j_3\} = q}} \prod_{p=1}^3 v(I^{i_p} \cap J^{j_p}), \quad q = 1, 2, 3.$$

If  $q = 1$ , then  $J^{j_1} = J^{j_2} = J^{j_3} := J$  and using the same arguments as for evaluating  $T_2$ , we get  $T_{11} = \sum_J v(J)^3$ . If  $q = 2$ , then  $j_1 = j_2 \neq j_3$  or  $j_1 = j_3 \neq j_2$  or  $j_1 \neq j_2 = j_3$ . Because of the symmetry, it suffices to consider one of these cases. Let  $j_1 = j_2 \neq j_3$  and

set  $J = J^{j_1}$  and  $J' = J^{j_3}$ . The relations  $(i, j) \in \mathcal{C}_3$  implies that both  $J$  and  $J'$  have non-empty intersections with both  $I^{i_1}$  and  $I^{i_3}$ . Therefore,  $I^{i_1} = I^{i_3} := I$  and setting  $I^{i_2} = I'$  we get

$$\begin{aligned} T_{12} &= 3 \sum_{J \neq J', I, I'} v(I \cap J)v(I' \cap J)v(I \cap J') = 3 \sum_{J \neq J', I} v(I \cap J)v(J)v(I \cap J') \\ &= 3 \sum_{I, J} v(I \cap J)v(J)[v(I) - v(I \cap J)]. \end{aligned}$$

In the case when all indices  $j_1, j_2$  and  $j_3$  are different, it is easily seen that  $(i, j) \in \mathcal{C}_3$  entails  $i_1 = i_2 = i_3$ . Therefore,

$$\begin{aligned} T_{13} &= \sum_{\substack{I, J, J', J'' \\ \#\{J, J', J''\}=3}} v(I \cap J)v(I \cap J')v(I \cap J'') \\ &= \sum_{I, J, J', J \neq J'} v(I \cap J)v(I \cap J')[v(I) - v(I \cap J') - v(I \cap J)] \\ &= \sum_{I, J, J', J \neq J'} v(I \cap J)v(I \cap J')v(I) - 2 \sum_{I, J \neq J'} v(I \cap J)^2v(I \cap J'). \end{aligned}$$

Using the identity  $\sum_{J': J \neq J'} v(I \cap J') = v(I) - v(I \cap J)$  we get  $T_{13} = \sum_I v(I)^3 - \sum_{I, J} v(I \cap J)^2[3v(I) - 2v(I \cap J)]$ . Summing up the terms  $T_{11}, T_{12}, T_{1,3}$  and  $T_2$  we get equality (14). Equality (13) can be proved along the same lines.  $\square$

**Remark 2.** If the observations are synchronous, that is  $\Pi^1 = \Pi^2 = \Pi$ , then  $\mu_2$  and  $\mu_3$  have the following simple expressions:

$$\begin{aligned} 2\mu_2 &= \sum_{I \in \Pi} [v(I)^2 + v_1(I)v_2(I)], \\ 4\mu_3 &= \sum_{I \in \Pi} [v(I)^3 + 3v_1(I)v_2(I)v(I)]. \end{aligned}$$

**Lemma 2.** Assume that we are given two sequences of partitions  $\Pi_n^1 = \{I_n^i, i \leq N_{1,n}\}$  and  $\Pi_n^2 = \{J_n^j, j \leq N_{2,n}\}$  of the interval  $[0, T]$ . Define the matrices  $A_n$  and  $\Sigma_n$  by (9) and (10). If the functions  $\sigma_1$  and  $\sigma_2$  are bounded on  $[0, T]$  by some constant  $\sigma$ , then

$$\max_{\ell} \lambda_{\ell, n}^2 = \|(\Sigma_n^{1/2} A_n \Sigma_n^{1/2})^2\| \leq 3\sigma^4 r_n^2,$$

where

$$r_n = [(\max_i |I_n^i|) \vee (\max_j |J_n^j|)].$$

*Proof.* Let us define a new partition  $\tilde{\Pi}_n^1$  as follows:  $I \in \tilde{\Pi}_n^1$  if and only if either  $I \in \Pi_n^1$  and it has non-empty intersection with two distinct intervals from  $\Pi_n^2$  or there is  $J \in \Pi_n^2$  such that  $I$  is the union of all intervals from  $\Pi_n^1$  included in  $J$ . The partition  $\tilde{\Pi}_n^2$  is defined analogously. It is easy to check that the estimator  $\hat{\theta}_n$  based on  $(\tilde{\Pi}_n^1, \tilde{\Pi}_n^2)$  is equal to the one based on  $(\Pi_n^1, \Pi_n^2)$ . It follows that  $\mu_{p, n} = \tilde{\mu}_{p, n}$  for every  $p \in \mathbb{N}$ . Therefore, the relation  $\max_{\ell} \lambda_{\ell, n}^2 = \lim_{p \rightarrow \infty} \mu_{2p, n}^{1/p}$  implies that  $\max_{\ell} \lambda_{\ell, n}^2 = \max_{\ell} \tilde{\lambda}_{\ell, n}^2$ . It is clear that  $r_n = \tilde{r}_n$ , but the advantage of working with  $(\tilde{\Pi}_n^1, \tilde{\Pi}_n^2)$  is that

$$\max_{J \in \tilde{\Pi}^2} \sum_{I \in \tilde{\Pi}^1} K_{IJ} \leq 3, \quad \max_{I \in \tilde{\Pi}^1} \sum_{J \in \tilde{\Pi}^2} K_{IJ} \leq 3. \quad (16)$$

In the remaining of this proof, without loss of generality we assume that (16) is fulfilled for partitions  $(\Pi^1, \Pi^2)$ . The estimate  $\|(\Sigma_n^{1/2} A_n \Sigma_n^{1/2})^2\| \leq \|\Sigma_n\|^2 \|A_n\|^2$  implies that it suffices to estimate  $\|A_n\|$  and  $\|\Sigma_n\|$ . To bound from above  $\|A_n\|^2$ , we use  $\|A_n\|^2 = \max_{u:|u|=1} |A_n u|^2$  and

$$|A_n u|^2 = \frac{1}{4} \sum_i \left( \sum_j K_{I^i J^j} u_{N_1+j} \right)^2 + \frac{1}{4} \sum_j \left( \sum_i K_{I^i J^j} u_i \right)^2.$$

Applying the Cauchy-Schwarz inequality and changing the order of summation, we get

$$\sum_i \left( \sum_j K_{I^i J^j} u_{N_1+j} \right)^2 \leq \frac{3}{4} \sum_j u_{N_1+j}^2$$

and  $\sum_j \left( \sum_i K_{I^i J^j} u_i \right)^2 \leq \frac{3}{4} \sum_i u_i^2$ . These inequalities yield  $\|A_n\|^2 \leq 3/4$ .

On the other hand,

$$\begin{aligned} \|\Sigma_n\| &= \max_{u:|u|=1} u^T \Sigma_n u = \max_{u:|u|=1} \sum_{\ell, \ell'=1}^N \sigma_{\ell, \ell'} u_\ell u_{\ell'} \\ &= \max_{u:|u|=1} \left( \sum_{\ell=1}^N \sigma_{\ell, \ell} u_\ell^2 + 2 \sum_{i,j} v(I_n^i \cap J_n^j) u_i u_{N_1+j} \right). \end{aligned}$$

Since  $\sigma_{\ell, \ell'}$ s are given by (10), the first sum in the right hand side is bounded by  $\sigma^2 (\max_i |I_n^i|) \vee (\max_j |J_n^j|)$ , whereas the second sum can be bounded using the inequality between geometrical and arithmetical means :

$$\begin{aligned} 2 \sum_{i,j} v(I_n^i \cap J_n^j) u_i u_{N_1+j} &\leq \sum_{i,j} v(I_n^i \cap J_n^j) u_i^2 + \sum_{i,j} v(I_n^i \cap J_n^j) u_{N_1+j}^2 \\ &= \sum_i v(I_n^i) u_i^2 + \sum_j v(J_n^j) u_{N_1+j}^2 \\ &\leq |u|^2 \sigma^2 (\max_i |I_n^i|) \vee (\max_j |J_n^j|). \end{aligned}$$

This completes the proof of the lemma.  $\square$

As a by-product of the preceding lemma, we give below a simple sufficient condition for the asymptotic normality of  $\hat{\theta}_n$ .

**Corollary 2.** *In the notation of Lemma 2, if*

$$\lim_{n \rightarrow \infty} \frac{r_n^2}{\mu_{2,n}} = 0, \quad (17)$$

then  $(\hat{\theta}_n - \theta) / \sqrt{2\mu_{2,n}}$  converges in distribution to a standard Gaussian random variable.

*Proof.* According to Proposition 2, it is enough to show that

$$\lim_{n \rightarrow \infty} \frac{\|(\Sigma_n^{1/2} A_n \Sigma_n^{1/2})^2\|}{\text{Tr}[(\Sigma_n A_n)^2]} = 0.$$

This convergence follows from assumption (17) and Lemma 2.  $\square$

### 3.3. Expansion of the characteristic function for random sampling schemes.

We assume now that the partitions  $\Pi_n^1$  and  $\Pi_n^2$  are random and independent of  $\{X_{1,t} - X_{1,0}, X_{2,t} - X_{2,0}\}_{t \in [0, T]}$ . We denote by  $\mathbf{E}^\Pi$  the conditional expectation given  $\Pi_n$ , where  $\Pi_n = (\Pi_n^1, \Pi_n^2)$ . Since in this setup the quantities  $r_n$  and  $\mu_{2,n}$  — introduced in Lemma 2 and in Proposition 2, respectively — are random, Corollary 2 can not be applied directly. The following result gives a sufficient condition for the convergence in distribution of Corollary 2 to hold in the setup of random sampling scheme.

**Proposition 4.** *Let  $r_n$  be defined as in Lemma 2. If  $r_n^2/\mu_{2,n}$  tends to zero in probability as  $n \rightarrow \infty$ , then  $(\hat{\theta}_n - \theta_n)/\sqrt{2\mu_{2,n}}$  converges in distribution to a standard normal random variable. If moreover,  $2\mu_{2,n}/b_n \xrightarrow[n \rightarrow \infty]{P} \mathbf{c}$  for some deterministic sequence  $\{b_n\}$  and some positive constant  $\mathbf{c}$ , then  $(\hat{\theta}_n - \theta)/\sqrt{b_n} \xrightarrow[n \rightarrow \infty]{D} N(0, \mathbf{c})$ .*

*Proof.* Denote  $\sigma[\Pi] = \sigma[\Pi_n, n \in \mathbb{N}]$ . We may represent random variables on a product space. By subsequent argument, we may assume  $r_n^2/\mu_{2,n} \rightarrow 0$  a.s. Then by Corollary 2 and Fubini's theorem,

$$E\left[\psi \exp\left(iu(\hat{\theta}_n - \theta_n)/\sqrt{2\mu_{2,n}}\right)\right] \rightarrow E[\psi e^{-\frac{1}{2}u^2}]$$

for every  $u \in \mathbb{R}$  and every bounded  $\sigma[\Pi]$ -measurable random variable  $\psi$ . The desired results follow from this convergence.  $\square$

From now on, we assume that the assumptions of Proposition 4 are fulfilled and aim at finding the asymptotic expansion of the distribution of the random variable  $\mathcal{X}_n = (\hat{\theta}_n - \theta)/\sqrt{b_n}$  as  $n \rightarrow \infty$ . The first step in deriving the asymptotic expansion of a distribution is the expansion of the characteristic function. As usual, the desired expansion involves the  $r$ -th conditional cumulant of  $\mathcal{X}_n$  given  $\Pi$ , henceforth denoted by  $\kappa_r[\mathcal{X}_n]$ . Let  $\bar{\lambda}_{r,n}$  be the normalized  $r$ -th conditional cumulant of  $\mathcal{X}_n$ :

$$\bar{\lambda}_{r,n} = b_n^{-\frac{r-2}{2}} \kappa_r[\mathcal{X}_n] = 2^{r-1}(r-1)! b_n^{-r+1} \mu_{r,n}.$$

Denote  $\alpha_n = \sqrt{3}\sigma^2 r_n b_n^{-1/2}$ .

**Lemma 3.** *Let  $r$  be a positive integer.*

(a) *We have*

$$|\mu_{r,n}| \leq \sum_{\ell} |\lambda_{\ell,n}|^r \leq \max_{\ell} |\lambda_{\ell}|^{r-2} \mu_{2,n} \leq (\alpha_n \sqrt{b_n})^{r-2} \mu_{2,n} \quad (18)$$

(b) *The conditional cumulants admit  $|\kappa_r[\mathcal{X}_n]| \leq c(r)\sigma^{2(r-2)}(b_n^{-\frac{1}{2}}r_n)^{r-2} \bar{\lambda}_{2,n}$ , where  $c(r) = (2\sqrt{3})^{r-2}(r-1)!$  is constant.*

*Proof.* Lemma 2 yields (a), and (a) entails (b).  $\square$

**Proposition 5.** *Let the sequence  $\{b_n\}$  be as in Proposition 4. For some fixed  $\mathbf{c}_1 > 0$ , let*

$$\mathcal{P}_n(\delta) = \{\Pi : \alpha_n < \delta, \bar{\lambda}_{2,n} < \mathbf{c}_1\}, \quad \forall \delta > 0.$$

*Then, for every  $j \in \mathbb{Z}_+$ , there exist some positive constants  $C$  and  $q$  such that*

$$\frac{d^j}{du^j} \left( \mathbf{E}^\Pi [e^{iu\mathcal{X}_n}] \right) = \frac{d^j}{du^j} \left\{ e^{-\frac{\bar{\lambda}_{2,n}u^2}{2}} \left( 1 + \frac{(iu)^3 b_n^{1/2}}{6} \bar{\lambda}_{3,n} \right) \right\} + \mathcal{O}(\delta^2)(1 + |u|^q) e^{-\frac{\bar{\lambda}_{2,n}}{2}u^2}$$

for every  $u$  satisfying  $|u| \leq C\delta^{-1/3}$  and for every  $\Pi_n \in \mathcal{P}_n(\delta)$ . In this formula  $\mathcal{O}(\delta^p)$  stands for a random variable depending only on partitions  $\Pi_n = (\Pi_n^1, \Pi_n^2)$  and verifying  $\limsup_{\delta \rightarrow 0} \sup_n \sup_{\Pi_n \in \mathcal{P}_n(\delta)} |\mathcal{O}(\delta^p)| \delta^{-p} < \infty$ .

*Proof.* Using (4) and the fact that in our setup  $\max_\ell |\lambda_\ell|$  is bounded by  $\sqrt{3}\sigma^2 r_n$ , we get

$$\mathbf{E}^\Pi[e^{iu\mathcal{X}_n}] = \exp \left\{ \sum_{k=2}^{\infty} \frac{(2iu)^k \mu_{k,n}}{2kb_n^{k/2}} \right\} = \exp \left\{ -\frac{\bar{\lambda}_{2,n}u^2}{2} + \frac{(iu)^3 \sqrt{b_n} \bar{\lambda}_{3,n}}{6} + r_n \right\}$$

for every  $u \in \mathbb{R}$  such that  $|u| < 1/(2\delta)$ . Let us define  $a_0(u) = -\bar{\lambda}_{2,n}u^2/2$ ,  $a_{1,n}(u) = \frac{(iu)^3 \sqrt{b_n} \bar{\lambda}_{3,n}}{6}$  and  $r_n(u) = \sum_{k=4}^{\infty} \frac{(2iu)^k \mu_{k,n}}{2kb_n^{k/2}}$ . One easily checks that

$$\begin{aligned} \mathbf{E}^\Pi[e^{iu\mathcal{X}_n}] - e^{a_0(u)}(1 + a_{1,n}(u)) &= e^{a_0(u)}(a_{1,n}(u) + r_n(u))^2 \int_0^1 \int_0^1 v e^{tv(a_{1,n}(u) + r_n(u))} dt dv \\ &\quad + r_n(u)e^{a_0(u)}. \end{aligned} \tag{19}$$

The inequalities (18) imply that there exists some constant  $C > 0$  such that for every  $\ell \leq j$  and for every  $\Pi_n \in \mathfrak{P}_n$ , it holds

$$\left| \frac{d^\ell r_n(u)}{du^\ell} \right| \leq C \frac{(1+u^4)\alpha_n^2 \mu_{2,n}}{b_n} \leq C_1(1+u^4)\delta^2,$$

as soon as  $|u| \leq 1/(4\alpha_n)$ . Similarly, for every  $\ell \in \mathbb{N}$ ,

$$\left| \frac{d^\ell}{du^\ell} a_{1,n}(u) \right| \leq C_2(1+|u|^3) \frac{\alpha_n \mu_{2,n}}{3b_n} \leq C_2(1+|u|^3)\delta, \quad \text{if } \Pi_n \in \mathcal{P}_n.$$

These inequalities in conjunction with Eq. (19) yield the estimate

$$\frac{d^j}{du^j} \left( \mathbf{E}^\Pi[e^{iu\mathcal{X}_n}] - e^{a_0(u)}(1 + a_{1,n}(u)) \right) = \mathcal{O}(\delta^2)(1 + |u|^q)e^{a_0(u)}.$$

This completes the proof of the proposition  $\square$

**Remark 3.** As usual in asymptotic expansions, the coefficient of the second order term (i.e., the coefficient of  $(iu)^3 b_n^{1/2}$ ) in the obtained decomposition is given by the normalized third cumulant divided by 6. It also admits the following representations:

$$\frac{b_n^{1/2}}{6} \bar{\lambda}_{3,n} = \frac{1}{6} \kappa_3[\mathcal{X}_n] = \frac{4\mu_{3,n}}{3b_n \sqrt{b_n}}$$

where  $\mu_{3,n}$  is defined by (14).

#### 4. ASYMPTOTIC EXPANSION OF THE DISTRIBUTION

In this section, we will derive the second-order asymptotic expansion for the distribution of  $\mathcal{X}_n$  for the model (1) without drifts. We will treat a model with drifts in Section 6, and there the Malliavin calculus plays an essential role to handle general nonlinear Wiener functionals.

Given positive numbers  $M$  and  $\gamma$ , let  $\mathcal{E}(M, \gamma)$  denote the set of measurable functions  $f : \mathbb{R} \rightarrow \mathbb{R}$  satisfying  $|f(x)| \leq M(1 + |x|^\gamma)$  for all  $x \in \mathbb{R}$ . For positive numbers  $C, \eta, r_0$  and  $\mathfrak{c}^*$  we set

$$\mathcal{E}^0 = \mathcal{E}^0(C, \eta, r_0, \mathfrak{c}^*) = \left\{ f : \int_{\mathbb{R}} \bar{\omega}_f(z, r) \phi(z; \mathfrak{c}^*) dz \leq Cr^\eta, \forall r \leq r_0 \right\},$$

where

$$\bar{\omega}_f(z, r) = \sup_{x:|x|\leq r} |f(z+x) - f(z)|$$

and  $\phi(z; \Sigma)$  is the centered normal density with variance  $\Sigma$ . Note that this class is large enough to contain most functions that are encountered in practice. In particular, all functions satisfying the generalized Hölder condition  $|f(z+x) - f(z)| \leq F(z)|x|^\eta$  with some function  $F$  such that  $\int F(z)\phi(z; \mathbf{c}^*) dz \leq C$  belong to  $\mathcal{E}^0(\mathbf{C}, \eta, \infty, \mathbf{c}^*)$ . It is also easy to check that the set of all indicator functions of intervals of  $\mathbb{R}$  is included in  $\mathcal{E}^0(\sqrt{2\pi}\mathbf{c}^*, 1, \infty, \mathbf{c}^*)$  for any  $\mathbf{c}^* > 0$ .

Our aim is now to get uniform in  $f \in \mathcal{E}^*$  asymptotic expansion for the sequence  $\mathbf{E}[f(b_n^{-1/2}(\hat{\theta}_n - \theta))]$  with  $\mathcal{E}^* = \mathcal{E}(M, \gamma) \cap \mathcal{E}^0(\mathbf{C}, \eta, r_0, \mathbf{c}^*)$ . To this end, define a  $\sigma[\Pi]$ -dependent random signed-measure  $\Psi_n^\Pi$  on  $\mathbb{R}$  by the density

$$p_{3,n}(z) = \phi(z; \bar{\lambda}_{2,n}) + \frac{b_n^{1/2}}{6} \bar{\lambda}_{3,n} h_3(z; \bar{\lambda}_{2,n}) \phi(z; \bar{\lambda}_{2,n}).$$

Here  $h_r(z; \Sigma)$  is the  $r$ -th Hermite polynomial defined by

$$h_r(z; \Sigma) = (-1)^r \phi(z; \Sigma)^{-1} \partial_z^r \phi(z; \Sigma)$$

and in particular,  $h_3(z; \bar{\lambda}_{2,n}) = (z^3 - 3\bar{\lambda}_{2,n}z)/\bar{\lambda}_{2,n}^3$ . The Fourier transform of  $\Psi_n^\Pi$  is given by

$$\hat{\Psi}_n^\Pi(u) = e^{-\frac{1}{2}\bar{\lambda}_{2,n}u^2} \left[ 1 + \frac{b_n^{1/2}}{6} \bar{\lambda}_{3,n}(iu)^3 \right].$$

**Theorem 1.** *Let  $M, \gamma, \eta, \mathbf{C}, r_0 > 0$ ,  $a \in (\frac{3}{4}, 1)$  and  $0 < \mathbf{c}_0 < \mathbf{c}_1 < \mathbf{c}^*$  and let*

$$\mathcal{P}_n(\mathbf{c}_0, \mathbf{c}_1, a) = \{\mathbf{c}_0 < \bar{\lambda}_{2,n} < \mathbf{c}_1, r_n \leq b_n^a\},$$

where  $r_n$  is the maximal lag of the time points defined in Lemma 2 and  $\bar{\lambda}_{2,n} = 2\mu_{2,n}/b_n$  with  $\mu_{2,n}$  as in (13). Then, there exists a sequence  $\epsilon_n = \epsilon_n(M, \gamma, \eta, \mathbf{C}, r_0, a, \mathbf{c}_0, \mathbf{c}_1)$  such that  $\epsilon_n = O(b_n^{2a-1})$  and for every  $\Pi_n \in \mathcal{P}_n(\mathbf{c}_0, \mathbf{c}_1, a)$ , it holds

$$\sup_{f \in \mathcal{E}(M, \gamma) \cap \mathcal{E}^0(\mathbf{C}, \eta, r_0, \mathbf{c}^*)} |\mathbf{E}^\Pi[f(\mathcal{X}_n)] - \Psi_n^\Pi[f]| \leq \epsilon_n,$$

where  $\mathcal{X}_n = (\hat{\theta}_n - \theta)/\sqrt{b_n}$ .

*Proof of Theorem 1.* Let  $\mathbf{h}(x) = 1 + x^{p_0}$  for a positive even integer  $p_0$ ; we will take a large  $p_0$  according to  $\gamma$ . Let  $\mathcal{K}$  be a probability measure on  $\mathbb{R}$  such that  $\text{supp}(\hat{\mathcal{K}})$  is compact, and let  $\mathbf{a}$  be a number such that

$$\tilde{\alpha} := \mathcal{K}(\{x; |x| < \mathbf{a}\}) > \frac{1}{2}.$$

Let  $K_1 > K > 0$  and set  $\zeta = (K_1 - K)/2$ . For  $\epsilon > 0$ , define the measure  $\mathcal{K}_\epsilon$  by  $\mathcal{K}_\epsilon(B) = \mathcal{K}(\epsilon^{-1}B)$  for Borel sets  $B$ .

Sweeting's smoothing lemma provides the inequality:

$$|\mathbf{E}^\Pi[f(\mathcal{X}_n)] - \Psi_n^\Pi[f]| \leq CM[\mathbb{A}_0 + \mathbb{A}_1 + \mathbb{A}_2] + (2\tilde{\alpha} - 1)^{-1}\mathbb{A}_3, \quad (20)$$

where  $C$  is a constant depending on  $\gamma$  and  $\tilde{\alpha}$ ,

$$\begin{aligned}\mathbb{A}_0 &= \left| \int_{\mathbb{R}} \mathbf{h}(x) \mathcal{K}_{b_n^{K_1}} * (P^{\mathcal{X}_n|\Pi} - \Psi_n^\Pi)(dx) \right|, \\ \mathbb{A}_1 &= (P^{\mathcal{X}_n|\Pi} + |\Psi_n^\Pi|)[\mathbf{h}] \int_{x:|x| \geq 2ab_n^{-\zeta}} \mathbf{h}(x) \mathcal{K}(dx) \\ \mathbb{A}_2 &= (P^{\mathcal{X}_n|\Pi} + |\Psi_n^\Pi|)[\mathbf{h}] \delta(\tilde{\alpha})^{[a^{-1/2}b_n^{-\zeta}]}, \quad \delta(\tilde{\alpha}) = (1 - \tilde{\alpha})/\tilde{\alpha} < 1, \\ \mathbb{A}_3 &= \int_{\mathbb{R}} \sup_{x:|x| \leq ab_n^{K+\zeta}[(2a)^{-1}b_n^{-\zeta}]} \omega_f(x+y, 2ab_n^{K_1})(\Psi_n^\Pi)^+(dy).\end{aligned}$$

See Sweeting [28], Bhattacharya and Rao [5].

If  $n$  is sufficiently large,  $[x+y-2ab_n^{K_1}, x+y+2ab_n^{K_1}] \subset [y-b_n^K, y+b_n^K]$  and therefore

$$\mathbb{A}_3 \leq \int_{\mathbb{R}} \omega_f(z, b_n^K)(\Psi_n^\Pi)^+(dz) \leq C^\circ \int_{\mathbb{R}} \omega_f(z, b_n^K) \phi(z; \mathbf{c}^*) dz \leq Cb_n^{K\eta}.$$

On the other hand,  $\mathbb{A}_0$  admits the estimate

$$\mathbb{A}_0 \leq \sum_{\alpha=0}^{2+p_0} \int_{\mathbb{R}} \left| \partial_u^\alpha \left[ \left( \varphi_{\mathcal{X}_n}^\Pi(u) - \hat{\Psi}_n^\Pi(u) \right) \hat{\mathcal{K}}(n^{-K_1}u) \right] \right| du,$$

where  $\varphi_{\mathcal{X}_n}^\Pi(u) = \mathbf{E}^\Pi[e^{iu\mathcal{X}_n}]$ .

Let  $\delta_n = b_n^{a-1/2}$ . By virtue of Proposition 5 and Lemma 1, we have

$$\begin{aligned}& \int_{\mathbb{R}} \left| \partial_u^\alpha \left[ \left( \varphi_{\mathcal{X}_n}^\Pi(u) - \hat{\Psi}_n^\Pi(u) \right) \hat{\mathcal{K}}(n^{-K_1}u) \right] \right| du \\ & \leq \int_{u:|u| \leq C\delta_n^{-1/3}} \left| \partial_u^\alpha \left[ \left( \varphi_{\mathcal{X}_n}^\Pi(u) - \hat{\Psi}_n^\Pi(u) \right) \hat{\mathcal{K}}(n^{-K_1}u) \right] \right| du \\ & \quad + \int_{u:|u| > C\delta_n^{-1/3}} \left| \partial_u^\alpha \left[ \left( \varphi_{\mathcal{X}_n}^\Pi(u) - \hat{\Psi}_n^\Pi(u) \right) \hat{\mathcal{K}}(n^{-K_1}u) \right] \right| du \\ & \leq \int_{u:|u| \leq C\delta_n^{-1/3}} \mathcal{O}(\delta_n^2)(1+|u|^q)e^{-\frac{\mu_{2,n}u^2}{2}} du + \int_{u:|u| > C\delta_n^{-1/3}} \frac{C_2}{1+|u|^L} du \\ & \quad + \sum_{\alpha'=0}^{2+p_0} \int_{u:|u| > C\delta_n^{-1/3}} |\partial_u^{\alpha'} \hat{\Psi}_n^\Pi(u)| \mathbf{1}_{A_n} du \\ & \leq C_3[\mathcal{O}(\delta_n^2) + \delta_n^{(L-1)/3}] \leq C_4\delta_n^2,\end{aligned}$$

where  $L$  can be chosen as large as we need, therefore  $\mathbb{A}_0 \leq C_5\delta_n^2$ . As we already mentioned, the Rosenthal inequality yields the uniform in  $n$  boundedness of  $P^{\mathcal{X}_n|\Pi}[\mathbf{h}] = 1 + \mathbf{E}^\Pi[\mathcal{X}_n^{p_0}]$ . In conjunction with the boundedness of  $|\Psi_n^\Pi|[\mathbf{h}]$ , this implies that the term  $\mathbb{A}_2$  decreases exponentially fast as  $n$  goes to infinity. We choose the kernel  $\mathcal{K}$  so that its derivative  $\mathcal{K}'(x)$  tends to zero as  $|x| \rightarrow \infty$  at least as fast as the polynomial  $|x|^{-1-\zeta^{-1}}$ . Such a choice of  $\mathcal{K}$  can be achieved, for example, by taking a multiple convolution of the sinc kernel. For such a  $\mathcal{K}$ , the term  $\mathbb{A}_2$  becomes  $O(b_n)$  and the desired inequality follows.  $\square$

**Remark 4.** The approximating measure  $\Psi_n$  provided by Theorem 1 contains the Gaussian density with variance  $\bar{\lambda}_{2,n}$ , which depends on  $n$ . One easily deduce from that result

that the distribution of  $(b_n \bar{\lambda}_{2,n})^{-1/2}(\hat{\theta}_n - \theta)$  can be approximated by the measure

$$\left\{ \phi(z; 1) + \frac{\sqrt{b_n}}{6} \left( z^3 - \frac{3}{\bar{\lambda}_{2,n}} z \right) \phi(z; 1) \right\} dz.$$

As we have already seen, under mild assumptions including the convergence of  $\bar{\lambda}_{2,n}$  to some constant  $\mathfrak{c}$ , the estimator  $\hat{\theta}_n$  is asymptotically normal. It is therefore natural to address the issue of approximating the distribution of  $\mathcal{X}_n$  by a measure similar to  $\Psi_n$  but based on the Gaussian density with variance  $\mathfrak{c}$ . To this end, we define the signed measure  $\tilde{\Psi}_n^\Pi$  on  $\mathbb{R}$  by the density

$$\tilde{p}_{3,n}(z) = \phi(z; \mathfrak{c}) \left[ 1 + \frac{1}{2}(\bar{\lambda}_{2,n} - \mathfrak{c})h_2(z; \mathfrak{c}) + \frac{b_n^{1/2}}{6}\bar{\lambda}_{3,n} h_3(z; \mathfrak{c}) \right].$$

Note that  $h_2(z; \mathfrak{c}) = (z^2 - \mathfrak{c})/\mathfrak{c}^2$ .

**Proposition 6.** *Let  $M, \gamma, \eta, \mathbf{C}, r_0 > 0$ ,  $\mathfrak{c}^* > \mathfrak{c} > 0$ ,  $a \in (\frac{3}{4}, 1)$  and let*

$$A_n(a) = \{ r_n \leq b_n^a, (\bar{\lambda}_{2,n} - \mathfrak{c})^2 \leq b_n^{2a-1} \}$$

where  $r_n$  is the maximal lag of the time points defined in Lemma 2 and  $\bar{\lambda}_{2,n} = 2\mu_{2,n}/b_n$  with  $\mu_{2,n}$  as in (13). Then, there exists a sequence  $\epsilon_n = \epsilon_n(M, \gamma, \eta, \mathbf{C}, r_0, a, \mathfrak{c})$  such that  $\epsilon_n = O(b_n^{2a-1})$  and for every  $\Pi_n \in A_n(a)$ , it holds

$$\sup_{f \in \mathcal{E}(M, \gamma) \cap \mathcal{E}^0(\mathbf{C}, \eta, r_0, \mathfrak{c}^*)} \left| \mathbf{E}^\Pi[f(\mathcal{X}_n)] - \tilde{\Psi}_n^\Pi[f] \right| \leq \epsilon_n.$$

*Proof.* We notice that  $|\bar{\lambda}_{3,n}(\bar{\lambda}_{2,n} - \mathfrak{c})| = O(b_n^{-1}r_n \times b_n^{a-1/2}) = O(b_n^{2a-\frac{3}{2}}) = o(1)$  uniformly on the event  $A_n$ . Expanding  $\phi(z; \bar{\lambda}_{2,n})$  in  $\Psi_n^\Pi$  around  $\mathfrak{c}$  we get the desired inequality.  $\square$

The following result is a trivial consequence of Proposition 6 and provides an unconditional asymptotic expansion for the distribution of  $\mathcal{X}_n = b_n^{-1/2}(\hat{\theta}_n - \theta)$ .

**Theorem 2.** *Under the notation of Proposition 6, if  $\mathbf{P}(A_n(a)^c) = o(b_n^p)$  for every  $p > 1$ , and  $\mathbf{E}[\bar{\lambda}_{2,n} - \mathfrak{c}] = O(b_n^{2a-1})$ , then*

$$\sup_{f \in \mathcal{E}(M, \gamma) \cap \mathcal{E}^0(\mathbf{C}, \eta, r_0, \mathfrak{c}^*)} \left| \mathbf{E}[f(\mathcal{X}_n)] - \int_{\mathbb{R}} f(z) p_n^*(z) dz \right| = O(b_n^{2a-1}), \quad (21)$$

where  $p_n^*(z) = \phi(z; \mathfrak{c}) \left[ 1 + \frac{b_n^{1/2}}{6} \mathbf{E}[\bar{\lambda}_{3,n}] h_3(z; \mathfrak{c}) \right]$ . Moreover, if  $\sup_{n \in \mathbb{N}} \mathbf{E}[\bar{\lambda}_{3,n}] < \infty$ , then inequality (21) holds with  $p_n^*$  replaced by

$$p_n^\dagger(z) = \frac{\max(0, p_n^*(z))}{\int_{\mathbb{R}} \max(0, p_n^*(u)) du},$$

which is a probability density.

## 5. POISSON SAMPLING SCHEME

As an application of previous results let us consider the case when the partitions  $\Pi_n^1$  and  $\Pi_n^2$  are generated by Poisson point processes. Let  $\mathcal{P}^{i,n} = (\mathcal{P}_t^{i,n}, t \geq 0)$ ,  $i = 1, 2$ , be two independent homogeneous Poisson processes with intensities  $np_i$ ,  $i = 1, 2$ . Moreover, assume that these processes are independent of  $\mathbf{B}$ . Let the observation times  $S^1, \dots, S^{N_1}$  and  $T^1, \dots, T^{N_2}$  be the time instants corresponding to the jumps of  $\mathcal{P}^{1,n}$  and  $\mathcal{P}^{2,n}$



occurred before the instant  $T$ . Note that  $S^i$ 's and  $T^j$ 's depend also on  $n$ . However, for simplicity of exposition this dependence will not be reflected in our notation.

The estimation of the covariance  $\theta$  in the setup of Poisson sampling scheme has been studied by Hayashi and Yoshida. Let us state their result on the asymptotic normality of the estimator  $\hat{\theta}_n$ .

**Theorem 3** (Hayashi and Yoshida (2005)). *Let  $\hat{\theta}_n$  be defined by (2). If the functions  $\sigma_1$ ,  $\sigma_2$  and  $\rho$  are continuous, then the sequence  $\sqrt{n}(\hat{\theta}_n - \theta)$  converges in distribution to a centered Gaussian random variable with variance*

$$\mathfrak{c} = \left( \frac{2}{p_1} + \frac{2}{p_2} \right) \int_0^T \sigma_{1,t}^2 \sigma_{2,t}^2 (1 + \rho_t^2) dt - \frac{2}{p_1 + p_2} \int_0^T (\sigma_{1,t} \sigma_{2,t} \rho_t)^2 dt.$$

In the notation of the present work, this result entails that  $2n\mu_{2,n}$  converges in probability to  $\mathfrak{c}$ . Since we are interested in deriving the second order expansion of  $\sqrt{n}(\hat{\theta}_n - \theta)$ , we need a result providing the rate of convergence of  $2n\mu_{2,n}$  to  $\mathfrak{c}$ . It is done in the following

**Proposition 7.** *If the functions  $\sigma_1$ ,  $\sigma_2$  and  $\rho$  are Lipschitz continuous, then there exists a constant  $C > 2$  depending only on  $p_1$  and  $p_2$  such that, for every  $x > C \log n$  and for every  $n \geq 2$ , it holds*

$$\mathbf{P} \left( \left| 2n\mu_{2,n} - \mathfrak{c} \right| > \frac{C \log^3 n}{n} + \frac{x}{\sqrt{n}} \right) \leq C n e^{-x/C}.$$

*Proof.* Let us recall the relations

$$\begin{aligned} n \sum_{I,J} v_1(I) v_2(J) K_{IJ} &\xrightarrow[n \rightarrow \infty]{P} 2(p_1^{-1} + p_2^{-1}) \int_0^T \sigma_{1,t}^2 \sigma_{2,t}^2 dt, \\ n \sum_{I \in \Pi_i} v(I)^2 &\xrightarrow[n \rightarrow \infty]{P} 2p_i^{-1} \int_0^T (\sigma_{1,t} \sigma_{2,t} \rho_t)^2 dt, \quad i = 1, 2 \\ n \sum_{I,J} v(I \cap J)^2 &\xrightarrow[n \rightarrow \infty]{P} 2(p_1 + p_2)^{-1} \int_0^T (\sigma_{1,t} \sigma_{2,t} \rho_t)^2 dt \end{aligned}$$

proved in Hayashi and Yoshida (2005). The aim of the present lemma is to show that the rate of convergence in these relations is  $1/\sqrt{n}$  and to get an exponential control of probabilities of large deviations. Thus, let us denote  $\mathcal{T}_1 = n \sum_{I,J} v_1(I) v_2(J) K_{IJ}$  and show now that

$$\mathbf{P} \left( \left| \mathcal{T}_1 - 2(p_1^{-1} + p_2^{-1}) \int_0^T \sigma_{1,t}^2 \sigma_{2,t}^2 dt \right| \geq \frac{x}{\sqrt{n}} \right) \leq C n e^{-x/C}.$$

Let us denote by  $N(x) = \lceil nT/x \rceil$  the smallest positive integer such that  $N(x)x > nT$  and let us set  $L_i = [iT/N(x), (i+1)T/N(x)]$ . The intervals  $L_i$  define a uniform deterministic partition of  $[0, T]$  with a mesh-size of order  $x/n$ . Let  $\mathcal{E}$  be the event “for every  $i = 1, \dots, 4N(x)$ , the interval  $[\frac{iT}{4N(x)}, \frac{(i+1)T}{4N(x)}]$  contains at least one point from  $\Pi_n^1$  and one point from  $\Pi_n^2$ ”. The total probability formula implies that

$$\begin{aligned} &\mathbf{P} \left( \left| \mathcal{T}_1 - 2(p_1^{-1} + p_2^{-1}) \int_0^T \bar{h}(t) dt \right| \geq \frac{x}{\sqrt{n}} \right) \\ &\leq \mathbf{P} \left( \left| \mathcal{T}_1 - 2(p_1^{-1} + p_2^{-1}) \int_0^T \bar{h}(t) dt \right| \geq \frac{x}{\sqrt{n}} \mid \mathcal{E} \right) + \mathbf{P}(\mathcal{E}^c), \end{aligned}$$

where  $\mathcal{E}^c$  denotes the complementary event of  $\mathcal{E}$  and  $\bar{h}(t) = \sigma_{1,t}^2 \sigma_{2,t}^2$ . Easy computations show that  $\mathbf{P}(\mathcal{E}^c) \leq Cnx^{-1}e^{-x/C}$  for some  $C > 0$ .

Let now  $l_i$  be a point in  $L_i$  such that  $\int_{L_i} \bar{h}(t) dt = \bar{h}(l_i)|L_i|$ . Let us denote by  $a_I$  the left endpoint of the interval  $I$  and let us define the random variables

$$\eta_i^\circ = n\bar{h}(l_i) \sum_{I,J} |I||J|K_{IJ}\mathbf{1}_{\{a_I \in L_i\}}, \quad i = 1, \dots, N(x),$$

and write  $\mathcal{T}_1 = \mathcal{T}_{11} + \mathcal{T}_{12} + \mathcal{T}_{13} + \mathcal{O}(n|L_1|^2)$  on  $\mathcal{E}$ , where

$$\begin{aligned} \mathcal{T}_{11} &= \mathbf{E}^\mathcal{E} \left[ \sum_{i=1}^{N(x)} \eta_i^\circ \right] - 2(p_1^{-1} + p_2^{-1}) \int_0^T \bar{h}(t)^2 dt \\ \mathcal{T}_{12} &= \sum_{i=1}^{[N(x)/2]} (\eta_{2i}^\circ - \mathbf{E}^\mathcal{E}[\eta_{2i}^\circ]) \\ \mathcal{T}_{13} &= \sum_{i=0}^{[N(x)/2]} (\eta_{2i+1}^\circ - \mathbf{E}^\mathcal{E}[\eta_{2i+1}^\circ]). \end{aligned}$$

For evaluating the remainder term in  $\mathcal{T}_1$ , we have used the Lipschitz continuity of  $\sigma_1$  and  $\sigma_2$ , as well as the fact that  $r_n \leq |L_1|/2$  on  $\mathcal{E}$ .

On the one hand, since  $|\sum_{i=1}^{N(x)} \eta_i^\circ| \leq Cnr_n$ , we have

$$\left| \mathbf{E}^\mathcal{E} \left[ \sum_{i=1}^{N(x)} \eta_i^\circ \right] - \mathbf{E} \left[ \sum_{i=1}^{N(x)} \eta_i^\circ \right] \right| \leq \frac{n\mathbf{E}[r_n \mathbf{1}_{\mathcal{E}^c}]}{\mathbf{P}(\mathcal{E})}.$$

Using the inequality of Cauchy-Schwarz, as well as the bounds  $\mathbf{P}(\mathcal{E}^c) \leq Cne^{-x/C}$  and (22), we get  $|\mathbf{E}^\mathcal{E}[\sum_{i=1}^{N(x)} \eta_i^\circ] - \mathbf{E}[\sum_{i=1}^{N(x)} \eta_i^\circ]| \leq Cne^{-x/C}$ , for some constant  $C$  and for every  $x > C \log n$ .

On the other hand, in view of Lemma 15, we have

$$\mathbf{E}[\eta_i^\circ] \leq n\bar{h}(l_i) \mathbf{E} \left[ \sum_{I:a_I \in L_i} \left( |I|^2 + \frac{2|I|}{np_2} \right) \right] \leq Cn\mathbf{E}[(r_n + n^{-1})(|L_i| + r_n)].$$

Therefore, using (22), we get  $\mathbf{E}[\eta_i^\circ] = \mathcal{O}(n^{-1} \log^3 n)$  for every  $i \leq N(x)$ . Using once again Lemma 15, we get

$$\begin{aligned} \mathbf{E} \left[ \sum_{i=1}^{N(x)} \eta_i^\circ \right] &= \sum_{i=2}^{N(x)-1} n\bar{h}(l_i) \mathbf{E} \left[ \sum_{I:a_I \in L_i} |I| \cdot \mathbf{E}^{\Pi^1} \left( \sum_{J \in \Pi^2} |J|K_{IJ} \right) \right] + \mathcal{O} \left( \frac{\log^3 n}{n} \right) \\ &= \sum_{i=2}^{N(x)-1} n\bar{h}(l_i) \mathbf{E} \left[ \sum_{I:a_I \in L_i} (|I|^2 + 2|I|/(np_2)) \right] + \mathcal{O} \left( \frac{\log^3 n}{n} \right). \end{aligned}$$

Wald's equality yields

$$\mathbf{E} \left[ \sum_{I:a_I \in L_i} |I|^k \right] = \mathbf{E}[N_1(L_i)] \cdot \mathbf{E}[\zeta^k / (np_1)^k] + \mathcal{O}(e^{-\log^2 n/C}),$$

for every  $k > 0$  and for every  $i \leq N(x) - 1$ . Here,  $N_1(L_i)$  is the number of points of  $\mathcal{P}^{1,n}$  lying in  $L_i$  and  $\zeta \sim \mathcal{E}(1)$ . Putting all these estimates together, we get

$$\begin{aligned} \mathbf{E} \left[ \sum_{i=1}^{N(x)} \eta_i^\circ \right] &= \sum_{i=2}^{N(x)-1} n \bar{h}(l_i) \left( \frac{2|L_i|}{np_1} + \frac{2|L_i|}{np_2} \right) + \mathcal{O} \left( \frac{\log^3 n}{n} \right) \\ &= \left( \frac{2}{p_1} + \frac{2}{p_2} \right) \sum_{i=1}^{N(x)} \bar{h}(l_i) |L_i| + \mathcal{O} \left( \frac{\log^3 n}{n} \right). \end{aligned}$$

Since  $l_i$  has been chosen such that  $\bar{h}(l_i) |L_i| = \int_{L_i} \bar{h}(t) dt$ , the last relation implies that  $\mathcal{T}_{11} = \mathcal{O}(n^{-1} \log^3 n)$ .

The advantage of working with  $\eta_i^\circ$ s is that, conditionally to  $\mathcal{E}$ , the random variables  $\eta_{2i}^\circ$ ,  $i = 1, \dots, [N(x)/2]$ , are independent. Indeed, one easily checks that conditionally to  $\mathcal{E}$ ,  $\eta_{2i}^\circ$  depends only on the restrictions of  $\mathcal{P}^{1,n}$  and  $\mathcal{P}^{2,n}$  onto the interval  $[\frac{(4i-1)T}{2N(x)}, \frac{(4i+3)T}{2N(x)}]$ . Since these intervals are disjoint for different values of  $i \in \mathbb{N}$ , the restrictions of Poisson processes  $\mathcal{P}^{k,n}$ ,  $k = 1, 2$ , onto these intervals are independent. Therefore,  $\eta_{2i}^\circ$ ,  $i = 1, \dots, [N(x)/2]$ , form a sequence of random variables that are independent conditionally to  $\mathcal{E}$ . Moreover, they verify  $|\eta_i^\circ| \leq Cn|L_i|^2 = \frac{C \log^4 n}{n}$ .

These features enable us to use the Bernstein inequality in order to bound large deviations of  $\mathcal{T}_{12}$  as follows:

$$\mathbf{P}^{\mathcal{E}} \left( |\mathcal{T}_{12}| \geq x/\sqrt{n} \right) \leq 2 \exp \left( - \frac{x^2}{C(1 + xn^{-1/2} \log^4 n)} \right) \leq 2e^{-x/C},$$

for every  $x > 1$ . Obviously, the same inequality holds true for the term  $\mathcal{T}_{13}$ . These inequalities combined with the bound on the deterministic error term  $\mathcal{T}_{11}$  complete the proof.  $\square$

**Remark 5.** Since  $\mathcal{T}_{12}$  and  $\mathcal{T}_{13}$  are zero mean random variables, conditionally to  $\mathcal{E}$ , and  $\mathcal{E}^c$  has a probability bounded by  $Cne^{-x/C}$ , it follows from the computations of the preceding proof that

$$\mathbf{E}[\mathcal{T}_1] = 2(p_1^{-1} + p_2^{-1}) \int_0^T \sigma_{1,t}^2 \sigma_{2,t}^2 dt + \mathcal{O}(n^{-1} \log^3 n).$$

Similar arguments entail that  $\mathbf{E}[2n\mu_{2,n}] = \mathfrak{c} + \mathcal{O}(n^{-1} \log^3 n)$ .

**Lemma 4.** *There exists a constant  $C$  depending only on  $p_1$  and  $p_2$  such that, for every  $x > 0$ , it holds  $\mathbf{P}(nr_n > x) \leq Cne^{-x/C}$ .*

*Proof.* We start with bounding  $\mathbf{P}(\max_{I \in \Pi_n^1} n|I| > x)$ . According to the Markov inequality, for every  $u > 0$ ,

$$\mathbf{P}(\max_{I \in \Pi_n^1} n|I| > x) \leq e^{-ux} \mathbf{E} \left[ \sum_{I \in \Pi_n^1} e^{un|I|} \right].$$

The last sum can be bounded by the sum of  $N_1$  independent random variables each of which has the same law as  $e^{u\zeta/p_1}$ , with  $\zeta$  being exponential with mean 1. In view of the Wald equation, this yields  $\mathbf{E} \left[ \sum_{I \in \Pi_n^1} e^{un|I|} \right] = np_1 T \mathbf{E}[e^{u\zeta/p_1}]$ . Choosing  $u$  smaller than  $p_1$  and repeating the same arguments for  $\max_{J \in \Pi_n^2} n|J|$ , we obtain the desired result.  $\square$

**Proposition 8.** *In addition to the assumptions of Theorem 3, let  $\sigma_1$ ,  $\sigma_2$  and  $\rho$  be Lipshitz continuous. Then  $\mathbf{E}[\mu_{3,n}] = \frac{3}{2}\kappa n^{-2} + \mathcal{O}(\log^3 n/n^3)$ , where*

$$\kappa = \left( \frac{1}{p_1^2} + \frac{1}{p_2^2} \right) \int_0^T h(t)^3 dt + \frac{3p_1^2 + 2p_1p_2 + 3p_2^2}{p_1^2p_2^2} \int_0^T \sigma_{1,t}^2 \sigma_{2,t}^2 h(t) dt$$

and  $h(t) = \sigma_{1,t}\sigma_{2,t}\rho_t$ .

*Proof.* The assertion of the theorem follows from the following relations:

$$\begin{aligned} \mathbf{E} \left[ \sum_{I \in \Pi_i} v(I)^3 \right] &= \frac{6}{n^2 p_i^2} \int_0^T h(t)^3 dt + \mathcal{O} \left( \frac{\log^3 n}{n^3} \right), \quad i = 1, 2, \\ \mathbf{E} \left[ \sum_{I, J} v(I \cap J)^3 \right] &= \frac{6}{n^2 (p_1 + p_2)^2} \int_0^T h(t)^3 dt + \mathcal{O} \left( \frac{\log^3 n}{n^3} \right), \\ \mathbf{E} \left[ \sum_{I, J} v(I \cap J)^2 v(I) \right] &= \frac{18p_1 + 12p_2}{n^2 p_1 (p_1 + p_2)^2} \int_0^T h(t)^3 dt + \mathcal{O} \left( \frac{\log^3 n}{n^3} \right), \\ \mathbf{E} \left[ \sum_{I, J} v(I \cap J)^2 v(J) \right] &= \frac{18p_2 + 12p_1}{n^2 p_2 (p_1 + p_2)^2} \int_0^T h(t)^3 dt + \mathcal{O} \left( \frac{\log^3 n}{n^3} \right), \\ \mathbf{E} \left[ \sum_{I, J} v(I \cap J) v(I) v(J) \right] &= \frac{4}{n^2 p_2 p_1} \int_0^T h(t)^3 dt + \mathcal{O} \left( \frac{\log^3 n}{n^3} \right), \\ \mathbf{E} \left[ \sum_{I, J} v(I \cup J) v_1(I) v_2(J) \right] &= \frac{6p_1^2 + 4p_1p_2 + 6p_2^2}{n^2 p_1^2 p_2^2} \int_0^T \frac{h(t)^3}{\rho_t^2} dt + \mathcal{O} \left( \frac{\log^3 n}{n^3} \right). \end{aligned}$$

Let us prove in detail the fifth relation. The proofs of the other relations are based on similar arguments and are easier than that of fifth relation.

Using the Lipshitz continuity of the function  $h$ , one can check that  $v(I \cap J)v(I)v(J) = h(a_I)^3 |I| \cdot |J| \cdot |I \cap J| + \mathcal{O}(r_n^3) |I \cap J|$ , where  $a_I$  stands for the left endpoint of the interval  $I$ .

On the one hand, for any  $p > 0$ , in view of Lemma 4, ii), we have

$$\mathbf{E}[r_n^p] = n^{-p} \int_0^\infty \mathbf{P}((nr_n)^p \geq t) dt \leq Cn^{-p} \int_0^\infty (ne^{-t^{1/p}}) \wedge 1 dt. \quad (22)$$

Simple computations show that  $\int_0^\infty (ne^{-t^{1/p}}) \wedge 1 dt = \mathcal{O}(\log^p n)$ . Therefore,

$$\mathbf{E} \left[ \sum_{I, J} r_n^3 |I \cap J| \right] \leq T \mathbf{E}[r_n^3] = \mathcal{O} \left( \frac{\log^3 n}{n^3} \right).$$

On the other hand

$$\mathbf{E} \left[ \sum_{I \in \Pi^1} h(a_I)^3 |I| \sum_{J \in \Pi^2} |J| |I \cap J| \right] = \mathbf{E} \left[ \sum_{I \in \Pi^1} h(a_I)^3 |I| \mathbf{E}^I \left( \sum_{J \in \Pi^2} |J| |I \cap J| \right) \right],$$

where  $\mathbf{E}^I$  stands for the conditional expectation given  $I$ . According to Lemmas 13 and 15,

$$\mathbf{E}^I \left( \sum_{J \in \Pi^2} |J| |I \cap J| \right) = \frac{2|I|}{np_2} - \frac{(1 - e^{-np_2|I|})(e^{-np_2a_I} + e^{-np_2(T-b_I)})}{n^2 p_2^2}.$$

Let us show now that

$$\begin{aligned}\mathcal{T}_1 &:= \frac{2}{np_2} \mathbf{E} \left[ \sum_{I \in \Pi^1} h(a_I)^3 |I|^2 \right] = \frac{4}{n^2 p_1 p_2} \int_0^T h^3(t) dt + \mathcal{O}(n^{-3}), \\ \mathcal{T}_2 &:= \mathbf{E} \left[ \sum_{I \in \Pi^1} h(a_I)^3 |I| \frac{(1 - e^{-np_2|I|})e^{-np_2 a_I}}{n^2 p_2^2} \right] = \mathcal{O}(n^{-3}), \\ \mathcal{T}_3 &:= \mathbf{E} \left[ \sum_{I \in \Pi^1} h(a_I)^3 |I| \frac{(1 - e^{-np_2|I|})e^{-np_2(T-b_I)}}{n^2 p_2^2} \right] = \mathcal{O}(n^{-3}).\end{aligned}$$

To this end, we use the characterization of a Poisson process as a renewal process with exponential waiting times. Let  $(\zeta_k, k \geq 1)$  be a sequence of i.i.d. random variables of exponential law with mean  $1/(np_1)$ . Then  $N_1, S^i$  can be defined by  $N_1 = \inf\{k \geq 1 : \zeta_1 + \dots + \zeta_k \geq T\}$  and  $S^i = (\zeta_1 + \dots + \zeta_i) \wedge T$  for  $i = 1, \dots, N_1$ . In this notation,

$$\begin{aligned}\mathcal{T}_1 &= \frac{2}{np_2} \mathbf{E} \left[ \sum_{i=1}^{N_1-1} h(S^i)^3 \zeta_{i+1}^2 \right] + \mathcal{O}(n^{-3}), \\ |\mathcal{T}_2| &\leq \|h\|_\infty^3 \mathbf{E} \left[ \sum_{i=1}^{N_1-1} \frac{\zeta_{i+1} e^{-np_2 S^i}}{n^2 p_2^2} \right] + \mathcal{O}(n^{-3}),\end{aligned}$$

where  $\|h\|_\infty = \max_{t \in [0, T]} |h(t)|$ . Remark that  $N_1$  is a stopping time with respect to the filtration  $\mathcal{F}_k = \sigma(\zeta_1, \dots, \zeta_k)$ ,  $k \geq 1$ . It is easily seen that

$$\begin{aligned}M_k &= \sum_{i=1}^{k-1} h(S^i)^3 (\zeta_{i+1}^2 - \mathbf{E}[\zeta_{i+1}^2]), \\ M'_k &= \sum_{i=1}^{k-1} (\zeta_{i+1} - \mathbf{E}[\zeta_1]) e^{-np_2 S^i}\end{aligned}$$

are  $\mathcal{F}_k$ -martingales for which the conditions of the optional stopping theorem are fulfilled. Therefore

$$\begin{aligned}\mathcal{T}_1 &= \frac{2}{np_2} \mathbf{E}[\zeta_1^2] \mathbf{E} \left[ \sum_{i=1}^{N_1-1} h(S^i)^3 \right] + \mathcal{O}(n^{-3}), \\ \mathcal{T}_2 &\leq \frac{\|h\|_\infty^3}{n^2 p_2^2} \mathbf{E}[\zeta_1] \mathbf{E} \left[ \sum_{i=1}^{N_1-1} e^{-np_2 S^i} \right] + \mathcal{O}(n^{-3}).\end{aligned}$$

These relations imply that

$$\begin{aligned}\mathcal{T}_1 &= \frac{4}{n^2 p_1 p_2} \int_0^T h(t)^3 dt + \mathcal{O}(n^{-3}), \\ |\mathcal{T}_2| &\leq \frac{\|h\|_\infty^3}{n^2 p_2^2} \int_0^T e^{-np_2 t} dt + \mathcal{O}(n^{-3}) = \mathcal{O}(n^{-3}).\end{aligned}$$

In the above inequalities we used the fact that for any integrable function  $f$  on  $[0, T]$ , it holds  $\mathbf{E}[\sum_{i=1}^{N_1-1} f(S^i)] = np_1 \int_0^T f(t) dt$ .

The term  $\mathcal{T}_3$  can be bounded in the same way as  $\mathcal{T}_2$  by using the fact that if  $\{t_1, \dots, t_N\}$  is a realization of a homogeneous Poisson point process in  $[0, T]$ , then  $\{T - t_1, \dots, T - t_N\}$

can be seen as a realization of the same Poisson point process. This completes the proof of the proposition.  $\square$

Prior to stating the main result of this section, let us recall several notation. We denote by  $h_t$  the function  $\sigma_{1,t}\sigma_{2,t}\rho_t$  and by  $x_+$  the positive part of a real  $x$ . Finally, we write  $g_1(z) \propto g_2(z)$  if for some  $C_g \in \mathbb{R}$  the equality  $g_1(z) = C_g g_2(z)$  holds for every  $z$ .

**Theorem 4.** *Let the sampling scheme be generated by two independent Poisson processes with intensities  $np_1$  and  $np_2$ , independent of the driving BM  $\mathbf{B}$ . If the functions  $\sigma_1$ ,  $\sigma_2$  and  $\rho$  are Lipschitz continuous then, for every  $a \in (\frac{3}{4}, 1)$ , it holds*

$$\sup_{f \in \mathcal{E}(M, \gamma) \cap \mathcal{E}^0(\mathbf{C}, \eta, r_0, \mathbf{c}^*)} \left| \mathbf{E}[f(\sqrt{n}(\hat{\theta}_n - \theta))] - \int_{\mathbb{R}} f(z) p_n^\circ(z) dz \right| = O(n^{1-2a}), \quad (23)$$

where

$$p_n^\circ(z) \propto \frac{1}{\sqrt{2\pi\mathbf{c}}} \left[ 1 + \frac{2\kappa(z^3 - 3\mathbf{c}z)}{\sqrt{n}\mathbf{c}^3} \right]_+ e^{-z^2/(2\mathbf{c})}$$

is a probability density with

$$\begin{aligned} \mathbf{c} &= \left( \frac{2}{p_1} + \frac{2}{p_2} \right) \int_0^T \sigma_{1,t}^2 \sigma_{2,t}^2 (1 + \rho_t^2) dt - \frac{2}{p_1 + p_2} \int_0^T (\sigma_{1,t} \sigma_{2,t} \rho_t)^2 dt, \\ \kappa &= \left( \frac{1}{p_1^2} + \frac{1}{p_2^2} \right) \int_0^T h(t)^3 dt + \frac{3p_1^2 + 2p_1 p_2 + 3p_2^2}{p_1^2 p_2^2} \int_0^T \sigma_{1,t}^2 \sigma_{2,t}^2 h(t) dt. \end{aligned}$$

*Proof.* Propositions 7 and 8, in conjunction with Lemma 4 and Remark 5 imply that Theorem 2 can be applied with  $b_n = 1/n$  and any  $a < 1$ . This yields the desired result.  $\square$

## 6. STOCHASTIC DECOMPOSITION FOR $\hat{\theta}_n$ IN A MODEL WITH DRIFT TERMS

So far we have considered a Gaussian system  $(X_{1,t} - X_{1,0}, X_{2,t} - X_{2,0})$  as the underlying model and essentially finite dimensional Gaussian calculus served as a tool. In this section, we will treat a system that has random drift terms. It will be seen that the principal part of the estimator is the same as in the case without drifts. Thus, the contribution of the principal part to the asymptotic expansion of the estimator has already been assessed in the previous section.

Beyond being a useful tool for deriving asymptotic expansions of the distribution of  $\hat{\theta}_n$ , the stochastic decomposition of the HY-estimator that we obtain below bridges the problem of estimating the covariance and that of signal detection in Gaussian white noise. The latter problem has been extensively studied in the statistical literature and we believe that the methodology developed for the problem of signal detection may be of interest for our problem.

To state the main result of this section, let us recall that we deal with processes  $X_1$  and  $X_2$  given by

$$\begin{cases} dX_{1,t} = \beta_{1,t} dt + \sigma_{1,t} dB_{1,t}, & t \in [0, T], \\ dX_{2,t} = \beta_{2,t} dt + \sigma_{2,t} dB_{2,t}, & t \in [0, T], \end{cases}$$

where  $\beta_{i,t}$  are progressively measurable processes and assumed to be unknown to the observer. We will assume that these drift processes admit the following stochastic decompositions:

$$d\beta_{i,t} = \beta_{i,t}^{[0]} dt + \beta_{i1,t}^{[1]} dB_{1,t} + \beta_{i2,t}^{[1]} dB_{2,t}, \quad i = 1, 2,$$

where  $\beta_i^{[0]}$ ,  $\beta_{ij}^{[1]}$ ,  $i, j = 1, 2$  are progressively measurable processes with respect to the filtration  $\{\sigma(\mathbf{B}_s, s \leq t)\}_{t \in [0, T]}$ .

In this section, we will separate the assumptions on the sampling scheme from those on  $\rho$  and on the drifts and volatilities of  $X_1$  and  $X_2$ . For this reason, let us introduce the following measures on  $([0, T]^2, \mathcal{B}_{[0, T]^2})$ :

$$\begin{aligned} \mathcal{V}_n^I(\cdot) &= b_n^{-1} |\cdot \cap \{\cup_I I \times I\}|, & \mathcal{V}_n^J(\cdot) &= b_n^{-1} |\cdot \cap \{\cup_J J \times J\}|, \\ \mathcal{V}_n^{I \cap J}(\cdot) &= b_n^{-1} |\cdot \cap \{\cup_{I, J} (I \cap J) \times (I \cap J)\}|, \\ \mathcal{V}_n^{I, J}(\cdot) &= b_n^{-1} \sum_{I, J} K_{IJ} |\cdot \cap (I \times J)|. \end{aligned}$$

Note that these measures depend on the sampling schemes and, therefore, they are random if the sampling schemes are random. Similarly, let  $\mathcal{V}_n^{I, I', J}(\cdot) = b_n^{-2} |\cdot \cap \{\cup_J J \times I(J) \times I(J)\}|$ ,  $\mathcal{V}_n^{I, J, J'}(\cdot) = b_n^{-2} |\cdot \cap \{\cup_I I \times J(I) \times J(I)\}|$  and  $\mathcal{V}_n^{J(I), I(J), J \cap I}(\cdot) = b_n^{-2} |\cdot \cap \{\cup_{I, J} J(I) \times I(J) \times J \cap I\}|$  be (random) measures defined on  $([0, T]^3, \mathcal{B}_{[0, T]^3})$ .

**Assumption P1:** The random measures  $\mathcal{V}_n^I$ ,  $\mathcal{V}_n^J$ ,  $\mathcal{V}_n^{I \cap J}$  and  $\mathcal{V}_n^{I, J}$  converge weakly to some deterministic measures  $\mathcal{V}^I$ ,  $\mathcal{V}^J$ ,  $\mathcal{V}^{I \cap J}$  and  $\mathcal{V}^{I, J}$  in probability, as  $n \rightarrow \infty$ . These measures are concentrated on the diagonal  $\mathcal{D}_T^2 = \{(s, t) \in [0, T]^2 : s = t\}$  and absolutely continuous w.r.t. the Lebesgue measure on the line.

**Assumption P2:** As  $n \rightarrow \infty$ , the random measures  $\mathcal{V}_n^{I, I', J}$ ,  $\mathcal{V}_n^{I, J, J'}$  and  $\mathcal{V}_n^{J(I), I(J), I \cap J}$  converge weakly to some deterministic measures  $\mathcal{V}^{I, I', J}$ ,  $\mathcal{V}^{I, J, J'}$  and  $\mathcal{V}^{J(I), I(J), I \cap J}$  in probability. These measures are concentrated on the diagonal  $\mathcal{D}_t^3 = \{(s, t, u) \in [0, T]^3 : s = t = u\}$  and absolutely continuous w.r.t. the Lebesgue measure on the line.

The weak convergence of  $\mathcal{V}_n^I$  to  $\mathcal{V}^I$  in probability should be understood as follows: for every continuous function  $\varphi : [0, T]^2 \rightarrow \mathbb{R}$ , the sequence of random variables  $\int_{[0, T]^2} \varphi d\mathcal{V}_n^I$  converges in probability to  $\int_{[0, T]^2} \varphi d\mathcal{V}^I$  as  $n$  tends to infinity.

Recall that according to our assumptions  $\Pi$  is independent of  $\mathbf{B}$ , where  $\Pi$  is the collection of random intervals  $I^i := (S^{i-1} \wedge T, S^i \wedge T]$ ,  $J^j := (T^{j-1} \wedge T, T^j \wedge T]$  with  $i = 1, \dots, N_1$  and  $j = 1, \dots, N_2$ . In what follows, the following notation will be used: for two functions  $f, g : [0, T] \rightarrow \mathbb{R}$ , we denote by  $f \cdot g$  the function  $t \mapsto \int_0^t f_s dg_s$  and we often write  $I$  or  $J$  instead of  $\mathbf{1}_I$  or  $\mathbf{1}_J$ . Thus the estimator  $\hat{\theta}_n$  can be rewritten as

$$\hat{\theta}_n = \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} K_{ij} \{I^i \cdot X_1\}_T \times \{J^j \cdot X_2\}_T.$$

We want to derive an asymptotic expansion of this estimator by applying Theorem 5. The role of  $\mathcal{X}_n$  will play the sequence

$$M_T^n = b_n^{-1/2} \left( \sum_{i, j} K_{ij} \{(I^i \sigma_1) \cdot B_1\}_T \{(J^j \sigma_2) \cdot B_2\}_T - \theta \right),$$

the asymptotic expansion of which has already been obtained in preceding sections. Note that using the Itô formula, one can write  $M^n$  in the stochastic integral form as follows:

$$M^n = \mathbb{H}^{1,n} \cdot B_1 + \mathbb{H}^{2,n} \cdot B_2, \quad (24)$$

where

$$\mathbb{H}^{1,n} = \sum_{I,J} b_n^{-1/2} K_{IJ} (J\sigma_2 \cdot B_2) I\sigma_1, \quad \mathbb{H}^{2,n} = \sum_{I,J} b_n^{-1/2} K_{IJ} (I\sigma_1 \cdot B_1) J\sigma_2,$$

The role of  $s_n$  is played by  $b_n^{1/2}$ .

**Lemma 5.** *Assume that  $\sigma_1, \sigma_2$  and  $\rho$  are bounded and  $\beta_{ij}^{[\ell-1]} s$  are bounded in  $L^4$  uniformly in  $[0, T]$  for every  $i, j, \ell \in \{1, 2\}$ . If  $r_n^3 = o_p(b_n^2)$ , then*

$$b_n^{-1/2}(\hat{\theta}_n - \theta) = M_T^n + b_n^{1/2}(N_T^n + A_T^n) + o_p(b_n^{1/2}),$$

where  $dN_t^n = \mathbb{G}_t^{1,n} dB_{1,t} + \mathbb{G}_t^{2,n} dB_{2,t}$  is a local martingale with

$$\begin{aligned} \mathbb{G}^{1,n} &= b_n^{-1} \sum_{i,j} K_{ij} \{((J^j \beta_2) \cdot t)(I^i \sigma_1)\} + b_n^{-1} \sum_{i,j} K_{ij} \{(T^j - T^{j-1} \vee \cdot)_+ I^i \sigma_1 \beta_{2,S^{i-1}}\}, \\ \mathbb{G}^{2,n} &= b_n^{-1} \sum_{i,j} K_{ij} \{((I^i \beta_1) \cdot t)(J^j \sigma_2)\} + b_n^{-1} \sum_{i,j} K_{ij} \{(S^i - S^{i-1} \vee \cdot)_+ J^j \sigma_2 \beta_{1,T^{j-1}}\}, \end{aligned}$$

and  $A^n$  is a bounded variation process defined by

$$\begin{aligned} A^n &= b_n^{-1} \sum_{i,j} K_{ij} \left\{ J^j \{ [I^i \sigma_1 (\beta_{21}^{[1]} + \beta_{22}^{[1]} \rho)] \cdot s \} + I^i \{ [J^j \sigma_2 (\beta_{11}^{[1]} \rho + \beta_{12}^{[1]})] \cdot s \} \right\} \cdot t \\ &\quad + b_n^{-1} \sum_{i,j} K_{ij} \{ (I^i \beta_1) \cdot t \} \times \{ (J^j \beta_2) \cdot t \}. \end{aligned}$$

*Proof.* Let us denote by  $\Phi_n^1$  the difference  $b_n^{-1/2}(\hat{\theta}_n - \theta) - M_T^n$  and write it in the form

$$\Phi_n^1 = b_n^{-1/2}(\Phi_n^2 + \Phi_n^3),$$

where

$$\begin{aligned} \Phi_n^2 &= \sum_{I,J} K_{IJ} \left( \{ (I\beta_1) \cdot t \}_T \times \{ (J\sigma_2) \cdot B_2 \}_T + \{ (J\beta_2) \cdot t \}_T \times \{ (I\sigma_1) \cdot B_1 \}_T \right), \\ \Phi_n^3 &= \sum_{I,J} K_{IJ} \{ (I\beta_1) \cdot t \}_T \times \{ (J\beta_2) \cdot t \}_T \end{aligned}$$

Since we will be interested in applying martingale limit theorems, it is convenient to decompose  $\Phi_n^\ell$ s in a sum of a martingale and a bounded variation process. This is achieved by the Itô formula,

$$\begin{aligned} \Phi_n^2 &= \sum_{I,J} K_{IJ} \left( \left\{ \{ ((I\beta_1) \cdot t)(J\sigma_2) \} \cdot B_2 \right\}_T + \left\{ \{ ((J\beta_2) \cdot t)(I\sigma_1) \} \cdot B_1 \right\}_T \right) \\ &\quad + \sum_{I,J} K_{IJ} \left( \left\{ \{ ((J\sigma_2) \cdot B_2)(I\beta_1) \} \cdot t \right\}_T + \left\{ \{ ((I\sigma_1) \cdot B_1)(J\beta_2) \} \cdot t \right\}_T \right). \end{aligned}$$

The last two terms in this expression need some further analysis. Let us denote

$$\Phi_n^{21} = \sum_{I,J} K_{IJ} \left\{ \{ ((J\sigma_2) \cdot B_2)(I\beta_1) \} \cdot t \right\}_T.$$



Since  $((J^j \sigma_2) \cdot B_2)_s = 0$  for  $s \in (0, T^{j-1})$ , we see that

$$\begin{aligned}
\Phi_n^{21} &= \sum_{i,j} K_{ij} \left\{ \left( (J^j \sigma_2) \cdot B_2 \right) I^i [\beta_{1,T^{j-1}} + (1_{(T^{j-1}, \infty)} \beta_1^{[0]}) \cdot t \right. \\
&\quad \left. + (1_{(T^{j-1}, \infty)} \beta_{11}^{[1]}) \cdot B_1 + (1_{(T^{j-1}, \infty)} \beta_{12}^{[1]}) \cdot B_2 \right\} \cdot t \Big\}_T \\
&= \sum_{i,j} K_{ij} \left( \left\{ (J^j \sigma_2) \cdot B_2 \right\} I^i \beta_{1,T^{j-1}} \right) \cdot t_T + \sum_{I,J} K_{IJ} \left( I \left\{ (J \sigma_2 \beta_{11}^{[1]}) \cdot \langle B_2, B_1 \rangle \right\} \right) \cdot t_T \\
&\quad + \sum_{I,J} K_{IJ} \left( I \left\{ (J \sigma_2 \beta_{12}^{[1]}) \cdot \langle B_2, B_2 \rangle \right\} \right) \cdot t_T + o_p(b_n), \tag{25}
\end{aligned}$$

Let us explain why the remainder term in the last display is  $o_P(1)$ . In fact, the remainder contains five summands which can all be treated in a similar way. Let us do it for one of them, which has the form  $\Psi_n = \sum_{I,J} K_{IJ} \left\{ \left( \left\{ (J \sigma_2) \cdot B_2 \right\} \beta_{11}^{[1]} \right) \cdot B_1 \right\} I \Big\}_T$ . We first use that

$$\begin{aligned}
\Psi_n &= \sum_I \left\{ \left( \left\{ (J(I) \sigma_2) \cdot B_2 \right\} \beta_{11}^{[1]} \right) \cdot B_1 \right\} I \Big\}_T \\
&= \int_0^T \left( \int_0^s \left[ \int_s^T \sum_I \mathbf{1}_I(t) \mathbf{1}_{J(I)}(u) dt \right] \sigma_{2,u} dB_{2,u} \right) \beta_{11,s}^{[1]} dB_{1,s}.
\end{aligned}$$

Then, by the Cauchy-Schwarz inequality and the martingale property of the stochastic integral, we get

$$\begin{aligned}
\mathbf{E}^\Pi[\Psi_n^2] &\leq \int_0^T \left( \int_0^s \left[ \int_s^T \sum_I \mathbf{1}_I(t) \mathbf{1}_{J(I)}(u) dt \right]^2 \sigma_{2,u}^2 du \right) \left( \mathbf{E}^\Pi[(\beta_{11,s}^{[1]})^4] \right)^{1/2} ds \\
&\leq C \int_0^T \int_0^s \left[ \sum_I |I| \mathbf{1}_{J(I)}(u) \mathbf{1}_{J(I)}(s) \right]^2 du ds \leq C \sum_{I,I'} |I| |I'| |J(I) \cap J(I')|^2 \leq Cr_n^3
\end{aligned}$$

under the assumption that  $\max_{t \in [0, T]} \mathbf{E}^\Pi[(\beta_{11,t}^{[1]})^4]$  and  $\max_{t \in [0, T]} \sigma_{2,t}$  are finite. Now, interchanging the order of integrations, the first summand in the RHS of (25) can be rewritten as follows

$$\left\{ \left( (J \sigma_2) \cdot B_2 \right) I \beta_{1,T^{j-1}} \right\} \cdot t_T = \left\{ (S^i - S^{i-1} \vee \cdot)_+ J \sigma_2 \beta_{1,T^{j-1}} \right\} \cdot B_2 \Big\}_T. \tag{26}$$

Using the same kind of arguments, one can check that the term  $\Phi_n^{22} = \Phi_n^2 - \Phi_n^{21}$  admits the representation

$$\begin{aligned}
\Phi_n^{22} &:= \sum_{I,J} K_{IJ} \left\{ \left( (I \sigma_1) \cdot B_1 \right) (J \beta_2) \right\} \cdot t \Big\}_T \\
&= \sum_{i,j} K_{ij} \left( \left\{ (T^j - T^{j-1} \vee \cdot)_+ I \sigma_1 \beta_{2,S^{i-1}} \right\} \cdot B_1 \right) \Big\}_T \\
&\quad + \sum_{I,J} K_{IJ} \left( J \left\{ (I \sigma_1 \beta_{22}^{[1]}) \cdot \langle B_2, B_1 \rangle + (I \sigma_1 \beta_{21}^{[1]}) \cdot \langle B_1, B_1 \rangle \right\} \right) \cdot t_T + o_p(b_n). \tag{27}
\end{aligned}$$

Combining (25)-(27) and using that  $\langle B_1, B_1 \rangle_t = \langle B_2, B_2 \rangle_t = t$ ,  $\langle B_2, B_1 \rangle_t = \int_0^t \rho_s ds$  we get the desired result.  $\square$

Lemma 5 provides a stochastic decomposition of the HY-estimator with a RHS depending on  $n$ . Under the assumptions P1 and P2 of the convergence of random measures

associated to the sampling scheme, it is possible to obtain refinement of this result with a RHS depending on  $n$  exclusively through  $b_n$ . To this end, limit theorems for martingales will be used. An important step for proving limit theorems for martingales is the computation of the limits of their quadratic variations and covariations, which will be treated below.

**6.1. Convergence of quadratic variations and covariations.** In view of (24), for  $\nu = 1, 2$ , we have

$$\begin{aligned} \langle M^n, B_\nu \rangle &= \mathbb{H}^{1,n} \cdot \langle B_1, B_\nu \rangle + \mathbb{H}^{2,n} \cdot \langle B_2, B_\nu \rangle \\ &= \sum_{I,J} b_n^{-1/2} K_{IJ} \left[ \{(J\sigma_2 \cdot B_2)I\sigma_1\} \cdot \langle B_1, B_\nu \rangle + \{(I\sigma_1 \cdot B_1)J\sigma_2\} \cdot \langle B_2, B_\nu \rangle \right]. \end{aligned}$$

**Lemma 6.** *If  $\sigma_1, \sigma_2$  and  $\rho$  are bounded in  $[0, T]$  and  $r_n^2 = o_p(b_n)$ , then*

$$\begin{aligned} \sup_{\nu=1,2} \left| \sum_{I,J} b_n^{-1/2} K_{IJ} \left( \{(J\sigma_2 \cdot B_2)I\sigma_1\} \cdot \langle B_1, B_\nu \rangle \right)_t \right| \xrightarrow[n \rightarrow \infty]{p} 0, \\ \sup_{\nu=1,2} \left| \sum_{I,J} b_n^{-1/2} K_{IJ} \left( \{(I\sigma_1 \cdot B_1)J\sigma_2\} \cdot \langle B_2, B_\nu \rangle \right)_t \right| \xrightarrow[n \rightarrow \infty]{p} 0, \end{aligned}$$

for every  $t \in [0, T]$ . As a consequence, for every  $t \in [0, T]$ ,  $\max_{\nu=1,2} |\langle M^n, B_\nu \rangle_t|$  tends to zero in probability as  $n \rightarrow \infty$ .

*Proof.* We will prove only the first relation, the proof of the second being quite similar. Consider the case  $\nu = 1$ , the case  $\nu = 2$  can be treated similarly in view of the relation  $\langle B_1, B_2 \rangle_t = \int_0^t \rho_s ds$  and the boundedness of  $\rho$ . To simplify subsequent formulae, let us denote  $\xi^{[11]} = b_n^{-1/2} \sum_{ij} K_{IJ} \{(J\sigma_2 \cdot B_2)I\sigma_1\} \cdot \langle B_1, B_1 \rangle$ . In other words,  $\xi^{[11]}$  is a random process indexed by  $t \in [0, T]$  defined by

$$\begin{aligned} b_n^{1/2} \xi_t^{[11]} &= \sum_{I,J} K_{IJ} \int_0^t \mathbf{1}_I(u) \sigma_{1,u} \int_0^u \mathbf{1}_J(s) \sigma_{2,s} dB_{2,s} du \\ &= \sum_{I,J} K_{IJ} \int_0^t \mathbf{1}_J(s) \sigma_{2,s} \int_s^t \mathbf{1}_I(u) \sigma_{1,u} du dB_{2,s} \\ &= \int_0^t \sum_J \mathbf{1}_J(s) \sigma_{2,s} \int_s^t \mathbf{1}_{I(J)}(u) \sigma_{1,u} du dB_{2,s}. \end{aligned}$$

The latter expression implies that conditionally to  $\Pi_n$ ,  $\xi^{[11]}$  is a Gaussian process with zero mean. Moreover,

$$\begin{aligned} \mathbf{E}^\Pi [(\xi_t^{[11]})^2] &= b_n^{-1} \sum_J \int_0^t \mathbf{1}_J(s) \sigma_{2,s}^2 \left( \int_s^t \mathbf{1}_{I(J)}(u) \sigma_{1,u} du \right)^2 ds \\ &\leq b_n^{-1} \|\sigma_1^2\|_\infty \|\sigma_2^2\|_\infty \sum_J |J| |I(J)|^2 \leq C b_n^{-1} r_n^2, \end{aligned}$$

where  $C$  is a positive constant. This yields the desired result.  $\square$

We study now the behavior of the quadratic variation

$$\langle M^n, M^n \rangle_t = \sum_{c,d=1}^2 (\mathbb{H}^{c,n} \mathbb{H}^{d,n}) \cdot \langle B_c, B_d \rangle_t \quad (28)$$

as  $n$  tends to infinity. First, we note that

$$\begin{aligned} \mathbb{H}^{1,n} \mathbb{H}^{2,n} &= \sum_{i,j,i',j'} b_n^{-1} K_{ij} K_{i'j'} (J^j \sigma_2 \cdot B_2) I^i \sigma_1 (I^{i'} \sigma_1 \cdot B_1) J^{j'} \sigma_2 \\ &= \sum_{i,j,i',j'} b_n^{-1} K_{ij} K_{i'j'} (J^j \sigma_2 \cdot B_2) J^{j'} \sigma_2 1_{\{j \leq j'\}} (I^{i'} \sigma_1 \cdot B_1) I^i \sigma_1 1_{\{i' \leq i\}} \end{aligned}$$

Denote by  $R^n(i, i', j, j')$  the summand on the right-hand side of the last equation. This term is different from zero only if the conditions  $I^i \cap J^j \neq \emptyset$ ,  $I^{i'} \cap J^{j'} \neq \emptyset$ ,  $I^{i'} \cap J^j \neq \emptyset$ ,  $j \leq j'$  and  $i' \leq i$  are fulfilled. If  $i' < i$ , then these conditions are fulfilled only if  $j = j'$ . Similarly, the terms with  $j < j'$  are non-zero only if  $i = i'$ . This leads to

$$\begin{aligned} \mathbb{H}^{1,n} \mathbb{H}^{2,n} &= \sum_{i,j,j': j \leq j'} b_n^{-1} K_{ij} K_{ij'} (J^j \sigma_2 \cdot B_2) J^{j'} \sigma_2 (I^i \sigma_1 \cdot B_1) I \sigma_1 \\ &\quad + \sum_{i',j,i: i' \leq i} b_n^{-1} K_{ij} K_{i'j} (J^j \sigma_2 \cdot B_2) J^j \sigma_2 (I^{i'} \sigma_1 \cdot B_1) I^i \sigma_1 \\ &\quad - \sum_{I,J} b_n^{-1} K_{ij} (J^j \sigma_2 \cdot B_2) J^j \sigma_2 (I^i \sigma_1 \cdot B_1) I^i \sigma_1. \end{aligned}$$

Sum them up in  $j'$  and in  $i$  respectively and use

$$\begin{aligned} (J^j \sigma_2 \cdot B_2) I \sum_{j': j \leq j'} K_{ij'} J^{j'} &= (J^j \sigma_2 \cdot B_2) I \mathbf{1}_{[T^{j-1}, T]} = (J^j \sigma_2 \cdot B_2) I, \\ (I^{i'} \sigma_1 \cdot B_1) J \sum_{i: i' \leq i} K_{ij} I^i &= (I^{i'} \sigma_1 \cdot B_1) J \mathbf{1}_{[S^{i-1}, T]} = (I^{i'} \sigma_1 \cdot B_1) J. \end{aligned}$$

to obtain

$$\mathbb{H}^{1,n} \mathbb{H}^{2,n} = b_n^{-1} \sum_{I,J} \sigma_1 \sigma_2 K_{IJ} (J \sigma_2 \cdot B_2) (I \sigma_1 \cdot B_1) (I + J - IJ).$$

The last equality implies that

$$\mathbb{H}^{1,n} \mathbb{H}^{2,n} \cdot \langle B_1, B_2 \rangle_t = \int_0^t \sigma_{1,s} \sigma_{2,s} \sum_{I,J} \tilde{K}_{IJ}^n(s) (J \sigma_2 \cdot B_2)_s (I \sigma_1 \cdot B_1)_s d\langle B_1, B_2 \rangle_s,$$

where  $\tilde{K}_{IJ}^n(t) = b_n^{-1} K_{IJ} (I_t + J_t - I_t J_t)$ .

**Lemma 7.** *Assume that  $r_n^3 = o_p(b_n^2)$  and the functions  $\sigma_1$ ,  $\sigma_2$  and  $\rho$  are continuous. If Assumption P1 is fulfilled then, for any  $t \in [0, T]$ ,*

$$\begin{aligned} \int_0^t \mathbb{H}_s^{1,n} \mathbb{H}_s^{2,n} d\langle B_1, B_2 \rangle_s &\xrightarrow[n \rightarrow \infty]{p} \frac{1}{2} \int_0^t h_s^2 \{ \mathcal{V}^I(ds) + \mathcal{V}^J(ds) - \mathcal{V}^{I \cap J}(ds) \}, \\ \int_0^t (\mathbb{H}_s^{1,n})^2 d\langle B_1, B_1 \rangle_s + \int_0^t (\mathbb{H}_s^{2,n})^2 d\langle B_2, B_2 \rangle_s &\xrightarrow[n \rightarrow \infty]{p} \int_0^t \sigma_{1,s}^2 \sigma_{2,s}^2 \mathcal{V}^{I,J}(ds) \end{aligned}$$

and consequently

$$\langle M^n, M^n \rangle_t \xrightarrow[n \rightarrow \infty]{p} \int_0^t h_s^2 \{ \mathcal{V}^I(ds) + \mathcal{V}^J(ds) - \mathcal{V}^{I \cap J}(ds) \} + \int_0^t \sigma_{1,s}^2 \sigma_{2,s}^2 \mathcal{V}^{I,J}(ds).$$

*Proof.* One easily checks that

$$\int_0^t \mathbb{H}_s^{1,n} \mathbb{H}_s^{2,n} d\langle B_1, B_2 \rangle_s = \int_0^t \sum_{I,J} h_s \tilde{K}_{ij}^n(s) (J\sigma_2 \cdot B_2)_s (I\sigma_1 \cdot B_1)_s ds. \quad (29)$$

To prove the convergence of this expression, we apply the Itô formula to the product  $(J\sigma_2 \cdot B_2)_s (I\sigma_1 \cdot B_1)_s$ :

$$(J\sigma_2 \cdot B_2)_s (I\sigma_1 \cdot B_1)_s = \{(J\sigma_2 \cdot B_2)I\sigma_1 \cdot B_1\}_s + \{(I\sigma_1 \cdot B_1)J\sigma_2 \cdot B_2\}_s + \{(IJh) \cdot t\}_s.$$

One can show that the contribution of the first two terms is asymptotically negligible, that is

$$\int_0^t \sum_{I,J} h_s \tilde{K}_{ij}^n(s) (\{(J\sigma_2 \cdot B_2)I\sigma_1 \cdot B_1\}_s + \{(I\sigma_1 \cdot B_1)J\sigma_2 \cdot B_2\}_s) ds \xrightarrow[n \rightarrow \infty]{p} 0, \quad (30)$$

Thus, the main term is

$$\int_0^t \sum_{I,J} h_s \tilde{K}_{ij}^n(s) \{(IJh) \cdot t\}_s ds. \quad (31)$$

To prove (30), we show the convergence in  $L^2$ . More rigorously, using the notation  $\tilde{K}_{IJ}^n(s) = \int_s^t \tilde{K}_{IJ}^n(u) h_u ds$  and interchanging the order of integrals, we get

$$\begin{aligned} \mathbf{E}^\Pi \left( \int_0^t \sum_{I,J} h_s \tilde{K}_{ij}^n(s) \{(J\sigma_2 \cdot B_2)I\sigma_1 \cdot B_1\}_s ds \right)^2 &= \mathbf{E}^\Pi \left( \left\{ \sum_{I,J} \tilde{K}_{ij}^n(J\sigma_2 \cdot B_2)I\sigma_1 \cdot B_1 \right\}_t \right)^2 \\ &\leq \int_0^t \mathbf{E}^\Pi \left[ \left( \sum_{I,J} \tilde{K}_{IJ}^n(u) (J\sigma_2 \cdot B_2)_u I_u \sigma_{1,u} \right)^2 \right] du \\ &= \int_0^t \int_0^u \left( \sum_{I,J} \tilde{K}_{IJ}^n(u) I_u \sigma_{1,u} J_v \sigma_{2,v} \right)^2 dv du \\ &\leq C b_n^{-2} \int_0^T \int_0^T r_n^2 \sum_{I,J} \left( K_{IJ} \mathbf{1}_I(u) \mathbf{1}_J(v) \right) dv du \leq C b_n^{-2} r_n^3. \end{aligned}$$

Let us show now that the term (31) converges in probability. Simple algebra allows us to rewrite that term in the form

$$\frac{1}{2b_n} \sum_I \left( \int_0^t h_s \mathbf{1}_I(s) ds \right)^2 + \frac{1}{2b_n} \sum_J \left( \int_0^t h_s \mathbf{1}_J(s) ds \right)^2 - \frac{1}{2b_n} \sum_{I,J} \left( \int_0^t h_s \mathbf{1}_{I \cap J}(s) ds \right)^2,$$

which in turn is nothing else but

$$\int_{[0,t]^2} h_s h_{s'} \mathbf{1}_{\{s \vee s' \leq t\}} \{ \mathcal{V}_n^I(ds, ds') + \mathcal{V}_n^J(ds, ds') - \mathcal{V}_n^{I \cap J}(ds, ds') \}.$$

The weak convergence of measures stated in Assumption P1 completes the proof of the first assertion of the lemma. The proof of the second assertion is quite similar and therefore is omitted.  $\square$

**Lemma 8.** *If  $\sigma_1, \sigma_2$  and  $\rho$  are bounded,  $\sup_{t \in [0, T]} \mathbf{E}[\beta_{i,t}^2] < \infty$ ,  $i = 1, 2$  and  $r_n^4 = o_p(b_n^3)$  as  $n \rightarrow \infty$ , then for any  $t \in [0, T]$  the sequence of random variables  $\langle M^n, N^n \rangle_t$  tends in probability to zero as  $n$  tends to infinity.*

*Proof.* Using the representations of  $M^n$  and  $N^n$  as stochastic integrals, we get

$$\langle M^n, N^n \rangle_t = \int_0^t (\mathbb{H}_s^{1,n} \mathbb{G}_s^{1,n} + \mathbb{H}_s^{1,n} \mathbb{G}_s^{2,n} \rho_s + \mathbb{H}_s^{2,n} \mathbb{G}_s^{1,n} \rho_s + \mathbb{H}_s^{2,n} \mathbb{G}_s^{2,n}) ds. \quad (32)$$

Let us denote by  $\mathbb{G}^{11,n}$  the first summand  $b_n^{-1} \sum_{I,J} K_{IJ} \{((J\beta_2) \cdot t)(I\sigma_1)\}$  in  $\mathbb{G}^{1,n}$  and let us show that  $\int_0^t \mathbb{H}_s^{1,n} \mathbb{G}_s^{11,n} ds$  tends to zero in probability as  $n \rightarrow \infty$ . Simple algebra yields

$$\begin{aligned} \int_0^t \mathbb{H}_s^{1,n} \mathbb{G}_s^{11,n} ds &= b_n^{-3/2} \int_0^t \sum_I I_s \sigma_{1,s}^2 \int_0^s J(I)_u \beta_{2,u} du \int_0^s J(I)_u \sigma_{2,u} dB_{2,u} ds \\ &= b_n^{-3/2} \int_0^t \sum_I I_s \sigma_{1,s}^2 \beta_{2,a_{J(I)}}(s - a_{J(I)}) \int_0^s J(I)_u \sigma_{2,u} dB_{2,u} ds \\ &\quad + b_n^{-3/2} \int_0^t \sum_I I_s \sigma_{1,s}^2 \int_0^s J(I)_u (\beta_{2,u} - \beta_{2,a_{J(I)}}) du \int_0^s J(I)_u \sigma_{2,u} dB_{2,u} ds \\ &:= \mathcal{T}_{1,n} + \mathcal{T}_{2,n}, \end{aligned}$$

where we denoted by  $a_{J(I)}$  the left endpoint of the interval  $J(I)$ . Let us show that both  $\mathcal{T}_{1,n}$  and  $\mathcal{T}_{2,n}$  tend to zero in probability. Indeed,

$$\begin{aligned} \mathbf{E}^\Pi[\mathcal{T}_{1,n}^2] &= b_n^{-3} \mathbf{E}^\Pi \left[ \int_0^t \sigma_{2,u}^2 \left( \sum_I J(I)_u \int_u^t I_s \sigma_{1,s}^2 (s - a_{J(I)}) ds \beta_{2,a_{J(I)}} \right)^2 du \right] \\ &\leq C b_n^{-3} \mathbf{E}^\Pi \left[ \int_0^t \left( \sum_I J(I)_u |I| |J(I)| |\beta_{2,a_{J(I)}}| \right)^2 du \right] \leq C b_n^{-3} r_n^4 \sup_{t \in [0, T]} \mathbf{E}[\beta_{2,t}^2], \end{aligned}$$

and, after applying the Cauchy-Schwarz inequality several times,

$$\begin{aligned} \mathbf{E}^\Pi[\mathcal{T}_{2,n}^2] &\leq b_n^{-3} \mathbf{E}^\Pi \left[ \int_0^t \sum_I I_s \sigma_{1,s}^4 \left( \int_0^s J(I)_u (\beta_{2,u} - \beta_{2,a_{J(I)}}) du \int_0^s J(I)_u \sigma_{2,u} dB_{2,u} \right)^2 ds \right] \\ &\leq C b_n^{-3} \mathbf{E}^\Pi \left[ \int_0^t \sum_I I_s |J(I)|^4 ds \right] \leq b_n^{-3} r_n^4. \end{aligned}$$

Similar arguments yield the convergence to zero of the sequence  $\mathbf{E}[(\int_0^t \mathbb{H}_s^{1,n} \mathbb{G}_s^{12,n} ds)^2]$ . Thus  $\int_0^t \mathbb{H}_s^{1,n} \mathbb{G}_s^{1,n} ds$  tends to zero in probability as  $n \rightarrow \infty$ . The convergence to zero of the other terms of the sum in the right-hand side of (32) can be shown similarly. This completes the proof of the lemma.  $\square$

**Lemma 9.** *Assume that  $r_n^3 = o_p(b_n^2)$  and that  $\sup_{t \in [0, T]} \mathbf{E}[(\beta_{ij,t}^{[\ell-1]})^2] < \infty$  for every  $i, j, \ell \in \{1, 2\}$ . Then, under Assumption P1, for every fixed  $t \in [0, T]$ , we have*

$$\begin{aligned} \langle N^n, B_1 \rangle_t &\xrightarrow[n \rightarrow \infty]{p} \int_0^t (\beta_{2,s} \sigma_{1,s} + \beta_{1,s} \sigma_{2,s} \rho_s) \mathcal{V}^{I,J}(ds), \\ \langle N^n, B_2 \rangle_t &\xrightarrow[n \rightarrow \infty]{p} \int_0^t (\beta_{1,s} \sigma_{2,s} + \beta_{2,s} \sigma_{1,s} \rho_s) \mathcal{V}^{I,J}(ds) \end{aligned}$$

*Proof.* Let us prove the first assertion, the proof of the second one being completely similar. Since  $N^n = \mathbb{G}^{1,n} \cdot B_1 + \mathbb{G}^{2,n} \cdot B_2$  with  $\mathbb{G}^{1,n}$  and  $\mathbb{G}^{2,n}$  defined in Lemma 5, we

have  $\langle N^n, B_1 \rangle_t = \int_0^t (\mathbb{G}_s^{1,n} + \mathbb{G}_s^{n,2} \rho_s) ds$ . It is easily seen that

$$\begin{aligned} \int_0^t \mathbb{G}_s^{1,n} ds &= b_n^{-1} \sum_{i,j} K_{ij} \int_0^t \left( ((J^j \beta_2) \cdot t)_s I_s^i \sigma_{1,s} + (T^j - T^{j-1} \vee s)_+ I_s^i \sigma_{1,s} \beta_{2,S^{i-1}} \right) ds \\ &= \int_0^t \int_0^t (\beta_{2,u} \sigma_{1,s} \mathbf{1}(u \leq s) + \mathbf{1}(u > s) \sigma_{1,s} \beta_{2,s}) \mathcal{V}_n^{I,J}(ds, du) \\ &\quad + b_n^{-1} \sum_{i,j} K_{ij} \int_0^t (T^j - T^{j-1} \vee s)_+ I_s^i \sigma_{1,s} (\beta_{2,S^{i-1}} - \beta_{2,s}) ds. \end{aligned}$$

Since  $\beta_2$  is supposed to be an Itô process with  $\beta_2^{[0]}$ ,  $\beta_{21}^{[1]}$  and  $\beta_{22}^{[1]}$  being uniformly bounded in  $L^2$ -norm, the expectation  $\mathbf{E}^{\Pi}[\|\beta_{2,S^{i-1}} - \beta_{2,s}\|]$  is bounded up to a constant factor by  $|I|^{1/2}$ . This implies that the second term in the last display is  $o_p(b_n^{-1} \sum_{I,J} K_{IJ} |I|^{3/2} |J|) = o_p(r_n^{3/2} b_n^{-1})$ , while the first term converges to  $\int_0^t \beta_{2,s} \sigma_{1,s} \mathcal{V}^{I,J}(ds)$  by virtue of Assumption P1.

Identical arguments imply the convergence of  $\int_0^t \mathbb{G}_s^{n,2} \rho_s ds$  to  $\int_0^t \beta_{1,s} \sigma_{2,s} \rho_s \mathcal{V}^{I,J}(ds)$  and the assertion of the lemma follows.  $\square$

**Lemma 10.** *If Assumption P2 is fulfilled, then for every  $t \in [0, T]$ , it holds*

$$\begin{aligned} \langle N^n, N^n \rangle_t &\xrightarrow[n \rightarrow \infty]{p} \int_0^t \beta_{2,s}^2 \sigma_{1,s}^2 \mathcal{V}^{I,J,J'}(ds) + \int_0^t \beta_{1,s}^2 \sigma_{2,s}^2 \mathcal{V}^{I,I',J}(ds) \\ &\quad + 2 \int_0^t \beta_{2,s} \beta_{1,s} \sigma_{1,s} \sigma_{2,s} \rho_s \mathcal{V}^{J(I),I(J),I \cap J}(ds). \end{aligned}$$

*Proof.* Since  $N^n = \mathbb{G}^{1,n} \cdot B_1 + \mathbb{G}^{2,n} \cdot B_2$ , its quadratic variation is given by  $\langle N^n, N^n \rangle = [(\mathbb{G}^{1,n})^2 + 2\mathbb{G}^{1,n} \mathbb{G}^{2,n} \rho + (\mathbb{G}^{2,n})^2] \cdot t$ . Using the semimartingale decomposition of  $\beta_2$ , one checks that

$$\begin{aligned} \int_0^t (\mathbb{G}_s^{1,n})^2 ds &= b_n^{-2} \int_0^t \left( \sum_{I,J} K_{IJ} \int_0^t J_u \beta_{2,u} du I_s \sigma_{1,s} \right)^2 ds + o_p(r_n^3 b_n^{-2}) \\ &= b_n^{-2} \int_0^t \sum_I I_s \sigma_{1,s}^2 \left( \int_0^t J(I)_u \beta_{2,u} du \right)^2 ds + o_p(r_n^3 b_n^{-2}) \\ &= \int_{[0,t]^3} \sigma_{1,s}^2 \beta_{2,u} \beta_{2,u'} \mathcal{V}_n^{I,J,J'}(ds, du, du') + o_p(r_n^3 b_n^{-2}). \end{aligned}$$

Analogous computations show that

$$\begin{aligned} \int_0^t (\mathbb{G}_s^{2,n})^2 ds &= \int_{[0,t]^3} \beta_{1,s} \beta_{1,s'} \sigma_{2,u}^2 \mathcal{V}_t^{I,I',J}(ds, ds', du) + o_p(r_n^3 b_n^{-2}), \\ \int_0^t \mathbb{G}_s^{1,n} \mathbb{G}_s^{2,n} \rho_s ds &= \int_{[0,t]^3} \beta_{2,u} \beta_{1,s'} \sigma_{1,s} \sigma_{2,s} \rho_s \mathcal{V}_t^{J(I),I(J),I \cap J}(du, ds', ds) + o_p(r_n^3 b_n^{-2}). \end{aligned}$$

Now, the desired result follows from Assumption P2.  $\square$

The process

$$\begin{aligned} A^n &= b_n^{-1} \sum_{I,J} K_{IJ} \left\{ J \{ [I\sigma_1(\beta_{21}^{[1]} + \beta_{22}^{[1]}\rho)] \cdot s \} + I \{ [J\sigma_2(\beta_{11}^{[1]}\rho + \beta_{12}^{[1]})] \cdot s \} \right\} \cdot t \\ &\quad + b_n^{-1} \sum_{I,J} K_{IJ} \{ (I\beta_1) \cdot t \} \times \{ (J\beta_2) \cdot t \}. \end{aligned}$$

can be written as  $A_t^n = A_t^{1,n} + A_t^{2,n}$ , where

$$\begin{aligned} A_t^{1,n} &= b_n^{-1} \sum_{I,J} K_{IJ} \int_I \int_J \left\{ \sigma_{1,u}(\beta_{21,u}^{[1]} + \beta_{22,u}^{[1]}\rho_u) + \sigma_{2,s}(\beta_{11,s}^{[1]}\rho_s + \beta_{12,s}^{[1]}) \right\} \mathbf{1}_{\{u \leq s \leq t\}} du ds \\ A_t^{2,n} &= b_n^{-1} \sum_{I,J} K_{IJ} \int_I \int_J \beta_{1,u}\beta_{2,s} \mathbf{1}_{\{u \vee s \leq t\}} du ds = \int_{[0,t]^2} \beta_{1,u}\beta_{2,s} \mathcal{V}_n^{I,J}(du, ds). \end{aligned}$$

Using Assumption P1 and the fact that the measures  $\mathcal{V}_n^{I,J}$  are concentrated on the diagonal of the square  $[0, t]^2$ , we get

$$\begin{aligned} A_t^n &= \frac{1}{2} \int_0^t \{ \sigma_{1,u}(\beta_{21,u}^{[1]} + \beta_{22,u}^{[1]}\rho_u) + \sigma_{2,u}(\beta_{11,u}^{[1]}\rho_u + \beta_{12,u}^{[1]}) + 2\beta_{1,u}\beta_{2,u} \} \mathcal{V}_n^{I,J}(du) + o_p(1) \\ &:= A_t^\infty + o_p(1). \end{aligned} \tag{33}$$

**Proposition 9.** *Assume that the functions  $\sigma_1$ ,  $\sigma_2$  and  $\rho$  are continuous in  $[0, T]$  and that  $\sup_{t \in [0, T]} \mathbf{E}[(\beta_{ij}^{[\ell-1]})^4] < \infty$  for every  $i, j, \ell \in \{1, 2\}$ . If assumptions P and P1 are fulfilled, then the sequence of two dimensional processes  $(M^n, N^n + A^n)$  converges weakly to a process  $(M^\infty, N^\infty + A^\infty)$ . Furthermore,  $N^\infty + A^\infty$  is independent of  $M^\infty$ .*

*Proof.* We already did the major part of the proof by showing the convergence in probability of the sequences of quadratic variations-covariations and that of  $A_t^n$ . Now, if we apply Theorem 2-1 from [16] to the semimartingale  $Z^n = (M^n, N^n + A^n)^\mathbf{T}$  with  $\mathbf{B}$  serving as a martingale of reference (denoted by  $M^n$  in [16]), we obtain the weak convergence of  $Z^n$  to a process  $Z$ . Moreover, it follows from (ii) of the aforementioned theorem that  $Z$  may be constructed on an enlargement of the original probability space on which there is a two-dimensional Brownian motion  $\tilde{\mathbf{B}}$  independent of  $\mathbf{B}$  such that

$$Z_t = \begin{pmatrix} 0 \\ A_t^\infty \end{pmatrix} + \int_0^t \frac{d\mathcal{V}^{I,J}}{dt}(s) \begin{pmatrix} 0 & 0 \\ \beta_{2,s}\sigma_{1,s} & \beta_{1,s}\sigma_{2,s} \end{pmatrix} d\mathbf{B}_s + \int_0^t \begin{pmatrix} \mathbf{m}_s & 0 \\ 0 & \mathbf{w}_s \end{pmatrix} d\tilde{\mathbf{B}}_s,$$

where

$$\mathbf{m}_s^2 = h_s^2 \left\{ \frac{d\mathcal{V}^I}{ds} + \frac{d\mathcal{V}^J}{ds} - \frac{d\mathcal{V}^{I \cap J}}{ds} \right\} + \sigma_{1,s}^2 \sigma_{2,s}^2 \frac{d\mathcal{V}^{I,J}}{ds}$$

stands for the Radon-Nikodym derivative of  $\lim_{n \rightarrow \infty} \langle M^n, M^n \rangle_t$  with respect to the Lebesgue measure (cf. Lemma 7) and  $\mathbf{w}_s$  is a predictable process (hence independent of  $\tilde{\mathbf{B}}$ ). If we denote  $(M^\infty, N^\infty) = Z^\mathbf{T} - (0, A^\infty)$ , we get  $M_t^\infty = \int_0^t \mathbf{m}_s d\tilde{\mathbf{B}}_{1,s}$  and  $N_t^\infty = \int_0^t \mathbf{n}_{1,s} dB_{1,s} + \int_0^t \mathbf{n}_{2,s} dB_{2,s} + \int_0^t \mathbf{w}_{1,s} d\tilde{\mathbf{B}}_{2,s}$  with a predictable process  $\mathbf{n}_s = (\mathbf{n}_1, \mathbf{n}_2)$ , and the assertion of the proposition follows.  $\square$

This result implies in particular that  $\mathbf{E}[N_t^\infty + A_t^\infty | M_t^\infty] = \mathbf{E}[N_t^\infty + A_t^\infty] = \mathbf{E}[A_t^\infty]$  for every  $t \in [0, T]$ . Therefore, using (33), we get

$$\begin{aligned} \mathbf{A} &= \mathbf{E}[N_T^\infty + A_T^\infty | M_T^\infty] \\ &= \frac{1}{2} \int_0^T \{ \sigma_{1,u} \mathbf{E}(\beta_{21,u}^{[1]} + \beta_{22,u}^{[1]} \rho_u) + \sigma_{2,u} \mathbf{E}(\beta_{11,u}^{[1]} \rho_u + \beta_{12,u}^{[1]}) + 2\mathbf{E}[\beta_{1,u} \beta_{2,u}] \} \mathcal{V}^{L,J}(du). \end{aligned}$$

## 7. EXPANSION OF THE DISTRIBUTION FOR A MODEL WITH DRIFT TERMS

The aim of this section is to obtain an asymptotic expansion for the distribution of the HY-estimator in the case where the diffusions  $X_1$  and  $X_2$  have non-zero drifts. As shows the stochastic expansion of  $\hat{\theta}_n$  obtained in Lemma 5, the main term in the expansion of  $b_n^{-1/2}(\hat{\theta}_n - \theta)$  is independent of the drifts. Therefore, asymptotic expansions for its distribution are already obtained in Sections 4 and 5. This indicates that the influence of the drifts on the distribution of  $\hat{\theta}_n$  can be regarded as a small perturbation of the distribution in the case where there is no drift. Before stating the main result of this section, let us give a theorem that allows to derive the second-order expansion of the distribution of a random variable defined on the Wiener space in presence of a random perturbation.

**7.1. Perturbation.** Since the drift terms are possibly non-linear functionals of the Brownian motion  $\mathbf{B}$ , we need the Malliavin calculus to carry out computations on the infinite-dimensional Gaussian space.

The basis of our arguments is a perturbation method for deriving asymptotic expansion. It was used in [31] for the perturbation of a martingale but the proof was written inseparably from the martingale structure. In order to apply this methodology to the present situation, we will begin with generalizing Theorem 2.1 of Sakamoto and Yoshida [26].

We consider a probability space equipped with a differential calculus in Malliavin's sense, an integration-by-parts formula and the Sobolev spaces  $\mathbb{D}_{p,\ell}$  equipped with the norm  $\|\cdot\|_{p,\ell}$ . For positive numbers  $M$  and  $\gamma$ , let  $\mathcal{E}(M, \gamma)$  be the set of all measurable functions  $f: \mathbb{R}^d \rightarrow \mathbb{R}$  satisfying  $|f(x)| \leq M(1 + |x|^\gamma)$  for all  $x \in \mathbb{R}^d$ . Let  $\mathcal{E}'$  be a subset of  $\mathcal{E}(M, \gamma)$ .

Let  $\mathcal{X}_n$  and  $\mathcal{Y}_n$  be  $\mathbb{R}^d$ -valued Wiener functionals and put

$$\mathcal{Z}_n = \mathcal{X}_n + s_n \mathcal{Y}_n$$

for some sequence of positive numbers  $s_n$  tending to 0 as  $n \rightarrow \infty$ . We write  $G_n(f) = \bar{o}(s_n)$  if  $s_n^{-1} \sup_{f \in \mathcal{E}} |G_n(f)| \rightarrow 0$  as  $n \rightarrow \infty$ .

**Theorem 5.** *Let  $\ell$  be an integer such that  $\ell > d+2$ . Suppose that the following conditions are satisfied:*

- (1)  $\sup_n \|\mathcal{X}_n\|_{p,\ell} + \sup_n \|\mathcal{Y}_n\|_{p,\ell} < \infty$  for any  $p > 1$ ,
- (2)  $(\mathcal{X}_n, \mathcal{Y}_n) \xrightarrow{d} (\mathcal{X}_\infty, \mathcal{Y}_\infty)$  for some random variables  $\mathcal{X}_\infty$  and  $\mathcal{Y}_\infty$ .

*In addition, assume that there exists a functional  $\tau_n$  such that*

- (3)  $\sup_n \|\tau_n\|_{p,\ell-1} < \infty$  for any  $p > 1$ .
- (4)  $P[|\tau_n| > 1/2] = o(s_n^\alpha)$  for some  $\alpha > 1$ .
- (5)  $\sup_n E[1_{\{|\tau_n| < 1\}} (\det \sigma_{\mathcal{X}_n})^{-p}] < \infty$  for any  $p > 1$ .



- (6) *There is a sequence of signed measures  $\Psi_n$  on  $\mathbb{B}_d$  such that for any positive numbers  $M$  and  $\gamma$ ,  $\mathbf{E}[f(\mathcal{X}_n)] = \Psi_n[f] + \bar{o}(s_n)$  as  $n \rightarrow \infty$  for  $f \in \mathcal{E}'$ . Moreover, for every polynomial  $\pi(x)$  in  $x$ , there exists a constant  $c_\pi$  such that  $|\Psi_n[e^{iu \cdot x} \pi(x)]| \leq c_\pi(1 + |u|^{\ell-1})^{-1}$  for all  $u \in \mathbb{R}^d$ .*

Then  $\mathcal{X}_\infty$  has a density  $p^{\mathcal{X}_\infty}$  with respect to the Lebesgue measure and, for any positive numbers  $M$  and  $\gamma$ ,

$$\mathbf{E}[f(\mathcal{Z}_n)] = \Psi_n[f] + s_n \int_{\mathbb{R}} f(x) g_\infty(x) dx + \bar{o}(s_n) \quad (34)$$

for  $f \in \mathcal{E}(M, \gamma)$ , where  $g_\infty(x) = -\operatorname{div}_x(\mathbb{E}[\mathcal{Y}_\infty \mid \mathcal{X}_\infty = x] p^{\mathcal{X}_\infty}(x))$ .

*Proof.* Let  $\psi_n$  be some truncation functional to be defined later and let  $\zeta(x) = 1 + |x|^{2m}$  ( $x \in \mathbb{R}^d$ ), where  $m$  is an integer such that  $2m > \gamma + d$ . We have

$$\begin{aligned} \mathbf{E}[f(\mathcal{Z}_n)] &= \mathbf{E}[f(\mathcal{Z}_n)\psi_n] + \mathbf{E}[f(\mathcal{Z}_n)(1 - \psi_n)] \\ &= \int_{\mathbb{R}^d} f(x) \tilde{p}_n(x) dx + \mathbf{E}[f(\mathcal{Z}_n)(1 - \psi_n)], \end{aligned}$$

where  $\tilde{p}_n(x) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{-iu \cdot x} \hat{g}_n^0(u) du$ , with  $\hat{g}_n^0(u) = \mathbf{E}[e^{iu \cdot \mathcal{Z}_n} \psi_n]$ .

We will show below (cf. (36)) that the term  $\mathbf{E}[f(\mathcal{Z}_n)(1 - \psi_n)]$  is  $\bar{o}(s_n)$  and is negligible with respect to  $\mathbf{E}[f(\mathcal{Z}_n)\psi_n]$ . To deal with this latter term, let us introduce the notation

$$\begin{aligned} h_n^0(x) &= \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{-iu \cdot x} \hat{h}_n^0(u) du, \\ \hat{h}_n^0(u) &= \Psi_n[e^{iu \cdot x}] + s_n \mathbb{E}\left[e^{iu \cdot \mathcal{X}_\infty} iu \cdot \mathcal{Y}_\infty\right], \\ \hat{g}_n(u) &= \mathbf{E}[e^{iu \cdot \mathcal{Z}_n} \psi_n \zeta(\mathcal{Z}_n)], \\ \hat{h}_n(u) &= \zeta(-i\partial_u) \hat{h}_n^0(u) = \Psi_n[e^{iu \cdot x} \zeta(x)] + s_n \mathbb{E}\left[\zeta(-i\partial_u)(e^{iu \cdot y} iu) \Big|_{y=\mathcal{X}_\infty} \cdot \mathcal{Y}_\infty\right]. \end{aligned}$$

Using the integration by parts formula, we get

$$\begin{aligned} \zeta(x) \tilde{p}_n(x) &= \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{-iu \cdot x} \hat{g}_n(u) du, \\ \zeta(x) h_n^0(x) &= \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{-iu \cdot x} \hat{h}_n(u) du. \end{aligned}$$

Further, there is a linear form  $\zeta_2(x, y)[\cdot]$  of polynomial elements such that

$$\zeta(x + y) = \zeta(x) + \partial\zeta(x)[y] + \zeta_2(x, y)[y^{\otimes 2}]$$

for  $x, y \in \mathbb{R}^d$ . We also notice that

$$\begin{aligned} \zeta(-i\partial_u)(e^{iu \cdot y} iu) &= \zeta(-i\partial_u) \partial_y e^{iu \cdot y} = \partial_y(\zeta(-i\partial_u) e^{iu \cdot y}) \\ &= \partial_y(e^{iu \cdot y} \zeta(y)) = e^{iu \cdot y} \zeta(y) iu + e^{iu \cdot y} \partial\zeta(y) \end{aligned}$$

for  $u, y \in \mathbb{R}^d$ .

Let  $\varphi(x) = f(x)/\zeta(x)$  and  $\Lambda_n = \{u \in \mathbb{R}^d; |u| \leq s_n^{-1}\}$ . Then

$$(2\pi)^d \int_{\mathbb{R}^d} f(x) \{\tilde{p}_n(x) - h_n^0(x)\} dx = A(n) + s_n B(n) + s_n C(n) + s_n^2 D(n) + E(n),$$

where

$$\begin{aligned}
A(n) &= \int_{\mathbb{R}^d} dx \varphi(x) \int_{\Lambda_n} e^{-iu \cdot x} \left\{ \mathbf{E} \left[ e^{iu \cdot \mathcal{X}_n} \psi_n \zeta(\mathcal{X}_n) \right] - \Psi_n \left[ e^{iu \cdot x} \zeta(x) \right] \right\} du, \\
B(n) &= \int_{\mathbb{R}^d} dx \varphi(x) \int_{\Lambda_n} e^{-iu \cdot x} \left\{ \mathbf{E} \left[ e^{iu \cdot \mathcal{X}_n} iu \cdot \mathcal{Y}_n \int_0^1 \exp(is_n u \cdot \mathcal{Y}_n s) ds \psi_n \zeta(\mathcal{X}_n) \right] \right. \\
&\quad \left. - \mathbf{E} \left[ e^{iu \cdot \mathcal{X}_\infty} iu \cdot \mathcal{Y}_\infty \zeta(\mathcal{X}_n) \right] \right\} du, \\
C(n) &= \int_{\mathbb{R}^d} dx \varphi(x) \int_{\Lambda_n} e^{-iu \cdot x} \left\{ \mathbf{E} \left[ e^{iu \cdot \mathcal{Z}_n} \psi_n \partial \zeta(\mathcal{X}_n) [\mathcal{Y}_n] \right] - \mathbb{E} \left[ e^{iu \cdot \mathcal{X}_\infty} \partial \zeta(\mathcal{X}_\infty) [\mathcal{Y}_\infty] \right] \right\} du, \\
D(n) &= \int_{\mathbb{R}^d} dx \varphi(x) \int_{\Lambda_n} e^{-iu \cdot x} \mathbf{E} \left[ e^{iu \cdot \mathcal{Z}_n} \psi_n \zeta_2(\mathcal{X}_n, s_n \mathcal{Y}_n) [\mathcal{Y}_n^{\otimes 2}] \right] du \\
E(n) &= \int_{\mathbb{R}^d} dx \varphi(x) \int_{\Lambda_n^c} e^{-iu \cdot x} \left( \hat{g}_n(u) - \hat{h}_n(u) \right) du.
\end{aligned}$$

Since

$$\int_{\mathbb{R}^d} dx \varphi(x) \int_{\mathbb{R}^d} e^{-iu \cdot x} \mathbf{E} \left[ e^{iu \cdot \mathcal{X}_n} \psi_n \zeta(\mathcal{X}_n) \right] du = (2\pi)^d \mathbf{E} \left[ \varphi(\mathcal{X}_n) \psi_n \zeta(\mathcal{X}_n) \right]$$

and  $\int_{\mathbb{R}^d} dx \varphi(x) \int_{\mathbb{R}^d} e^{-iu \cdot x} \Psi_n \left[ e^{iu \cdot x} \zeta(x) \right] du = (2\pi)^d \Psi_n[\varphi \zeta]$ , we have

$$\begin{aligned}
|A(n)| &\leq (2\pi)^d \left| \mathbf{E} \left[ \varphi(\mathcal{X}_n) \psi_n \zeta(\mathcal{X}_n) \right] - \Psi_n[\varphi \zeta] \right| + F(n) \\
&\leq (2\pi)^d \left| \mathbf{E} \left[ \varphi(\mathcal{X}_n) (1 - \psi_n) \zeta(\mathcal{X}_n) \right] \right| + F(n) + \bar{o}(s_n)
\end{aligned}$$

from Condition (6), where

$$F(n) = (2\pi)^d \int_{\mathbb{R}^d} |\varphi(x)| dx \times \int_{\Lambda_n^c} \left\{ \left| \mathbf{E} \left[ e^{iu \cdot \mathcal{X}_n} \psi_n \zeta(\mathcal{X}_n) \right] \right| + \left| \Psi_n \left[ e^{iu \cdot x} \zeta(x) \right] \right| \right\} du.$$

In what follows  $C$  denotes a generic constant independent of  $n$  and  $u$  and it varies from line to line.

To evaluate  $F(n)$ , we need the explicit form of  $\psi_n$ . Let us denote by  $\psi$  a smooth function from  $\mathbb{R}$  into  $[0, 1]$  such that  $\psi(t) = 1$  if  $|t| \leq 1/2$  and  $\psi(t) = 0$  if  $|t| \geq 1$ . We can write

$$\det \left[ I_d + s_n \sigma_{\mathcal{X}_n}^{-1} (\langle \mathcal{X}_n, \mathcal{Y}_n \rangle + \langle \mathcal{Y}_n, \mathcal{X}_n \rangle) \right] = 1 + s_n \det \sigma_{\mathcal{X}_n}^{-d} K_n$$

with a certain functional  $K_n$  satisfying  $\sup_n \|K_n\|_{p, \ell-1} < \infty$  for every  $p > 1$ . Let

$$\psi_n = \psi(\tau_n) \psi \left( 2s_n \det \sigma_{\mathcal{X}_n}^{-d} K_n \right).$$

Obviously,  $\psi_n \in \cap_{p>1} \mathbb{D}_{p, \ell-1}$ ; in order to prove it, replace  $\sigma_{\mathcal{X}_n}$  by  $\sigma_{\mathcal{X}_n} + k^{-1} I_d$ , differentiate, and take limits in  $L^p$ -spaces as  $k \rightarrow \infty$ . Furthermore, we infer that  $\sup_n \|\psi_n\|_{p, \ell-1} < \infty$  for every  $p > 1$ . If  $\psi_n > 0$ , then  $\det(\sigma_{\mathcal{X}_n}^{-1} \sigma_{\mathcal{Z}_n}) \geq 1/2$  leading to

$$\det \sigma_{\mathcal{Z}_n} \geq \frac{1}{2} \det \sigma_{\mathcal{X}_n}. \tag{35}$$

By applying the Integration-by-parts (IBP) formula and the non-degeneracy assumption for  $\mathcal{X}_n$  under truncation, we find that

$$\sup_n \left| \mathbf{E} \left[ e^{iu \cdot \mathcal{X}_n} \psi_n \zeta(\mathcal{X}_n) \right] \right| \leq \frac{C}{1 + |u|^{\ell-1}}$$

for all  $u \in \mathbb{R}^d$ . This together with Condition (6) implies that  $F(n) = \bar{O}(s_n^2) = \bar{o}(s_n)$ . Besides,

$$\left| \mathbf{E} \left[ \varphi(\mathcal{X}_n)(1 - \psi_n)\zeta(\mathcal{X}_n) \right] \right| \leq C_q \|1 - \psi_n\|_q = \bar{o}(s_n). \quad (36)$$

Here  $q$  is arbitrary constant such that  $q \in (0, 1)$ . Consequently,  $A(n) = \bar{o}(s_n)$ .

Taking the limit of

$$\sup_n \left| \mathbf{E} \left[ \zeta(-i\partial_u)(e^{iu \cdot y} iu) \Big|_{y=\mathcal{X}_n} \cdot \mathcal{Y}_n \right] \right| \leq \frac{C}{1 + |u|^{\ell-2}},$$

we have

$$\left| \mathbf{E} \left[ \zeta(-i\partial_u)(e^{iu \cdot y} iu) \Big|_{y=\mathcal{X}_\infty} \cdot \mathcal{Y}_\infty \right] \right| \leq \frac{C}{1 + |u|^{\ell-2}}$$

for all  $u \in \mathbb{R}^d$ . On the other hand, from the IBP formula in view of the uniform nondegeneracy of  $\mathcal{Z}_n$  under truncation deduced from that of  $\mathcal{X}_n$  by (35), it follows that

$$\sup_n |\hat{g}_n(u)| \leq \frac{C}{1 + |u|^{\ell-1}}$$

for all  $u \in \mathbb{R}^d$ . From these estimates, we have  $E(n) = \bar{O}(s_n^2) = \bar{o}(s_n)$ . Similar argument yields the estimate  $\sup_n |D(n)| < \infty$ .

To obtain  $C(n) = \bar{o}(1)$ , Lebesgue's dominated convergence theorem applies with the aid of the estimate

$$\sup_n \left| \mathbf{E} \left[ e^{iu \cdot \mathcal{Z}_n} \psi_n \partial \zeta(\mathcal{X}_n) [\mathcal{Y}_n] \right] - \mathbf{E} \left[ e^{iu \cdot \mathcal{X}_\infty} \partial \zeta(\mathcal{X}_\infty) [\mathcal{Y}_\infty] \right] \right| \leq \frac{C}{1 + |u|^{\ell-1}}$$

for all  $u \in \mathbb{R}^d$ . In the same way, we can obtain  $B(n) = \bar{o}(1)$ . However, we have to use more elaborately the estimate

$$\sup_n 1_{\Lambda_n}(u) \left| \mathbf{E} \left[ e^{iu \cdot \mathcal{X}_n} iu \cdot \mathcal{Y}_n \int_0^1 \exp(is_n u \cdot \mathcal{Y}_n s) ds \psi_n \zeta(\mathcal{X}_n) \right] \right| \leq \frac{C}{1 + |u|^{\ell-2}}$$

for all  $u \in \mathbb{R}^d$  ( $C$  is independent of  $u$ ) and its limit

$$\left| \mathbf{E} \left[ e^{iu \cdot \mathcal{X}_\infty} iu \cdot \mathcal{Y}_\infty \zeta(\mathcal{X}_\infty) \right] \right| \leq \frac{C}{1 + |u|^{\ell-2}}$$

for all  $u \in \mathbb{R}^d$ .

After all,  $\int_{\mathbb{R}^d} f(x) \tilde{p}_n(x) dx - \int_{\mathbb{R}^d} f(x) h_n^0(x) dx = \bar{o}(s_n)$  as  $n \rightarrow \infty$ . From definition of  $h_n^0(x)$ , it is easy to show that  $\mathcal{X}_\infty$  has a differentiable density  $p^{\mathcal{X}_\infty}$  and that

$$h_n^0(x) = \frac{d\Psi_n}{dx}(x) - s_n \operatorname{div} \left\{ \mathbf{E}[\mathcal{Y}_\infty | \mathcal{X}_\infty = x] p^{\mathcal{X}_\infty}(x) \right\}.$$

The existence of the integral  $\int_{\mathbb{R}^d} f(x) h_n^0(x) dx$  is ensured as a consequence under our assumptions.  $\square$

**7.2. Asymptotic expansion of the distribution.** We are now in a position to state and to prove the main result of this section, which provides an unconditional asymptotic expansion of the distribution of the HY-estimator. It is also possible to derive asymptotic expansions conditionally to the processes generating the observation times, but they have more complicated form and are not presented here.

**Theorem 6.** *Suppose that Assumptions P1 and P2 are satisfied and*

$$\sup_{t \in [0, T]} \|\beta_{i,t}^{[l-1]}\|_{p,4} < \infty, \quad \text{for all } p > 1 \quad \text{and } i, l \in \{1, 2\}.$$

Let us define

$$\begin{aligned} \mathbf{c} &= \int_0^T \sigma_{1,t}^2 \sigma_{2,t}^2 \mathcal{V}^{L,J}(dt) + \int_0^T \sigma_{1,t} \sigma_{2,t} \rho_t \{ \mathcal{V}^I(dt) + \mathcal{V}^J(dt) - \mathcal{V}^{I \cap J}(dt) \}, \\ \mathbf{A} &= \frac{1}{2} \int_0^T \{ \sigma_{1,u} \mathbf{E}(\beta_{21,u}^{[1]} + \beta_{22,u}^{[1]} \rho_u) + \sigma_{2,u} \mathbf{E}(\beta_{11,u}^{[1]} \rho_u + \beta_{12,u}^{[1]}) + 2\mathbf{E}[\beta_{1,u} \beta_{2,u}] \} \mathcal{V}^{L,J}(du). \end{aligned}$$

Under the notation of Proposition 6, if for some  $a \in (3/4, 1)$ ,  $\mathbf{P}(A_n(a)^c) = o(b_n^p)$  for every  $p > 1$ , and  $\mathbf{E}[2\mu_{2,n} - \mathbf{c}] = O(b_n^{2a-1})$ , then

$$\sup_{f \in \mathcal{E}(M, \gamma) \cap \mathcal{E}^0(\mathbf{C}, \eta, r_0, \mathbf{c}^*)} \left| \mathbf{E}[f(b_n^{-1/2}(\hat{\theta}_n - \theta))] - \int_{\mathbb{R}} f(z) p_n^*(z) dz \right| = o(b_n^{1/2}), \quad (37)$$

where

$$p_n^*(z) = \frac{e^{-z^2/(2\mathbf{c})}}{\sqrt{2\pi\mathbf{c}}} \left[ 1 + \frac{b_n^{1/2}}{6\mathbf{c}^3} (\mathbf{E}[\bar{\lambda}_{3,n}](z^3 - 3\mathbf{c}z) + 6\mathbf{A}\mathbf{c}^2 z) \right].$$

Moreover, if  $\sup_{n \in \mathbb{N}} \mathbf{E}[\bar{\lambda}_{3,n}] < \infty$ , then inequality (37) holds with  $p_n^*$  replaced by

$$p_n^+(z) = \frac{\max(0, p_n^*(z))}{\int_{\mathbb{R}} \max(0, p_n^*(u)) du},$$

which is a probability density.

*Proof.* We apply Theorem 5 to  $\mathcal{Z}_n = b_n^{-1/2}(\hat{\theta}_n - \theta)$  with  $\ell = 4$ ,  $\mathcal{X}_n = M_T^n$  and  $\mathcal{Y}_n = b_n^{-1/2}(\mathcal{Z}_n - M_T^n)$ . Thus, we need to check that all 6 conditions of Theorem 5 are fulfilled. In view of Lemma 5 and Proposition 9, we have  $(\mathcal{X}_n, \mathcal{Y}_n)$  converges in distribution to some random vector  $(\mathcal{X}_\infty, \mathcal{Y}_\infty)$ . Thus the second condition of Theorem 5 is verified.

We have already seen in Section 3.1 that the principal part  $\mathcal{X}_n$  of  $b_n^{-1/2}(\hat{\theta}_n - \theta)$  can be written in the form

$$\mathcal{X}_n = b_n^{-1/2}(\boldsymbol{\xi}^T \mathbf{A} \boldsymbol{\xi} - \theta) = b_n^{-1/2} \sum_{\ell=1}^N \lambda_{\ell,n} (\zeta_{\ell,n}^2 - 1),$$

where

$$\boldsymbol{\xi} = (\{I^1 \sigma_1 \cdot B_1\}_T, \dots, \{I^{N_1} \sigma_1 \cdot B_1\}_T, \{J^1 \sigma_2 \cdot B_2\}_T, \dots, \{J^{N_2} \sigma_2 \cdot B_2\}_T)^T \sim \mathcal{N}_N(0, \Sigma)$$

and the entries of the matrices  $\Sigma$  and  $\mathbf{A}$  are given by (10) and (9) respectively. Recall that the vector  $\boldsymbol{\zeta} \in \mathbb{R}^N$  is obtained as a linear transformation of  $\boldsymbol{\xi}$  and is distributed according to  $\mathcal{N}(0, I)$ .

Let  $W = C_0([0, T], \mathbb{R}^2)$  be the Wiener space of continuous functions from  $[0, T]$  to  $\mathbb{R}^2$  starting at the origin. Recall that  $W$  is a measurable space equipped with the Borel  $\sigma$ -field induced by the uniform topology. The reference measure on  $W$  is the measure generated by the standard Wiener process (in our case, the two-dimensional Brownian motion).

Let  $w = (w_1, w_2)$  be the canonical process on  $W$ . Then,  $(B_1, B_2)$  can be defined by

$$B_{1,t} = w_{1,t}, \quad B_{2,t} = \int_0^t \rho_s dw_{1,s} + \int_0^t \sqrt{1 - \rho_s^2} dw_{2,s}.$$

Obviously, for every  $\ell = 1, \dots, N$ , there is some function  $\phi^\ell \in L^2([0, T], \mathbb{R}^2)$  such that  $\zeta_{\ell,n} = \int_0^T \phi_{1,t}^\ell dw_{1,t} + \int_0^T \phi_{2,t}^\ell dw_{2,t} := w(\phi^\ell)$ .

The process  $w$  is an isonormal Gaussian process on  $H = L^2([0, T], \mathbb{R}^2)$  (see [22, Def. 1.1.1]) Using the definition of the Malliavin derivative (see [22, Def. 1.2.1]) and the chain rule [22, Prop. 1.2.3], we get the following expression for the Malliavin derivative of  $\mathcal{X}_n$ :

$$D_t \mathcal{X}_n = 2b_n^{-1/2} \sum_{\ell=1}^N \lambda_{\ell,n} \zeta_{\ell,n} \phi_t^\ell.$$

Since the components of  $\zeta$  are non-correlated with variance equal to one, the family  $\{\phi^\ell\}_{\ell \leq N}$  is orthonormal. As a first consequence of this fact, we get that  $\sup_n \|\mathcal{X}_n\|_{p,4} < \infty$  for every  $p > 1$ . To show this, Rosenthal's inequality and the result of Lemma 3 can be used. As a second consequence, we obtain that the Malliavin covariance of  $\mathcal{X}_n$  is

$$\sigma_{\mathcal{X}_n} = 4b_n^{-1} \sum_{\ell=1}^n \lambda_{\ell,n}^2 \zeta_{\ell,n}^2 = 4b_n^{-1} \mu_{2,n} + 4b_n^{-1} \sum_{\ell=1}^n \lambda_{\ell,n}^2 (\zeta_{\ell,n}^2 - 1). \quad (38)$$

Let us introduce the random variable  $\tau_n$  that will play a role of truncation:

$$\tau_n = -(2 - 8\mu_{2,n}(\mathbf{c}b_n)^{-1})_+ + 8(\mathbf{c}b_n)^{-1} \sum_{\ell=1}^N \lambda_{\ell,n}^2 (\zeta_{\ell,n}^2 - 1).$$

In this notation, we have

$$\sigma_{\mathcal{X}_n} \geq \mathbf{c} + \frac{\mathbf{c}\tau_n}{2}$$

and, therefore,  $\mathbf{1}_{\{|\tau_n| < 1\}} |\sigma_{\mathcal{X}_n}^{-1}| < 2/\mathbf{c}$ . Thus, the condition (5) of Theorem 5 is obviously fulfilled. Let us check now that  $\tau_n$  satisfies conditions (3) and (4) of the aforementioned theorem.

To verify condition (3) of Theorem 5, we remark that

$$D\tau_n = 16(\mathbf{c}b_n)^{-1} \sum_{\ell=1}^N \lambda_{\ell,n}^2 \zeta_{\ell,n} \phi^\ell, \quad D^2\tau_n = 16(\mathbf{c}b_n)^{-1} \sum_{\ell=1}^N \lambda_{\ell,n}^2 \phi^\ell \otimes \phi^\ell$$

$D^k\tau_n \equiv 0$  for every  $k \geq 3$ . Therefore,

$$\|D\tau_n\|_H^2 = 256(\mathbf{c}b_n)^{-2} \sum_{\ell=1}^N \lambda_{\ell,n}^4 \zeta_{\ell,n}^2, \quad \|D^2\tau_n\|_{H \otimes H}^2 = 256(\mathbf{c}b_n)^{-2} \sum_{\ell=1}^N \lambda_{\ell,n}^4.$$

In view of the Rosenthal inequality, we get

$$\mathbf{E}^\Pi[\|D\tau_n\|_H^p] \leq C(p)b_n^{-p}(\mu_{4,n}^{p/2} + \mu_{2p,n} + \mu_{8,n}^{p/4}),$$

for every  $p \geq 2$ . Using the definition of  $\mu_{k,n}$ , one can check that  $\mu_{2k,n} \leq \mu_{4,n}^{k/2}$ . In view of inequality (18) and the obvious bound  $\mu_{2,n} \leq Cr_n$ , we get

$$\mathbf{E}^\Pi[\|D\tau_n\|_H^p] \leq Cb_n^{-p}r_n^{3p/2}, \quad \mathbf{E}^\Pi[\|D^2\tau_n\|_{H \otimes H}^p] \leq Cb_n^{-p}r_n^{3p/2}, \quad \forall p \geq 4.$$

Similar arguments yield

$$\mathbf{E}[\tau_n^p] = \mathbf{E}^\Pi[\tau_n^p] \leq C(1 + b_n^{-p}\mathbf{E}[r_n^{3p/2}]) \leq C(1 + b_n^{-p}b_n^{9p/8} + T^{3p/2}b_n^{-p}\mathbf{P}(A_n(a)^c)) < \infty.$$

To check condition (4) of Theorem 5, we use the inequality

$$\mathbf{P}(|\tau_n| > 1/2) \leq \mathbf{P}(2 - 8\mu_{2,n}(\mathbf{c}b_n)^{-1} > 0) + \mathbf{P}\left(8(\mathbf{c}b_n)^{-1} \left| \sum_{\ell=1}^N \lambda_{\ell,n}^2 (\zeta_{\ell,n}^2 - 1) \right| > 1/2\right).$$

On the one hand, since the event  $\{2 - 8\mu_{2,n}(\mathbf{c}b_n)^{-1} > 0\} = \{\bar{\lambda}_{2,n} - \mathbf{c} < -\mathbf{c}/2\}$  is included in  $A_n(a)^c$ , its probability is  $o(b_n^p)$  for every  $p > 1$ . On the other hand, combining the Tchebychev and the Rosenthal inequalities, for every  $k \geq 16$  we get

$$\begin{aligned} \mathbf{P}\left(8(\mathbf{c}b_n)^{-1} \left| \sum_{\ell=1}^N \lambda_{\ell,n}^2 (\zeta_{\ell,n}^2 - 1) \right| > 1/2\right) &\leq Cb_n^{-k} \mathbf{E}[\mu_{4,n}^{k/2} + \mu_{2k,n}] \leq Cb_n^{-k} \mathbf{E}[r_n^{3k/2}] \\ &\leq Cb_n^{-k+9k/8} + Cb_n^{-k} \mathbf{P}(A_n(a)^c) = O(b_n^2). \end{aligned}$$

Thus, we proved that conditions (2)-(5) of Theorem 5 are fulfilled and that  $\sup_n \|\mathcal{X}_n\|_{p,4} < \infty$ . Condition (6) is ensured by Theorem 2. To complete the proof, it remains to check that  $\sup_n \|\mathcal{Y}_n\|_{p,4} < \infty$ . This inequality can be proved using the identity  $\mathcal{Y}_n = b_n^{-1}(\Phi_n^2 + \Phi_n^3)$ , where  $\Phi_n^2$  and  $\Phi_n^3$  are the random variables defined in the proof of Lemma 5. The proof is rather technical, but is based on the arguments that we have already used several times in this and the previous sections. Therefore it will be omitted.  $\square$

In the case when the sampling scheme is generated by two Poisson processes, we get the following consequence of the last theorem.

**Proposition 10.** *Let the observation times of processes  $X_1$  and  $X_2$  be generated by two independent Poisson processes with intensities  $np_1$  and  $np_2$ ,  $p_1p_2 > 0$ . If*

- the observation times are independent of the process  $\mathbf{X}$ ,
- the functions  $\sigma_1$ ,  $\sigma_2$  and  $\rho$  are Lipschitz continuous,
- $\sup_{t \in [0, T]} \|\beta_{i,t}^{[l-1]}\|_{p,4} < \infty$  for all  $p > 1$ ,  $i, l \in \{1, 2\}$ .

Then then

$$\sup_{f \in \mathcal{E}(M, \gamma) \cap \mathcal{E}^0(\mathbf{C}, \eta, r_0, \mathbf{c}^*)} \left| \mathbf{E}[f(n^{1/2}(\hat{\theta}_n - \theta))] - \int_{\mathbb{R}} f(z) p_n^\circ(z) dz \right| = o(n^{-1/2}), \quad (39)$$

where

$$p_n^\circ(z) \propto \frac{e^{-z^2/(2\mathbf{c})}}{\sqrt{2\pi\mathbf{c}}} \left[ 1 + \frac{1}{\sqrt{n\mathbf{c}^3}} (2\kappa z^3 - 6\kappa\mathbf{c}z + \mathbf{A}\mathbf{c}^2z) \right]_+$$

is a probability density with

$$\begin{aligned} \mathbf{c} &= \left( \frac{2}{p_1} + \frac{2}{p_2} \right) \int_0^T \sigma_{1,t}^2 \sigma_{2,t}^2 (1 + \rho_t^2) dt - \frac{2}{p_1 + p_2} \int_0^T (\sigma_{1,t} \sigma_{2,t} \rho_t)^2 dt, \\ \kappa &= \left( \frac{1}{p_1^2} + \frac{1}{p_2^2} \right) \int_0^T h(t)^3 dt + \frac{3p_1^2 + 2p_1p_2 + 3p_2^2}{p_1^2 p_2^2} \int_0^T \sigma_{1,t}^2 \sigma_{2,t}^2 h(t) dt, \\ \mathbf{A} &= \left( \frac{1}{p_1} + \frac{1}{p_2} \right) \int_0^T \{ \sigma_{1,t} \mathbf{E}(\beta_{21,t}^{[1]} + \beta_{22,t}^{[1]} \rho_t) + \sigma_{2,t} \mathbf{E}(\beta_{11,t}^{[1]} \rho_t + \beta_{12,t}^{[1]}) + 2\mathbf{E}[\beta_{1,t} \beta_{2,t}] \} dt. \end{aligned}$$

*Proof.* Lemmas 16-19 (cf. Appendix) imply that the partitions generated by independent Poisson processes satisfy Assumptions P1 and P2. Therefore, using Theorems 6 and 4, we get the desired result.  $\square$

## 8. APPENDIX

**Lemma 11.** *For every  $\lambda > 0$  it holds  $\sum_{k=0}^{\infty} \frac{\lambda^k}{k!(k+2)} = \lambda^{-2}(\lambda e^\lambda - e^\lambda + 1)$ .*

*Proof.* It follows from the equality  $1/(k!(k+2)) = 1/((k+1)!) - 1/((k+2)!)$  and the power series expansion of the exponential function.  $\square$

**Lemma 12.** *Let  $\mathcal{P}$  be a homogeneous Poisson point process on  $\mathbb{R}$  with intensity  $\lambda > 0$  and let  $a \in \mathbb{R}$  be some point. For every  $\omega$ , let us denote by  $I_a(\omega)$  the interval of the partition of  $\mathbb{R}$  generated by  $\mathcal{P}$  containing  $a$ . Then  $|I_a|$  is distributed according to the law  $\text{Gamma}(2, \lambda)$ .*

*Proof.* W.l.o.g. we can assume that  $a = 0$ . Since the restrictions of  $\mathcal{P}$  on  $(-\infty, 0)$  and  $[0, \infty)$  are two independent Poisson processes, the law of  $|I_a|$  coincides with the law of the sum of two iid random variables distributed according to the exponential law with parameter  $\lambda$ . Thus the assertion of the lemma follows from the well known properties of Gamma laws.  $\square$

**Lemma 13.** *Let  $\mathcal{P}$  be a homogeneous Poisson point process on  $\mathbb{R}$  with intensity  $\lambda > 0$  and let  $I = [a, b] \subset \mathbb{R}$  be some interval. For every  $\omega$ , let us denote by  $N = N(\omega)$  the number of points of  $\mathcal{P}(\omega)$  lying in  $I$  and by  $t_i = t_i(\omega)$ ,  $i = 1, \dots, N$  the ordered sequence of these points. Then*

$$\mathbf{E} \left[ \sum_{i=0}^N (t_{i+1} - t_i)^2 \right] = \frac{2(|I|\lambda - 1 + e^{-|I|\lambda})}{\lambda^2},$$

where we used  $t_0 = a$  and  $t_{N+1} = b$ .

*Proof.* Without loss of generality, we assume that  $I = [0, 1]$ . We use the fact that conditionally to  $N(\omega) = k$ , the random vector  $(t_1, \dots, t_k)$  have the same distribution as  $(U_{(1)}, \dots, U_{(k)})$ , where  $U_1, \dots, U_k$  are independent uniformly in  $[0, 1]$  distributed random variables and  $U_{(1)}, \dots, U_{(k)}$  are the corresponding order statistics. Since the joint density of  $(U_{(i)}, U_{(i+1)})$  is given by

$$f_{(U_{(i)}, U_{(i+1)})}(x, y) = \frac{k!}{(i-1)!(k-i-1)!} x^{i-1} (1-y)^{k-i-1} \mathbf{1}_{\{x \leq y\}},$$

the expectation  $\mathbf{E}[(U_{(i+1)} - U_{(i)})^2]$  is equal to  $2/[(k+1)(k+2)]$ . It is easily seen that  $\mathbf{E}[U_{(1)}^2] = \mathbf{E}[(1 - U_{(k)})^2] = 2/[(k+1)(k+2)]$ . Therefore,

$$\mathbf{E} \left[ \sum_{i=0}^N (t_{i+1} - t_i)^2 \right] = \sum_{k=0}^{\infty} \left( \sum_{i=0}^k \frac{2}{(k+1)(k+2)} \right) \mathbf{P}(N = k) = \sum_{k=0}^{\infty} \frac{2e^{-\lambda} \lambda^k}{k!(k+2)}.$$

The desired result follows now from Lemma 11.  $\square$

**Lemma 14.** *Let  $\zeta_1 \sim \mathcal{E}(\lambda_1)$  and  $\mathcal{P}^2$  be a Poisson process with intensity  $\lambda_2$  independent of  $\zeta_1$ . Let us denote by  $\Pi_\zeta^2$  the partition of  $[0, \zeta_1]$  generated by  $\mathcal{P}^2$ . Then*

$$\mathbf{E} \left[ \zeta_1 \sum_{J \in \Pi_\zeta^2} |J|^2 \right] = \frac{6\lambda_1 + 4\lambda_2}{\lambda_1^2(\lambda_1 + \lambda_2)^2}.$$

*Proof.* By rescaling and by using Lemma 13, we get

$$\mathbf{E} \left[ \sum_{J \in \Pi_\zeta^2} |J|^2 \mid \zeta_1 \right] = \frac{2\zeta_1^2(\lambda_2\zeta_1 - 1 + e^{-\lambda_2\zeta_1})}{\lambda_2^2\zeta_1^2}.$$

Therefore,

$$\begin{aligned} \mathbf{E} \left[ \zeta_1 \sum_{J \in \Pi_\zeta^2} |J|^2 \right] &= \frac{2}{\lambda_2} \mathbf{E}[\zeta_1^2] - \frac{2}{\lambda_2^2} \mathbf{E}[\zeta_1] + \frac{2}{\lambda_2^2} \mathbf{E}[\zeta_1 e^{-\lambda_2\zeta_1}] \\ &= \frac{4}{\lambda_2\lambda_1^2} - \frac{2}{\lambda_2^2\lambda_1} + \frac{2}{\lambda_2^2} \frac{\lambda_1}{(\lambda_1 + \lambda_2)^2} = \frac{6\lambda_1 + 4\lambda_2}{\lambda_1^2(\lambda_1 + \lambda_2)^2}. \end{aligned}$$

This completes the proof of the lemma.  $\square$

**Lemma 15.** *Let  $I = [a, b]$  be an interval of  $[0, T]$ . If  $\mathcal{P}$  is a Poisson point process with intensity  $\lambda$  and  $\Pi$  is the partition of  $[0, T]$  generated by  $\mathcal{P}$ , then*

$$\begin{aligned} \mathbf{E} \left[ \sum_{J \in \Pi} |J|K_{IJ} \right] &= |I| + 2\lambda^{-1} - \lambda^{-1}(e^{-\lambda a} + e^{-\lambda(T-b)}), \\ \mathbf{E} \left[ \sum_{J \in \Pi} |J \setminus I| \cdot |J \cap I| \right] &= \lambda^{-2}(1 - e^{-\lambda|I|})(2 - e^{-\lambda a} - e^{-\lambda(T-b)}). \end{aligned}$$

*Proof.* We can consider the Poisson point process  $\mathcal{P}$  on  $[0, T]$  as the union of three independent Poisson point processes:  $\mathcal{P}_a$  on  $[0, a]$ ,  $\mathcal{P}_I$  on  $I = [a, b]$  and  $\mathcal{P}_b$  on  $[b, T]$ . Let  $t_1 \leq \dots \leq t_{N_a}$  (resp.  $t'_1 \leq \dots \leq t''_{N_b}$ ) be the points of  $\mathcal{P}_a$  (resp.  $\mathcal{P}_b$ ). Then  $\mathbf{E}[\sum_J |J|K_{IJ}] = \mathbf{E}[(a - t_{N_a}) + |I| + (t''_1 - b)]$ . For every integer  $k \geq 0$ , conditionally to  $N_a = k$ , the random variable  $t_{N_a}$  has the same law as the last order statistic  $U_{(k)}$  of a sequence  $U_1, \dots, U_k$  of i.i.d. uniformly in  $[0, a]$  distributed random variables. Therefore,  $\mathbf{E}[a - t_{N_a} | N_a = k] = a/(k + 1)$  and

$$\mathbf{E}[a - t_{N_a}] = \sum_{k=0}^{\infty} \frac{(a\lambda)^k a}{k!(k+1)} e^{-a\lambda} = \frac{1 - e^{-a\lambda}}{\lambda}.$$

The same arguments yield  $\mathbf{E}[t''_1 - b] = \lambda^{-1}(1 - e^{-(T-b)\lambda})$  and the first assertion of the lemma follows.

Using the same notation, we have  $\sum_J |J \setminus I| \cdot |J \cap I| = (a - t_{N_a})(t'_1 - a) + (b - t'_{N_I})(t'_1 - b)$ , where  $t'_1 \leq \dots \leq t'_{N_I}$  are the points of  $\mathcal{P}$  lying in  $I$ . Using the conditional independence of  $t_{N_a}$ ,  $(t'_1, t'_{N_I})$  and  $t''_1$  given  $N_a$ ,  $N_I$  and  $N_b$ , as well as the representation by means of order statistics of the uniform distribution we get the second assertion of the lemma.  $\square$

**Lemma 16.** *Let  $t > 0$  and let  $\mathcal{P}$  be a Poisson process on  $[0, t]$  with intensity  $\lambda$ . We denote by  $\Pi$  the random partition of  $[0, t]$  generated by  $\mathcal{P}$ . For every continuous function  $h : [0, t]^2 \rightarrow \mathbb{R}$ , it holds*

$$\lambda \sum_{I \in \Pi} \int_{I \times I} h(s, s') ds ds' \xrightarrow[\lambda \rightarrow \infty]{L^1(P)} 2 \int_0^t h(s, s) ds.$$

*Proof.* Let  $K$  be a positive integer and let us denote by

$$w_h(\delta) = \max\{|h(s, s') - h(u, u')| : (s, s', u, u') \in [0, T]^4 \text{ and } |s - u| \leq \delta, |s' - u'| \leq \delta\} \quad (40)$$



the modulus of continuity of  $h$ . Since  $h$  is continuous and  $[0, t]^2$  is compact, we have  $w_h(t/K) \rightarrow 0$  as  $K \rightarrow \infty$ .

It holds  $\lambda \sum_{I \in \Pi} \int_{I \times I} h(s, s') ds ds' = 2 \int_0^t h(s, s) ds + \mathcal{T}_1 + \mathcal{T}_2 + \mathcal{T}_3$  with

$$\begin{aligned} \mathcal{T}_1 &= \lambda \sum_{I \in \Pi} \int_{I \times I} h(s, s') ds ds' - \lambda \sum_{i=1}^K h\left(\frac{it}{K}, \frac{it}{K}\right) \sum_{I \in \Pi_i^K} |I|^2, \\ \mathcal{T}_2 &= \sum_{i=1}^K h\left(\frac{it}{K}, \frac{it}{K}\right) \left( \lambda \sum_{I \in \Pi_i^K} |I|^2 - \frac{2t}{K} \right), \\ \mathcal{T}_3 &= 2 \sum_{i=1}^K \frac{t}{K} h\left(\frac{it}{K}, \frac{it}{K}\right) - 2 \int_0^t h(s, s) ds, \end{aligned}$$

where  $\Pi_i^K$  is the restriction of the Poisson process  $\mathcal{P}$  on the interval  $[(i-1)t/K, it/K]$ . For the first term, easy algebra yields

$$\mathbf{E}[|\mathcal{T}_1|] \leq \lambda \|h\|_\infty \mathbf{E} \left[ \sum_{I \in \Pi} |I|^2 - \sum_{i=1}^K \sum_{I \in \Pi_i^K} |I|^2 \right] + \lambda w_h(t/K) \sum_{i=1}^K \mathbf{E} \left[ \sum_{I \in \Pi_i^K} |I|^2 \right].$$

This inequality combined with Lemma 13 implies that

$$\limsup_{\lambda \rightarrow \infty} \mathbf{E}[|\mathcal{T}_1|] \leq \limsup_{\lambda \rightarrow \infty} \left( \lambda \|h\|_\infty \frac{K}{\lambda^2} + \lambda w_h(t/K) \frac{2t}{\lambda} \right) = 2t w_h(t/K).$$

In order to bound  $\mathbf{E}[|\mathcal{T}_2|]$ , we evaluate  $\mathbf{E}[|\lambda \sum_{I \in \Pi_i^K} |I|^2 - \frac{2t}{K}|]$ . The value of this term being independent of  $i$ , we only evaluate the term corresponding to  $i = 1$ . Let  $\{\zeta_j, j \in \mathbb{N}\}$  be a family of iid exponential random variables with scaling parameter one and let  $N = \min\{k : \zeta_1 + \dots + \zeta_k \geq npt/K\}$ . Then

$$\left| \lambda \sum_{I \in \Pi_1^K} |I|^2 - \frac{2t}{K} \right| \leq \frac{1}{\lambda} \left| \sum_{j=1}^N (\zeta_j^2 - 2) \right| + \left| \frac{2(N-1)}{\lambda} - \frac{2t}{K} \right| + \frac{\zeta_N^2 + 2}{\lambda}.$$

Note that  $\mathbf{E}[\zeta_N^2] = 6$  by virtue of Lemma 12. In view of the Cauchy-Schwarz inequality and Wald's identity [27, Ch. VII, Thm. 3, Eq. (15)], we get  $\mathbf{E}[|\sum_{j=1}^N (\zeta_j^2 - 2)|] \leq [\text{Var}(\zeta_1^2) \mathbf{E}(N)]^{1/2} = O(\lambda^{1/2})$ . Finally, it is clear that  $|\mathcal{T}_3| \leq 2t w_h(t/K)$ . Putting these estimates together, we get  $\limsup_{\lambda \rightarrow \infty} \mathbf{E}[|\mathcal{T}_1 + \mathcal{T}_2 + \mathcal{T}_3|] \leq 4t w_h(t/K)$ . Using the fact that  $w_h(t/K)$  tends to zero as  $K \rightarrow \infty$ , we arrive at the desired result.  $\square$

**Lemma 17.** *Let  $t > 0$  and let  $\mathcal{P}^i, i = 1, 2$ , be two Poisson processes on  $[0, t]$  with intensities  $\lambda_i, i = 1, 2$ . Let  $\Pi^i$  be the random partition of  $[0, t]$  generated by  $\mathcal{P}^i, i = 1, 2$  and let  $\lambda_0 = \lambda_1 \lambda_2 / (\lambda_1 + \lambda_2)$ . For every continuous function  $h : [0, t]^2 \rightarrow \mathbb{R}$  there exists a constant  $C > 0$  such that for every  $x \in [C \log \lambda_0, C \lambda_0^{1/6}]$  the inequality*

$$\mathbf{P} \left( \left| \lambda_0 \sum_{I, J} K_{IJ} \int_{I \times J} h(s, s') - 2 \int_0^t h(s, s) ds \right| \geq \frac{x}{\sqrt{\lambda_0}} + Cx \left( \frac{1}{\lambda_0} + w_h\left(\frac{x}{\lambda_0}\right) \right) \right) \leq C \lambda_0 e^{-x/C}$$

holds for sufficiently large  $\lambda_0$ , with  $w_h(\cdot)$  being defined by (40).

*Proof.* W.l.o.g. we assume that  $t = 1$ . Set  $\mathcal{T} = \lambda_0 \sum_{I \in \Pi^1, J \in \Pi^2} K_{IJ} \int_{I \times J} h(s, s') ds ds'$  and  $\bar{h}(s) = h(s, s)$ . Let us denote by  $N(x) = \lceil \lambda_0/x \rceil$  the smallest positive integer such

that  $N(x)x > \lambda_0$  and let us set  $L_i = [iN(x)^{-1}, (i+1)N(x)^{-1}]$ . The intervals  $L_i$  define a uniform deterministic partition of  $[0, 1]$  with a mesh-size of order  $x/\lambda_0$ . Let  $\mathcal{E}$  be the event “for every  $i = 1, \dots, 4N(x)$ , the interval  $[\frac{i}{4N(x)}, \frac{(i+1)}{4N(x)}]$  contains at least one point from  $\Pi^1$  and one point from  $\Pi^2$ ”. The total probability formula implies that

$$\mathbf{P}\left(\left|\mathcal{T} - 2 \int_0^1 \bar{h}(s) ds\right| \geq \frac{x}{\sqrt{\lambda_0}}\right) \leq \mathbf{P}\left(\left|\mathcal{T} - 2 \int_0^1 \bar{h}(s) ds\right| \geq \frac{x}{\sqrt{\lambda_0}} \mid \mathcal{E}\right) + \mathbf{P}(\mathcal{E}^c),$$

where  $\mathcal{E}^c$  denotes the complementary event of  $\mathcal{E}$ . Easy computations show that  $\mathbf{P}(\mathcal{E}^c) \leq C\lambda_0 x^{-1} e^{-x/C}$  for some  $C > 0$ .

Let now  $l_i$  be a point in  $L_i$  such that  $\int_{L_i} \bar{h}(t) dt = \bar{h}(l_i)|L_i|$  and let  $a_I$  be the left endpoint of  $I$ . We define the random variables

$$\eta_i^\circ = \lambda_0 \bar{h}(l_i) \sum_{I, J} |I||J| K_{IJ} \mathbf{1}_{\{a_I \in L_i\}}, \quad i = 1, \dots, N(x),$$

and write  $\mathcal{T}_1 = \mathcal{T}_{11} + \mathcal{T}_{12} + \mathcal{T}_{13} + \mathcal{O}(\lambda_0 |L_1| w_h(|L_1|))$  on  $\mathcal{E}$ , where

$$\begin{aligned} \mathcal{T}_{11} &= \mathbf{E}^\mathcal{E} \left[ \sum_{i=1}^{N(x)} \eta_i^\circ \right] - 2 \int_0^1 \bar{h}(s) ds \\ \mathcal{T}_{12} &= \sum_{i=1}^{[N(x)/2]} (\eta_{2i}^\circ - \mathbf{E}^\mathcal{E}[\eta_{2i}^\circ]) \\ \mathcal{T}_{13} &= \sum_{i=0}^{[N(x)/2]} (\eta_{2i+1}^\circ - \mathbf{E}^\mathcal{E}[\eta_{2i+1}^\circ]). \end{aligned}$$

For evaluating the remainder term in  $\mathcal{T}_1$ , we have used the fact that  $r = \max_{I \in \Pi^1} |I| \vee \max_{J \in \Pi^2} |J| \leq |L_1|/2$  on  $\mathcal{E}$ .

On the one hand, since  $|\sum_{i=1}^{N(x)} \eta_i^\circ| \leq C\lambda_0 r$ , we have

$$\left| \mathbf{E}^\mathcal{E} \left[ \sum_{i=1}^{N(x)} \eta_i^\circ \right] - \mathbf{E} \left[ \sum_{i=1}^{N(x)} \eta_i^\circ \right] \right| \leq \frac{\lambda_0 \mathbf{E}[r \mathbf{1}_{\mathcal{E}^c}]}{\mathbf{P}(\mathcal{E})}.$$

Using the inequality of Cauchy-Schwarz, as well as the bounds  $\mathbf{P}(\mathcal{E}^c) \leq C\lambda_0 e^{-x/C}$  and (22), we get  $|\mathbf{E}^\mathcal{E}[\sum_{i=1}^{N(x)} \eta_i^\circ] - \mathbf{E}[\sum_{i=1}^{N(x)} \eta_i^\circ]| \leq C\lambda_0 e^{-x/C}$ , for some constant  $C$  and for every  $x > C \log \lambda_0$ .

On the other hand, in view of Lemma 15, we have

$$\mathbf{E}[\eta_i^\circ] \leq \lambda_0 \bar{h}(l_i) \mathbf{E} \left[ \sum_{I: a_I \in L_i} \left( |I|^2 + \frac{2|I|}{\lambda_2} \right) \right] \leq C\lambda_0 \bar{h}(l_i) (\lambda_1^{-1} + \lambda_2^{-1}) |L_i| = \mathcal{O}(x\lambda_0^{-1}).$$

Using once again Lemma 15, we get

$$\begin{aligned} \mathbf{E} \left[ \sum_{i=1}^{N(x)} \eta_i^\circ \right] &= \sum_{i=2}^{N(x)-1} \lambda_0 \bar{h}(l_i) \mathbf{E} \left[ \sum_{I: a_I \in L_i} |I| \cdot \mathbf{E}^{\Pi^1} \left( \sum_{j \in \Pi^2} K_{IJ} |J| \right) \right] + \mathcal{O}(x\lambda_0^{-1}) \\ &= \sum_{i=2}^{N(x)-1} \lambda_0 \bar{h}(l_i) \mathbf{E} \left[ \sum_{I: a_I \in L_i} (|I|^2 + 2|I|/\lambda_2) \right] + \mathcal{O}(x\lambda_0^{-1}). \end{aligned}$$

Wald's equality yields

$$\mathbf{E} \left[ \sum_{I: a_I \in L_i} |I|^k \right] = k! |L_i| \lambda_1^{1-k} + \mathcal{O}(\lambda_1^{-k}), \quad (41)$$

for every  $k > 0$  and for every  $i \leq N(x) - 1$ . Putting all these estimates together, we get

$$\begin{aligned} \mathbf{E} \left[ \sum_{i=1}^{N(x)} \eta_i^\circ \right] &= \sum_{i=2}^{N(x)-1} n \lambda_0 \bar{h}(l_i) \left( \frac{2|L_i|}{\lambda_1} + \frac{2|L_i|}{\lambda_2} \right) + \mathcal{O}(x \lambda_0^{-1}) \\ &= \sum_{i=1}^{N(x)} 2 \bar{h}(l_i) |L_i| + \mathcal{O}(x \lambda_0^{-1}). \end{aligned}$$

Since  $l_i$  is chosen to verify  $\bar{h}(l_i) |L_i| = \int_{L_i} \bar{h}(t) dt$ , we get  $\mathcal{T}_{11} = \mathcal{O}(x \lambda_0^{-1})$ .

The advantage of working with  $\eta_i^\circ$ s is that, conditionally to  $\mathcal{E}$ , the random variables  $\eta_{2i}^\circ$ ,  $i = 1, \dots, [N(x)/2]$ , are independent. Indeed, one easily checks that conditionally to  $\mathcal{E}$ ,  $\eta_{2i}^\circ$  depends only on the restrictions of  $\mathcal{P}^1$  and  $\mathcal{P}^2$  onto the interval  $[\frac{(4i-1)}{2N(x)}, \frac{(4i+3)}{2N(x)}]$ . Since these intervals are disjoint for different values of  $i \in \mathbb{N}$ , the restrictions of Poisson processes  $\mathcal{P}^k$ ,  $k = 1, 2$ , onto these intervals are independent. Therefore,  $\eta_{2i}^\circ$ ,  $i = 1, \dots, [N(x)/2]$ , form a sequence of random variables that are independent conditionally to  $\mathcal{E}$ . Moreover, conditionally to  $\mathcal{E}$ , they verify  $|\eta_i^\circ| \leq C \lambda_0 r |L_i| \leq C x^2 / \lambda_0$ . One can also check that  $\mathbf{E}^\mathcal{E}[(\eta_i^\circ)^2] = \mathcal{O}(x^2 \lambda_0^{-2})$ .

These features enable us to use the Bernstein inequality in order to bound large deviations of  $\mathcal{T}_{12}$  as follows:

$$\mathbf{P}^\mathcal{E} (|\mathcal{T}_{12}| \geq x / \sqrt{\lambda_0}) \leq 2 \exp \left( - \frac{x^2 / (2\lambda_0)}{C(N(x)x^2\lambda_0^{-2} + x^3\lambda_0^{-3/2})} \right) \leq 2e^{-x/C}, \quad \forall x \in [1, \lambda_0^{1/6}]$$

Obviously, the same inequality holds true for the term  $\mathcal{T}_{13}$ . These inequalities combined with the bound on the deterministic error term  $\mathcal{T}_{11}$  complete the proof.  $\square$

**Lemma 18.** *Let  $T > 0$  and let  $\mathcal{P}_n^i$ ,  $i = 1, 2$ , be two Poisson processes on  $[0, T]$  with intensities  $np_i$ ,  $i = 1, 2$ . For every continuous function  $h : [0, T]^3 \rightarrow \mathbb{R}$ , it holds*

$$n^2 \sum_{I \in \Pi_n^1} \int_{I \times J(I) \times J(I)} h(s, t, u) ds dt du \xrightarrow[n \rightarrow \infty]{p} \left( \frac{6}{p_1^2} + \frac{8}{p_1 p_2} + \frac{6}{p_2^2} \right) \int_0^T h(s, s, s) ds.$$

*Proof.* Let us denote  $\mathcal{T}_n = n^2 \sum_{I \in \Pi_n^1} \int_{I \times J(I) \times J(I)} h(s, t, u) ds dt du$  and let us consider the uniform partition  $\{L_i = [(i-1)/N, i/N), i = 1, \dots, N\}$  with  $N = [n^{1-\varepsilon}]$  slightly smaller than  $n$  ( $\varepsilon$  is a small positive number). For every integer  $i$  smaller than  $[n^{1-\varepsilon}]$ , we define  $l_i$  as the real number such that  $\bar{h}(l_i) = |L_i|^{-1} \int_{L_i} \bar{h}_s ds$ , where  $\bar{h}_s = h(s, s, s)$ . The continuity of  $h$  implies that

$$\mathcal{T}_n = n^2 (1 + o(1)) \sum_{i=1}^N \sum_I \bar{h}(l_i) |I| |J(I)|^2 \mathbf{1}_{L_i}(a_I).$$

For every  $i$ , we set  $\eta_i^\circ = n^2 \sum_I \bar{h}(l_i) |I| |J(I)|^2 \mathbf{1}_{L_i}(a_I)$ . We first remark that

$$\mathbf{E} \left[ \sum_I \bar{h}(l_i) |I| |J(I)|^2 \mathbf{1}_{L_i}(a_I) \right] = N^{-1} \mathcal{O}(\mathbf{E}[r_n^2]), \quad \forall i = 1, \dots, N.$$

Let now  $i \in 2, \dots, N-1$  and  $I$  be an interval of  $\Pi^1$  satisfying  $a_I \in L_i$ , then  $||J(I)| - |I| - \xi_1^\circ - \xi_2^\circ| \leq (\xi_1^\circ - N^{-1})_+ + (\xi_2^\circ - N^{-1})_+$ , where  $\xi_1^\circ$  and  $\xi_2^\circ$  are two random variables distributed according to the exponential distribution with parameters  $np_2$  conditionally to  $\Pi^1$ . Moreover, conditionally to  $\Pi^1$ ,  $\xi_1^\circ$  and  $\xi_2^\circ$  are independent. Since  $N = O(n^{1-\varepsilon})$  and  $\mathbf{E}^{\Pi^1}[(\xi_j^\circ)^4] = O(n^{-4})$ , by the Cauchy-Schwarz inequality we have  $\mathbf{E}^{\Pi^1}[(\xi_j^\circ - N^{-1})_+^2] = O(n^{-2-4\varepsilon})$  for  $j = 1, 2$ . This implies that  $\mathbf{E}^{\Pi^1}[|J(I)|^2] = |I|^2 + 4|I|(np_2)^{-1} + 6(np_2)^2 + O(|I|n^{-1-2\varepsilon})$ . Combining this estimate with (41), we get

$$\mathbf{E}[\eta_i^\circ] = \bar{h}(l_i)|L_i| \left( \frac{6}{p_1^2} + \frac{8}{p_1 p_2} + \frac{6}{p_2^2} \right) + n^2 |L_i| O(n^{-1-2\varepsilon}) = \left( \frac{6}{p_1^2} + \frac{8}{p_1 p_2} + \frac{6}{p_2^2} \right) \int_{L_i} \bar{h}(s) ds + o(1).$$

By reasoning in a similar way, we get  $\mathbf{E}[\eta_i^\circ \eta_j^\circ] - \mathbf{E}[\eta_i^\circ] \mathbf{E}[\eta_j^\circ] = o(|L_i|^2)$  as soon as  $|i-j| > 2$ . Standard arguments imply that  $\text{Var}[\sum_i \eta_i^\circ] = O(N \max_i \text{Var}(\eta_i^\circ)) + o(N^2 |L_1|^2)$ . Since  $|\eta_i^\circ| \leq C(nr_n)^2 |L_1|$  for every  $i$ , we get  $\text{Var}[\sum_i \eta_i^\circ] = O(N |L_1|^2 \mathbf{E}[(nr_n)^4]) + o(1) = o(1)$  and the desired convergence property follows from the convergence of  $\mathcal{T}_n$  in  $L^2$ .  $\square$

**Lemma 19.** *Let  $T > 0$  and let  $\mathcal{P}_n^i$ ,  $i = 1, 2$ , be two Poisson processes on  $[0, T]$  with intensities  $np_i$ ,  $i = 1, 2$ . There is a constant  $\nu(p_1, p_2)$  depending only on  $p_1$  and  $p_2$  such that for every continuous function  $h : [0, T]^3 \rightarrow \mathbb{R}$*

$$n^2 \sum_{I \in \Pi_n^1} \sum_{J \in \Pi_n^2} \int_{I(J) \times J(I) \times I \cap J} h(s, t, u) ds dt du \xrightarrow[n \rightarrow \infty]{p} \nu(p_1, p_2) \int_0^T h(s, s, s) ds.$$

*Proof.* The proof of this lemma follows from the invariance of the law of a Poisson process under scaling and translation, as well as from the independence of disjoint sets' measures. It is similar to the proofs of preceding lemmas and therefore will be omitted.  $\square$

## REFERENCES

- [1] ANDERSEN, T.G. AND BOLLERSLEV, T. Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review*, **39** (1998), 885-905.
- [2] ANDERSEN, T.G., BOLLERSLEV, T., DIEBOLD, F.X. AND EBENS, H. The distribution of realized stock return volatility. *J. Financial Economics*, **61** (2001), 43-76.
- [3] ANDERSEN, T.G., BOLLERSLEV, T., DIEBOLD, F.X. AND LABYS, P. The distribution of realized exchange rate volatility. *J. Amer. Statist. Assoc.*, **96** (2001), 42-55.
- [4] BARNDORFF-NIELSEN, O.E. AND SHEPHARD, N. Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *J. Royal Statist. Soc., Ser. B*, **64** (2002), 253-280.
- [5] BHATTACHARYA, R.N. AND RANGA RAO, R. *Normal approximation and asymptotic expansions*. 2nd ed. New York: Wiley 1986.
- [6] COMTE, F. AND RENAULT, E. Long memory in continuous-time stochastic volatility models. *Math. Finance*, **8** (1998), 291-323.
- [7] DACUNHA-CASTELLE, D. AND FLORENS-ZMIROU, D. Estimation of the coefficients of diffusion from discrete observations. *Stochastics*, **19** (1986), 263-284.
- [8] FLORENS-ZMIROU, D. On estimating the diffusion coefficient from discrete observations. *J. Appl. Probab.* **30** (1993), no. 4, 790-804.
- [9] GENON-CATALOT, V. AND JACOD, J. On the estimation of the diffusion coefficient for multi-dimensional diffusion processes. *Ann. Inst. Henri Poincaré* **29** (1993), 1, 119-151.
- [10] GRIFFIN, J.E. AND OOMEN, R.C.A. Covariance measurement in the presence of non-synchronous trading and market microstructure noise. Preprint (2006), available at SSRN: <http://ssrn.com/abstract=912541>.

- [11] HAYASHI, T. AND KUSUOKA, S. Nonsynchronous cvariation measurement for continuous semimartingales. preprint (2004).
- [12] HAYASHI, T. AND YOSHIDA, N. On covariance estimation of nonsynchronously observed diffusion processes. preprint (2003). *Bernoulli*, 11, 359-379 (2005).
- [13] HAYASHI, T. AND YOSHIDA, N. Asymptotic normality of nonsynchronous covariance estimators for diffusion processes. preprint (2004).
- [14] HAYASHI, T. AND YOSHIDA, N. Nonsynchronous covariance estimator and limit theorem. preprint (2006).
- [15] HOSHIKAWA, T., KANATANI, T., NAGAI, K. AND NISHIYAMA, Y. Nonparametric estimation methods of integrated multivariate volatilities. Working Paper (2006).
- [16] JACOD, J. On processes with conditional independent increments and stable convergence in law. *Seminaire de probabilites de Strasbourg*, 36 (2002), 383-401.
- [17] JACOD, J. AND SHIRYAEV, A. *Limit theorems for stochastic processes*. Second edition. Springer-Verlag, Berlin, 2003.
- [18] KESSLER, M. Estimation of diffusion processes from discrete observations. *Scand. J. Statist.* **24** (1997), 211-229.
- [19] MALLIAVIN, P. AND MANCINO, M.E. Fourier Series Method for Measurement of Multivariate Volatilities. *Finance Stoch.* **6** (2002), 49-61.
- [20] MYKLAND, P.A. A Gaussian Calculus for Inference from High Frequency Data. Technical Report 563, Dept of Statistics, The Univ. of Chicago.
- [21] MYKLAND, P.A. AND ZHANG, L. Anova for diffusions and Ito processes. to appear in *Ann. Statist.* **34** (2006), 1931-1963.
- [22] NUALART, D. *The Malliavin calculus and related topics*. Second edition. Probability and its Applications (New York). Springer-Verlag, Berlin, 2006.
- [23] PRAKASA-RAO, B. L. S. Asymptotic theory for non-linear least square estimator for diffusion processes. *Math. Operationsforsch. Statist. Ser. Stat.*, **14** (1983), 195-209.
- [24] PRAKASA-RAO, B. L. S. Statistical inference from sampled data for stochastic processes. *Contemp. Math.*, **80** (1988), 249-284.
- [25] REVUZ, D. AND YOR, M. *Continuous martingales and Brownian motion*. Third edition. Springer-Verlag, Berlin, 1999.
- [26] SAKAMOTO, Y. AND YOSHIDA, N. Asymptotic expansion under degeneracy. *J. Japan Stat. Soc.*, **33** (2003), 145-156.
- [27] SHIRYAEV, A. N. *Probability*. Second edition. Graduate Texts in Mathematics, 95. Springer-Verlag, New York, 1996.
- [28] SWEETING, T.J. Speeds of convergence for the multidimensional central limit theorem. *Ann. Probab.* **5** (1977), 28-41.
- [29] VOEV, V. AND LUNDE, A. Integrated covariance estimation using high-frequency data in the presence of noise. Working Paper, presented at *CIREQ Conference on Realized Volatility* 2006.
- [30] YOSHIDA, N. Estimation for diffusion processes from discrete observation. *J. Multivariate Anal.* **41** (1992), 220-242.
- [31] YOSHIDA, N. Malliavin calculus and asymptotic expansion for martingales. *Probab. Theory Related Fields*, 109 (1997), 301-342.