Second-order asymptotic expansion for a non-synchronous covariation estimator

Arnak S. Dalalyan, Nakahiro Yoshida

Submitted to the Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
arXiv: 0804.0676

Second-order asymptotic expansion for a non-synchronous covariation estimator

Arnak Dalalyan[a] and Nakahiro Yoshida[b]

[a] *LIGM/IMAGINE, Ecole des Ponts ParisTech, Université Paris-Est, E-mail:* dalalyan@imagine.enpc.fr

[b] *University of Tokyo and Japan Science and Technology Agency, E-mail:* nakahiro@ms.u-tokyo.ac.jp

Abstract. In this paper, we consider the problem of estimating the covariation of two diffusion processes when observations are subject to non-synchronicity. Building on recent papers [18, 19], we derive second-order asymptotic expansions for the distribution of the Hayashi-Yoshida estimator in a fairly general setup including random sampling schemes and non-anticipative random drifts. The key steps leading to our results are a second-order decomposition of the estimator's distribution in the Gaussian set-up, a stochastic decomposition of the estimator itself and an accurate evaluation of the Malliavin covariance. To give a concrete example, we compute the constants involved in the resulting expansions for the particular case of sampling scheme generated by two independent Poisson processes.

Résumé. Dans cet article, nous considérons le problème d'estimation de la covariation de deux processus de diffusion observés de façon asynchrone. Nous nous plaçons dans le cadre présenté dans [18, 19] et établissons un développement asymptotique au second ordre de la loi de l'estimateur de Hayashi-Yoshida. Ce développement est valable pour les drifts aléatoires non-anticipatifs et pour des pas d'échantillonnage irréguliers, éventuellement aléatoires, mais indépendant des processus observés. L'approche utilisée pour obtenir les principaux résultats peut être décomposée en trois étapes. La première consiste à établir un développement au second-ordre de la loi de l'estimateur dans le cadre Gaussien. La deuxième est l'obtention d'une décomposition stochastique de l'estimateur lui-même et la dernire est l'évaluation de la covariance de Malliavin. A titre d'exemple, nous calculons les constantes du développement au second ordre dans le cas où l'échantillonnage est obtenu par deux processus de Poisson indépendants.

Keywords: Edgeworth expansion, covariation estimation, diffusion process, asynchronous observations, Poisson sampling.

1. Introduction

In the last decade, studies on covariance estimation has attracted considerable attention thanks to the applications in mathematical finance and econometrics; see *e.g.* Andersen and Bollerslev [1], Comte and Renault [9], Andersen *et al.* [2, 3], Barndorff-Nielsen and Shephard [6]. All these papers consider the situation where two diffusion processes are observed at the same discrete instants. In contrast with this, covariance estimation under a "non-synchronous" sampling scheme has rarely been treated theoretically in spite of its importance in the analysis of high-frequency financial data [39, 27, 41]. The first contributions to the statistical inference for covariance estimation with non-synchronous data have been made by Hayashi and Yoshida [18, 19]. They proposed an estimator of the covariation and explored its statistical properties such as the consistency

and the asymptotic normality. Interestingly, it follows from the results in [19] that the drifts of the observed diffusions do not affect the asymptotic variance of the covariance estimator. The aim of the present paper is to complement the results in [18, 19] by establishing a second-order asymptotic expansion for the distribution of the covariance estimator. In particular, we get explicit expressions that have the advantage of reflecting the impact of drifts on the asymptotic distribution of the estimator.

One common approach to cope with non-synchronicity is the following. First, two regularly spaced time series are generated by interpolating the observed non-synchronous data. Then the realized covariance estimator is computed for the interpolated time series. However, it is known that such a synchronization technique causes estimation bias, which is often referred to as the *Epps effect* [11]. Another estimator of the covariance, based on the harmonic analysis, has been proposed by Malliavin and Mancino [28]. In the case where in addition to the non-synchronicity the data is contaminated by a microstructure noise, estimators of the covariance have been proposed by Palandri [33], Barndorff-Nielsen *et al.* [5] and Zhang [46]. A detailed account on covariance estimation for non-synchronous data can be found in [20] and [46].

In order to present the framework and to describe our contributions, we need some notation. Let $\boldsymbol{X} = (X_1, X_2)$ be a two dimensional diffusion process given by

$$d\boldsymbol{X}_t = \boldsymbol{\beta}_t\, dt + \mathrm{diag}(\boldsymbol{\sigma}_t)\, d\boldsymbol{B}_t, \tag{1}$$

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

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

In (1), $\boldsymbol{\beta} = (\beta_1, \beta_2)^{\mathbf{T}}$ is a progressively measurable process, $\boldsymbol{\sigma} = (\sigma_1, \sigma_2)^{\mathbf{T}}$ is a deterministic function and $\mathrm{diag}(\boldsymbol{\sigma})$ stands for the diagonal matrix having σ_i as i^{th} diagonal entry, $i = 1, 2$. In what follows, we restrict our attention to the case when σ_1, σ_2 and ρ are deterministic functions; the functions σ_i, $i = 1, 2$ take positive values while ρ 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 the time instants $0 = S^0 < S^1 < \ldots < S^{N_1} = T$ and $0 = T^0 < \ldots < T^{N_2} = T$. Let us denote $I^i = (S^{i-1}, S^i]$ and $J^j = (T^{j-1}, T^j]$. The families $\Pi^1 = \{I^i, i = 1, \ldots, N_1\}$ and $\Pi^2 = \{J^j, j = 1, \ldots, N_2\}$ 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 $\Delta_j X_2 = X_{2,T^j} - X_{2,T^{j-1}}$.

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 $(X_{1,S^i}, X_{2,T^j}, i = 0, \ldots, N_1, j = 0, \ldots, N_2)$. The parameter θ represents the covariance between the 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 processes X_1 and X_2 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 θ. Indeed, it converges in probability to θ when the maximum lag of the sampling times tends to 0 in probability. In the field of statistical inference for stochastic processes, this fact has been applied to estimating the volatility and the covariation between semimartingales. The asymptotic distributions are well investigated; see Dacunha-Castelle and Florens-Zmirou [10], Florens-Zmirou [12], Prakasa Rao [34, 35], Yoshida [44], Genon-Catalot and Jacod [14], Kessler [26], and Mykland and Zhang [31].

An estimator of θ, which is unbiased when the drift $\boldsymbol{\beta}$ is identically zero, has been proposed in [18]. Henceforth called HY-estimator, it is defined as follows:

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

It is established in [18] that under mild assumptions, $\hat{\theta}$ is consistent as the maximum lag of the sampling times tends to 0 in probability. Kusuoka and Hayashi [17] extended the consistency result to a more general sampling scheme. Asymptotic normality of the HY-estimator was proved in Hayashi and Yoshida [19] under the assumption that the sampling times are independent of the process $\boldsymbol{X}$. For related literature, see Hoshikawa *et al.* [22], Griffin and Oomen [15], Robert and Rosenbaum [37] and Voev and Lunde [43]. The general case of a sampling scheme depending on the process $\boldsymbol{X}$ has been studied in Hayashi and Yoshida [20, 21], where a stochastic analytic proof of the asymptotic mixed normality of the HY-estimator is presented. An estimator for the variance of the HY-estimator under the assumption that the observed process $\boldsymbol{X}$ has no drift has been recently proposed by Mykland [30].

In the present work, the main emphasis is put on the higher-order asymptotic behavior of the HY-estimator. Note that the theory of asymptotic expansions is one of chapters of statistics that received a revival of interest owing to its usefulness for exploring properties of bootstrap-based statistical methods. For a comprehensive introduction to this subject we refer the reader to Hall [16]. Results on asymptotic expansions in other contexts can be found in Bose [8], Mykland [29], Koul and Surgailis [25], Bertail and Clémençon [7], Zhang et al. [47], Fukasawa [13] and the references therein.

Section 3 contains an asymptotic expansion of the distribution of the HY-estimator. As a first step for deriving asymptotic expansions for the distribution of the HY-estimator, we give in Section 3.2 a representation of the cumulants of $\hat{\theta}$ as functionals of the sampling times, and obtain asymptotic estimates for them. This is used to derive a second-order asymptotic expansion of the characteristic function of the estimator while the asymptotic normality is also proved as an application of those estimates.

The application of these results in the setup of Poisson sampling schemes is presented in Section 4. We assume that the Poisson processes generating the sampling times have constant intensities np_1 and np_2, where n is a parameter guaranteeing the high-frequency of the observations ($n \to \infty$). This setup has the advantage of making it possible to compute all the quantities involved in the asymptotic expansion. We show that the residual term in the proposed asymptotic expansion of the distribution of $\sqrt{n}(\hat{\theta}_n - \theta)$ behaves nearly like n^{-1}, as n goes to infinity.

When there are (possibly random) drift terms in the stochastic differential equation of $\boldsymbol{X}_t$, some additional terms appear in the asymptotic expansion. In order to identify these terms, we derive in Section 5 a stochastic decomposition of the HY-estimator and explore the asymptotic behavior of the variables appearing in the second-order terms. Since the asymptotics we get is non-Gaussian, the classical techniques leading to Edgeworth expansions can not be used. Instead, our arguments rely on the limit theory for semimartingales.

The asymptotic expansion of the distribution of the HY-estimator is carried out in Section 6 using a perturbation method. We apply the Malliavin calculus first to ensure the regularity of the distribution of the principal part—a quadratic form of Gaussian random variables—and then to extend this property to the model under the perturbation. To enhance the legibility, we postpone the most technical proofs to the last three sections.

2. Elementary properties of $\hat{\theta}$

As noticed by Mykland [30], the estimator $\hat{\theta}$ is the Maximum Likelihood Estimator (MLE) of θ. Let us present here some computations that not only show that $\hat{\theta}$ is the MLE of θ, 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 Σ. The entries of the matrix Σ 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 to use results on the exponential family, it is convenient to consider the parametrization by the entries of the inverse, denoted by $V = \Sigma^{-1}$, of the covariance matrix Σ. 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:

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

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. Furthermore, 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 θ 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

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

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

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

for every $\ell \le \ell'$ 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 θ. Moreover, it is the estimator having the smallest quadratic risk among all unbiased estimators of θ.*

This proposition advocates for using the HY-estimator in the case where $\boldsymbol{\beta} \equiv 0$. If the latter condition is not satisfied, $\hat\theta$ is not necessarily unbiased, but under very mild assumptions it is consistent [18] and asymptotically normal [19] as the maximum lag of the sampling times tends to 0. This explains the popularity of the HY-estimator motivating our interest in its second-order asymptotic expansion. At a heuristical level, the construction of the HY-estimator can be derived

from the decomposition $\theta = \sum_{i,j} \mathbf{1}(I^i \cap J^j \neq \varnothing) \int_{I^i \cap J^j} \sigma_{1,t}\sigma_{2,t}\rho_t \, dt$. Indeed, each term of that decomposition is nearly equal to the covariance of the increments $\Delta_i X_1$ and $\Delta_j X_2$, since the martingale part of a small increment of a semi-martingale dominates the increment of the bounded-variation part. Hence, if I^i and J^j are small, it is reasonable to estimate $\int_{I^i \cap J^j} \sigma_{1,t}\sigma_{2,t}\rho_t \, dt$ by the product $\Delta_i X_1 \cdot \Delta_j X_2$ and, therefore, to estimate θ by the HY-estimator $\hat{\theta}$.

3. Asymptotic expansion of the distribution in Gaussian setup

3.1. Notation and main results

In this section, we will derive the second-order asymptotic expansion of the distribution of $b_n^{-1/2}(\hat{\theta}_n - \theta)$, where b_n is a suitably chosen normalization factor, for the model (1) without drifts. We will treat a model with drifts in Section 5, where we will resort to the Malliavin calculus for dealing with general nonlinear Wiener functionals.

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

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

where

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

and $\phi(z;\Sigma)$ is the density of the centered normal distribution with variance Σ. 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 $\int F(z)\phi(z;\mathfrak{c}^*)\,dz \leq \mathsf{C}$ belong to $\mathcal{E}^0(\mathsf{C},\eta,\infty,\mathfrak{c}^*)$. It is also easy to check that the set of all indicator functions of intervals of $\mathbb{R}$ is included in $\mathcal{E}^0(\sqrt{2\pi}\mathfrak{c}^*,1,\infty,\mathfrak{c}^*)$ for any $\mathfrak{c}^* > 0$.

Our aim is now to get uniformly in $f \in \mathcal{E}^*$ an asymptotic expansion for the sequence $\mathbf{E}[f(b_n^{-1/2}(\hat{\theta}_n - \theta))]$ with $\mathcal{E}^* = \mathcal{E}(M,\gamma) \cap \mathcal{E}^0(\mathsf{C},\eta,r_0,\mathfrak{c}^*)$. To this end, define $h_r(z;\Sigma)$ as the r-th Hermite polynomial given by

$$h_r(z;\Sigma) = (-1)^r \phi(z;\Sigma)^{-1}\partial_z^r \phi(z;\Sigma), \qquad \forall z \in \mathbb{R}.$$

In particular, $h_2(z;\Sigma) = (z^2 - \Sigma)/\Sigma^2$ and $h_3(z;\Sigma) = (z^3 - 3\Sigma z)/\Sigma^3$. Along with the Hermite polynomials, it is customary to express the second-order asymptotic expansion of a distribution in terms of the first-order and the second-order cumulants. To define this quantities in the present framework, let us denote, for any Borel set $S \subset \mathbb{R}$,

$$v(S) = \int_S \rho_t\sigma_{1,t}\sigma_{2,t}\,dt, \quad v_1(S) = \int_S \sigma_{1,t}^2\,dt, \quad v_2(S) = \int_S \sigma_{2,t}^2\,dt, \tag{4}$$

and introduce

$$\mu_2 = \frac{1}{2}\Big\{ \sum_{I,J} v_1(I)v_2(J)K_{IJ} + \sum_{I\in\Pi^1} v(I)^2 + \sum_{J\in\Pi^2} v(J)^2 - \sum_{I,J} v(I\cap J)^2 \Big\}, \tag{5}$$

$$\begin{aligned}\mu_3 = \frac{1}{4}\Big\{ &\sum_{I\in\Pi^1} v(I)^3 + \sum_{J\in\Pi^2} v(J)^3 + 2\sum_{I,J} v(I\cap J)^3 + 3\sum_{I,J} v_1(I)v_2(J)v(I\cup J)K_{IJ} \\ &- 3\sum_{I,J}[v(I\cap J)^2(v(I)+v(J)) - v(I\cap J)v(I)v(J)]\Big\},\end{aligned} \tag{6}$$

where $K_{IJ} = \mathbf{1}(I\cap J \neq \varnothing)$ and $\sum_{I,J} = \sum_{I\in\Pi^1}\sum_{J\in\Pi^2}$. Since we are dealing with the asymptotics of high frequency data, we will assume that all the intervals $I^i = I_n^i$ and $J^j = J_n^j$ depend on some

parameter n—representing the frequency of the sampling—that is large. To make the dependence on n explicit, we will write $\mu_{2,n}$ and $\mu_{3,n}$ instead of μ_2 and μ_3. Furthermore, as the time interval $[0,T]$ is fixed, the maximal sampling step $r_n = [(\max_i |I_n^i|) \vee (\max_j |J_n^j|)]$ is assumed to tend to zero as $n \to \infty$. Using this notation, we define

$$\bar\lambda_{2,n} = 2\, b_n^{-1}\mu_{2,n}, \qquad \text{and} \qquad \bar\lambda_{3,n} = 8\, b_n^{-2}\mu_{3,n}, \tag{7}$$

for some deterministic sequence b_n, tending to zero as $n \to \infty$. To some extent, one can think of b_n as the rate of convergence of $\mu_{2,n}$ to zero. This point will become clearer in Section 4, where the concrete example of the Poisson sampling scheme is analyzed.

We introduce a $\sigma[\Pi]$-dependent random signed-measure Ψ_n^Π on $\mathbb{R}$ by the density

$$p_{3,n}(z) \quad = \quad \phi(z;\bar\lambda_{2,n})\Big[1 + \frac{b_n^{1/2}}{6}\bar\lambda_{3,n}\, h_3(z;\bar\lambda_{2,n})\Big].$$

It is not hard to check that the Fourier transform of Ψ_n^Π is given by

$$\hat\Psi_n^\Pi(u) = e^{-\frac{1}{2}\bar\lambda_{2,n}u^2}\Big[1 + \frac{b_n^{1/2}}{6}\bar\lambda_{3,n}(\mathrm{i}u)^3\Big].$$

In the case where no assumption on the convergence of $\mu_{2,n}$ is made, the measure Ψ_n^Π will serve as the second-order approximation to the distribution of $\mathcal{X}_n = b_n^{-1/2}(\hat\theta_n - \theta)$. However, for many sampling schemes one can prove the convergence of $\bar\lambda_{2,n}$ to some constant $\mathfrak{c}$, implying that the estimator $\hat\theta_n$ is asymptotically normal with asymptotic variance $\mathfrak{c}$. It is therefore natural to address the issue of approximating the distribution of $\mathcal{X}_n$ by a measure similar to Ψ_n^Π 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

$$\tilde p_{3,n}(z) = \phi(z;\mathfrak{c})\Big[1 + \frac{1}{2}(\bar\lambda_{2,n} - \mathfrak{c})h_2(z;\mathfrak{c}) + \frac{b_n^{1/2}}{6}\bar\lambda_{3,n}\, h_3(z;\mathfrak{c})\Big].$$

The following result, the proof of which is deferred to Section 7, asserts that $p_{3,n}$ and $\tilde p_{3,n}$ are good approximations to the density of $(\hat\theta_n - \theta)/\sqrt{b_n}$.

Theorem 1. *Let $M, \gamma, \eta, \mathsf{C}, r_0, \mathfrak{c}^* > 0$ be the parameters describing the set of functions of interest. For $a \in (\frac{3}{4}, 1)$ and $\mathfrak{c}, \mathfrak{c}_0, \mathfrak{c}_1 \in (0, \mathfrak{c}^*)$ set*

$$\mathcal{P}_n(\mathfrak{c}_0,\mathfrak{c}_1,a) = \{\, \mathfrak{c}_0 < \bar\lambda_{2,n} < \mathfrak{c}_1,\ r_n \le b_n^a \,\},$$
$$A_n(a) = \{\, (\bar\lambda_{2,n} - \mathfrak{c})^2 \le b_n^{2a-1},\ r_n \le b_n^a \,\},$$

where r_n is the maximal lag of the sampling times and $\bar\lambda_{2,n} = 2b_n^{-1}\mu_{2,n}$. Then, there exists a sequence $\epsilon_n = \epsilon_n(M,\gamma,\eta,\mathsf{C},r_0,a,\mathfrak{c}_0,\mathfrak{c}_1)$ such that $\epsilon_n = O(b_n^{2a-1})$ and the inequalities

$$\sup_{f\in\mathcal{E}(M,\gamma)\cap\mathcal{E}^0(\mathsf{C},\eta,r_0,\mathfrak{c}^*)} \Big|\, \mathbf{E}^\Pi[f(\mathcal{X}_n)] - \Psi_n^\Pi[f]\,\Big| \le \epsilon_n, \qquad \forall \Pi_n \in \mathcal{P}_n(\mathfrak{c}_0,\mathfrak{c}_1,a), \tag{8}$$

$$\sup_{f\in\mathcal{E}(M,\gamma)\cap\mathcal{E}^0(\mathsf{C},\eta,r_0,\mathfrak{c}^*)} \Big|\, \mathbf{E}^\Pi[f(\mathcal{X}_n)] - \tilde\Psi_n^\Pi[f]\,\Big| \le \epsilon_n, \qquad \forall \Pi_n \in A_n(a), \tag{9}$$

hold true, where $\mathcal{X}_n = b_n^{-1/2}(\hat\theta_n - \theta)$.

Remark 1. The approximating measure Ψ_n^Π provided by Theorem 1 contains the Gaussian density with variance $\bar\lambda_{2,n}$, which depends on n. One can 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

$$\Big[1 + \frac{\sqrt{b_n}\bar\lambda_{3,n}}{6}\Big(z^3 - \frac{3}{\bar\lambda_{2,n}}z\Big)\Big]\phi(z;1)\,dz.$$

The following result is an immediate consequence of (9) 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 Theorem 1, 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*

$$\sup_{f \in \mathcal{E}(M,\gamma) \cap \mathcal{E}^0(\mathsf{C},\eta,r_0,\mathfrak{c}^*)} \left| \mathbf{E}[f(\mathcal{X}_n)] - \int_{\mathbb{R}} f(z)\, p_n^*(z)\, dz \right| = O(b_n^{2a-1}), \tag{10}$$

where $p_n^*(z) = \phi(z;\mathfrak{c})\Big[1 + \frac{b_n^{1/2}}{6}\mathbf{E}[\bar{\lambda}_{3,n}]\, h_3(z;\mathfrak{c})\Big]$. *Moreover, if* $\sup_{n\in\mathbb{N}} \mathbf{E}[\bar{\lambda}_{3,n}] < \infty$, *then relation (10) 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.

3.2. Gaussian analysis and expansion of the characteristic function

The goal of this section is to prepare the ground for the proof of Theorem 1. To this end, we present in Section 3.2.1 general results on the characteristic function of a random variable that can be written as a quadratic functional of a standard Gaussian vector. As usual, this characteristic function involves the cumulants that take a simplified form in the context of the HY-estimator. Section 3.2.2 is devoted to proving that the second and the third cumulants for the HY-estimator can be computed using formulae (5) and (6). These results lead to a second-order expansion of the characteristic function of the HY-estimator, which is rigorously stated and proved in Section 3.2.3. Finally, the proof of Theorem 1 is presented in Section 3.3.

3.2.1. General Gaussian setup

In order to determine the asymptotic expansion of the distribution of $\hat{\theta}$, we start with expanding its characteristic function. 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.

Recall that

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

In other terms, $\hat{\theta}$ is a quadratic form of a centered Gaussian vector. The aim of the present subsection is twofold. Firstly, we compute the cumulants of any quadratic form Q of a Gaussian vector $\boldsymbol{\xi}$ as functions of the matrix associated to the quadratic form Q and the covariance matrix of $\boldsymbol{\xi}$. Among other things, this computation allows us to give a simple condition implying the weak convergence of a series of quadratic forms of Gaussian vectors. The second goal of the present subsection is to show that the tails of the characteristic function of a quadratic form of a Gaussian vector have at least polynomial decay. To achieve this second goal, we establish an explicit upper bound for the characteristic function of interest. It should be pointed out that most results and conditions are stated in terms of the spectral characteristics of the matrix $\Sigma^{1/2} A \Sigma^{1/2}$.

Since A is a symmetric matrix, the N-by-N matrix $\Sigma^{1/2} A \Sigma^{1/2}$ is symmetric and therefore diagonalizable. Let Λ 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 Σ 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 ζ_ℓ^2s are independent and distributed according to the χ^2_1 distribution. Hence $\mathbf{E}[e^{\mathrm{i}u\zeta_\ell^2}] = (1-2\mathrm{i}u)^{-1/2}$ and

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

By taking the logarithm and using its Taylor series we get

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

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

$$\log \varphi_{\hat\theta}(u) = \sum_{k=1}^\infty \frac{(2\mathrm{i}u)^k}{2k}\mu_k, \qquad |u| < 1/(2\|\boldsymbol{\lambda}\|_\infty), \tag{11}$$

with $\|\boldsymbol{\lambda}\|_\infty = \max_\ell |\lambda_\ell|$ and $\mu_k = \sum_{\ell=1}^N \lambda_\ell^k = \mathrm{Tr}[(\Sigma^{1/2} A \Sigma^{1/2})^k] = \mathrm{Tr}[(\Sigma \cdot A)^k]$, where the last equality follows from the property $\mathrm{Tr}(M_1 \cdot M_2) = \mathrm{Tr}(M_2 \cdot M_1)$ provided that both products are well defined. Separating the first two terms in the RHS of (11), we arrive at

$$\log \varphi_{\hat\theta}(u) = \mathrm{i}\theta u - u^2\mu_2 + \sum_{k=3}^\infty \frac{(2\mathrm{i}u)^k}{2k}\mu_k, \qquad |u| < 1/(2\|\boldsymbol{\lambda}\|_\infty). \tag{12}$$

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

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

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

$$|\log \varphi_{\hat\theta-\theta}(v/\sqrt{2\mu_2}) + \frac{v^2}{2}| \le -v^2 \log(1-\sqrt{2}|v|\bar\alpha), \tag{13}$$

for every $|v| < (\sqrt{2}\bar\alpha)^{-1}$. As a 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 $\lim_{n\to\infty} \|\boldsymbol{\lambda}_n\|_\infty^2/\mu_{2,n} = 0$, then*

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

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

Proof. Set $\mu_{k,n} = \mathrm{Tr}[(\Sigma_n A_n)^k] = \sum_\ell \lambda_{\ell,n}^k$ and $\eta_n = (\hat\theta_n - \theta_n)/\sqrt{2\mu_{2,n}}$. The inequality (13) and the condition $\lim_{n\to\infty} \|\boldsymbol{\lambda}_n\|_\infty^2/\mu_{2,n} = 0$ imply that the characteristic function of η_n converges pointwise to the characteristic function of a standard Gaussian distribution. This completes the proof of the proposition. □

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 $\|\boldsymbol{\lambda}\|_\infty^2 \le \mu_2/(2p)$ holds, then for every $j \in \mathbb{N}$*

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

Proof. Thanks to the fact that ζ_ℓ^2 is distributed according to the χ_1^2 distribution, one easily checks that $\left|\varphi_{\hat{\theta}}(u)\right| = \left|\prod_{\ell=1}^N (1-2\mathrm{i}u\lambda_\ell)^{-1/2}\right| = \prod_{\ell=1}^N (1+4u^2\lambda_\ell^2)^{-1/4}$. In view of the assumptions of the lemma, for every $i = 1, \dots, p$, there exists an integer ℓ_i verifying $\mu_2^{-1}\sum_{\ell=1}^{\ell_i}\lambda_\ell^2 < i/p$ and $\mu_2^{-1}\sum_{\ell=1}^{\ell_i+1}\lambda_\ell^2 \ge i/p$. For this sequence ℓ_i, we get $\mu_2^{-1}\sum_{\ell=\ell_i+1}^{\ell_{i+1}}\lambda_\ell^2 \ge (i+1)/p - 1/(2p) - i/p = 1/(2p)$ and therefore

$$\prod_{\ell=1}^N (1+4u^2\lambda_\ell^2)^{-1/4} \le \prod_{i=1}^p \Big(1+4u^2\sum_{\ell=\ell_i+1}^{\ell_{i+1}}\lambda_\ell^2\Big)^{-1/4} \le (p/2)^{p/4}(1+\mu_2 u^2)^{-p/4}. \tag{14}$$

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

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

Simple computations yield

$$\Big|\frac{d^{j_\ell}}{du^{j_\ell}}(1-2\mathrm{i}u\lambda_\ell)^{-1/2}\Big| \le \Big|\frac{j_\ell!\,(2\mathrm{i}\lambda_\ell)^{j_\ell}}{(1-2\mathrm{i}u\lambda_\ell)^{j_\ell+1/2}}\Big| \le \frac{j_\ell!\|2\lambda\|_\infty^{j_\ell}}{(1+4u^2\lambda_\ell^2)^{1/4}}.$$

Therefore,

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

and the desired inequality for $\theta = 0$ follows from (14). For θ different from zero, it suffices to use the relation $|\varphi_{\hat{\theta}-\theta}^{(j)}(u)| \le \sum_{k=0}^j C_j^k |\mathrm{i}\theta|^k |\varphi_{\hat{\theta}}^{(j-k)}(u)|$ and the obtained estimate for $|\varphi_{\hat{\theta}}^{(j-k)}(u)|$. □

Remark 2. 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 the absolute value of u is 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 suitable decay properties for the tails of the derivatives of $\varphi_{\hat{\theta}-\theta}$.

3.2.2. Computation of μ_k in our setup

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 Σ have special forms. In particular, they contain only a small number of nonzero entries and, therefore, the expression of μ_k takes a simplified form.

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

Definition 1. *We call chain of length k, any vector $(\boldsymbol{i},\boldsymbol{j}) \in \{1,\dots,N_1\}^k \times \{1,\dots,N_2\}^k$ such that $I^{i_p} \cap J^{j_p} \neq \varnothing$ and $J^{j_p} \cap I^{i_{p+1}} \neq \varnothing$ 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 $\mathscr{C}_k$.*

In the definition of $\mathscr{C}_k$, i_p (resp. j_p) stands for the pth coordinate of $\boldsymbol{i}$ (resp. $\boldsymbol{j}$).

Proposition 3. *The coefficients μ_2 and μ_3 can be computed by the formulae*

$$\mu_2 = \frac{1}{2}\sum_{(\boldsymbol{i},\boldsymbol{j})\in\mathscr{C}_2}\prod_{p=1}^{2} v(I^{i_p}\cap J^{j_p}) + \frac{1}{2}\sum_{(i,j)\in\mathscr{C}_1} v_1(I^i)v_2(J^j),$$

$$\mu_3 = \frac{1}{4}\sum_{(\boldsymbol{i},\boldsymbol{j})\in\mathscr{C}_3}\prod_{p=1}^{3} v(I^{i_p}\cap J^{j_p}) + \frac{3}{4}\sum_{(\boldsymbol{i},\boldsymbol{j})\in\mathscr{C}_2} v_1(I^{i_1})v_2(J^{j_1})v(I^{i_2}\cap J^{j_2}),$$

where v, v_1 and v_2 are defined by (4).

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 $\mu_3 = \mathrm{Tr}[(\Sigma\cdot A)^3]$, we have

$$\mu_3 = \sum_{\ell_1,\dots,\ell_6=1}^{N} \sigma_{\ell_1\ell_2}a_{\ell_2\ell_3}\sigma_{\ell_3\ell_4}a_{\ell_4\ell_5}\sigma_{\ell_5\ell_6}a_{\ell_6\ell_1}. \tag{15}$$

In our setup, the entries of the matrix A are

$$a_{\ell,\ell'} = \frac{1}{2}\cdot\mathbf{1}(\ell\le N_1, \ell' > N_1, I^{\ell}\cap J^{\ell'-N_1}\neq\varnothing) + \frac{1}{2}\cdot\mathbf{1}(\ell > N_1, \ell'\le N_1, I^{\ell'}\cap J^{\ell-N_1}\neq\varnothing), \tag{16}$$

and those of Σ are

$$\sigma_{\ell,\ell'} = \begin{cases} v(I^{\ell}\cap J^{\ell'-N_1}), & \text{if } \ell\le N_1, \ell' > N_1,\\ v(I^{\ell'}\cap J^{\ell-N_1}), & \text{if } \ell'\le N_1, \ell > N_1,\\ v_1(I^{\ell}), & \text{if } \ell=\ell'\le N_1,\\ v_2(J^{\ell-N_1}), & \text{if } \ell=\ell' > N_1,\\ 0, & \text{otherwise.}\end{cases} \tag{17}$$

To compute the sum in the right hand side of (15), we consider different cases separately.
Case A: $\ell_1 \le N_1$ Our aim now is to compute

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

This can be done by considering the following four subcases:

Case A.1 $\ell_1\neq\ell_2$ and $\ell_3\neq\ell_4$, Case A.2 $\ell_1=\ell_2$ and $\ell_3=\ell_4$,
Case A.3 $\ell_1\neq\ell_2$ and $\ell_3=\ell_4$, Case A.4 $\ell_1=\ell_2$ and $\ell_3\neq\ell_4$.

In the case A.1, in order that the corresponding term in (15) be nonzero, the indices $\ell_i, i\le 6$, should satisfy $\ell_1\le N_1$, $\ell_2 > N_1$, $\ell_3\le N_1$, $\ell_4 > N_1$, $\ell_5\le N_1$ and $\ell_6 > N_1$. Moreover, if we set $\boldsymbol{i} = (\ell_1,\ell_3,\ell_5)$ and $\boldsymbol{j} = (\ell_2,\ell_4,\ell_6)$, then $(\boldsymbol{i},\boldsymbol{j})$ should belong to $\mathscr{C}_3$. Therefore, $\sigma_{i_pj_p} = v(I^{i_p}\cap J^{j_p})$ for $p = 1,2,3$ and

$$\sigma_{\ell_1\ell_2}a_{\ell_2\ell_3}\sigma_{\ell_3\ell_4}a_{\ell_4\ell_5}\sigma_{\ell_5\ell_6}a_{\ell_6\ell_1} = \frac{1}{8}\mathbf{1}((\boldsymbol{i},\boldsymbol{j})\in\mathscr{C}_3)\prod_{p=1}^{3} v(I^{i_p}\cap J^{j_p}). \tag{18}$$

In the case A.2, in order to get nonzero term in (15), the indices $\ell_i, i \le 6$, should satisfy $\ell_1 = \ell_2 \le N_1$, $\ell_3 = \ell_4 > N_1$, $\ell_5 \le N_1$ and $\ell_6 > N_1$. Moreover, if we set $\boldsymbol{i} = (\ell_1, \ell_5)$ and $\boldsymbol{j} = (\ell_3, \ell_6)$, then $(\boldsymbol{i}, \boldsymbol{j})$ should belong to $\mathscr{C}_2$. Therefore,

$$\begin{aligned}\sigma_{\ell_1\ell_2}a_{\ell_2\ell_3}\sigma_{\ell_3\ell_4}a_{\ell_4\ell_5}\sigma_{\ell_5\ell_6}a_{\ell_6\ell_1} &= \sigma_{i_1i_1}a_{i_1j_1}\sigma_{j_1j_1}a_{j_1i_2}\sigma_{i_2j_2}a_{j_2i_1}\\ &= \frac{1}{8}\mathbf{1}((\boldsymbol{i},\boldsymbol{j}) \in \mathscr{C}_2)v_1(I^{i_1})v_2(J^{j_1})v(I^{i_2} \cap J^{j_2}). \end{aligned} \tag{19}$$

In the cases A.3 and A.4, it is easily seen that the corresponding summand in the right hand side of (15) is $\neq 0$ 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 Σ, 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 $\ell_6 \le N_1$, and $a_{\ell_1\ell_6} \neq 0$ and $\ell_6 \le N_1$ entails $\ell_1 > N_1$, we get $\mu_{3,B} = \sum_{\ell_6 \le N_1} \sum_{\ell_1,\ell_2,\dots,\ell_5=1}^{N} \sigma_{\ell_6\ell_5}a_{\ell_5\ell_4}\sigma_{\ell_4\ell_3}a_{\ell_3\ell_2}\sigma_{\ell_2\ell_1}a_{\ell_1\ell_6}$. By reordering the indices we get $\mu_{3,B} = \mu_{3,A}$ and the assertion of the proposition follows. □

Corollary 1. *The terms μ_2 and μ_3 may alternatively be computed by formulae (5)-(6).*

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 μ_3 given in Proposition 3. In this notation, $4\mu_3 = T_1 + 3T_2$.

On the one hand, $(\boldsymbol{i}, \boldsymbol{j}) \in \mathscr{C}_2$ 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,

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

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

$$T_2 = \sum_{I,J} v_1(I)v_2(J)[(v(I) + v(J))K_{IJ} - v(I \cap J)] = v(I \cup J)K_{IJ}. \tag{20}$$

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

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

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 $(\boldsymbol{i}, \boldsymbol{j}) \in \mathcal{C}_3$ 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

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

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

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

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 (6). Equality (5) can be proved along the same lines. □

Remark 3. If the observations are synchronous, that is $\Pi^1 = \Pi^2 = \Pi$, then μ_2 and μ_3 have the following simple expressions:

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

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 Σ_n by (16) and (17). If the functions σ_1 and σ_2 are bounded on $[0,T]$ by some constant σ, 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 Π_n^2 or there is $J \in \Pi_n^2$ such that I is the union of all intervals from Π_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 (Π_n^1, Π_n^2). It follows that $\mu_{p,n} = \tilde{\mu}_{p,n}$ for every $p \in \mathbb{N}$. Therefore, the relation $\max_\ell \lambda_{\ell,n}^2 = \lim_{p\to\infty} \mu_{2p,n}^{1/p}$ 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

$$\max_{J \in \tilde{\Pi}^2} \sum_{I \in \tilde{\Pi}^1} K_{IJ} \leq 3, \quad \max_{I \in \tilde{\Pi}^1} \sum_{J \in \tilde{\Pi}^2} K_{IJ} \leq 3. \tag{21}$$

In the remaining of this proof, without loss of generality we assume that (21) is fulfilled for partitions (Π^1, Π^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

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

Applying the Cauchy-Schwarz inequality and changing the order of summation, we get the inequalities $\sum_i \big(\sum_j K_{I^i J^j} u_{N_1+j}\big)^2 \le \frac{3}{4}\sum_j u_{N_1+j}^2$ and $\sum_j \big(\sum_i K_{I^i J^j} u_i\big)^2 \le \frac{3}{4}\sum_i u_i^2$, which imply that $\|A_n\|^2 \le 3/4$.

On the other hand,

$$\|\Sigma_n\| = \max_{u:|u|=1} \sum_{\ell,\ell'=1}^N \sigma_{\ell,\ell'} u_\ell u_{\ell'} = \max_{u:|u|=1} \Big(\sum_{\ell=1}^N \sigma_{\ell,\ell} u_\ell^2 + 2\sum_{i,j} v(I_n^i \cap J_n^j) u_i u_{N_1+j}\Big).$$

Since $\sigma_{\ell,\ell'}$s are given by (17), 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 relating the geometrical and the arithmetical means :

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

This completes the proof of the lemma. □

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*

$$\lim_{n\to\infty} \frac{r_n^2}{\mu_{2,n}} = 0, \tag{22}$$

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 $\lim_{n\to\infty} \frac{\|(\Sigma_n^{1/2} A_n \Sigma_n^{1/2})^2\|}{\mathrm{Tr}[(\Sigma_n A_n)^2)]} = 0$. This convergence follows from assumption (22) and Lemma 2. □

3.2.3. Expansion of the characteristic function for random sampling schemes

We assume now that the partitions Π_n^1 and Π_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 Π_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 $n \to \infty$, then $(\hat\theta_n - \theta_n)/\sqrt{2\mu_{2,n}}$ converges in distribution to a standard normal random variable. If moreover, $2\mu_{2,n}/b_n \xrightarrow[n\to\infty]{P} \mathfrak{c}$ for some deterministic sequence $\{b_n\}$ and some positive constant $\mathfrak{c}$, then $(\hat\theta_n - \theta)/\sqrt{b_n} \xrightarrow[n\to\infty]{D} N(0, \mathfrak{c})$.*

Proof. Denote $\sigma[\Pi] = \sigma[\Pi_n,\ n \in \mathbb{N}]$. Our aim is to show that for every $u \in \mathbb{R}$ the convergence $\mathbf{E}\Big[\exp\Big(\mathrm{i}u(\hat\theta_n - \theta_n)/\sqrt{2\mu_{2,n}}\Big)\Big] \xrightarrow[n\to\infty]{} \mathbf{E}[e^{-\frac{1}{2}u^2}]$ holds. Let us denote $a_n = \mathbf{E}\Big[\exp\Big(\mathrm{i}u(\hat\theta_n - \theta_n)/\sqrt{2\mu_{2,n}}\Big)\Big]$ and $a = \mathbf{E}[e^{-\frac{1}{2}u^2}]$. To show the desired convergence, it suffices to check that every convergent subsequence of $\{a_n\}$ converges to a. For checking this property, one can simply remark that for any subsequence $\{a_{n_k}\}$, there is a sub-subsequence $\{n_{k_j}\}$ such that $r^2_{n_{k_j}}/\mu_{2,n_{k_j}}$ converges almost surely. Then, Corollary 2 implies that $a_{n_{k_j}}$ converges to a as $j \to \infty$. Therefore, a is also the limit of the sequence $\{a_{n_k}\}$ and the first assertion of the proposition follows. The second assertion follows from the first one by a simple application of the Slutsky lemma. □

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 $n \to \infty$. 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 Π, henceforth denoted by $\kappa_r[\mathcal{X}_n]$. Let $\bar\lambda_{r,n}$ be the normalized r-th conditional cumulant of $\mathcal{X}_n$:

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

Note that this notation is consistent with those introduced in (7).

Lemma 3. *For every positive integer r, we have*

$$|\mu_{r,n}| \le \sum_\ell |\lambda_{\ell,n}|^r \le \max_\ell |\lambda_\ell|^{r-2}\mu_{2,n} \le (\alpha_n\sqrt{b_n})^{r-2}\mu_{2,n}, \tag{23}$$

where $\alpha_n = \sqrt{3}\sigma^2 r_n b_n^{-1/2}$. In terms of the conditional cumulants, this is equivalent to $|\kappa_r[\mathcal{X}_n]| \le c_r\alpha_n^{r-2}\,\bar\lambda_{2,n}$, where $c_r = 2^{r-2}(r-1)!$.

Proof. This is an immediate consequence of Lemma 2. □

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

$$\mathcal{P}_n(\delta) = \big\{\Pi\ :\ \alpha_n < \delta,\ \bar\lambda_{2,n} < \mathfrak{c}_1\big\}, \qquad \forall \delta > 0.$$

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

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

for every u satisfying $|u| \le C\delta^{-1/3}$ 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 satisfying the condition $\limsup_{\delta\to 0}\sup_n \sup_{\Pi_n\in\mathcal{P}_n(\delta)} |\mathcal{O}(\delta^p)|\delta^{-p} < \infty$.

Proof. Let us define $a_0(u) = -\bar\lambda_{2,n}u^2/2$, $a_{1,n}(u) = \frac{(\mathrm{i}u)^3 b_n^{1/2}\bar\lambda_{3,n}}{6}$ and $\mathsf{r}_n(u) = \sum_{k=4}^\infty \frac{(2\mathrm{i}u)^k\mu_{k,n}}{2kb_n^{k/2}}$. Using (11) and the fact that in our setup $\max_\ell|\lambda_\ell|$ is bounded by $\sqrt{3}\sigma^2 r_n$, we get

$$\mathbf{E}^\Pi[e^{\mathrm{i}u\mathcal{X}_n}] = \exp\Big\{\sum_{k=2}^\infty \frac{(2\mathrm{i}u)^k\mu_{k,n}}{2kb_n^{k/2}}\Big\} = \exp\Big\{-\frac{\bar\lambda_{2,n}u^2}{2} + \frac{(\mathrm{i}u)^3 b_n^{1/2}\bar\lambda_{3,n}}{6} + \mathsf{r}_n(u)\Big\}$$

for every $u \in \mathbb{R}$ such that $|u| < 1/(2\delta)$. One easily checks that

$$\mathbf{E}^\Pi[e^{\mathrm{i}u\mathcal{X}_n}] - e^{a_0(u)}(1 + a_{1,n}(u)) = e^{a_0(u)}(a_{1,n}(u) + \mathsf{r}_n(u))^2\int_0^1\int_0^1 v e^{tv(a_{1,n}(u)+\mathsf{r}_n(u))}\,dtdv + \mathsf{r}_n(u)e^{a_0(u)}. \tag{24}$$

Inequalities (23) imply that there exists some constant $C > 0$ such that for every $\ell \le j$ and for every $\Pi_n \in \mathcal{P}_n(\delta)$, it holds that

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

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

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

These inequalities in conjunction with Eq. (24) yield the estimate

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

This completes the proof of the proposition □

Remark 4. As usual in asymptotic expansions, the coefficient of the second order term (i.e., the coefficient of $(\mathrm{i}u)^3 b_n^{1/2}$) 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 (6).

3.3. Proof of Theorem 1

Let us start by proving relation (8). Let $\mathsf{h}(x) = 1 + |x|^\gamma$. Let $\mathcal{K}$ be a probability density on $\mathbb{R}$ such that the Fourier transform $\hat{\mathcal{K}}$ of $\mathcal{K}$ is compactly supported, $\int_{\mathbb{R}} |x|^{\gamma+2}\mathcal{K}(x)\,dx < \infty$ and $\int_{-1}^{1}\mathcal{K}(x)\,dx \ge 2/3$. Let $K > 0$. For $\epsilon > 0$, define the measure $\mathcal{K}_\epsilon$ by $\mathcal{K}_\epsilon(x) = \mathcal{K}(\epsilon^{-1}x)$ for all $x \in \mathbb{R}$. Using the modified version of the Sweeting lemma [42] stated in Babu and Singh [4, Lemma 1], we get:

$$|\mathbf{E}^\Pi[f(\mathcal{X}_n)] - \Psi_n^\Pi[f]| \quad \le \quad 9^\gamma M(P^{\mathcal{X}_n|\Pi} + |\Psi_n^\Pi|)[\mathsf{h}]\left(\mathbb{A}_0 + \mathbb{A}_1 + \mathbb{A}_2\right) + \mathbb{A}_3, \tag{25}$$

where

$$\mathbb{A}_0 = \int_{\mathbb{R}} \mathsf{h}(x)\big|\mathcal{K}_{b_n^K} * (P^{\mathcal{X}_n|\Pi} - \Psi_n^\Pi)\big|(dx), \quad \mathbb{A}_1 = b_n^K\int_{\mathbb{R}} |x|^{\gamma+2}\mathcal{K}(x)\,dx, \quad \mathbb{A}_2 = 2^{1-b_n^{-K/4}}$$
$$\mathbb{A}_3 = \sup_{|x|\le b_n^K}\int_{\mathbb{R}} \omega_f(x-y, 2b_n^K)|\Psi_n^\Pi|(dy).$$

As we already mentioned, the Rosenthal inequality yields that $P^{\mathcal{X}_n|\Pi}[\mathsf{h}] = 1 + \mathbf{E}^\Pi[|\mathcal{X}_n|^\gamma]$ is bounded uniformly in n. Furthermore, it is obvious that the term $|\Psi_n^\Pi|[\mathsf{h}]$ is bounded uniformly in n.

If n is sufficiently large, $[x-y-2b_n^K, x-y+2b_n^K] \subset [-y-3b_n^{K/4}, -y+3b_n^{K/4}]$ and therefore

$$\mathbb{A}_3 \le 2\int_{\mathbb{R}} \omega_f(-y, 3b_n^{K/4})|\Psi_n^\Pi|(dy) \le \mathsf{C}^\circ\int_{\mathbb{R}} \omega_f(y, 3b_n^{K/4})\phi(y;\mathfrak{c}^*)\,dy \le \mathsf{C}b_n^{K\eta/4}.$$

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

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

where $\varphi^{\Pi}_{\mathcal{X}_n}(u) = \mathbf{E}^{\Pi}[e^{\mathrm{i}u\mathcal{X}_n}]$. Let $\delta_n = b_n^{a-1/2}$. By virtue of Proposition 5 and Lemma 1, we have

$$\begin{aligned}
&\int_{\mathbb{R}} \Big| \partial_u^\alpha \Big[\Big(\varphi^{\Pi}_{\mathcal{X}_n}(u) - \hat{\Psi}^{\Pi}_n(u) \Big) \hat{\mathcal{K}}(b_n^K u) \Big] \Big| \, du \\
&\qquad \le \int_{u:|u|\le C\delta_n^{-1/3}} \Big| \partial_u^\alpha \Big[\Big(\varphi^{\Pi}_{\mathcal{X}_n}(u) - \hat{\Psi}^{\Pi}_n(u) \Big) \hat{\mathcal{K}}(b_n^K u) \Big] \Big| \, du \\
&\qquad\quad + \int_{u:|u|>C\delta_n^{-1/3}} \Big| \partial_u^\alpha \Big[\Big(\varphi^{\Pi}_{\mathcal{X}_n}(u) - \hat{\Psi}^{\Pi}_n(u) \Big) \hat{\mathcal{K}}(b_n^K u) \Big] \Big| \, du \\
&\qquad \le \int_{u:|u|\le C\delta_n^{-1/3}} \mathcal{O}(\delta_n^2)(1+|u|^q) e^{-\mu_{2,n} u^2/2} \, du + \int_{u:|u|>C\delta_n^{-1/3}} \frac{C_2}{1+|u|^L} \, du \\
&\qquad\quad + \sum_{\alpha'=0}^{2+\gamma} \int_{u:|u|>C\delta_n^{-1/3}} |\partial_u^{\alpha'} \hat{\Psi}^{\Pi}_n(u)| \, du \\
&\qquad \le C_3[\mathcal{O}(\delta_n^2) + \delta_n^{(L-1)/3}] \le C_4 \delta_n^2,
\end{aligned}$$

where L can be chosen as large as we need, therefore $\mathbb{A}_0 \le C_5 \delta_n^2$. Combining all these estimates, we get

$$|\mathbf{E}^{\Pi}[f(\mathcal{X}_n)] - \Psi^{\Pi}_n[f]| \le \mathsf{C}\big(b_n^{2a-1} + b_n^K + 2^{-b_n^{-K/4}} + b_n^{K\eta/4}\big).$$

Choosing $K > \max(2a-1, 4(2a-1)/\eta)$, we get the relation stated in (8).

To prove (9), 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 Ψ^{Π}_n around $\mathfrak{c}$ we get the desired result.

4. Poisson sampling scheme

As an application of previous results let us consider the case when the partitions Π^1_n and Π^2_n are generated by Poisson point processes. Let $\mathscr{P}^{i,n} = (\mathscr{P}^{i,n}_t, t \ge 0)$, $i = 1, 2$, be two independent homogeneous Poisson processes with intensities np_i, $i = 1, 2$. Moreover, assume that these processes are independent of $\boldsymbol{B}$. Let the sampling times $S^1, \dots, S^{N_1}$ and $T^1, \dots, T^{N_2}$ be the time instants corresponding to the jumps of $\mathscr{P}^{1,n}$ and $\mathscr{P}^{2,n}$ occurred before the instant T. Note that S^is and T^js depend also on n. However, for simplicity of exposition this dependence will not be reflected in our notation.

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 3. *Let the sampling scheme be generated by two independent Poisson processes with intensities np_1 and np_2, independent of the driving BM $\boldsymbol{B}$. If the functions σ_1, σ_2 and ρ are Lipschitz continuous then, for every $a \in (\frac{3}{4}, 1)$, it holds that*

$$\sup_{f \in \mathcal{E}(M,\gamma) \cap \mathcal{E}^0(\mathsf{C},\eta,r_0,\mathfrak{c}^*)} \left| \mathbf{E}[f(\sqrt{n}(\hat{\theta}_n - \theta))] - \int_{\mathbb{R}} f(z)\, p^\circ_n(z)\, dz \right| = O(n^{1-2a}), \tag{26}$$

where

$$p^\circ_n(z) \propto \frac{1}{\sqrt{2\pi\mathfrak{c}}} \Big[1 + \frac{2\kappa(z^3 - 3\mathfrak{c}z)}{\sqrt{n}\,\mathfrak{c}^3} \Big]_+ e^{-z^2/(2\mathfrak{c})}$$

is a probability density with

$$\begin{aligned}
\mathfrak{c} &= \left(\frac{2}{p_1} + \frac{2}{p_2} \right) \int_0^T \sigma_{1,t}^2 \sigma_{2,t}^2 (1+\rho_t^2) dt - \frac{2}{p_1+p_2} \int_0^T (\sigma_{1,t}\sigma_{2,t}\rho_t)^2 dt, \\
\kappa &= \left(\frac{1}{p_1^2} + \frac{1}{p_2^2} \right) \int_0^T h(t)^3 \, dt + \frac{3p_1^2 + 2p_1p_2 + 3p_2^2}{p_1^2 p_2^2} \int_0^T \sigma_{1,t}^2 \sigma_{2,t}^2 h(t) \, dt.
\end{aligned}$$

Before proceeding with the proof of this theorem, let us note that it extends the asymptotic normality result proved in Hayashi and Yoshida [19], providing the second-order term in the asymptotic expansion of the distribution of $\hat{\theta}_n$. Note however that the price to pay for getting this expansion is a slightly stronger assumption on the functions σ_1, σ_2 and ρ. Indeed, we assume in Theorem 3 that these functions are Lipschitz, while in [19] only the continuity of these functions was required.

Remark also that the constant of proportionality in the definition of p_n° can be replaced by one. Indeed, $p_n^\circ(z)$ is the positive part of the function

$$z \mapsto \frac{1}{\sqrt{2\pi\mathfrak{c}}}\Big[1 + \frac{2\kappa(z^3 - 3\mathfrak{c}z)}{\sqrt{n}\,\mathfrak{c}^3}\Big]e^{-z^2/(2\mathfrak{c})}, \tag{27}$$

whose integral over $\mathbb{R}$ is equal to one. Moreover, for some $c > 0$, the function (27) is positive on the interval $[-cn^{1/6}, cn^{1/6}]$ and its absolute value is bounded by an exponentially decreasing function outside the interval $[-cn^{1/6}, cn^{1/6}]$. This implies that the proportionality constant in the definition of p_n° is $1 + O(e^{-n^{1/3}/(4\mathfrak{c})})$ and, consequently, its exact value is unimportant.

Proof of Theorem 3. We want to apply Theorem 2. To this end, we have to accomplish the following tasks:

[T1] prove that $\bar{\lambda}_{2,n} = 2n\mu_{2,n}$ is very close to $\mathfrak{c}$ in expectation and in probability,
[T2] check that the maximal sampling step r_n is smaller than b_n^a with high probability,
[T3] determine the asymptotic behavior of $\mathbf{E}[\bar{\lambda}_{3,n}]$,

with $b_n = 1/n$ and some $a < 1$. In fact, we will show that any $a < 1$ can be used.

Concerning the task [T1], it is proved in [19] that $2n\mu_{2,n}$ converges in probability to $\mathfrak{c}$. In the present work, we need a result providing the rate of convergence of $2n\mu_{2,n}$ to $\mathfrak{c}$. It is done in the following

Proposition 6. *If the functions σ_1, σ_2 and ρ 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 that*

$$\mathbf{P}\left(|2n\mu_{2,n} - \mathfrak{c}| > \frac{C\log^3 n}{n} + \frac{x}{\sqrt{n}}\right) \leq Cne^{-x/C}. \tag{28}$$

Furthermore, $\mathbf{E}[2n\mu_{2,n}] = \mathfrak{c} + \mathcal{O}(n^{-1}\log^3 n)$ as n goes to infinity.

The proof of this proposition is deferred to Section 7.

The task [T2], consisting in bounding the probability of the event $r_n > b_n^a = n^{-a}$ is done using the following lemma.

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

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

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

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 ζ being exponentially distributed with mean 1. In view of the Wald equation, this yields $\mathbf{E}\big[\sum_{I\in\Pi_n^1} e^{un|I|}\big] = np_1T\mathbf{E}[e^{u\zeta/p_1}]$. 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. □

Replacing x by $n^{\frac{1}{2}-a}$ in (28) and by n^{1-a} in Lemma 4, we obtain that the probability of the event $A_n(a)^c$ is exponentially small as $n \to \infty$. Therefore, $\mathbf{P}(A_n(a)^c) = o(b_n^p) = o(n^{-p})$ for every $p > 0$. One also deduces from Proposition 6 that $\mathbf{E}[\bar{\lambda}_{2,n}] - \mathfrak{c} = o(n^{1-2a})$ as $n \to \infty$. Thus, it remains to accomplish the task [T3], which is done using the following proposition, the proof of this proposition is deferred to Section 7.

Proposition 7. *Under the assumptions of Theorem 3, it holds that* $\mathbf{E}[\mu_{3,n}] = \frac{3}{2}\kappa n^{-2} + \mathcal{O}(\frac{\log^3 n}{n^3})$.

Combining these results, we get the assertion of Theorem 3. □

5. 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:

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

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

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

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

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 $([0,T]^3, \mathscr{B}_{[0,T]^3})$.

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 $n \to \infty$. 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 $n \to \infty$, 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 $\varphi : [0,T]^2 \to \mathbb{R}$, 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. For the purposes of the present work, it is probably possible to slightly relax Assumption P2 by replacing the weak convergence by the tightness condition. However, to avoid additional technicalities we assume that the weak convergence of measures stated in Assumption P2 holds.

Recall that according to our assumptions Π is independent of $\boldsymbol{B}$, where Π 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 $f, g : [0,T] \to \mathbb{R}$, 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 the distribution of this estimator using a perturbation method based on a stochastic expansion of the estimator $\hat{\theta}_n$ itself. The main term in this stochastic expansion is

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

Note that the asymptotic expansion of the distribution of M_T^n has already been obtained in preceding sections. In this section, we will need a representation of M^n as a stochastic integral with respect to the BM (B_1, B_2) that can be written—using the Itô formula—as follows:

$$M^n = \mathbb{H}^{1,n} \cdot B_1 + \mathbb{H}^{2,n} \cdot B_2, \tag{29}$$

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

Lemma 5. *Assume that σ_1, σ_2 and ρ 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

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

and A^n is a bounded variation process defined by

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

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 a 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.

5.1. Convergence of quadratic variations and covariations

To establish an asymptotic expansion of $b_n^{-1/2}(\hat{\theta}_n - \theta)$ that is more explicit than the one given by Lemma 5, we need to identify the limiting distribution of the martingale (B_1, B_2, M^n, N^n) as n goes to infinity. The convergence of the quadratic variation-matrix is a classical tool for proving the convergence of a martingale. Most results of the present section being quite technical, we postponed their proofs to Section 8.

We start with the cross terms $\langle M^n, B_1\rangle$ and $\langle M^n, B_2\rangle$. In view of (29), for $\nu = 1, 2$, we have

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

Lemma 6. *If σ_1, σ_2 and ρ are bounded in $[0,T]$ and $r_n^2 = o_p(b_n)$, then*

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

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

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 $n\to\infty$.

We study now the behavior of the quadratic variation

$$\langle M^n, M^n\rangle_t = \sum_{c,d=1}^{2}(\mathbb{H}^{c,n}\mathbb{H}^{d,n})\cdot\langle B_c, B_d\rangle_t \tag{30}$$

as n tends to infinity. First, we note that

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

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 $I^i\cap J^j\neq\varnothing$, $I^i\cap J^{j'}\neq\varnothing$, $I^{i'}\cap J^{j'}\neq\varnothing$, $j\le j'$ and $i'\le i$ 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

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

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

$$(J^j\sigma_2\cdot B_2)I\sum_{j'\colon j\le j'}K_{ij'}J^{j'}=(J^j\sigma_2\cdot B_2)I\mathbf{1}_{[T^{j-1},T]}=(J^j\sigma_2\cdot B_2)I,$$
$$(I^{i'}\sigma_1\cdot B_1)J\sum_{i\colon i'\le i}K_{ij}I^i=(I^{i'}\sigma_1\cdot B_1)J\mathbf{1}_{[S^{i-1},T]}=(I^{i'}\sigma_1\cdot B_1)J.$$

to obtain $\mathbb{H}^{1,n}\mathbb{H}^{2,n}=b_n^{-1}\sum_{I,J}\sigma_1\sigma_2K_{IJ}(J\sigma_2\cdot B_2)(I\sigma_1\cdot B_1)(I+J-IJ)$. This implies that

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

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

Lemma 7. *Assume that $r_n^3=o_p(b_n^2)$ and the functions σ_1, σ_2 and ρ are continuous. If Assumption P1 is fulfilled then, for any $t\in[0,T]$,*

$$\int_0^t\mathbb{H}^{1,n}_s\mathbb{H}^{2,n}_s\,d\langle B_1,B_2\rangle_s\xrightarrow[n\to\infty]{P}\frac{1}{2}\int_0^t h_s^2\left\{\mathcal{V}^I(ds)+\mathcal{V}^J(ds)-\mathcal{V}^{I\cap J}(ds)\right\},$$
$$\int_0^t(\mathbb{H}^{1,n}_s)^2\,d\langle B_1,B_1\rangle_s+\int_0^t(\mathbb{H}^{2,n}_s)^2\,d\langle B_2,B_2\rangle_s\xrightarrow[n\to\infty]{P}\int_0^t\sigma_{1,s}^2\sigma_{2,s}^2\,\mathcal{V}^{I,J}(ds)$$

and consequently

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

Using the claims of two last lemmas, one can already derive the asymptotic distribution of the martingale (B_1,B_2,M^n) as $n\to\infty$. However, for our purposes, it is crucial to know the asymptotics of the joint distribution of the triplet (B_1,B_2,M^n) with the martingale N^n.

Lemma 8. *If σ_1,σ_2 and ρ 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 $n\to\infty$, 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.*

An interesting fact revealed by this lemma is the orthogonality of M^n and N^n in terms of quadratic covariation. This indicates that the limiting distribution of (M^n,N^n) is that of two independent martingales. This statement will be rigorously proved at the end of this section. Prior to presenting that proof, we wish to investigate the structure of the limiting distribution of N^n and how it relates to the BM $\boldsymbol{B}$.

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*

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

This lemma describes the parts of the limit of N^n that can be described or explained by B_1 and B_2. This is however not enough. One also needs to evaluate the limiting quadratic variation of the process N^n.

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

$$\langle N^n,N^n\rangle_t\xrightarrow[n\to\infty]{P}\int_0^t\beta_2^2\sigma_1^2\,d\mathcal{V}^{I,J,J'}+\int_0^t\beta_1^2\sigma_2^2\,d\mathcal{V}^{I,I',J}+2\int_0^t\beta_2\beta_1\sigma_1\sigma_2\rho\,d\mathcal{V}^{J(I),I(J),I\cap J}.$$

The last step before stating the main result on the convergence of the processes involved in the stochastic decomposition presented in Lemma 5 is the proof of the convergence of the bounded variation process A^n. Recall that the latter is defined by

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

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

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

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 $A_t^n = A_t^\infty + o_p(1)$ with

$$A_t^\infty = \frac{1}{2} \int_0^t \{\sigma_{1,u}(\beta_{21,u}^{[1]} + \beta_{22,u}^{[1]}\rho_u) + \sigma_{2,u}(\beta_{11,u}^{[1]}\rho_u + \beta_{12,u}^{[1]}) + 2\beta_{1,u}\beta_{2,u}\} \mathcal{V}^{I,J}(du). \tag{31}$$

Proposition 8. *Assume that the functions σ_1, σ_2 and ρ 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^∞.*

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 [23] to the semimartingale $Z^n = (M^n, N^n + A^n)^{\mathbf{T}}$ with $\boldsymbol{B}$ serving as a martingale of reference (denoted by M^n in [23]), 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 $\tilde{\boldsymbol{B}}$ independent of $\boldsymbol{B}$ such that

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

where

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

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

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 (31), we get

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

As we see in the next section, this expression of A appears in the asymptotic expansion of the distribution function of $b_n^{-1/2}(\hat{\theta}_n - \theta)$.

6. 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 3 and 4. 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.

6.1. Perturbation

Since the drift terms are possibly non-linear functionals of the Brownian motion $\boldsymbol{B}$, 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 [45] 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 [38].

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 γ, let $\mathcal{E}(M,\gamma)$ be the set of all measurable functions $f : \mathbb{R}^d \to \mathbb{R}$ satisfying $|f(x)| \le M(1+|x|^\gamma)$ 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 $n \to \infty$. We write $G_n(f) = \bar{o}(s_n)$ if $s_n^{-1}\sup_{f\in\mathcal{E}}|G_n(f)| \to 0$ as $n \to \infty$.

Theorem 4. *Let ℓ 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 τ_n such that

(3) $\sup_n ||\tau_n||_{p,\ell-1} < \infty$ *for any* $p > 1$.
(4) $\mathbf{P}[|\tau_n| > 1/2] = o(s_n^\alpha)$ *for some* $\alpha > 1$.
(5) $\sup_n \mathbf{E}[1_{\{|\tau_n|<1\}}(\det\sigma_{\mathcal{X}_n})^{-p}] < \infty$ *for any* $p > 1$.
(6) There is a sequence of signed measures Ψ_n on $\mathbb{B}_d$ such that for any positive numbers M and γ, $\mathbf{E}[f(\mathcal{X}_n)] = \Psi_n[f] + \bar{o}(s_n)$ as $n \to \infty$ for $f \in \mathcal{E}'$. Moreover, for every polynomial $\pi(x)$ in x, there exists a constant c_π such that $|\Psi_n[e^{\mathrm{i}u\cdot x}\pi(x)]| \le c_\pi(1+|u|^{\ell-1})^{-1}$ 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 γ,

$$\mathbf{E}[f(\mathcal{Z}_n)] = \Psi_n[f] + s_n\int_{\mathbb{R}} f(x)g_\infty(x)\,dx + \bar{o}(s_n) \tag{32}$$

for $f \in \mathcal{E}'$, where $g_\infty(x) = -\mathrm{div}_x\big(\mathbf{E}[\mathcal{Y}_\infty \mid \mathcal{X}_\infty = x]\,p^{\mathcal{X}_\infty}(x)\big)$.

6.2. 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 sampling times, but they have more complicated form and are not presented here.

Theorem 5. *Suppose that Assumptions* P1 *and* P2 *are satisfied and*

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

Let us define

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

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

$$\sup_{f\in\mathcal{E}(M,\gamma)\cap\mathcal{E}^0(\mathsf{C},\eta,r_0,\mathfrak{c}^*)} \left|\mathbf{E}[f(b_n^{-1/2}(\hat\theta_n-\theta))] - \int_{\mathbb{R}} f(z)\,p_n^*(z)\,dz\right| = o(b_n^{1/2}), \tag{33}$$

where

$$p_n^*(z) = \frac{e^{-z^2/(2\mathfrak{c})}}{\sqrt{2\pi\mathfrak{c}}}\Big[1+\frac{b_n^{1/2}}{6\mathfrak{c}^3}\big(\mathbf{E}[\bar\lambda_{3,n}](z^3-3\mathfrak{c}z)+6\mathsf{A}\mathfrak{c}^2 z\big)\Big].$$

Moreover, if $\sup_{n\in\mathbb{N}}\mathbf{E}[\bar\lambda_{3,n}]<\infty$, then inequality (33) 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 4 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 the 6 conditions of Theorem 4 are fulfilled. In view of Lemma 5 and Proposition 8, $(\mathcal{X}_n,\mathcal{Y}_n)$ converges in distribution to some random vector $(\mathcal{X}_\infty,\mathcal{Y}_\infty)$. Thus the second condition of Theorem 4 is verified.

We have already seen in Section 3.2.1 that the principal part $\mathcal{X}_n$ of $b_n^{-1/2}(\hat\theta_n-\theta)$ can be written in the form $\mathcal{X}_n = b_n^{-1/2}(\boldsymbol{\xi}^{\mathbf{T}}A\boldsymbol{\xi}-\theta) = b_n^{-1/2}\sum_{\ell=1}^N \lambda_{\ell,n}(\zeta_{\ell,n}^2-1)$, where

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

and the entries of the matrices Σ and A are given by (17) and (16) 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 $\mathbb{R}^2$ vanishing at the origin. Recall that W is a measurable space equipped with the Borel σ-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

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

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

The process w is an isonormal Gaussian process on $H = L^2([0,T], \mathbb{R}^2)$ (see [32, Def. 1.1.1]) Using the definition of the Malliavin derivative (see [32, Def. 1.2.1]) and the chain rule [32, Prop. 1.2.3], we get the following expression for the Malliavin derivative of $\mathcal{X}_n$:

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

Since the components of $\boldsymbol{\zeta}$ are non-correlated with variance equal to one, the family $\{\phi^\ell\}_{\ell \le N}$ 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

$$\sigma_{\mathcal{X}_n} = 4 b_n^{-1} \sum_{\ell=1}^n \lambda^2_{\ell,n} \zeta^2_{\ell,n} = 4 b_n^{-1} \mu_{2,n} + 4 b_n^{-1} \sum_{\ell=1}^n \lambda^2_{\ell,n} (\zeta^2_{\ell,n} - 1). \tag{34}$$

Let us introduce the random variable τ_n that will play a role of truncation:

$$\tau_n = -\big(2 - 8\mu_{2,n} (\mathfrak{c} b_n)^{-1}\big)_+ + 8 (\mathfrak{c} b_n)^{-1} \sum_{\ell=1}^N \lambda^2_{\ell,n} (\zeta^2_{\ell,n} - 1).$$

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

To verify condition (3) of Theorem 4, we remark that

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

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

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

In view of the Rosenthal inequality, we get

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

for every $p \ge 2$. Using the definition of $\mu_{k,n}$, one can check that $\mu_{2k,n} \le \mu^{k/2}_{4,n}$. In view of inequality (23) and the obvious bound $\mu_{2,n} \le C r_n$, we get

$$\mathbf{E}^\Pi [\|D\tau_n\|^p_H] \le C b_n^{-p} r_n^{3p/2}, \qquad \mathbf{E}^\Pi [\|D^2\tau_n\|^p_{H\otimes H}] \le C b_n^{-p} r_n^{3p/2}, \quad \forall p \ge 4.$$

Similar arguments yield

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

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

$$\mathbf{P}(|\tau_n| > 1/2) \le \mathbf{P}\big(2 - 8\mu_{2,n}(\mathfrak{c} b_n)^{-1} > 0\big) + \mathbf{P}\left(8 (\mathfrak{c} b_n)^{-1} \Big| \sum_{\ell=1}^N \lambda^2_{\ell,n} (\zeta^2_{\ell,n} - 1) \Big| > 1/2\right).$$

On the one hand, since the event $\{2 - 8\mu_{2,n}(\mathfrak{c}b_n)^{-1} > 0\} = \{\bar{\lambda}_{2,n} - \mathfrak{c} < -\mathfrak{c}/2\}$ 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

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

Thus, we proved that conditions (2)-(5) of Theorem 4 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 Φ_n^2 and Φ_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. □

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

Proposition 9. *Let the sampling 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 sampling times are independent of the process $\boldsymbol{X}$,*
- *the functions σ_1, σ_2 and ρ 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

$$\sup_{f\in\mathcal{E}(M,\gamma)\cap\mathcal{E}^0(\mathsf{C},\eta,r_0,\mathfrak{c}^*)}\left|\mathbf{E}[f(n^{1/2}(\hat{\theta}_n-\theta))] - \int_{\mathbb{R}} f(z)\,p_n^\circ(z)\,dz\right| = o(n^{-1/2}), \tag{35}$$

where

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

is a probability density with

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

Proof. Lemmas 16-19 (cf. Section 9) imply that the partitions generated by independent Poisson processes satisfy Assumptions P1 and P2. Therefore, using Theorems 5 and 3, we get the desired result. □

7. Proofs of theorems and propositions

Proof of Proposition 6. Let us recall the relations

$$n\sum_{I,J} v_1(I)v_2(J)K_{IJ} \xrightarrow[n\to\infty]{P} 2(p_1^{-1}+p_2^{-1})\int_0^T \sigma_{1,t}^2\sigma_{2,t}^2 dt,$$

$$n\sum_{I\in\Pi_i} v(I)^2 \xrightarrow[n\to\infty]{P} 2p_i^{-1}\int_0^T (\sigma_{1,t}\sigma_{2,t}\rho_t)^2\,dt, \quad i=1,2$$

$$n\sum_{I,J} v(I\cap J)^2 \xrightarrow[n\to\infty]{P} 2(p_1+p_2)^{-1}\int_0^T (\sigma_{1,t}\sigma_{2,t}\rho_t)^2\,dt$$

proved in Hayashi and Yoshida [19]. The aim of the present proposition is to show that the rate of convergence in these relations is $1/\sqrt{n}$ and to get an exponential control of the 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 that

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

Let $N(x) = \lceil nT/x\rceil$ be the smallest positive integer such that $N(x)x > nT$ and let us set $L_i = [iTN(x)^{-1}, (i+1)TN(x)^{-1}]$. 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 Π_n^1 and one point from Π_n^2". The total probability formula implies that

$$\mathbf{P}\left(\left|\mathcal{T}_1 - \int_0^T \bar{h}(t)dt\right| \geq \frac{x}{\sqrt{n}}\right) \leq \mathbf{P}\left(\left|\mathcal{T}_1 - \int_0^T \bar{h}(t)\,dt\right| \geq \frac{x}{\sqrt{n}}\;\middle|\;\mathcal{E}\right) + \mathbf{P}(\mathcal{E}^c),$$

where $\mathcal{E}^c$ denotes the complementary event of $\mathcal{E}$ and $\bar{h}(t) = 2(p_1^{-1}+p_2^{-1})\sigma_{1,t}^2\sigma_{2,t}^2$. 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 define the random variables

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

In what follows, we denote by $\mathbf{E}^{\mathcal{E}}$ the conditional expectation given $\mathcal{E}$. It holds that $\mathcal{T}_1 - \int_0^T \bar{h}(t)\,dt = \mathcal{T}_{11}+\mathcal{T}_{12}+\mathcal{T}_{13}+\mathcal{O}(n|L_1|^2)$ on $\mathcal{E}$, where

$$\mathcal{T}_{11} = \mathbf{E}^{\mathcal{E}}\Big[\sum_{i=1}^{N(x)}\eta_i^\circ\Big] - \int_0^T \bar{h}(t)\,dt, \quad \mathcal{T}_{1s} = \sum_{i=1}^{[N(x)/2]}(\eta_{2i+s-2}^\circ - \mathbf{E}^{\mathcal{E}}[\eta_{2i+s-2}^\circ]), \quad s=2,3.$$

For evaluating the remainder term in $\mathcal{T}_1$, we have used the Lipschitz continuity of σ_1 and σ_2, as well as the fact that $r_n \leq |L_1|/2$ on $\mathcal{E}$.

Remark that in view of Lemma 4, for any $p>0$, we have

$$\mathbf{E}[r_n^p] = n^{-p}\int_0^\infty \mathbf{P}((nr_n)^p \geq t)\,dt \leq Cn^{-p}\int_0^\infty (ne^{-t^{1/p}})\wedge 1\,dt = Cn^{-p}\mathcal{O}(\log^p n). \tag{36}$$

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

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

Using the inequality of Cauchy-Schwarz, as well as the bounds $\mathbf{P}(\mathcal{E}^c) \le Cne^{-x/C}$ and (36), we get $\big|\mathbf{E}^{\mathcal{E}}\big[\sum_{i=1}^{N(x)} \eta_i^\circ\big] - \mathbf{E}\big[\sum_{i=1}^{N(x)} \eta_i^\circ\big]\big| \le Cne^{-x/C}$, for some constant C and for every $x > C\log n$.

On the other hand, in view of Lemma 15 presented in Section 9 below, we have

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

Therefore, using (36), we get $\mathbf{E}[\eta_i^\circ] = \mathcal{O}(n^{-1}\log^3 n)$ for every $i \le N(x)$. Using once again Lemma 15, we get

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

Wald's equality yields $\mathbf{E}\Big[\sum_{I:a_I\in L_i} |I|^k\Big] = \mathbf{E}[N_1(L_i)]\cdot\mathbf{E}[\zeta^k/(np_1)^k] + \mathcal{O}(e^{-\log^2 n/C})$, for every $k > 0$ and for every $i \le N(x) - 1$. Here, $N_1(L_i)$ is the number of points of $\mathscr{P}^{1,n}$ lying in L_i and $\zeta \sim \mathscr{E}(1)$, the exponential distribution with parameter one. Putting all these estimates together, we get

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

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 η_i°s is that, conditionally to $\mathcal{E}$, the random variables η_{2i}°, $i = 1,\dots,[N(x)/2]$, are independent. Indeed, one easily checks that conditionally to $\mathcal{E}$, η_{2i}° depends only on the restrictions of $\mathscr{P}^{1,n}$ and $\mathscr{P}^{2,n}$ 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 $\mathscr{P}^{k,n}$, $k = 1, 2$, onto these intervals are independent. Therefore, η_{2i}°, $i = 1,\dots,[N(x)/2]$, form a sequence of random variables that are independent conditionally to $\mathcal{E}$. Moreover, they verify $|\eta_i^\circ| \le Cn|L_i|^2 = \frac{C\log^4 n}{n}$.

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

$$\mathbf{P}^{\mathcal{E}}\big(|\mathcal{T}_{12}| \ge x/\sqrt{n}\big) \le 2\exp\Big(-\frac{x^2}{C(1 + xn^{-1/2}\log^4 n)}\Big) \le 2e^{-x/C},$$

for every $x > 1$. Obviously, the same inequality holds true for the term $\mathcal{T}_{13}$. These inequalities combined with the bound on the error term $\mathcal{T}_{11}$ complete the proof of (28).

Moreover, 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 above 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 $\mathbf{E}[2n\mu_{2,n}] = \mathfrak{c} + \mathcal{O}(n^{-1}\log^3 n)$. □

Proof of Proposition 7. The assertion of the theorem follows from the following relations:

$$\mathbf{E}\Big[\sum_{I\in\Pi_i} v(I)^3\Big] = \frac{6}{n^2p_i^2}\int_0^T h(t)^3\,dt + \mathcal{O}\Big(\frac{\log^3 n}{n^3}\Big), \quad i=1,2,$$

$$\mathbf{E}\Big[\sum_{I,J} v(I\cap J)^3\Big] = \frac{6}{n^2(p_1+p_2)^2}\int_0^T h(t)^3\,dt + \mathcal{O}\Big(\frac{\log^3 n}{n^3}\Big),$$

$$\mathbf{E}\Big[\sum_{I,J} v(I\cap J)^2 v(I)\Big] = \frac{18p_1+12p_2}{n^2p_1(p_1+p_2)^2}\int_0^T h(t)^3\,dt + \mathcal{O}\Big(\frac{\log^3 n}{n^3}\Big),$$

$$\mathbf{E}\Big[\sum_{I,J} v(I\cap J)^2 v(J)\Big] = \frac{18p_2+12p_1}{n^2p_2(p_1+p_2)^2}\int_0^T h(t)^3\,dt + \mathcal{O}\Big(\frac{\log^3 n}{n^3}\Big),$$

$$\mathbf{E}\Big[\sum_{I,J} v(I\cap J)v(I)v(J)\Big] = \frac{4}{n^2p_2p_1}\int_0^T h(t)^3\,dt + \mathcal{O}\Big(\frac{\log^3 n}{n^3}\Big),$$

$$\mathbf{E}\Big[\sum_{I,J} v(I\cup J)v_1(I)v_2(J)\Big] = \frac{6p_1^2+4p_1p_2+6p_2^2}{n^2p_1^2p_2^2}\int_0^T \frac{h(t)^3}{\rho_t^2}\,dt + \mathcal{O}\Big(\frac{\log^3 n}{n^3}\Big).$$

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 Lipschitz 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 is the left endpoint of the interval I.

In view of (36), we have $\mathbf{E}\Big[\sum_{I,J} r_n^3|I\cap J|\Big] \le T\mathbf{E}[r_n^3] = \mathcal{O}\Big(\frac{\log^3 n}{n^3}\Big)$. On the other hand

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

where $\mathbf{E}^I$ is the conditional expectation given I. According to Lemmas 13 and 15, presented in Section 9 below,

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

Now, let us show that

$$\mathcal{T}_1 := \frac{2}{np_2}\,\mathbf{E}\Big[\sum_{I\in\Pi^1} h(a_I)^3|I|^2\Big] = \frac{4}{n^2p_1p_2}\int_0^T h^3(t)\,dt + \mathcal{O}(n^{-3}),$$

$$\mathcal{T}_2 := \mathbf{E}\Big[\sum_{I\in\Pi^1} h(a_I)^3|I|\frac{(1-e^{-np_2|I|})e^{-np_2a_I}}{n^2p_2^2}\Big] = \mathcal{O}(n^{-3}),$$

$$\mathcal{T}_3 := \mathbf{E}\Big[\sum_{I\in\Pi^1} h(a_I)^3|I|\frac{(1-e^{-np_2|I|})e^{-np_2(T-b_I)}}{n^2p_2^2}\Big] = \mathcal{O}(n^{-3}).$$

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

$$\mathcal{T}_1 = \frac{2}{np_2}\,\mathbf{E}\Big[\sum_{i=1}^{N_1-1} h(S^i)^3\zeta_{i+1}^2\Big] + \mathcal{O}(n^{-3}), \quad |\mathcal{T}_2| \le \|h\|_\infty^3\mathbf{E}\Big[\sum_{i=1}^{N_1-1}\frac{\zeta_{i+1}e^{-np_2S^i}}{n^2p_2^2}\Big] + \mathcal{O}(n^{-3}),$$

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

$$M_k = \sum_{i=1}^{k-1} h(S^i)^3(\zeta_{i+1}^2 - \mathbf{E}[\zeta_{i+1}^2]), \quad M_k' = \sum_{i=1}^{k-1} \big(\zeta_{i+1} - \mathbf{E}[\zeta_1]\big)e^{-np_2 S^i}$$

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

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

These relations imply that

$$\mathcal{T}_1 = \frac{4}{n^2p_1p_2}\int_0^T h(t)^3\,dt + \mathcal{O}(n^{-3}),\ |\mathcal{T}_2| \leq \frac{\|h\|_\infty^3}{n^2p_2^2}\int_0^T e^{-np_2t}\,dt + \mathcal{O}(n^{-3}) = \mathcal{O}(n^{-3}).$$

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

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. □

Proof of Theorem 4. Let ψ_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

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

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

We will show below (cf. (38)) 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

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

Using the Integration By Parts (IBP) formula, we get

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

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

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

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

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

Since

$$\int_{\mathbb{R}^d} dx\varphi(x) \int_{\mathbb{R}^d} e^{-\mathrm{i}u\cdot x}\mathbf{E}\Big[e^{\mathrm{i}u\cdot\mathcal{X}_n}\psi_n\zeta(\mathcal{X}_n)\Big]\,du = (2\pi)^d\mathbf{E}\Big[\varphi(\mathcal{X}_n)\psi_n\zeta(\mathcal{X}_n)\Big]$$

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

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

from condition (6) of Theorem 4, where

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

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 ψ_n. Let us denote by ψ a smooth function from $\mathbb{R}$ into $[0,1]$ such that $\psi(t) = 1$ if $|t| \le 1/2$ and $\psi(t) = 0$ if $|t| \ge 1$. We can write

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

with a certain functional K_n satisfying, for every $p > 1$, the condition $\sup_n \|K_n\|_{p,\ell-1} < \infty$. Let $\psi_n = \psi(\tau_n)\psi\Big(2s_n \det\sigma_{\mathcal{X}_n}^{-d}\,K_n\Big)$. 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 $k \to \infty$. 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

$$\det \sigma_{\mathcal{Z}_n} \geq \frac{1}{2} \det \sigma_{\mathcal{X}_n}. \tag{37}$$

By applying the IBP formula and the non-degeneracy assumption for $\mathcal{X}_n$ under truncation, we find that $\sup_n \left| \mathbf{E}\left[e^{\mathrm{i}u \cdot \mathcal{X}_n} \psi_n \zeta(\mathcal{X}_n) \right] \right| \leq \frac{\mathsf{C}}{1+|u|^{\ell-1}}$ for all $u \in \mathbb{R}^d$. Combined with condition (6), this implies that $F(n) = \bar{O}(s_n^2) = \bar{o}(s_n)$. Besides,

$$\left| \mathbf{E}\left[\varphi(\mathcal{X}_n)(1 - \psi_n)\zeta(\mathcal{X}_n) \right] \right| \leq \mathsf{C}_q \|1 - \psi_n\|_q = \bar{o}(s_n). \tag{38}$$

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

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

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

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 (37), it follows that $\sup_n |\hat{g}_n(u)| \leq \frac{\mathsf{C}}{1+|u|^{\ell-1}}$ 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 $C(n) = \bar{o}(1)$, we apply Lebesgue's dominated convergence theorem in conjunction with the estimate

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

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

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

(C is independent of u) and its limiting version $\left| \mathbf{E}\left[e^{\mathrm{i}u \cdot \mathcal{X}_\infty} \mathrm{i}u \cdot \mathcal{Y}_\infty \zeta(\mathcal{X}_\infty) \right] \right| \leq \frac{\mathsf{C}}{1+|u|^{\ell-2}}$.

Combining all the estimates, we get $\int_{\mathbb{R}^d} f(x) \tilde{p}_n(x) \, dx - \int_{\mathbb{R}^d} f(x) h_n^0(x) \, dx = \bar{o}(s_n)$ as $n \to \infty$. From the 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 $h_n^0(x) = \frac{d\Psi_n}{dx}(x) - s_n \mathrm{div}\Big\{ \mathbf{E}[\mathcal{Y}_\infty \,|\, \mathcal{X}_\infty = x] p^{\mathcal{X}_\infty}(x) \Big\}$. The existence of the integral $\int_{\mathbb{R}^d} f(x) h_n^0(x) dx$ is ensured as a consequence under the assumptions of Theorem 4. □

8. Convergence of martingales and quadratic variations

This section collects the proofs of technical results stated in Section 5. The major part of them make use of stochastic analysis and aim at controlling quadratic variations and covariations of some martingales.

Proof of Lemma 5. Let us denote by Φ_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

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

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

$$\Phi_n^2=\sum_{I,J}K_{IJ}\Big(\Big\{\{((I\beta_1)\cdot t)(J\sigma_2)\}\cdot B_2\Big\}_T+\Big\{\{((J\beta_2)\cdot t)(I\sigma_1)\}\cdot B_1\Big\}_T\Big)$$
$$+\sum_{I,J}K_{IJ}\Big(\Big\{\{((J\sigma_2)\cdot B_2)(I\beta_1)\}\cdot t\Big\}_T+\Big\{\{((I\sigma_1)\cdot B_1)(J\beta_2)\}\cdot t\Big\}_T\Big).$$

The last two terms in this expression need some further analysis. Let us introduce the notation $\Phi_n^{21}=\sum_{I,J}K_{IJ}\big\{\{((J\sigma_2)\cdot B_2)(I\beta_1)\}\cdot t\big\}_T$. Since $((J^j\sigma_2)\cdot B_2)_s=0$ for $s\in(0,T^{j-1})$,

$$\begin{aligned}\Phi_n^{21}&=\sum_{i,j}K_{ij}\Big\{\{((J^j\sigma_2)\cdot B_2)I^i[\beta_{1,T^{j-1}}+(1_{(T^{j-1},\infty)}\beta_1^{[0]})\cdot t\\&\quad+(1_{(T^{j-1},\infty)}\beta_{11}^{[1]})\cdot B_1+(1_{(T^{j-1},\infty)}\beta_{12}^{[1]})\cdot B_2]\}\cdot t\Big\}_T\\&=\sum_{i,j}K_{ij}\big(\{(J^j\sigma_2)\cdot B_2\}I^i\beta_{1,T^{j-1}}\big)\cdot t_T+\sum_{I,J}K_{IJ}\big(I\{(J\sigma_2\beta_{11}^{[1]})\cdot\langle B_2,B_1\rangle\}\big)\cdot t_T\\&\quad+\sum_{I,J}K_{IJ}\big(I\{(J\sigma_2\beta_{12}^{[1]})\cdot\langle B_2,B_2\rangle\}\big)\cdot t_T+o_P(b_n).\end{aligned}\tag{39}$$

Let us explain how the last $o_P(b_n)$ is obtained. In fact, the remainder term in the last equation contains five summands which can all be treated similarly. Let us do it for one of them, which has the form $\Psi_n=\sum_{I,J}K_{IJ}\Big\{\big([\{(J\sigma_2)\cdot B_2\}\beta_{11}^{[1]}]\cdot B_1\big)I\Big\}\cdot t_T$. We first use that

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

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

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

under the assumption that $\max_{t\in[0,T]}\mathbf{E}[(\beta_{11,t}^{[1]})^4]$ and $\max_{t\in[0,T]}\sigma_{2,t}$ are finite. Now, interchanging the order of integrations, the first summand in the RHS of (39) can be rewritten as follows

$$\{((J\sigma_2)\cdot B_2)I\beta_{1,T^{j-1}}\}\cdot t_T=\Big\{\{(S^i-S^{i-1}\vee\cdot)_+J\sigma_2\beta_{1,T^{j-1}}\}\cdot B_2\Big\}_T.\tag{40}$$

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

$$\begin{aligned}\Phi_n^{22} &:= \sum_{I,J} K_{IJ}\Big\{\{((I\sigma_1)\cdot B_1)(J\beta_2)\}\cdot t\Big\}_T \\ &= \sum_{i,j} K_{ij}\big(\big\{\big((T^j - T^{j-1}\vee\cdot)_+ I\sigma_1\beta_{2,S^{i-1}}\big)\cdot B_1\big\}_T \\ &\quad + \sum_{I,J} K_{IJ}\big(J\big\{(I\sigma_1\beta_{22}^{[1]})\cdot\langle B_2,B_1\rangle + (I\sigma_1\beta_{21}^{[1]})\cdot\langle B_1,B_1\rangle\big\}\big)\cdot t_T + o_p(b_n). \qquad (41)\end{aligned}$$

Combining (39)-(41) 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. □

Proof of Lemma 6. 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 ρ. 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 Π_n, $\xi^{[11]}$ is a Gaussian process with zero mean. Moreover,

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

where C is a positive constant. This yields the desired result. □

Proof of Lemma 7. One easily checks that

$$\int_0^t \mathbb{H}_s^{1,n}\mathbb{H}_s^{2,n}\,d\langle B_1,B_2\rangle_s = \int_0^t \sum_{I,J} h_s\tilde{K}_{ij}^n(s)(J\sigma_2\cdot B_2)_s(I\sigma_1\cdot B_1)_s\,ds. \qquad (42)$$

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

$$\int_0^t \sum_{I,J} h_s\tilde{K}_{ij}^n(s)(\{(J\sigma_2\cdot B_2)I\sigma_1\cdot B_1\}_s + \{(I\sigma_1\cdot B_1)J\sigma_2\cdot B_2\}_s)\,ds \xrightarrow[n\to\infty]{p} 0, \qquad (43)$$

Thus, the main term is

$$\int_0^t \sum_{I,J} h_s \tilde{K}^n_{ij}(s)\{(IJh)\cdot t\}_s\, ds. \tag{44}$$

To prove (43), we show the convergence in L^2. More rigorously, using the notation $\check{K}^n_{IJ}(s) = \int_s^t \tilde{K_{IJ}}^n(s)h_s\, ds$ and interchanging the order of integrals, we get

$$\begin{aligned}
\mathbf{E}^\Pi\Big(\int_0^t \sum_{I,J} h_s\tilde{K}^n_{ij}(s)\{(J\sigma_2\cdot B_2)I\sigma_1\cdot B_1\}_s\, ds\Big)^2 &= \mathbf{E}^\Pi\Big(\Big\{\sum_{I,J}\check{K}^n_{ij}(J\sigma_2\cdot B_2)I\sigma_1\cdot B_1\Big\}_t\Big)^2\\
&\le \int_0^t \mathbf{E}^\Pi\Big[\Big(\sum_{I,J}\check{K}^n_{IJ}(u)(J\sigma_2\cdot B_2)_u I_u\sigma_{1,u}\Big)^2\Big]\, du\\
&= \int_0^t\int_0^u \Big(\sum_{I,J}\check{K}^n_{IJ}(u)I_u\sigma_{1,u}J_v\sigma_{2,v}\Big)^2 dvdu\\
&\le Cb_n^{-2}\int_0^T\int_0^T r_n^2\sum_{I,J}\Big(K_{IJ}\mathbf{1}_I(u)\mathbf{1}_J(v)\Big)\, dvdu \le Cb_n^{-2}r_n^3.
\end{aligned}$$

Let us show now that the term (44) converges in probability. Simple algebra allows us to rewrite that term in the form

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

which in turn is nothing else but $\int_{[0,t]^2} h_sh_{s'}\mathbf{1}_{\{s\vee s'\le t\}}\{\mathcal{V}^I_n + \mathcal{V}^J_n - \mathcal{V}^{I\cap J}_n\}(ds, ds')$. The weak convergence of measures stated in Assumption P1 completes the proof of the first assertion. The proof of the second assertion is quite similar and therefore is omitted. □

Proof of Lemma 8. Using the representations of M^n and N^n as stochastic integrals, we get

$$\langle M^n, N^n\rangle_t = \int_0^t \big(\mathbb{H}^{1,n}_s\mathbb{G}^{1,n}_s + \mathbb{H}^{1,n}_s\mathbb{G}^{2,n}_s\rho_s + \mathbb{H}^{2,n}_s\mathbb{G}^{1,n}_s\rho_s + \mathbb{H}^{2,n}_s\mathbb{G}^{2,n}_s\big)\, ds. \tag{45}$$

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 $\int_0^t \mathbb{H}^{1,n}_s\mathbb{G}^{11,n}_s\, ds$ tends to zero in probability as $n\to\infty$. Simple algebra yields

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

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,

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

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

$$\mathbf{E}^{\Pi}[\mathcal{T}_{2,n}^2] \le b_n^{-3}\mathbf{E}^{\Pi}\Big[\int_0^t \sum_I I_s \sigma_{1,s}^4 \Big(\int_0^s J(I)_u(\beta_{2,u}-\beta_{2,a_{J(I)}})\,du \int_0^s J(I)_u\sigma_{2,u}\,dB_{2,u}\Big)^2 ds\Big]$$
$$\le Cb_n^{-3}\mathbf{E}^{\Pi}\Big[\int_0^t \sum_I I_s|J(I)|^4\,ds\Big] \le b_n^{-3}r_n^4.$$

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 $n\to\infty$. The convergence to zero of the other terms of the sum in the right-hand side of (45) can be shown similarly. □

Proof of Lemma 9. Let us prove the first assertion, the proof of the second one being completely similar. Since $N^n = \mathbb{G}^{1,n}\cdot B_1 + \mathbb{G}^{n,2}\cdot B_2$ 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

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

Since β_2 is 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 formula 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 $\int_0^t \beta_{2,s}\sigma_{1,s}\,\mathcal{V}^{I,J}(ds)$ in view 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. □

Proof of Lemma 10. Since $N^n = \mathbb{G}^{1,n}\cdot B_1 + \mathbb{G}^{2,n}\cdot B_2$, its quadratic variation is given by $\langle N^n, N^n\rangle = \big[(\mathbb{G}^{1,n})^2 + 2\mathbb{G}^{1,n}\mathbb{G}^{2,n}\rho + (\mathbb{G}^{2,n})^2\big]\cdot t$. Using the semimartingale decomposition of β_2, one checks that

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

Analogous computations show that

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

Now, the desired result follows from Assumption P2. □

9. Technical results on Poisson point processes

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

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. □

Lemma 12. *Let $\mathscr{P}$ be a homogeneous Poisson point process on $\mathbb{R}$ with intensity $\lambda > 0$ and let $a \in \mathbb{R}$. For every ω, let $I_a(w)$ be the interval that contains a and that is an element of the partition of $\mathbb{R}$ generated by $\mathscr{P}$. Then $|I_a|$ is distributed according to the law* Gamma$(2, \lambda)$.

Proof. W.l.o.g. we can assume that $a = 0$. Since the restrictions of $\mathscr{P}$ 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 i.i.d. random variables exponentially distributed with parameter λ. Thus the assertion of the lemma follows from the well known properties of the Gamma distribution. □

Lemma 13. *Let $\mathscr{P}$ 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 ω, let us denote by $N = N(\omega)$ the number of points of $\mathscr{P}(\omega)$ lying in I and by $t_i = t_i(\omega)$, $i = 1, \dots, N$ the ordered sequence of these points. Then*

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

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 $U_{(1)}$, …, $U_{(k)}$ are the corresponding order statistics. Since the joint density of $(U_{(i)}, U_{(i+1)})$ is given by

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

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,

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

The desired result follows now from Lemma 11. □

Lemma 14. *Let $\zeta_1 \sim \mathscr{E}(\lambda_1)$ and $\mathscr{P}^2$ be a Poisson process with intensity λ_2 independent of ζ_1. Let us denote by Π_ζ^2 the partition of $[0, \zeta_1]$ generated by $\mathscr{P}^2$. Then*

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

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

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

Therefore,

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

This completes the proof of the lemma. □

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

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

Proof. We can consider the Poisson point process $\mathscr{P}$ on $[0,T]$ as the union of three independent Poisson point processes: $\mathscr{P}_a$ on $[0,a]$, $\mathscr{P}_I$ on $I=[a,b]$ and $\mathscr{P}_b$ on $[b,T]$. Let $t_1\le\ldots\le t_{N_a}$ (resp. $t_1''\le\ldots\le t_{N_b}''$) be the points of $\mathscr{P}_a$ (resp. $\mathscr{P}_b$). Then $\mathbf{E}[\sum_J|J|K_{IJ}] = \mathbf{E}[(a-t_{N_a})+|I|+(t_1''-b)]$. For every integer $k\ge0$, 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 $U_1,\ldots,U_k$ 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 $t_1'\le\ldots\le t'_{N_I}$ are the points of $\mathscr{P}$ lying in I. Thanks to 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. □

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

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

Proof. Let K be a positive integer and let as denote by

$$w_h(\delta) = \max\{|h(s,s')-h(u,u')| : (s,s',u,u')\in[0,T]^4 \text{ and } |s-u|\le\delta, |s'-u'|\le\delta\} \quad (46)$$

the modulus of continuity of h. Since h is continuous and $[0,t]^2$ is compact, we have $w_h(t/K)\to0$ as $K\to\infty$.

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

$$\mathcal{T}_1 = \lambda\sum_{I\in\Pi}\int_{I\times I}h(s,s')\,ds\,ds' - \lambda\sum_{i=1}^K h\Big(\frac{it}{K},\frac{it}{K}\Big)\sum_{I\in\Pi_i^K}|I|^2,$$

$$\mathcal{T}_2 = \sum_{i=1}^K h\Big(\frac{it}{K},\frac{it}{K}\Big)\Big(\lambda\sum_{I\in\Pi_i^K}|I|^2-\frac{2t}{K}\Big),$$

$$\mathcal{T}_3 = 2\sum_{i=1}^K\frac{t}{K}\,h\Big(\frac{it}{K},\frac{it}{K}\Big) - 2\int_0^t h(s,s)\,ds,$$

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

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

This inequality combined with Lemma 13 implies that

$$\limsup_{\lambda\to\infty}\mathbf{E}[|\mathcal{T}_1|]\le\limsup_{\lambda\to\infty}\Big(\lambda\|h\|_\infty\frac{K}{\lambda^2}+\lambda w_h(t/K)\frac{2t}{\lambda}\Big)=2tw_h(t/K).$$

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 i.i.d. exponentially distributed random variables with scaling parameter one and let $N=\min\{k:\zeta_1+\ldots+\zeta_k\ge npt/K\}$. Then

$$\Big|\lambda\sum_{I\in\Pi_i^K}|I|^2-\frac{2t}{K}\Big|\le\frac{1}{\lambda}\Big|\sum_{j=1}^N(\zeta_j^2-2)\Big|+\Big|\frac{2(N-1)}{\lambda}-\frac{2t}{K}\Big|+\frac{\zeta_N^2+2}{\lambda}.$$

Note that $\mathbf{E}[\zeta_N^2]=6$ by virtue of Lemma 12. In view of the Cauchy-Schwarz inequality and Wald's identity [40, Ch. VII, Thm. 3, Eq. (15)], we get $\mathbf{E}[|\sum_{j=1}^N(\zeta_j^2-2)|]\le[\mathrm{Var}(\zeta_1^2)\,\mathbf{E}(N)]^{1/2}=O(\lambda^{1/2})$. Finally, it is clear that $|\mathcal{T}_3|\le 2t\,w_h(t/K)$. Putting these estimates together, we get $\limsup_{\lambda\to\infty}\mathbf{E}[|\mathcal{T}_1+\mathcal{T}_2+\mathcal{T}_3|]\le 4tw_h(t/K)$. Using the fact that $w_h(t/K)$ tends to zero as $K\to\infty$, we arrive at the desired result. □

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

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

holds for sufficiently large λ_0, with $w_h(\cdot)$ being defined by (46).

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/λ_0. Let $\mathcal{E}$ be the event "for every $i=1,\ldots,4N(x)$, the interval

$[\frac{i}{4N(x)}, \frac{(i+1)}{4N(x)}]$ contains at least one point from Π^1 and one point from Π^2". The total probability formula implies that

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

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

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

$$\mathcal{T}_{11} = \mathbf{E}^{\mathcal{E}}\Big[\sum_{i=1}^{N(x)} \eta_i^\circ\Big] - 2\int_0^1 \bar{h}(s)\,ds, \quad \mathcal{T}_{1s} = \sum_{i=1}^{[N(x)/2]} (\eta_{2i+s-2}^\circ - \mathbf{E}^{\mathcal{E}}[\eta_{2i+s-2}^\circ]), \quad s = 2, 3.$$

Let us emphasize that 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

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

Using the inequality of Cauchy-Schwarz, as well as the bounds $\mathbf{P}(\mathcal{E}^c) \leq C\lambda_0 e^{-x/C}$ and (36), we get $\big|\mathbf{E}^{\mathcal{E}}\big[\sum_{i=1}^{N(x)} \eta_i^\circ\big] - \mathbf{E}\big[\sum_{i=1}^{N(x)} \eta_i^\circ\big]\big| \leq C\lambda_0 e^{-x/C}$, for some constant C and for every $x > C\log\lambda_0$.

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

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

Using once again Lemma 15, we get

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

Wald's equality yields

$$\mathbf{E}\Big[\sum_{I: a_I \in L_i} |I|^k\Big] = k!|L_i|\lambda_1^{1-k} + \mathcal{O}(\lambda_1^{-k}), \tag{47}$$

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

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

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 η_i°s is that, conditionally to $\mathcal{E}$, the random variables η_{2i}°, $i = 1, \dots, [N(x)/2]$, are independent. Indeed, one easily checks that conditionally to $\mathcal{E}$, η_{2i}° depends only on the restrictions of $\mathscr{P}^1$ and $\mathscr{P}^1$ 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 $\mathscr{P}^k$, $k = 1, 2$, onto these intervals are independent. Therefore, η_{2i}°, $i = 1, \dots, [N(x)/2]$, form a sequence of random variables that are independent conditionally to $\mathcal{E}$. Moreover, conditionally to $\mathcal{E}$, they verify $|\eta_i^\circ| \le C\lambda_0 r|L_i| \le Cx^2/\lambda_0$. 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:

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

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. □

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

$$n^2 \sum_{I \in \Pi_n^1} \int_{I\times J(I)\times J(I)} h(s,t,u)\,ds\,dt\,du \xrightarrow[n\to\infty]{P} \Big(\frac{6}{p_1^2} + \frac{8}{p_1p_2} + \frac{6}{p_2^2}\Big)\int_0^T h(s,s,s)\,ds.$$

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 $\{L_i = \big[(i-1)/N, i/N\big),\ i = 1, \dots, N\}$ with $N = [n^{1-\varepsilon}]$ slightly smaller than n (ε 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

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

Let now $i \in 2, \dots, N-1$ and I be an interval of Π^1 satisfying $a_I \in L_i$, then $\big||J(I)| - |I| - \xi_1^\circ - \xi_2^\circ\big| \le (\xi_1^\circ - N^{-1})_+ + (\xi_2^\circ - N^{-1})_+$, where ξ_1° and ξ_2° are two random variables distributed according to the exponential distribution with parameters np_2 conditionally to Π^1. Moreover, conditionally to Π^1, ξ_1° and ξ_2° 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 (47), we get

$$\mathbf{E}[\eta_i^\circ] = \bar{h}(l_i)|L_i|\Big(\frac{6}{p_1^2} + \frac{8}{p_1p_2} + \frac{6}{p_2^2}\Big) + n^2|L_i|O(n^{-1-2\varepsilon}) = \Big(\frac{6}{p_1^2} + \frac{8}{p_1p_2} + \frac{6}{p_2^2}\Big)\int_{L_i} \bar{h}(s)\,ds + o(1).$$

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 $\mathrm{Var}[\sum_i \eta_i^\circ] = O(N \max_i \mathrm{Var}(\eta_i^\circ)) + o(N^2|L_1|^2)$. Since $|\eta_i^\circ| \le C(nr_n)^2|L_1|$ for every i, we get $\mathrm{Var}[\sum_i \eta_i^\circ] = O(N|L_1|^2\mathbf{E}[(nr_n)^4]) + o(1) = o(1)$ and the desired convergence property follows from the convergence of $\mathcal{T}_n$ in L^2. □

Lemma 19. *Let $T > 0$ and let $\mathscr{P}_n^i$, $i = 1, 2$, 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 $h : [0, T]^3 \to \mathbb{R}$*

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

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. □

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] Babu, G. J. and Singh, K. On one term Edgeworth correction by Efron's bootstrap. *Sankhya-Ser. A*, **46** (1984), no. 2, 219-232.

[5] Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A. and Shephard, N. Multivariate Realised Kernels: Consistent Positive Semi-Definite Estimators of the Covariation of Equity Prices with Noise and Non-Synchronous Trading. Manuscript available at `http://www.nuffield.ox.ac.uk/economics/papers/index2007and2008.aspx`

[6] 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.

[7] Bertail, P. and Clémencon, S. Edgeworth expansions of suitably normalized sample mean statistics for atomic Markov chains. *Probab. Theory Related Fields* **130** (2004), no. 3, 388-414.

[8] Bose, A. Edgeworth correction by bootstrap in autoregressions. *Ann. Statist.* **16** (1988), no. 4, 1709-1722.

[9] Comte, F. and Renault, E. Long memory in continuous-time stochastic volatility models. *Math. Finance*, **8** (1998), 291-323.

[10] Dacunha-Castelle, D. and Florens-Zmirou, D. Estimation of the coefficients of diffusion from discrete observations. *Stochastics*, **19** (1986), 263-284.

[11] Epps, T. Comovements in Stock Prices in the Very Short Run. *Journal of the American Statistical Association*, **74**, No. 366 (1979), 291-298.

[12] Florens-Zmirou, D. On estimating the diffusion coefficient from discrete observations. *J. Appl. Probab.* **30** (1993), no. 4, 790–804.

[13] Fukasawa, M. Edgeworth expansion for ergodic diffusions. *Probab. Theory Related Fields* **142** (2008), no. 1-2, 1-20.

[14] 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.

[15] 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`.

[16] Hall, P. *The bootstrap and Edgeworth expansion.* Springer Series in Statistics. Springer-Verlag, New York, 1992.

[17] Hayashi, T. and Kusuoka, S. Consistent estimation of covariation under non-synchronicity. *Stat. Inference Stoch. Process.* **11** (2008), no. 1, 93–106.

[18] Hayashi, T. and Yoshida, N. On covariance estimation of non-synchronously observed diffusion processes. preprint (2003). *Bernoulli*, 11, 359-379 (2005).

[19] Hayashi, T. and Yoshida, N. Asymptotic normality of a covariance estimator for non-synchronously observed diffusion processes. *Ann. Inst. Statist. Math.* **60** (2008), no. 2, 367–406.

[20] Hayashi, T. and Yoshida, N. Nonsynchronous covariance estimator and limit theorem. preprint (2006).
[21] Hayashi, T. and Yoshida, N. Nonsynchronous covariance estimator and limit theorem II. preprint (2008).
[22] Hoshikawa, T., Kanatani, T., Nagai, K. and Nishiyama, Y. Nonparametric estimation methods of integrated multivariate volatilities. Working Paper (2006).
[23] Jacod, J. On processes with conditional independent increments and stable convergence in law. *Seminaire de probabilites de Strasbourg*, 36 (2002), 383–401.
[24] Jacod, J. and Shiryaev, A. *Limit theorems for stochastic processes*. Second edition. Springer-Verlag, Berlin, 2003.
[25] Koul, H. and Surgailis, D. Asymptotic expansion of M-estimators with long-memory errors. Ann. Statist. **25** (1997), no. 2, 818-850.
[26] Kessler, M. Estimation of diffusion processes from discrete observations. *Scand. J. Statist.* **24** (1997), 211-229.
[27] Lo, A. W. and MacKinlay, A. C. An econometric analysis of non-synchronous trading. *J. Econometrics* **45** (1990), no. 1-2, 181–211.
[28] Malliavin, P. and Mancino, M.E. Fourier Series Method for Measurement of Multivariate Volatilities. *Finance Stoch.* **6** (2002), 49-61.
[29] Mykland, P.A. Asymptotic expansions for martingales. *Ann. Probab.* **21** (1993), 800-818.
[30] Mykland, P.A. A Gaussian Calculus for Inference from High Frequency Data. Technical Report 563, Dept of Statistics, The Univ. of Chicago.
[31] Mykland, P.A. and Zhang, L. Anova for diffusions and Ito processes. *Ann. Statist.* **34** (2006), 1931-1963.
[32] Nualart, D. *The Malliavin calculus and related topics*. Second edition. Probability and its Applications (New York). Springer-Verlag, Berlin, 2006.
[33] Palandri, A. Consistent realized covariance for asynchronous observations contaminated by market microstructure noise. Manuscript, available at `http://www.palandri.eu/research.html`.
[34] 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.
[35] Prakasa-Rao, B. L. S. Statistical inference from sampled data for stochastic processes. *Contemp. Math.*, **80** (1988), 249-284.
[36] Revuz, D. and Yor, M. *Continuous martingales and Brownian motion*. Third edition. Springer-Verlag, Berlin, 1999.
[37] Robert, C. and Rosenbaum, M. Ultra high frequency volatility and co-volatility estimation in a microstructure model with uncertainty zones. Submitted.
[38] Sakamoto, Y. and Yoshida, N. Asymptotic expansion under degeneracy. *J. Japan Stat. Soc.*, **33** (2003), 145–156.
[39] Shanken, J. Nonsynchronous Data and the Covariance-Factor Structure of Returns. *The Journal of Finance*, **42**, No. 2 (1987), 221-231.
[40] Shiryaev, A. N. *Probability*. Second edition. Graduate Texts in Mathematics, 95. Springer-Verlag, New York, 1996.
[41] Scholes, M. and Williams, J. Estimating betas from non-synchronous data. *Journal of Financial Economics*, **5** (1977), 309328.
[42] Sweeting, T.J. Speeds of convergence for the multidimensional central limit theorem. *Ann. Probab.* **5** (1977), 28-41.
[43] 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.
[44] Yoshida, N. Estimation for diffusion processes from discrete observation. *J. Multivariate Anal.* **41** (1992), 220-242.
[45] Yoshida, N. Malliavin calculus and asymptotic expansion for martingales. *Probab. Theory Related Fields*, 109 (1997), 301-342.
[46] Zhang, L. Estimating Covariation: Epps Effect, Microstructure Noise. *Journal of Econometrics*, forthcoming.
[47] Zhang, L., Mykland, P.A., and Ait-Sahalia, Y. Edgeworth expansions for realized volatility and related estimators. *Journal of Econometrics*, forthcoming.