Mixed-normal limit theorems for multiple Skorohod integrals in high-dimensions, with application to realized covariance

Yuta Koike[*]

Graduate School of Mathematical Sciences
The University of Tokyo
3-8-1 Komaba, Meguro-ku, Tokyo 153-8914
Japan
e-mail: kyuta@ms.u-tokyo.ac.jp

Abstract: This paper develops mixed-normal approximations for probabilities that vectors of multiple Skorohod integrals belong to random convex polytopes when the dimensions of the vectors possibly diverge to infinity. We apply the developed theory to establish the asymptotic mixed normality of the realized covariance matrix of a high-dimensional continuous semimartingale observed at a high-frequency, where the dimension can be much larger than the sample size. We also present an application of this result to testing the residual sparsity of a high-dimensional continuous-time factor model.

MSC 2010 subject classifications: 60F05, 60H07, 62H15.
Keywords and phrases: Bootstrap, central limit theorem, Chernozhukov-Chetverikov-Kato theory, high-frequency data, Malliavin calculus, multiple testing.

1. Introduction

Covariance matrix estimation of multiple assets is one of the most active research areas in high-frequency financial econometrics. Recently, many authors have been attacking the high-dimensionality in covariance matrix estimation from high-frequency data. A pioneering work on this topic is the paper by Wang & Zou [68], where the regularization methods (banding and thresholding) proposed in Bickel & Levina [7, 6] have been applied to estimating high-dimensional quadratic covariation matrices from noisy and non-synchronous high-frequency data. Subsequently, their approach has been enhanced by several papers such as [65, 44, 42]. Meanwhile, such methods require a kind of sparsity of the target quadratic covariation matrix itself, which seems unrealistic in financial data in view of the celebrated factor structure such as the Fama-French three-factor

model of [27]. To overcome this issue, Fan *et al.* [28] have proposed a covariance estimation method based on a continuous-time (approximate) factor model with observable factors, which can be seen as a continuous-time counterpart of the method introduced in Fan *et al.* [31]. The method has been further extended in various directions such as situations with unobservable factor, noisy and non-synchronous observations, heavy-tail errors and so on; see [1, 23, 43, 29, 59] for details. As an alternative approach to avoid assuming the sparsity of the target matrix itself, Brownlees *et al.* [9] have proposed applying the graphical Lasso, which imposes the sparsity on the inverse of the target matrix rather than the target matrix itself. On the empirical side, high-dimensional covariance matrix estimation from high-frequency financial data is particularly interesting in portfolio allocation. We refer to [67, 30, 49] for illustrations of relevant empirical work on this topic, in addition to the empirical results reported in the papers cited above.

To the best of the author's knowledge, however, there is no work to establish a statistical inference theory validating simultaneous hypothesis testing and construction of uniformly valid confidence regions for high-dimensional quadratic covariation estimation from high-frequency data. Such a theory is important in statistical applications as illustrated by the following example: Let $Y = (Y_t)_{t \in [0,1]}$ be a d-dimensional continuous semimartingale. We denote by Y^i the i-th component of Y for every $i = 1, \ldots, d$. If one attempts to apply a regularization procedure to estimating the quadratic covariation matrix $[Y, Y]_1 = ([Y^i, Y^j]_1)_{1 \le i, j \le d}$ of Y, it is important to understand whether the target matrix is really sparse or not, and if so, how sparse it is. This amounts to evaluating the following series of the statistical hypotheses *simultaneously*:

$$H_{(i,j)} : [Y^i, Y^j]_1 \equiv 0, \qquad i, j = 1, \ldots, d \text{ such that } i < j. \tag{1.1}$$

A natural way to construct test statistics for this problem is to estimate $[Y, Y]_1$ and test whether each of the entries is significantly away from 0 or not. Now suppose that Y is observed at the equidistant times $t_h = h/n$, $h = 0, 1, \ldots, n$. Then the most canonical estimator for $[Y, Y]_1$ would be the so-called *realized covariance matrix*:

$$\widehat{[Y, Y]}_1^n := \sum_{h=1}^n (Y_{t_h} - Y_{t_{h-1}})(Y_{t_h} - Y_{t_{h-1}})^\top. \tag{1.2}$$

If one wants to test the null hypothesis such that all the hypotheses in (1.1) is true, it is natural to consider the maximum type statistic

$$\max_{(i,j) \in \Lambda} \left| \widehat{[Y, Y]}_1^n \right|,$$

where $\Lambda := \{(i, j) \in \{1, \ldots, d\}^2 : i < j\}$. More generally, if one wants to control the family-wise error rate in multiple testing for the hypotheses (1.1), it is enough to approximate the distribution of $\max_{(i,j) \in \mathcal{L}} |\widehat{[Y, Y]}_1^n|$ for any $\mathcal{L} \subset \Lambda$, with the help of the stepdown procedure illustrated in Romano & Wolf [61].

Hence the problem amounts to approximating the distributions of such maximum type statistics in an appropriate sense. Using the test statistics considered in Bibinger & Mykland [5], this type of testing problem can be extended to the sparsity test for the residual processes of a continuous-time factor model with an observable factor and thus promising in applications to high-frequency financial data. In addition, such a problem will also be useful for covariance matrix modeling in a low-frequency setting because it often suffers from the curse of dimensionality due to the increase of the number of unknown parameters to be estimated, and thus it is a common practice to impose a certain structure on covariance matrices for reducing the number of unknown parameters in models. For example, Tao *et al.* [64] have proposed fitting a matrix factor model to daily covariance matrices which are estimated from high-frequency data using the methodology of [68], while Kurose & Omori [46, 47] have introduced a dynamic (multiple-block) equicorrelation structure to multivariate stochastic volatility models. The afore-mentioned testing will be useful for examining the validity of such specification. If the dimension d is fixed, the desired approximation can be obtained as a simple consequence of a multivariate mixed-normal limit theorem for $\sqrt{n}(\widehat{[Y,Y]}_1^n - [Y,Y]_1)$, which is well-studied in the literature and holds true under quite mild assumptions; see e.g. Theorem 5.4.2 of [35]. The problem here is how to establish an analogous result when the dimension d possibly diverges as n tends to infinity.

Indeed, even for the sum of independent random vectors, it is far from trivial to establish such a result in a situation where the dimension is possibly (much) larger than the sample size. This is not surprising because objective random vectors are typically not tight in the usual sense in such a high-dimensional setting, so any standard method to establish central limit theorems no longer works. A significant breakthrough in this subject was achieved by the seminal work of Chernozhukov, Chetverikov & Kato [14], where a Gaussian approximation of the maxima of the sum of independent random vectors in terms of the Kolmogorov distance has been established under quite mild assumptions which allow the dimension is (possibly exponentially) larger than the sample size. With the help of the Gaussian comparison theorem by Chernozhukov *et al.* [17], it enables us to construct feasible statistical inference procedures based on the maximum type statistics. Their theory, which we call the *Chernozhukov-Chetverikov-Kato theory*, or the *CCK theory* for short, has been developed in the subsequent work by Chernozhukov *et al.* [15, 18] and Chernozhukov *et al.* [19]: the first two papers have developed Gaussian approximation of the suprema of empirical processes, while the latter has extended the results of [14] to a central limit theorem for hyperrectangles, or sparsely convex sets in more general. Extension of the CCK theory to statistics other than the sum of independent random vectors has also been studied in many articles: Weakening the independence assumption has been studied in e.g. [70, 71, 16, 10]; Chen [11] and Chen & Kato [12, 13] have developed theories for U-statistics. Moreover, some authors have applied the CCK theory to statistical problems regarding high-frequency data; see Kato & Kurisu [40] and Koike [45]. Nevertheless, none of

the above studies is applicable to our problem due to its non-ergodic nature. That is, the asymptotic covariance matrix is random and depends on the σ-filed of the original probability space, so the asymptotic distribution is essentially non-Gaussian.

Meanwhile, inspection of the proofs of the CCK theory reveals that most the parts do not rely on *any* structure of the underlying statistics. To be precise, let S_n be the random vector corresponding to the objective statistic and suppose that we aim at approximating the distribution of S_n by its Gaussian analog $S_n^\dagger$ which has the same mean and covariance matrix as those of S_n. In the proofs of the CCK theory, the fact that S_n is the sum of independent random vectors is crucial only to obtain a good quantitative estimate for the quantities $|E[f(S_n)] - E[f(S_n^\dagger)]|$ for sufficiently smooth functions f. In the original CCK theory [14, 19], such an estimate has been established by the so-called *Stein's method*, especially *Slepian's interpolation* (also known as the *smart path method*) and *Stein's leave-one-out method*. Although their approach is not directly applicable to our problem, it suggests that we might alternatively use *Malliavin's integration by parts formula* because it can be viewed as an infinite-dimensional version of Stein's identity (cf. Sakamoto & Yoshida [63]). In fact, the recent active research in probabilistic literature shows a beautiful harmony between Malliavin calculus and Stein's method, which is nowadays called the *Malliavin-Stein method*; we refer to the monograph [54] for an introduction of this subject. Indeed, this idea has already been applied in [45] to a situation where S_n is a vector of smooth Wiener functionals (especially multiple Wiener-Itô integrals) and $S_n^\dagger$ is Gaussian, which has produced several impressive results. Our plan here is to apply this idea to a situation where S_n is a vector of multiple Skorohod integrals and $S_n^\dagger$ is conditionally Gaussian. In this regard, a relevant result has been given in Theorem 5.1 of Nourdin *et al.* [53]. However, this result is not directly applicable to the current situation because it assumes that the components of $S_n^\dagger$ are conditionally independent, which is less interesting to statistical applications (and especially not the case in the problem illustrated above). To remove such a restriction from the result of [53], we employ the novel interpolation method introduced in Nualart & Yoshida [57], instead of Slepian's interpolation used in [53] and the original CCK theory.

Another problem in the present context is validation of standardizing statistics by random variables. In a low-dimensional setting, this is typically achieved by proving the so-called *stable convergence in law* (see e.g. [60] for details). However, in a high-dimensional setting, the meaning of stable convergence is unclear and its naïve extension is not useful because of the lack of the continuous mapping theorem and the delta method (see Section 3 for a relevant discussion). So we also aim at developing a formulation appropriate to validating such an operation.

The remainder of the paper is organized as follows. Section 2 is devoted to some preliminaries on notation and concepts used in the paper. Section 3 presents the main results obtained in this paper. In Section 4 we apply the developed theory to establish the asymptotic mixed normality of realized covariance matrices in a high-dimensional setting and illustrate its application to

testing the residual sparsity of a continuous-time factor model. Section 5 provides a small simulation study as well as an empirical illustration using real data. All the proofs are collected in the Appendix.

2. Preliminaries

In this section we present some notation and concepts used throughout the paper.

2.1. Basic notation

We begin by introducing some basic notation which is more or less common in the literature. For a vector $x \in \mathbb{R}^d$, we write the i-th component of x as x^i for $i = 1, \ldots, d$. Also, we set $\min x := \min_{1 \leq i \leq d} x^i$. For two vectors $x, y \in \mathbb{R}^d$, the statement $x \leq y$ means $x^i \leq y^i$ for all $i = 1, \ldots, d$. For a vector $x \in \mathbb{R}^d$ and a scalar $a \in \mathbb{R}$, we set

$$x \pm a := (x^1 \pm a, \ldots, x^d \pm a)^\top.$$

Here, $\top$ stands for the transpose of a matrix.

For a matrix A, we write its (i, j)-th entry as A^{ij}. Also, $A^{i \cdot}$ and $A^{\cdot j}$ denote the i-th row vector and the j-th column vector, respectively. Here, we regard both the vectors $A^{i \cdot}$ and $A^{\cdot j}$ as column vectors. If A is an $m \times d$ matrix, we denote by $\|\|A\|\|_\infty$ the ℓ_∞-operator norm of A:

$$\|\|A\|\|_\infty = \max_{1 \leq i \leq m} \sum_{j=1}^d |A^{ij}|.$$

If B is another $m \times d$ matrix, we denote by $A \cdot B$ the Frobenius inner product of A and B. That is,

$$A \cdot B := \sum_{i=1}^m \sum_{j=1}^d A^{ij} B^{ij}.$$

For a $d \times d$ matrix A, we denote by $\mathrm{diag}(A)$ the d-dimensional vector consisting of the diagonal entries of A, i.e. $\mathrm{diag}(A) = (A^{11}, \ldots, A^{dd})^\top$.

For a random variable ξ and a number $p > 0$, we write $\|\xi\|_p = (E[|\xi|^p])^{1/p}$. We also use the notation $\|\xi\|_\infty$ to denote the essential supremum of ξ. We will denote by $L^{\infty-}$ the space of all random variables ξ such that $\|\xi\|_p < \infty$ for every $p \in [1, \infty)$. The notation $\to^p$ stands for convergence in probability.

If V is a real Hilbert space, we denote by $\langle \cdot, \cdot \rangle_V$ and $\| \cdot \|_V$ the inner product and norm of V, respectively. Also, we denote by $L^p(\Omega; V)$ the set of all V-valued random variables ξ such that $E[\|\xi\|_V^2] < \infty$.

Given real Hilbert spaces $V_1, \ldots, V_k$, we write their Hilbert space tensor product as $V_1 \otimes \cdots \otimes V_k$. For a real Hilbert space V, we write the kth tensor power of V as $V^{\otimes k}$, i.e.

$$V^{\otimes k} := \underbrace{V \otimes \cdots \otimes V}_{k}.$$

Note that the Hilbert space tensor product is uniquely determined up to isomorphism, and we often select a convenient realization case by case. For example, we identify the tensor product $V \otimes \mathbb{R}^d$ with the Hilbert space V^d equipped with the inner product $\langle (f_1, \ldots, f_d), (g_1, \ldots, g_d) \rangle_{V^d} = \sum_{i=1}^d \langle f_i, g_i \rangle_V$ for $f_1, \ldots, f_d$, $g_1, \ldots, g_d \in V$. This is possible because the latter is the Hilbert space tensor product of V and $\mathbb{R}^d$ in the sense of Definition E.8 in [37]. Namely, there is a bilinear map $T : V \times \mathbb{R}^d \to V^d$ such that the range of T is total in V^d and

$$\langle T(f, a), T(g, b) \rangle_{V^d} = \langle f, g \rangle_V \langle a, b \rangle_{\mathbb{R}^d}$$

for all $f, g \in V$ and $a, b \in \mathbb{R}^d$. In fact, we may define T by

$$T(f, a) = (a_1 f, \ldots, a_d f) \qquad (f \in V, \ a = (a_1, \ldots, a_d)^\top \in \mathbb{R}^d).$$

Evidently, T is bilinear and its range is total in $(H^{\otimes k})^d$. Moreover, for any $f, g \in V$ and $a = (a_1, \ldots, a_d)^\top \in \mathbb{R}^d$, $b = (b_1, \ldots, b_d)^\top \in \mathbb{R}^d$,

$$\langle T(f, a), T(g, b) \rangle_{V^d} = \sum_{i=1}^d \langle a_i f, b_i g \rangle_V = \sum_{i=1}^d a_i b_i \langle f, g \rangle_V = \langle f, g \rangle_V \langle a, b \rangle_{\mathbb{R}^d}.$$

For an element $f \in V^{\otimes k}$, we write the (canonical) symmetrization of f as $\mathrm{Sym}(f)$. Namely, the map $V^{\otimes k} \ni f \mapsto \mathrm{Sym}(f) \in V^{\otimes k}$ is characterized as the unique continuous linear operator on $V^{\otimes k}$ such that

$$\mathrm{Sym}(f_1 \otimes \cdots \otimes f_k) = \frac{1}{k!} \sum_{\tau \in \mathcal{S}_k} f_{\tau(1)} \otimes \cdots \otimes f_{\tau(k)}$$

for all $f_1, \ldots, f_k \in V$, where $\mathcal{S}_k$ denotes the set of all permutations of $\{1, \ldots, k\}$, i.e. the symmetric group of degree k. An element $f \in V^{\otimes k}$ is said to be *symmetric* if $\mathrm{Sym}(f) = f$. We refer to Appendix E of [37] for details on Hilbert space tensor products.

2.2. Multi-way arrays

In this subsection we introduce some notation related to multi-way arrays (or tensors) which are necessary to state our main results.

Given a positive integer N, we set $[N] := \{1, \ldots, N\}$ for short. We denote by $\mathbb{K}$ the real field $\mathbb{R}$ or the complex field $\mathbb{C}$ and consider a vector space V over $\mathbb{K}$. Given q positive integers $N_1, \ldots, N_q$, we denote by $V^{N_1 \times \cdots \times N_q}$ the set of all V-valued $N_1 \times \cdots \times N_q$ arrays, i.e. V-valued functions on $[N_1] \times \cdots \times [N_q]$. Note that $V^{N_1 \times N_2}$ corresponds to the set of all V-valued $N_1 \times N_2$ matrices. When $N_1 = \cdots = N_q = N$, we call an element of $V^{N_1 \times \cdots \times N_q}$ a V-valued N-dimensional q-way array. For an array $T \in V^{N_1 \times \cdots \times N_q}$ and indices $i_k \in [N_k]$ $(k = 1, \ldots, q)$, we write $T(i_1, \ldots, i_q)$ as $T^{i_1, \ldots, i_q}$ and T itself as $T = (T^{i_1, \ldots, i_q})_{(i_1, \ldots, i_q) \in \prod_{k=1}^q [N_k]}$. When $V = \mathbb{K}$, $V^{N_1 \times \cdots \times N_q}$ is naturally identified with the Hilbert space tensor product $\mathbb{K}^{N_1} \otimes \cdots \otimes \mathbb{K}^{N_q}$ by the unique linear isomorphism $\iota : \mathbb{K}^{N_1} \otimes \cdots \otimes$

$\mathbb{K}^{N_q} \to \mathbb{K}^{N_1 \times \cdots \times N_q}$ such that $\iota(x_1 \otimes \cdots \otimes x_q) = (x_1^{i_1} \cdots x_q^{i_q})_{(i_1,\ldots,i_q) \in \prod_{k=1}^{q}[N_k]}$ for $x_k \in \mathbb{K}^{N_k}$, $k = 1, \ldots, q$ (cf. Example E.10 of [37]).

For two $\mathbb{K}$-valued arrays $S, T \in \mathbb{K}^{N_1 \times \cdots \times N_q}$, we define their Hadamard-type product (i.e. entry-wise product) by

$$S \circ T := (S^{i_1,\ldots,i_q} T^{i_1,\ldots,i_q})_{(i_1,\ldots,i_q) \in \prod_{k=1}^{q}[N_k]} \in \mathbb{K}^{N_1 \times \cdots \times N_q}.$$

Also, we set

$$\|T\|_{\ell_p} := \begin{cases} \left\{ \displaystyle\sum_{(i_1,\ldots,i_q) \in \prod_{k=1}^{q}[N_k]} |T^{i_1,\ldots,i_q}|^p \right\}^{1/p} & \text{if } p \in (0,\infty), \\ \displaystyle\max_{(i_1,\ldots,i_q) \in \prod_{k=1}^{q}[N_k]} |T^{i_1,\ldots,i_q}| & \text{if } p = \infty. \end{cases}$$

Now suppose that V is a real Hilbert space. For $T \in V^{N_1 \times \cdots \times N_q}$ and $x \in V$, we define

$$\langle T, x \rangle_V := (\langle T^{i_1,\ldots,i_q}, x \rangle_V)_{(i_1,\ldots,i_q) \in \prod_{k=1}^{q}[N_k]} \in \mathbb{R}^{N_1 \times \cdots \times N_q}. \qquad (2.1)$$

Let m be a positive integer. For each $j = 1, \ldots, m$, let V_j be a real Hilbert space, $p_j \in \mathbb{N}$, $N_1^{(j)}, \ldots, N_{p_j}^{(j)} \in \mathbb{N}$ and $T_j \in V_j^{N_1^{(j)} \times \cdots \times N_{p_j}^{(j)}}$. Then we define

$$T_1 \otimes \cdots \otimes T_m$$
$$:= (T_1^{i_1,\ldots,i_{p_1}} \otimes \cdots \otimes T^{i_{p_1+\cdots+p_{m-1}+1},\ldots,i_{p_1+\cdots+p_m}})_{(i_1,\ldots,i_{p_1+\cdots+p_m}) \in \prod_{k=1}^{p_1+\cdots+p_m}[N_k]}$$
$$\in (V_1 \otimes \cdots \otimes V_m)^{N_1^{(1)} \times \cdots \times N_{p_1}^{(1)} \times \cdots \times N_1^{(m)} \times \cdots \times N_{p_m}^{(m)}}. \qquad (2.2)$$

In particular, we write

$$T^{\otimes m} := \underbrace{T \otimes \cdots \otimes T}_{m}.$$

2.3. Malliavin calculus

This subsection introduces some notation and concepts from Malliavin calculus used throughout the paper. We refer to Nualart [55], Chapter 2 of Nourdin & Peccati [54] and Chapter 15 of Janson [37] for further details on this subject.

Given a probability space $(\Omega, \mathcal{F}, P)$, let $\mathbb{W} = (\mathbb{W}(h))_{h \in H}$ be an isonormal Gaussian process over a real separable Hilbert space H.

Let V be another real separable Hilbert space. For any real number $p \geq 1$ and any integer $k \geq 1$, $\mathbb{D}_{k,p}(V)$ denotes the stochastic Sobolev space of V-valued random variables which are k times differentiable in the Malliavin sense and the derivatives up to order k have finite moments of order p. If $F \in \mathbb{D}_{k,p}(V)$, we denote by $D^k F$ the kth Malliavin derivative of F, which is an element of the space $L^p(\Omega; H^{\otimes k} \otimes V)$. We write DF instead of $D^1 F$ for short. We set $\mathbb{D}_{k,\infty}(V) = \bigcap_{p=1}^{\infty} \mathbb{D}_{k,p}(V)$. If $V = \mathbb{R}$, we simply write $\mathbb{D}_{k,p}(V)$ as $\mathbb{D}_{k,p}$.

For a d-dimensional random vector $F \in \mathbb{D}_{k,p}(\mathbb{R}^d)$, we identify the kth Malliavin derivative $D^k F$ of F as the $(H^{\otimes k})^d$-valued random variable $(D^k F^1, \ldots, D^k F^d)$ by identifying $H^{\otimes k} \otimes \mathbb{R}^d$ with $(H^{\otimes k})^d$ as in Section 2.1. Similarly, for a $d \times d'$ matrix valued random variable $F \in \mathbb{D}_{k,p}(\mathbb{R}^{d \times d'})$, we identify $D^k F$ as the $(H^{\otimes k})^{d \times d'}$-valued random variable $(D^k F^{ij})_{(i,j) \in [d] \times [d']}$.

For a positive integer q, we denote by δ^q the q-th multiple Skorohod integral, which is the adjoint operator of the densely defined operator $L^2(\Omega) \supset \mathbb{D}_{q,2} \ni F \mapsto D^q F \in L^2(\Omega; H^{\otimes q})$. That is, the domain $\mathrm{Dom}(\delta^q)$ of δ^q is defined as the set of all $H^{\otimes q}$-valued random variables u such that there is a constant $C > 0$ satisfying $|E[\langle u, D^q F \rangle_{H^{\otimes q}}]| \leq C\|F\|_2$ for all $F \in \mathbb{D}_{q,2}$, and the following duality formula holds for any $u \in \mathrm{Dom}(\delta^q)$ and $F \in \mathbb{D}_{q,2}$:

$$E[F\delta^q(u)] = E[\langle u, D^q F \rangle_{H^{\otimes q}}].$$

2.4. Multi-indices

This subsection collects some notation related to multi-indices.

Let q be a positive integer. We denote by $\mathbb{Z}_+$ the set of all non-negative integers. We define

$$\mathcal{A}(q) := \{\alpha \in \mathbb{Z}_+^q : \alpha_1 + 2\alpha_2 + \cdots + q\alpha_q = q\}.$$

For a multi-index $\alpha = (\alpha_1, \ldots, \alpha_q) \in \mathbb{Z}_+^q$, we set $|\alpha| = \alpha_1 + \cdots + \alpha_q$ as usual. Given another positive integer r, we define

$$\mathcal{N}_r(\alpha) := \left\{ \nu = (\nu_{ij})_{(i,j) \in [q] \times [r]} : \nu_{ij} \in \mathbb{Z}, \sum_{j=1}^r \nu_{ij} = \alpha_i \right\}$$

and

$$\mathcal{N}_r^*(\alpha) = \{\nu = (\nu_{ij}) \in \mathcal{N}_r(\alpha) : \nu_{q1} = 0\}.$$

Moreover, we define

$$\overline{\mathcal{A}}(q) := \bigcup_{p=1}^q \mathcal{A}(p) \quad \text{and} \quad \overline{\mathcal{N}}_r^*(q) := \bigcup_{\alpha \in \overline{\mathcal{A}}(q)} \mathcal{N}_r^*(\alpha).$$

Finally, for an element $\nu = (\nu_{ij}) \in \mathcal{N}_4(\alpha)$, we set $|\nu|_* := |\nu_{\cdot 1}| + 2|\nu_{\cdot 2}| + |\nu_{\cdot 3}|$ and $|\nu|_{**} := |\nu|_* + |\nu_{\cdot 4}|$.

3. Main results

Throughout the paper, we consider an asymptotic theory such that the parameter $n \in \mathbb{N}$ tends to infinity. For each $n \in \mathbb{N}$, we consider a probability space $(\Omega^n, \mathcal{F}^n, P^n)$, and we suppose that all the random variables at stage n are defined on $(\Omega^n, \mathcal{F}^n, P^n)$. We also suppose that an isonormal Gaussian

process $\mathbb{W}_n = (\mathbb{W}_n(h))_{h \in H_n}$ over a real separable Hilbert space H_n is defined on $(\Omega^n, \mathcal{F}^n, P^n)$. To keep the notation simple, we subtract the indices n from $(\Omega^n, \mathcal{F}^n, P^n)$, $\mathbb{W}_n$ and H_n, respectively. So we will write them simply as $(\Omega, \mathcal{F}, P)$, $\mathbb{W}$ and H, respectively. In particular, note that the spaces and the operators associated with $\mathbb{W}$ (which are introduced in Section 2.3) implicitly depend on n, although we do not attach the index n to them.

For each $n \in \mathbb{N}$, let M_n be a d-dimensional random vector consisting of multiple Skorohod integrals:

$$M_n^j = \delta^{q_j}(u_n^j), \qquad j = 1, \ldots, d,$$

where q_j is a positive integer and $u_n^j \in \mathrm{Dom}(\delta^{q_j})$ for every j. Here, we assume that the dimension d possibly depends on n as $d = d_n$, while q_j's do not depend on n. We also assume $d_n \geq 3$ for every n and $\overline{q} := \sup_j q_j < \infty$. Our aim is to study mixed-normal limit theorems for the following functionals:

$$Z_n = M_n + W_n, \qquad n = 1, 2, \ldots,$$

where W_n's are d-dimensional random vectors which represent the uncentered part of the functionals.

Let us introduce mixed-normal random vectors approximating the functionals Z_n in law as follows:

$$\mathfrak{Z}_n = \mathfrak{C}_n^{1/2} \zeta_n + W_n, \qquad n = 1, 2, \ldots.$$

Here, $\mathfrak{C}_n$ is a $d \times d$ symmetric positive semidefinite random matrix and ζ_n is a d-dimensional standard Gaussian vector independent of $\mathcal{F}$, which is defined on an extension of the probability space $(\Omega, \mathcal{F}, P)$ if necessary.

The main aim of this paper is to investigate reasonable regularity conditions under which the distribution of Z_n is well-approximated by that of $\mathfrak{Z}_n$. To be precise, we are interested in the following type of result:

$$\sup_{z \in \mathbb{R}^d} |P(Z_n \leq z) - P(\mathfrak{Z}_n \leq z)| \to 0 \qquad \text{as } n \to \infty.$$

It is well-recognized in statistic literature, however, that this type of result is usually insufficient for statistical applications because it does not ensure standardization by a *random* vector which is still random in the limit; such an operation is crucial for Studentization in the present context. In a low-dimensional setting, this issue is usually resolved by proving the stability of the convergence so that

$$(Z_n, X) \to^{\mathcal{L}} (\mathfrak{Z}_n, X) \qquad \text{as } n \to \infty$$

for any m-dimensional ($\mathcal{F}$-measurable) random variable X, where $\to^{\mathcal{L}}$ denotes the convergence in law. This statement is no longer meaningful in a high-dimensional setting such that $d \to \infty$ as $n \to \infty$, so we need to reformulate it appropriately. A naïve idea is to consider the following statement:

$$\sup_{z \in \mathbb{R}^d, x \in \mathbb{R}^m} |P(Z_n \leq z, X \leq x) - P(\mathfrak{Z}_n \leq z, X \leq x)| \to 0 \qquad \text{as } n \to \infty. \quad (3.1)$$

However, if m depends also on n, this type of statement is not attractive nei-
ther theoretical nor practical points of view due to the following reasons: From
a theoretical point of view, we need to assume a so-called *anti-concentration in-
equality* for X to prove this type of result by the CCK approach, but it is usually
hard to check such an inequality for general random variables, especially when
$m \to \infty$ as $n \to \infty$. Besides, from a practical point of view, it is still unclear
whether the convergence (3.1) ensures the validity of standardization of Z_n be-
cause no analog of the continuous mapping theorem has been established yet
for high-dimensional central limit theorems of the form (3.1). For these reasons
we choose the way to directly prove convergence results for normalized statistics
of Z_n. More formally, let Ξ_n be an $m \times d$ random matrix, where $m = m_n \geq 3$
possibly depends on n. Our aim is to establish

$$\sup_{y \in \mathbb{R}^m} |P(\Xi_n Z_n \leq y) - P(\Xi_n \mathfrak{Z}_n \leq y)| \to 0 \tag{3.2}$$

as $n \to \infty$ under reasonable regularity conditions on Z_n and Ξ_n. Mathematically
speaking, given a vector $y \in \mathbb{R}^m$, the set $\{z \in \mathbb{R}^d : \Xi_n z \leq y\}$ is a finite
intersection of hyperplanes in $\mathbb{R}^d$, i.e. convex polytopes in $\mathbb{R}^d$, so the convergence
(3.2) can be considered as a high-dimensional central limit theorem for random
convex polytopes. If we take Ξ_n as the $d \times d$ diagonal matrix whose diagonals are
the inverses of the "standard errors" of Z_n, the convergence (3.2) does ensures
the validity of (marginal) standardization of Z_n.

Now, our main theorem is stated as follows:

Theorem 3.1. *Suppose that* $M_n, W_n \in \mathbb{D}_{\overline{q}, \infty}(\mathbb{R}^d)$ *and* $\mathfrak{C}_n \in \mathbb{D}_{\overline{q}, \infty}(\mathbb{R}^{d \times d})$ *and
that* u_n^j *is symmetric for all* n *and* j. *Suppose also that* Ξ_n *can be written as* $\Xi_n =
\Upsilon_n \circ \boldsymbol{X}_n$ *with* Υ_n *being an* $m \times d$ *(deterministic) matrix such that* $\|\|\Upsilon_n\|\|_\infty \geq 1$
and $\boldsymbol{X}_n \in \mathbb{D}_{\overline{q}, \infty}(\mathbb{R}^{m \times d})$. *Assume that the following convergences hold true:*

$$\|\|\Upsilon_n\|\|_\infty^2 E\left[\|\boldsymbol{X}_n\|_{\ell_\infty}^2 \|\Delta_n\|_{\ell_\infty}\right] (\log m)^2 \to 0 \tag{3.3}$$

and

$$\|\|\Upsilon_n\|\|_\infty^{|\nu|_{**}+1} E\left[\left(1 + \|\boldsymbol{X}_n\|_{\ell_\infty}^{|\nu|_*+1}\right)\left(1 + \|Z_n\|_{\ell_\infty}^{|\nu.4|} + \|\mathfrak{Z}_n\|_{\ell_\infty}^{|\nu.4|}\right)\right.$$
$$\left. \times \max_{1 \leq j \leq d} \|\Delta_{n,j}(\nu)\|_{\ell_\infty}\right] (\log m)^{\frac{3}{2}|\nu|_{**}+\frac{1}{2}} \to 0 \tag{3.4}$$

as $n \to \infty$ *for every* $\nu \in \overline{\mathcal{N}}_4^*(\overline{q})$, *where*

$$\Delta_n = \left(\langle D^{q_j} M_n^i, u_n^j\rangle_{H^{\otimes q_j}} - \mathfrak{C}_n^{ij}\right)_{1 \leq i,j \leq d} \tag{3.5}$$

and

$$\Delta_{n,j}(\nu) :=$$
$$\left\langle \bigotimes_{k=1}^{q_j} (D^k M_n)^{\otimes \nu_{k1}} \otimes (D^k \mathfrak{C}_n)^{\otimes \nu_{k2}} \otimes (D^k W_n)^{\otimes \nu_{k3}} \otimes (D^k \boldsymbol{X}_n)^{\otimes \nu_{k4}}, u_n^j \right\rangle_{H^{\otimes q_j}}$$
$$\tag{3.6}$$

if $\nu \in \bigcup_{\alpha \in \mathcal{A}(q_j)} \mathcal{N}_4^*(\alpha)$ and $\Delta_{n,j}(\nu) = 0$ otherwise. Assume also that the following condition is satisfied:

$$\lim_{b \downarrow 0} \limsup_{n \to \infty} P(\min \operatorname{diag}(\Xi_n \mathfrak{C}_n \Xi_n^\top) < b) = 0. \tag{3.7}$$

Then we have (3.2) as $n \to \infty$.

Remark 3.1. The variable Δ_n defined in (3.5) is called the *quasi-tangent* in [57].

Remark 3.2. The variable $\Delta_{n,j}(\nu)$ defined in (3.6) takes values in

$$\mathbb{R}^{\overbrace{d \times \cdots \times d}^{|\nu|_*} \times \overbrace{m \times \cdots \times m}^{|\nu_{.4}|} \times \overbrace{d \times \cdots \times d}^{|\nu_{.4}|}}$$

when $\nu = (\nu_{kl}) \in \bigcup_{\alpha \in \mathcal{A}(q_j)} \mathcal{N}_4^*(\alpha)$. To see this, let us recall that $D^k M_n$, $D^k \mathfrak{C}_n$, $D^k W_n$ and $D^k \boldsymbol{X}_n$ take values in $(H^{\otimes k})^d$, $(H^{\otimes k})^{d \times d}$, $(H^{\otimes k})^d$ and $(H^{\otimes k})^{m \times d}$, respectively (cf. Section 2.3). Therefore, according to the notation defined by (2.2), the variable

$$\bigotimes_{k=1}^{q_j} (D^k M_n)^{\otimes \nu_{k1}} \otimes (D^k \mathfrak{C}_n)^{\otimes \nu_{k2}} \otimes (D^k W_n)^{\otimes \nu_{k3}} \otimes (D^k \boldsymbol{X}_n)^{\otimes \nu_{k4}}$$

takes values in

$$\left(H^{\otimes q_j} \right)^{\overbrace{d \times \cdots \times d}^{|\nu|_*} \times \overbrace{m \times \cdots \times m}^{|\nu_{.4}|} \times \overbrace{d \times \cdots \times d}^{|\nu_{.4}|}},$$

where we use the relation $\sum_{k=1}^{q_j} k(\nu_{k1} + \nu_{k2} + \nu_{k3} + \nu_{k4}) = q_j$. Hence, according to the notation defined by (2.1), we obtain

$$\Delta_{n,j}(\nu)$$
$$= \left\langle \bigotimes_{k=1}^{q_j} (D^k M_n)^{\otimes \nu_{k1}} \otimes (D^k \mathfrak{C}_n)^{\otimes \nu_{k2}} \otimes (D^k W_n)^{\otimes \nu_{k3}} \otimes (D^k \boldsymbol{X}_n)^{\otimes \nu_{k4}}, u_n^j \right\rangle_{H^{\otimes q_j}}$$
$$\in \mathbb{R}^{\overbrace{d \times \cdots \times d}^{|\nu|_*} \times \overbrace{m \times \cdots \times m}^{|\nu_{.4}|} \times \overbrace{d \times \cdots \times d}^{|\nu_{.4}|}}.$$

Remark 3.3. In Theorem 3.1, we require all the variables appearing there to have finite moments of all orders just for simplicity. It would be enough for them to have finite moments up to order p only, where p would be a function of $\overline{q}$.

Remark 3.4 (Quantitative bound). As in the original CCK theory, it is possible to give a quantitative version of the convergence (3.2), but we do not implement it here to make the statement of the theorem simpler.

Let us write down conditions (3.3)–(3.4) in the special case that $q_j \in \{1,2\}$ for all j. In this case, setting $\mathcal{J}_q = \{j \in \{1, \ldots, d\} : q_j = q\}$ for $q = 1, 2$, we can rewrite these conditions as follows:

$$\|\!|\Upsilon_n|\!\|_\infty^2 \max_{q=1,2} E \left[\|\boldsymbol{X}_n\|_{\ell^\infty}^2 \max_{1 \le i \le d} \max_{j \in \mathcal{J}_q} \left| \langle D^q M_n^i, u_n^j \rangle_H - \mathfrak{C}_n^{ij} \right| \right] (\log m)^2 \to 0,$$

$$\||\Upsilon_n\||_\infty^3 \max_{q=1,2} E\left[(1+\|\boldsymbol{X}_n\|_{\ell_\infty}^3) \max_{1\le i,j\le d} \max_{k\in\mathcal{J}_q} \left|\langle D^q\mathfrak{C}_n^{ij}, u_n^k\rangle_{H^{\otimes q}}\right|\right](\log m)^{\frac{7}{2}} \to 0,$$

$$\||\Upsilon_n\||_\infty^2 \max_{q=1,2} E\left[(1+\|\boldsymbol{X}_n\|_{\ell_\infty}^2) \max_{1\le i\le d} \max_{j\in\mathcal{J}_q} \left|\langle D^q W_n^i, u_n^j\rangle_{H^{\otimes q}}\right|\right](\log m)^2 \to 0,$$

$$\||\Upsilon_n\||_\infty^2 \max_{q=1,2} E\left[(1+\|Z_n\|_{\ell_\infty}+\|\mathfrak{Z}_n\|_{\ell_\infty}) \max_{1\le i\le m} \max_{1\le j\le d} \max_{k\in\mathcal{J}_q} \left|\langle D^q\boldsymbol{X}_n^{ij}, u_n^k\rangle_{H^{\otimes q}}\right|\right]$$
$$\times(\log m)^2 \to 0,$$

$$\||\Upsilon_n\||_\infty^3 E\left[(1+\|\boldsymbol{X}_n\|_{\ell_\infty}^3) \max_{1\le i,j\le d} \max_{k\in\mathcal{J}_2} \left|\langle DF_n^i\otimes DG_n^j, u_n^k\rangle_{H^{\otimes 2}}\right|\right](\log m)^{\frac{7}{2}} \to 0,$$

$$\||\Upsilon_n\||_\infty^5 E\left[(1+\|\boldsymbol{X}_n\|_{\ell_\infty}^5) \max_{1\le i,j,k,l\le d} \max_{h\in\mathcal{J}_2} \left|\langle D\mathfrak{C}_n^{ij}\otimes D\mathfrak{C}_n^{kl}, u_n^h\rangle_{H^{\otimes 2}}\right|\right](\log m)^{\frac{13}{2}} \to 0,$$

$$\||\Upsilon_n\||_\infty^3 E\left[(1+\|Z_n\|_{\ell_\infty}^2+\|\mathfrak{Z}_n\|_{\ell_\infty}^2)\right.$$
$$\left.\times \max_{1\le i,k\le m} \max_{1\le j,l\le d} \max_{h\in\mathcal{J}_2} \left|\langle D\boldsymbol{X}_n^{ij}\otimes D\boldsymbol{X}_n^{kl}, u_n^h\rangle_{H^{\otimes 2}}\right|\right](\log m)^{\frac{7}{2}} \to 0,$$

$$\||\Upsilon_n\||_\infty^4 E\left[(1+\|\boldsymbol{X}_n\|_{\ell_\infty}^4) \max_{1\le i,j,k\le d} \max_{l\in\mathcal{J}_2} \left|\langle D\mathfrak{C}_n^{ij}\otimes DF_n^k, u_n^l\rangle_{H^{\otimes 2}}\right|\right](\log m)^5 \to 0,$$

$$\||\Upsilon_n\||_\infty^3 E\left[(1+\|\boldsymbol{X}_n\|_{\ell_\infty}^2)(1+\|Z_n\|_{\ell_\infty}+\|\mathfrak{Z}_n\|_{\ell_\infty})\right.$$
$$\left.\times \max_{1\le j\le m} \max_{1\le i,k\le d} \max_{l\in\mathcal{J}_2} \left|\langle DF_n^i\otimes D\boldsymbol{X}_n^{jk}, u_n^l\rangle_{H^{\otimes 2}}\right|\right](\log m)^{\frac{7}{2}} \to 0,$$

$$\||\Upsilon_n\||_\infty^4 E\left[(1+\|\boldsymbol{X}_n\|_{\ell_\infty}^3)(1+\|Z_n\|_{\ell_\infty}+\|\mathfrak{Z}_n\|_{\ell_\infty})\right.$$
$$\left.\times \max_{1\le k\le m} \max_{1\le i,j,l\le d} \max_{h\in\mathcal{J}_2} \left|\langle D\mathfrak{C}_n^{ij}\otimes D\boldsymbol{X}_n^{kl}, u_n^h\rangle_{H^{\otimes 2}}\right|\right](\log m)^5 \to 0,$$

where $F_n, G_n \in \{M_n, W_n\}$. In particular, when $q_j = 1$ for all j, they consist of the following convergences:

$$\||\Upsilon_n\||_\infty^2 E\left[\|\boldsymbol{X}_n\|_{\ell_\infty}^2 \max_{1\le i,j\le d} \left|\langle DM_n^i, u_n^j\rangle_H - \mathfrak{C}_n^{ij}\right|\right](\log m)^2 \to 0,$$

$$\||\Upsilon_n\||_\infty^3 E\left[(1+\|\boldsymbol{X}_n\|_{\ell_\infty}^3) \max_{1\le i,j,k\le d} \left|\langle D\mathfrak{C}_n^{ij}, u_n^k\rangle_H\right|\right](\log m)^{\frac{7}{2}} \to 0,$$

$$\||\Upsilon_n\||_\infty^2 E\left[(1+\|\boldsymbol{X}_n\|_{\ell_\infty}^2) \max_{1\le i,j\le d} \left|\langle DW_n^i, u_n^j\rangle_H\right|\right](\log m)^2 \to 0,$$

$$\||\Upsilon_n\||_\infty^2 E\left[(1+\|Z_n\|_{\ell_\infty}+\|\mathfrak{Z}_n\|_{\ell_\infty}) \max_{1\le i\le m} \max_{1\le j,k\le d} \left|\langle D\boldsymbol{X}_n^{ij}, u_n^k\rangle_H\right|\right](\log m)^2 \to 0.$$

When $q_j = 2$ for all j, they consist of the following convergences:

$$\||\Upsilon_n\||_\infty^2 E\left[\|\boldsymbol{X}_n\|_{\ell_\infty}^2 \max_{1\le i,j\le d} \left|\langle D^2 M_n^i, u_n^j\rangle_{H^{\otimes 2}} - \mathfrak{C}_n^{ij}\right|\right](\log m)^2 \to 0, \qquad (3.8)$$

$$\||\Upsilon_n\||_\infty^3 E\left[(1+\|\boldsymbol{X}_n\|_{\ell_\infty}^3) \max_{1\le i,j,k\le d} \left|\langle D^2\mathfrak{C}_n^{ij}, u_n^k\rangle_{H^{\otimes 2}}\right|\right](\log m)^{\frac{7}{2}} \to 0, \qquad (3.9)$$

$$\|\|\Upsilon_n\|\|_\infty^2 E\left[(1+\|\boldsymbol{X}_n\|_{\ell_\infty}^2)\max_{1\le i,j\le d}\left|\langle D^2 W_n^i, u_n^j\rangle_{H^{\otimes 2}}\right|\right](\log m)^2 \to 0, \qquad (3.10)$$

$$\|\|\Upsilon_n\|\|_\infty^2 E\left[(1+\|Z_n\|_{\ell_\infty}+\|\mathfrak{Z}_n\|_{\ell_\infty})\max_{1\le i\le m}\max_{1\le j,k\le d}\left|\langle D^2 \boldsymbol{X}_n^{ij}, u_n^k\rangle_{H^{\otimes 2}}\right|\right]$$
$$\times (\log m)^2 \to 0, \qquad (3.11)$$

$$\|\|\Upsilon_n\|\|_\infty^3 E\left[(1+\|\boldsymbol{X}_n\|_{\ell_\infty}^3)\max_{1\le i,j,k\le d}\left|\langle DF_n^i \otimes DG_n^j, u_n^k\rangle_{H^{\otimes 2}}\right|\right](\log m)^{\frac{7}{2}} \to 0, \qquad (3.12)$$

$$\|\|\Upsilon_n\|\|_\infty^5 E\left[(1+\|\boldsymbol{X}_n\|_{\ell_\infty}^5)\max_{1\le i,j,k,l,h\le d}\left|\langle D\mathfrak{C}_n^{ij} \otimes D\mathfrak{C}_n^{kl}, u_n^h\rangle_{H^{\otimes 2}}\right|\right](\log m)^{\frac{13}{2}} \to 0, \qquad (3.13)$$

$$\|\|\Upsilon_n\|\|_\infty^3 E\Big[(1+\|Z_n\|_{\ell_\infty}^2+\|\mathfrak{Z}_n\|_{\ell_\infty}^2)$$
$$\times \max_{1\le i,k\le m}\max_{1\le j,l,h\le d}\left|\langle D\boldsymbol{X}_n^{ij} \otimes D\boldsymbol{X}_n^{kl}, u_n^h\rangle_{H^{\otimes 2}}\right|\Big](\log m)^{\frac{7}{2}} \to 0, \quad (3.14)$$

$$\|\|\Upsilon_n\|\|_\infty^4 E\left[(1+\|\boldsymbol{X}_n\|_{\ell_\infty}^4)\max_{1\le i,j,k,l\le d}\left|\langle D\mathfrak{C}_n^{ij} \otimes DF_n^k, u_n^l\rangle_{H^{\otimes 2}}\right|\right](\log m)^5 \to 0, \qquad (3.15)$$

$$\|\|\Upsilon_n\|\|_\infty^3 E\Big[(1+\|\boldsymbol{X}_n\|_{\ell_\infty}^2)(1+\|Z_n\|_{\ell_\infty}+\|\mathfrak{Z}_n\|_{\ell_\infty})$$
$$\times \max_{1\le j\le m}\max_{1\le i,k,l\le d}\left|\langle DF_n^i \otimes D\boldsymbol{X}_n^{jk}, u_n^l\rangle_{H^{\otimes 2}}\right|\Big](\log m)^{\frac{7}{2}} \to 0, \qquad (3.16)$$

$$\|\|\Upsilon_n\|\|_\infty^4 E\Big[(1+\|\boldsymbol{X}_n\|_{\ell_\infty}^3)(1+\|Z_n\|_{\ell_\infty}+\|\mathfrak{Z}_n\|_{\ell_\infty})$$
$$\times \max_{1\le k\le m}\max_{1\le i,j,l,h\le d}\left|\langle D\mathfrak{C}_n^{ij} \otimes D\boldsymbol{X}_n^{kl}, u_n^h\rangle_{H^{\otimes 2}}\right|\Big](\log m)^5 \to 0, \quad (3.17)$$

where $F_n, G_n \in \{M_n, W_n\}$.

As a special case of Theorem 3.1, we can deduce a high-dimensional central limit theorem for multiple Skorohod integrals in hyperrectangles as follows. Let $\mathcal{A}^{\mathrm{re}}(d)$ be the set of all hyperrectangles in $\mathbb{R}^d$, i.e. $\mathcal{A}^{\mathrm{re}}(d)$ consists of all sets A of the form

$$A = \{z \in \mathbb{R}^d : a_j \le z^j \le b_j \text{ for all } j = 1, \ldots, d\}$$

for some $-\infty \le a_j \le b_j \le \infty$, $j = 1, \ldots, d$. Taking Ξ_n as

$$\Xi_n = \begin{pmatrix} \mathsf{E}_d \\ -\mathsf{E}_d \end{pmatrix}$$

in Theorem 3.1, where E_d denotes the identity matrix of size d, we obtain the following result (note that (3.2) continues to hold true while $\mathbb{R}^d$ is replaced by $(-\infty,\infty]^d$):

Corollary 3.1. *Suppose that $M_n, W_n \in \mathbb{D}_{\overline{q},\infty}(\mathbb{R}^d)$ and $\mathfrak{C}_n \in \mathbb{D}_{\overline{q},\infty}(\mathbb{R}^{d\times d})$ and that u_n^j is symmetric for all n and j. Assume that the following convergences*

hold true:

$$E\left[\|\boldsymbol{X}_n\|_{\ell_\infty}^2\|\Delta_n\|_{\ell_\infty}\right](\log d)^2 \to 0$$

and

$$E\left[\left(1+\|\boldsymbol{X}_n\|_{\ell_\infty}^{|\nu|_*+1}\right)\left(1+\|Z_n\|_{\ell_\infty}^{|\nu_{\cdot 4}|}+\|\mathfrak{Z}_n\|_{\ell_\infty}^{|\nu_{\cdot 4}|}\right)\right.$$
$$\left.\times \max_{1\le j\le d}\|\Delta_{n,j}(\nu)\|_{\ell_\infty}\right](\log d)^{\frac{3}{2}|\nu|_{**}+\frac{1}{2}} \to 0$$

as $n \to \infty$ for every $\nu \in \overline{\mathcal{N}}_4^*(\overline{q})$. Assume also that the following condition is satisfied:

$$\lim_{b\downarrow 0}\limsup_{n\to\infty} P(\min\operatorname{diag}(\mathfrak{C}_n) < b) = 0.$$

Then we have

$$\sup_{A\in\mathcal{A}^{\mathrm{re}}(d)} |P(Z_n \in A) - P(\mathfrak{Z}_n \in A)| \to 0$$

as $n \to \infty$.

Some related results for statistical applications

In many applications, the objective variables are only approximately multiple Skorohod integrals. The following lemma is useful for such a situation.

Lemma 3.1. *For each $n \in \mathbb{N}$, let Y_n, Y_n' be m-dimensional random vectors such that*

$$\sqrt{\log m}\|Y_n' - Y_n\|_{\ell_\infty} \to^p 0$$

and

$$\sup_{y\in\mathbb{R}^m} |P(Y_n \le y) - P(\Xi_n\mathfrak{Z}_n \le y)| \to 0$$

as $n \to \infty$. Then we have

$$\sup_{y\in\mathbb{R}^m} |P(Y_n' \le y) - P(\Xi_n\mathfrak{Z}_n \le y)| \to 0$$

as $n \to \infty$, provided that (3.7) holds true.

In terms of statistical applications, the mixed-normal approximation given by Theorem 3.1 is often infeasible because the "asymptotic" covariance matrix $\mathfrak{C}_n$ usually contains unobservable quantities. In the following we give two auxiliary results bridging this gap. The first result ensures the validity of estimating the $\mathcal{F}$-conditional distribution of $\Xi_n\mathfrak{Z}_n$ while we replace $\mathfrak{C}_n, W_n$ and Ξ_n by their estimators.

Proposition 3.1. *For each n, let $\widehat{\mathfrak{C}}_n, \widehat{W}_n$ and $\widehat{\Xi}_n$ be a $d \times d$ symmetric positive semidefinite random matrix, a d-dimensional random vector and an $m \times d$ random matrix, respectively. Set $\widehat{\mathfrak{Z}}_n := \widehat{\mathfrak{C}}_n^{1/2}\zeta_n + \widehat{W}_n$. Suppose that*

$$\sqrt{\log m}\|\widehat{\Xi}_n\widehat{W}_n - \Xi_n W_n\|_{\ell_\infty} \to^p 0, \qquad (\log m)^2\|\widehat{\Xi}_n\widehat{\mathfrak{C}}_n\widehat{\Xi}_n^\top - \Xi_n\mathfrak{C}_n\Xi_n^\top\|_{\ell_\infty} \to^p 0 \tag{3.18}$$

as $n \to \infty$. Then we have

$$\sup_{y \in \mathbb{R}^m} |P(\widehat{\Xi}_n \widehat{\mathfrak{Z}}_n \leq y | \mathcal{F}) - P(\Xi_n \mathfrak{Z}_n \leq y | \mathcal{F})| \to^p 0$$

as $n \to \infty$, provided that (3.7) *holds true.*

We remark that the above proposition only gives a way to estimate the $\mathcal{F}$-conditional distribution of $\Xi_n \mathfrak{Z}_n$ when we have appropriate estimators for relevant variables: It says nothing about how to estimate the *unconditional* distribution of $\Xi_n \mathfrak{Z}_n$. Because of the non-ergodic nature of the problem, in general there seems no hope of consistently estimating the latter quantity even if we can consistently estimate unknown variables contained in $\Xi_n \mathfrak{Z}_n$. In a low-dimensional setting this issue is usually resolved by standardizing the objective statistic by a consistent estimator for its asymptotic covariance matrix, which is validated via the stability of convergence in law. In a high-dimensional setting, however, standardizing the (joint) distribution of the objective statistic is often difficult: Estimators for the conditional covariance matrix of the objective statistic are usually singular because the sample size is smaller than the dimension, and even if it is regular, computation of the inverse is typically time-consuming. Nevertheless, we can fortunately show that, in order to estimate quantiles of the unconditional distribution $\Xi_n \mathfrak{Z}_n$, it is sufficient to only estimate its $\mathcal{F}$-conditional distribution. We remark that this fact has already been known in high-frequency financial econometrics and typically been used to construct jump-related testing procedures; see [36, 48] for example. Formally, we can prove the following result:

Proposition 3.2. *For each $n \in \mathbb{N}$, let $T_n, T_n^\dagger, T_n^*$ be random variables defined on an extension of the probability space $(\Omega, \mathcal{F}, P)$. Suppose that*

$$\sup_{x \in \mathbb{R}} |P(T_n \leq x) - P(T_n^\dagger \leq x)| \to 0, \quad \sup_{x \in \mathbb{R}} |P(T_n^* \leq x | \mathcal{F}) - P(T_n^\dagger \leq x | \mathcal{F})| \to^p 0$$

as $n \to \infty$. Suppose also that there is a sequence (E_n) of elements in $\mathcal{F}$ such that the $\mathcal{F}$-conditional distribution of $T_n^\dagger$ has the density on E_n for every n and $\lim_{n \to \infty} P(E_n) = 1$. For each $n \in \mathbb{N}$, let q_n^* be the $\mathcal{F}$-conditional quantile function of T_n^*:

$$q_n^*(\alpha) = \inf\{x \in \mathbb{R} : P(T_n^* \leq x | \mathcal{F}) \geq \alpha\}, \qquad \alpha \in (0, 1).$$

Then we have

$$P(T_n \leq q_n^*(\alpha)) \to \alpha$$

as $n \to \infty$ for all $\alpha \in (0, 1)$.

4. Application to realized covariance

In this section we assume that the probability space $(\Omega, \mathcal{F}, P)$ admits the structure such that $\Omega = \Omega' \times \mathbf{W}$, $\mathcal{F} = \mathcal{F}' \otimes \mathbf{B}$ and $P = P' \times \mathbf{P}$ for some probability space $(\Omega', \mathcal{F}', P')$ and the r-dimensional Wiener space $(\mathbf{W}, \mathbf{B}, \mathbf{P})$ over time interval $[0, 1]$, and consider the partial Malliavin calculus with respect to the

r-dimensional Brownian motion $B = (B_t)_{t \in [0,1]}$ defined by $B_t(\omega', w) = w(t)$ for $\omega' \in \Omega'$, $w \in \mathbf{W}$ and $t \in [0, 1]$ (cf. Section 6.1 of [69]). In this setting the Hilbert space H coincides with the space $L^2([0, 1]; \mathbb{R}^r)$. We here allow the dimension $r = r_n$ to possibly depend on $n \in \mathbb{N}$, so $(\Omega, \mathcal{F}, P)$ and B may depend on n, but we subtract the index n from the notation. Let $(\mathcal{B}_t)_{t \in [0,1]}$ denote the filtration generated by the canonical process on $\mathbf{W}$, and define the filtration $(\mathcal{F}_t)_{t \in [0,1]}$ of $\mathcal{F}$ by $\mathcal{F}_t := \mathcal{F}' \otimes \mathcal{B}_t$ for $t \in [0, 1]$. On the stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$, we consider the d-dimensional continuous Itô semimartingale $Y = (Y_t)_{t \in [0,1]}$ given by the following:

$$Y_t = Y_0 + \int_0^t \mu_s ds + \int_0^t \sigma_s dB_s, \qquad t \in [0, 1].$$

Here, $\mu = (\mu_s)_{s \in [0,1]}$ is a d-dimensional $(\mathcal{F}_t)$-progressively measurable process and $\sigma = (\sigma_s)_{s \in [0,1]}$ is an $\mathbb{R}^{d \times r}$-valued $(\mathcal{F}_t)$-progressively measurable process such that

$$\int_0^1 \left(\|\mu_s\|_{\ell_1} + \|\sigma_s\|_{\ell_2}^2 \right) ds < \infty \quad \text{a.s.}$$

We remark that the processes μ and σ generally depend on n because d and r may depend on n. However, following the custom of high-dimensional statistics, we subtract the index n from the notation as above.

We observe the process Y at the discrete time points $t_h = t_h^n = h/n$, $h = 0, 1, \ldots, n$. In such a setting, the discretized quadratic covariation matrix

$$\widehat{[Y, Y]}_1^n := \sum_{h=1}^n (Y_{t_h} - Y_{t_{h-1}})(Y_{t_h} - Y_{t_{h-1}})^\top,$$

which is known as the *realized covariance matrix* in high-frequency financial econometrics, is a natural estimator for the quadratic covariance matrix of Y:

$$[Y, Y]_1 = \int_0^1 \Sigma_t dt, \qquad \Sigma_t := \sigma_t \sigma_t^\top.$$

The aim of this section is to establish the asymptotic mixed normality of the estimator $\widehat{[Y, Y]}_1^n$ in a high-dimensional setting such that the dimension d is possibly (much) larger than the sample size n.

Before stating the results, we introduce some notation. First, for a random variable F taking values in $\mathbb{R}^{N_1 \times \cdots \times N_q}$ for some $N_1, \ldots, N_q \in \mathbb{N}$, we set $\|F\|_{p,\ell_2} := \|\|F\|_{\ell_2}\|_p$ for every $p \in (0, \infty]$. Next, for a positive integer k, we identify the space $H^{\otimes k}$ with $L^2([0, 1]^k; (\mathbb{R}^r)^{\otimes k})$ in the canonical way (cf. Example E.10 in [37]). Therefore, if a univariate random variable F is k times differentiable in the Malliavin sense, the kth Malliavin derivative $D^k F$ of F takes values in $L^2([0, 1]^k; (\mathbb{R}^r)^{\otimes k})$, so we can consider the value $D^k F(t_1, \ldots, t_k)$ in $(\mathbb{R}^r)^{\otimes k}$ evaluated at $(t_1, \ldots, t_k) \in [0, 1]^k$. We denote this value by $D_{t_1, \ldots, t_k} F$. Moreover, for an index $(a_1, \ldots, a_k) \in \{1, \ldots, r\}^k$, we write the $(a_1, \ldots, a_k)$-th entry of $D_{t_1, \ldots, t_k} F$ as $D_{t_1, \ldots, t_k}^{(a_1, \ldots, a_k)} F$ (note that we identify $(\mathbb{R}^r)^{\otimes k}$ with $\mathbb{R}^{r \times \cdots \times r}$).

We remark that the variable $D_{t_1,\ldots,t_k}F$ is defined only a.e. on $[0,1]^k \times \Omega$ with respect to the measure $\mathsf{Leb}_k \times P$, where Leb_k denotes the Lebesgue measure on $[0,1]^k$. Therefore, if $D_{t_1,\ldots,t_k}F$ satisfies some property a.e. on $[0,1]^k \times \Omega$ with respect to the measure $\mathsf{Leb}_k \times P$, by convention we will always take a version of $D_{t_1,\ldots,t_k}F$ satisfying that property everywhere on $[0,1]^k \times \Omega$ if necessary. Also, note that if a d-dimensional random vector F is k times differentiable in the Malliavin sense, the kth Malliavin derivative $D^k F$ is first identified with the $(H^{\otimes k})^d$-valued random variable $(D^k F^1, \ldots, D^k F^d)$ according to the identification of $H^{\otimes k} \otimes \mathbb{R}^d$ with $(H^{\otimes k})^d$ (cf. Sections 2.1 and 2.3). Then, each $D^k F^j$ is identified with the $L^2([0,1]^k; (\mathbb{R}^r)^{\otimes k})$-valued random variable as above.

We define the $d^2 \times d^2$ random matrix $\mathfrak{C}_n$ by

$$\mathfrak{C}_n^{(i-1)d+j,(k-1)d+l} :=$$

$$n \sum_{h=1}^n \left\{ \left(\int_{t_{h-1}}^{t_h} \Sigma_s^{ik} ds \right) \left(\int_{t_{h-1}}^{t_h} \Sigma_s^{jl} ds \right) + \left(\int_{t_{h-1}}^{t_h} \Sigma_s^{il} ds \right) \left(\int_{t_{h-1}}^{t_h} \Sigma_s^{jk} ds \right) \right\},$$

$$i,j,k,l = 1,\ldots,d,$$

which plays the role of the conditional covariance matrix of the approximating mixed-normal distribution in our setting.

Remark 4.1. In the fixed dimensional setting, $\mathfrak{C}_n$ converges in probability as $n \to \infty$ to the random matrix $\bar{\mathfrak{C}}$ defined by

$$\bar{\mathfrak{C}}^{(i-1)d+j,(k-1)d+l} := \int_0^1 \left(\Sigma_t^{ik} \Sigma_t^{jl} + \Sigma_t^{il} \Sigma_t^{jk} \right) dt, \qquad i,j,k,l = 1,\ldots,d$$

under mild regularity assumptions, so $\bar{\mathfrak{C}}$ plays the role of the asymptotic covariance matrix in such a setting. However, in the high-dimensional setting the convergence rate of $\mathfrak{C}_n$ to $\bar{\mathfrak{C}}$ does matter and we usually need an additional condition like (4.7) to derive it. To avoid such an extra assumption, we use the "intermediate version" $\mathfrak{C}_n$ of $\bar{\mathfrak{C}}$ to state Theorem 4.1 below.

Theorem 4.1. *Suppose that $\mu_t \in \mathbb{D}_{1,\infty}(\mathbb{R}^d)$ and $\sigma_t \in \mathbb{D}_{2,\infty}(\mathbb{R}^{d \times r})$ for all $t \in [0,1]$. For every $n \in \mathbb{N}$, let $W_n \in \mathbb{D}_{2,\infty}(\mathbb{R}^{d^2})$, $\boldsymbol{X}_n \in \mathbb{D}_{2,\infty}(\mathbb{R}^{m \times d^2})$ and Υ_n be an $m \times d^2$ (deterministic) matrix such that $\||\Upsilon_n\||_\infty \geq 1$, where $m = m_n$ possibly depends on $n \in \mathbb{N}$. Define $\Xi_n := \Upsilon_n \circ \boldsymbol{X}_n$ and assume*

$$\lim_{b \downarrow 0} \limsup_{n \to \infty} P(\min \operatorname{diag}(\Xi_n \mathfrak{C}_n \Xi_n^\top) < b) = 0. \tag{4.1}$$

Then the following statements hold true:

(a) *Suppose that there is a constant $\varpi \in (0, \frac{1}{2})$ such that $\||\Upsilon_n\||_\infty^5 = O(n^\varpi)$ and*

$$\sup_{n \in \mathbb{N}} \max_{1 \leq i \leq d^2} \left(\|W_n^i\|_p + \sup_{0 \leq t \leq 1} \|D_t W_n^i\|_{p,\ell_2} + \sup_{0 \leq s,t \leq 1} \|D_{s,t} W_n^i\|_{p,\ell_2} \right) < \infty, \tag{4.2}$$

$$\sup_{n\in\mathbb{N}} \max_{1\leq i\leq m} \max_{1\leq j\leq d^2} \left(\|\boldsymbol{X}_n^{ij}\|_p + \sup_{0\leq t\leq 1} \|D_t\boldsymbol{X}_n^{ij}\|_{p,\ell_2} + \sup_{0\leq s,t\leq 1} \|D_{s,t}\boldsymbol{X}_n^{ij}\|_{p,\ell_2} \right) < \infty, \tag{4.3}$$

$$\sup_{n\in\mathbb{N}} \max_{1\leq i\leq d} \left(\sup_{0\leq t\leq 1} \|\mu_t^i\|_p + \sup_{0\leq s,t\leq 1} \|D_s\mu_t^i\|_{p,\ell_2} \right) < \infty, \tag{4.4}$$

$$\sup_{n\in\mathbb{N}} \max_{1\leq i\leq d} \left(\sup_{0\leq t\leq 1} \|\Sigma_t^{ii}\|_p + \sup_{0\leq s,t\leq 1} \|D_s\sigma_t^{i\cdot}\|_{p,\ell_2} + \sup_{0\leq s,t,u\leq 1} \|D_{s,t}\sigma_u^{i\cdot}\|_{p,\ell_2} \right) < \infty \tag{4.5}$$

for all $p \in [1,\infty)$. Suppose also that $d = O(n^{\mathfrak{c}})$ and $m = O(n^{\mathfrak{c}})$ as $n \to \infty$ for some $\mathfrak{c} > 0$. Then we have

$$\sup_{y\in\mathbb{R}^m} \left| P\left(\Xi_n\left(S_n + W_n\right) \leq y\right) - P(\Xi_n(\mathfrak{C}_n^{1/2}\zeta_n + W_n) \leq y) \right| \to 0 \tag{4.6}$$

as $n \to \infty$, where

$$S_n := \mathrm{vec}\left[\sqrt{n}\left(\widehat{[Y,Y]}_1^n - [Y,Y]_1\right)\right]$$

and ζ_n is a d^2-dimensional Gaussian vector independent of $\mathcal{F}$.

(b) *Suppose that $\|\|\Upsilon_n\|\|_\infty^5 (\log dm)^{\frac{13}{2}} = o(\sqrt{n})$ as $n \to \infty$ and (4.2)-(4.5) are satisfied for $p = \infty$. Then we have (4.6) as $n \to \infty$.*

Remark 4.2. We enumerate some remarks on the assumptions of Theorem 4.1 in the following:

(a) In typical applications of Theorem 4.1, we take $W_n \equiv 0$ and X_n a smooth functional of the volatility process σ. Hence only the assumptions on μ and σ do matter (see also Section 4.1). The Malliavin differentiability conditions on μ and σ are satisfied, for example, when μ and σ are respectively solutions of stochastic differential equations (SDEs) with sufficiently regular coefficients; see e.g. Section 2.2.2 of [55]. We remark that the (local) Malliavin differentiability has been known for solutions of some SDEs with irregular coefficients as well; see Section 4 of Alòs & Ewald [2] and Lemma 5.9 of Naganuma [52] for example.

(b) A major restriction imposed by the assumptions of Theorem 4.1 is that they require the arrays $(D_s^{(a)}\sigma_t^{ib})_{(a,b)\in[r]^2}$ and $(D_{s,t}^{(a,b)}\sigma_u^{ic})_{(a,b,c)\in[r]^2}$ are sufficiently "sparse" for all $s, t, u \in [0,1]$ so that

$$\sup_{0\leq s,t\leq 1} \|D_s\sigma_t^{i\cdot}\|_{p,\ell_2} \quad \text{and} \quad \sup_{0\leq s,t,u\leq 1} \|D_{s,t}\sigma_u^{i\cdot}\|_{p,\ell_2}$$

do not diverge as $n \to \infty$. This is a restriction because r typically diverges as $n \to \infty$ in a high-dimensional setting. Such a condition is satisfied e.g. when Y^i and $(\sigma_t^{i\cdot})_{t\in[0,1]}$ depend on only finitely many components of B for each i (they may vary with i, though). Therefore, it is satisfied if the price and volatility processes have a certain factor structure, which seems realistic in financial applications.

(c) The Malliavin differentiability condition on μ in Theorem 4.1 can be replaced by a continuity condition on μ analogous to (4.7). In fact, it is used only to prove Lemma B.6, where it is only crucial that μ is well-approximated by a "strongly predictable" process.

(d) Assumptions on the second Malliavin derivatives of the volatility process σ_t sometimes appear in high-frequency financial econometrics even for the fixed-dimensional case; see [22, 21] for example.

(e) The assumptions of Theorem 4.1 do not rule out the possibility of the presence of jumps in the volatility process σ; see Fukasawa [32].

(f) It would be enough in Theorem 4.1(a) to assume conditions (4.2)–(4.5) for some $p \in [1, \infty)$ only, where p depends on the value of $\mathfrak{c}$, i.e. the divergence rates of d and m.

By an analogous discussion to the one before Corollary 3.1, we can deduce a high-dimensional central limit theorem for realized covariance in hyperrectangles from Theorem 4.1:

Corollary 4.1. *Under the assumptions of Theorem 4.1 with replacing* (4.1) *by*

$$\lim_{b\downarrow 0} \limsup_{n\to\infty} P(\min \mathrm{diag}(\mathfrak{C}_n) < b) = 0,$$

we have

$$\sup_{A\in\mathcal{A}^{\mathrm{re}}(d^2)} \left| P\left(S_n + W_n \in A \right) - P\left(\mathfrak{C}_n^{1/2}\zeta_n + W_n \in A \right) \right| \to 0$$

as $n \to \infty$.

In some situations, it is more convenient to consider a localized version of the assumptions of Theorem 4.1 as follows:

Theorem 4.2. *For every $n \in \mathbb{N}$, let W_n be a d^2-dimensional random vector, $\boldsymbol{X}_n$ be an $m \times d^2$ random matrix and Υ_n be an $m \times d^2$ (deterministic) matrix such that $\|\|\Upsilon_n\|\|_\infty \geq 1$, where $m = m_n$ possibly depends on $n \in \mathbb{N}$. Moreover, for every $\nu \in \mathbb{N}$, let $\Omega_n(\nu) \in \mathcal{F}$, $\mu(\nu) = (\mu(\nu)_t)_{t\in[0,1]}$ be a d-dimensional $(\mathcal{F}_t)$-progressively measurable process, $\sigma(\nu) = (\sigma(\nu)_t)_{t\in[0,1]}$ be an $\mathbb{R}^{d\times r}$-valued $(\mathcal{F}_t)$-progressively measurable process, $W_n(\nu) \in \mathbb{D}_{2,\infty}(\mathbb{R}^{d^2})$ and $\boldsymbol{X}_n(\nu) \in \mathbb{D}_{2,\infty}(\mathbb{R}^{m\times d^2})$, and suppose that the following conditions are satisfied:*

(i) $\lim_{\nu\to\infty} \limsup_{n\to\infty} P(\Omega_n(\nu)^c) = 0.$

(ii) *For all $\nu \in \mathbb{N}$ and $t \in [0,1]$, $\mu_t = \mu(\nu)_t$ and $\sigma_t = \sigma(\nu)_t$ on $\Omega_n(\nu)$ as well as $\mu(\nu)_t \in \mathbb{D}_{1,\infty}(\mathbb{R}^d)$ and $\sigma(\nu)_t \in \mathbb{D}_{2,\infty}(\mathbb{R}^{d\times r})$.*

(iii) *For all $\nu \in \mathbb{N}$, $W_n = W_n(\nu)$ and $\boldsymbol{X}_n = \boldsymbol{X}_n(\nu)$ on $\Omega_n(\nu)$.*

(iv) *For all $\nu \in \mathbb{N}$, (4.1) holds true with replacing $\boldsymbol{X}_n$ and σ by $\boldsymbol{X}_n(\nu)$ and $\sigma(\nu)$ respectively.*

Then the following statements hold true:

(a) *Suppose that there are constants $\varpi \in (0, \frac{1}{2})$ and $\mathfrak{c} > 0$ such that $\|\|\Upsilon_n\|\|_\infty^5 = O(n^\varpi)$, $d = O(n^\mathfrak{c})$ and $m = O(n^\mathfrak{c})$ as $n \to \infty$. Suppose also that, for all*

$\nu \in \mathbb{N}$, (4.2)-(4.5) *are satisfied for all* $p \in [1, \infty)$ *with replacing* $W_n, \boldsymbol{X}_n, \mu, \sigma$
by $W_n(\nu), \boldsymbol{X}_n(\nu), \mu(\nu), \sigma(\nu)$ *respectively. Then we have* (4.6) *as* $n \to \infty$.

(b) *Suppose that* $\|\|\Upsilon_n\|\|_\infty^5 (\log dm)^{\frac{13}{2}} = o(\sqrt{n})$ *as* $n \to \infty$ *and, for all* $\nu \in$
$\mathbb{N}$, (4.2)-(4.5) *are satisfied for* $p = \infty$ *with replacing* $W_n, \boldsymbol{X}_n, \mu, \sigma$ *by*
$W_n(\nu), \boldsymbol{X}_n(\nu), \mu(\nu), \sigma(\nu)$ *respectively. Then we have* (4.6) *as* $n \to \infty$.

To make Theorems 4.1–4.2 statistically feasible, we need to estimate the
"asymptotic" covariance matrix $\mathfrak{C}_n$. We can construct a "consistent" estimator
for $\mathfrak{C}_n$ in the same way as in the low-dimensional setting of Barndorff-Nielsen
& Shephard [4]: Define the d^2-dimensional random vectors χ_h by

$$\chi_h := \mathrm{vec}\left[(Y_{t_h} - Y_{t_{h-1}})(Y_{t_h} - Y_{t_{h-1}})^\top\right], \qquad h = 1, \ldots, n.$$

Then we set

$$\widehat{\mathfrak{C}}_n := n \sum_{h=1}^{n} \chi_h \chi_h^\top - \frac{n}{2} \sum_{h=1}^{n-1} \left(\chi_h \chi_{h+1}^\top + \chi_{h+1} \chi_h^\top\right).$$

Proposition 4.1. *For all* $n \in \mathbb{N}$ *and* $\nu \in \mathbb{N}$, *let* $\Omega_n(\nu) \in \mathcal{F}$, $\mu(\nu) = (\mu(\nu)_t)_{t \in [0,1]}$
be a d-dimensional $(\mathcal{F}_t)$-*progressively measurable process and* $\sigma(\nu) = (\sigma(\nu)_t)_{t \in [0,1]}$
be an $\mathbb{R}^{d \times r}$-*valued* $(\mathcal{F}_t)$-*progressively measurable process, and suppose that the
following conditions are satisfied:*

(i) $\lim_{\nu \to \infty} \limsup_{n \to \infty} P(\Omega_n(\nu)^c) = 0$.

(ii) *For all* $\nu \in \mathbb{N}$ *and* $t \in [0,1]$, $\mu_t = \mu(\nu)_t$ *and* $\sigma_t = \sigma(\nu)_t$ *on* $\Omega_n(\nu)$ *as well
as* $\sigma(\nu)_t \in \mathbb{D}_{1,\infty}(\mathbb{R}^{d \times r})$.

(iii) *There is a constant* $\gamma \in (0, \frac{1}{2}]$ *such that*

$$\sup_{0 < t \le 1 - \frac{1}{n}} \left\|\max_{1 \le k,l \le d} \left|\Sigma(\nu)_{t+\frac{1}{n}}^{kl} - \Sigma(\nu)_t^{kl}\right|\right\|_2 = O(n^{-\gamma}) \qquad (4.7)$$

as $n \to \infty$, *where* $\Sigma(\nu)_t := \sigma(\nu)_t \sigma(\nu)_t^\top$.

Then the following statements hold true:

(a) *Suppose that*

$$\sup_{n \in \mathbb{N}} \max_{1 \le i \le d} \sup_{0 \le t \le 1} \left(\|\mu(\nu)_t^i\|_p + \|\Sigma(\nu)_t^{ii}\|_p + \sup_{0 \le s \le 1} \|D_s \sigma(\nu)_t^{i \cdot}\|_{p, \ell_2}\right) < \infty$$
$$(4.8)$$

for all $p \in [1, \infty)$ *and* $\nu \in \mathbb{N}$. *Suppose also that* $d = O(n^{\mathfrak{c}})$ *as* $n \to \infty$ *for
some* $\mathfrak{c} > 0$. *Then we have* $\|\widehat{\mathfrak{C}}_n - \mathfrak{C}_n\|_{\ell_\infty} = O_p(n^{-\varpi})$ *as* $n \to \infty$ *for any*
$\varpi \in (0, \gamma)$.

(b) *Suppose that* (4.8) *is satisfied for* $p = \infty$. *Then we have* $\|\widehat{\mathfrak{C}}_n - \mathfrak{C}_n\|_{\ell_\infty} =$
$O_p(n^{-1/2} \log^2 d + n^{-\gamma})$ *as* $n \to \infty$.

Remark 4.3. It is presumably possible to remove the (local) Malliavin differ-
entiability assumption on σ_t from Proposition 4.1 if we impose an additional
condition on d and $n^{-\gamma}$ (such an additional assumption will be even unneces-
sary to prove the part (a) only, but we keep that condition to prove two claims
in a unified way).

When the dimension d is very large, computation of $\widehat{\mathfrak{C}}_n^{1/2}$ is practically challenging, so it is better to employ a (wild) bootstrap to generate random vectors having the same distributions as that of $\widehat{\mathfrak{C}}_n^{1/2}\zeta_n$ as follows. Let $(e_h)_{h=1}^\infty$ be a centered Gaussian process independent of $\mathcal{F}$, which is defined on an extension of $(\Omega, \mathcal{F}, P)$ if necessary. Then we define

$$S_n^* := \sqrt{n} \sum_{h=1}^n e_h \chi_h.$$

The Gaussian process $(e_h)_{h=1}^\infty$ must have an appropriate covariance matrix so that the $\mathcal{F}$-conditional covariance matrix of S_n^* mimics $\widehat{\mathfrak{C}}_n$. As is well-known in the literature (see e.g. [34]), the standard i.i.d. wild bootstrap fails to approximate the joint distributions of statistics in the present context.[1] Alternatively, we assume that $(e_h)_{h=1}^\infty$ is stationary with auto-covariance function

$$E[e_h e_{h+\ell}] = \left\{ \begin{array}{cl} 1 & \text{if } \ell = 0, \\ -\frac{1}{2} & \text{if } \ell = 1, \\ 0 & \text{otherwise.} \end{array} \right.$$

Then we can easily check that the $\mathcal{F}$-conditional covariance matrix of S_n^* is equal to $\widehat{\mathfrak{C}}_n$, so S_n^* has the same distribution as that of $\widehat{\mathfrak{C}}_n^{1/2}\zeta_n$. We remark that such a sequence $(e_h)_{h=1}^\infty$ considered above can be generated by the following Gaussian MA(1) process:

$$e_h = \eta_h^* - \eta_{h-1}^*, \qquad h = 1, \ldots, n,$$

where $(\eta_h^*)_{h=0}^n$ is a sequence of i.i.d. centered Gaussian variables with variance $\frac{1}{2}$. Therefore, we can rewrite S_n^* as

$$S_n^* = \sqrt{n} \sum_{h=1}^{n-1} \eta_h^* (\chi_h - \chi_{h+1}) + \sqrt{n}(\eta_n^* \chi_n - \eta_0^* \chi_1).$$

The second term on the right side of the above equation is usually asymptotically negligible, so the bootstrap procedure considered here is essentially the same as the *wild blocks of blocks bootstrap* proposed in Hounyo [34].

4.1. Testing the residual sparsity of a continuous-time factor model

As an application of the theory developed above, we consider the problem of testing the correlation structure of the residual process of a continuous-time factor model. This problem was investigated in Section 4 of Bibinger & Mykland [5] for the case of two assets, and we are aim at extending their analysis to a multiple assets situation. Specifically, we suppose that the d-th asset Y^d is regarded as an observable factor and consider the following continuous-time

---

[1]It is also known that empirical bootstrap fails in the present context as well; see e.g. [25] for a discussion.

factor model:

$$Y^j = \beta^j Y^d + R^j, \qquad j = 1, \ldots, \underline{d} := d - 1. \tag{4.9}$$

Here, β^j is a constant and R^j is a semimartingale such that $[R^j, Y^d] \equiv 0$. Let us set $\Lambda_n := \{(i,j) : 1 \le i < j \le \underline{d}\}$. For each $(i,j) \in \Lambda_n$, we consider the following hypothesis testing problem:

$$H_0^{(i,j)} : [R^i, R^j]_1 = 0 \quad \text{a.s.} \qquad \text{vs} \qquad H_1^{(i,j)} : [R^i, R^j]_1 \ne 0 \quad \text{a.s.} \tag{4.10}$$

Our aim is to test the hypothesis (4.10) simultaneously for $(i,j) \in \Lambda_n$, but we start with constructing a test statistic for a fixed $(i,j) \in \Lambda_n$. For notational convenience, we construct the test statistic for every pair (i,j) in $\{1, \ldots, \underline{d}\}^2$.

Remark 4.4 (Sparsity test of the quadratic covariation matrix itself). Considering the case $Y^d \equiv 0$, we have $R^i = Y^i$ for all $i = 1, \ldots, \underline{d}$. Hence the problem turns to multiple testing for the hypotheses (1.1).

We follow [5] and consider the following statistic

$$\mathfrak{T}^{ij} := [Y^i, Y^d]_1 [Y^j, Y^d]_1 - [Y^i, Y^j]_1 [Y^d, Y^d]_1,$$

which is zero under $H_0^{(i,j)}$. Therefore, it is natural to consider the estimated version of $\mathfrak{T}^{ij}$ as follows:

$$\hat{\mathfrak{T}}_n^{ij} := [\widehat{Y^i, Y^d}]_1^n [\widehat{Y^j, Y^d}]_1^n - [\widehat{Y^i, Y^j}]_1^n [\widehat{Y^d, Y^d}]_1^n.$$

In order to make the test statistic scale invariant, we consider the Studentized version of $\hat{\mathfrak{T}}_n^{ij}$. According to [5], the "asymptotic variance" of $\hat{\mathfrak{T}}_n^{ij}$ is given by the following statistic:

$$\mathfrak{V}_n^{ij} :=$$

$$[Y^j, Y^d]_1^2 \mathfrak{C}_n^{id,id} + [Y^i, Y^d]_1^2 \mathfrak{C}_n^{jd,jd} + [Y^i, Y^j]_1^2 \mathfrak{C}_n^{d^2,d^2} + [Y^d, Y^d]_1^2 \mathfrak{C}_n^{(i-1)d+j,(i-1)d+j}$$

$$+ 2[Y^d, Y^d]_1 [Y^i, Y^j]_1 \mathfrak{C}_n^{(i-1)d+j,d^2} + 2[Y^i, Y^d]_1 [Y^j, Y^d]_1 \mathfrak{C}_n^{id,jd}$$

$$- 2[Y^i, Y^d]_1 [Y^d, Y^d]_1 \mathfrak{C}_n^{(i-1)d+j,jd} - 2[Y^j, Y^d]_1 [Y^d, Y^d]_1 \mathfrak{C}_n^{(i-1)d+j,id}$$

$$- 2[Y^i, Y^j]_1 [Y^i, Y^d]_1 \mathfrak{C}_n^{jd,d^2} - 2[Y^i, Y^j]_1 [Y^j, Y^d]_1 \mathfrak{C}_n^{id,d^2}.$$

Let us denote by $\hat{\mathfrak{V}}_n^{ij}$ the estimated version of $\mathfrak{V}_n^{ij}$, i.e. we define $\hat{\mathfrak{V}}_n^{ij}$ by the right side of the above equation with replacing $[Y, Y]_1$ and $\mathfrak{C}_n$ by $\widehat{[Y, Y]}_1$ and $\widehat{\mathfrak{C}}_n$, respectively. Then we define the test statistic by

$$T_n^{(i,j)} := \frac{\sqrt{n}\hat{\mathfrak{T}}_n^{ij}}{\sqrt{\hat{\mathfrak{V}}_n^{ij}}}.$$

The statistic $T_n^{(i,j)}$ is generally uncentered unless the null hypothesis $H_0^{(i,j)}$ is true, and it is convenient to consider the centered version of $T_n^{(i,j)}$ in the general situation as follows:

$$\tilde{T}_n^{(i,j)} := \frac{\sqrt{n}\left(\hat{\mathfrak{T}}_n^{ij} - \mathfrak{T}_n^{ij}\right)}{\sqrt{\hat{\mathfrak{V}}_n^{ij}}}.$$

Note that we can rewrite $\hat{\mathfrak{T}}_n^{ij} - \mathfrak{T}_n^{ij}$ as

$$\hat{\mathfrak{T}}_n^{ij} - \mathfrak{T}_n^{ij} = \left(\widehat{[Y^i, Y^d]}_1^n - [Y^i, Y^d]_1 \right) \widehat{[Y^j, Y^d]}_1^n + [Y^i, Y^d]_1 \left(\widehat{[Y^j, Y^d]}_1^n - [Y^j, Y^d]_1 \right)$$
$$- \left(\widehat{[Y^i, Y^j]}_1^n - [Y^i, Y^j]_1 \right) \widehat{[Y^d, Y^d]}_1^n - [Y^i, Y^j]_1 \left(\widehat{[Y^d, Y^d]}_1^n - [Y^d, Y^d]_1 \right).$$

Therefore, a bootstrapped version of $\tilde{T}_n^{(i,j)}$ is defined as

$$T_{n,*}^{(i,j)} := \frac{\sqrt{n}\hat{\mathfrak{T}}_{n,*}^{ij}}{\sqrt{\hat{\mathfrak{V}}_n^{ij}}},$$

where

$$\hat{\mathfrak{T}}_{n,*}^{ij} := \widehat{[Y^i, Y^d]}_1^{n,*} \widehat{[Y^j, Y^d]}_1^n + \widehat{[Y^i, Y^d]}_1^n \widehat{[Y^j, Y^d]}_1^{n,*}$$
$$- \widehat{[Y^i, Y^j]}_1^{n,*} \widehat{[Y^d, Y^d]}_1^n - \widehat{[Y^i, Y^j]}_1^n \widehat{[Y^d, Y^d]}_1^{n,*}$$

and

$$\widehat{[Y^i, Y^j]}_1^{n,*} := \sqrt{n} \sum_{h=1}^n e_h (Y_{t_h}^i - Y_{t_{h-1}}^i)(Y_{t_h}^j - Y_{t_{h-1}}^j), \qquad i, j = 1, \ldots, d.$$

Set $\tilde{T}_n = (\tilde{T}_n^{(i,j)})_{1 \le i,j \le \underline{d}}$ and $T_{n,*} = (T_{n,*}^{(i,j)})_{1 \le i,j \le \underline{d}}$. We derive mixed-normal approximations for $\text{vec}(\tilde{T}_n)$ and $\text{vec}(T_{n,*})$ by applying the theory developed above. For this purpose we define the $\underline{d}^2 \times d^2$ random matrix $\boldsymbol{X}_n$ by

$$\boldsymbol{X}_n^{(i-1)\underline{d}+j,(k-1)d+l} = \begin{cases} [Y^j, Y^d]_1 / \sqrt{\mathfrak{V}_n^{ij}} & \text{if } k = i, \ l = d, \\ [Y^i, Y^d]_1 / \sqrt{\mathfrak{V}_n^{ij}} & \text{if } k = j, \ l = d, \\ -[Y^d, Y^d]_1 / \sqrt{\mathfrak{V}_n^{ij}} & \text{if } k = l = d, \\ -[Y^i, Y^j]_1 / \sqrt{\mathfrak{V}_n^{ij}} & \text{if } k = i, l = j, \\ 0 & \text{otherwise} \end{cases}$$

for $i, j = 1, \ldots, \underline{d}$ and $k, l = 1, \ldots, d$. Note that the statistics $\text{vec}(\tilde{T}_n)$ and $\text{vec}(T_{n,*})$ can be approximated by $\boldsymbol{X}_n S_n$ and $\boldsymbol{X}_n S_n^*$, respectively. We then obtain the following result.

Proposition 4.2. *For all $n \in \mathbb{N}$ and $\nu \in \mathbb{N}$, let $\Omega_n(\nu) \in \mathcal{F}$, $\mu(\nu) = (\mu(\nu)_t)_{t \in [0,1]}$ be a d-dimensional $(\mathcal{F}_t)$-progressively measurable process and $\sigma(\nu) = (\sigma(\nu)_t)_{t \in [0,1]}$ be an $\mathbb{R}^{d \times r}$-valued $(\mathcal{F}_t)$-progressively measurable process, and suppose that the following conditions are satisfied:*

(i) *$\lim_{\nu \to \infty} \limsup_{n \to \infty} P(\Omega_n(\nu)^c) = 0$.*
(ii) *For all $\nu \in \mathbb{N}$ and $t \in [0, 1]$, $\mu_t = \mu(\nu)_t$ and $\sigma_t = \sigma(\nu)_t$ on $\Omega_n(\nu)$ as well as $\sigma(\nu)_t \in \mathbb{D}_{1,\infty}(\mathbb{R}^{d \times r})$.*
(iii) *For all $p \in [1, \infty)$, it holds that*

$$\sup_{n \in \mathbb{N}} \max_{1 \le i \le d} \left(\sup_{0 \le t \le 1} \|\mu_t^i\|_p + \sup_{0 \le s,t \le 1} \|D_s \mu(\nu)_t^i\|_{p,\ell_2} \right) < \infty,$$

$$\sup_{n \in \mathbb{N}} \max_{1 \le i \le d} \left(\sup_{0 \le t \le 1} \|\Sigma(\nu)_t^{ii}\|_p + \sup_{0 \le s,t \le 1} \|D_s \sigma(\nu)_t^{i\cdot}\|_{p,\ell_2} \right.$$

$$\left. + \sup_{0 \le s,t,u \le 1} \|D_{s,t} \sigma(\nu)_u^{i\cdot}\|_{p,\ell_2} \right) < \infty,$$

    where $\Sigma(\nu)_t := \sigma(\nu)_t \sigma(\nu)_t^\top$.

(iv) *There is a constant* $\gamma \in (0, \frac{1}{2}]$ *such that* (4.7) *holds true as* $n \to \infty$.

(v) *For all* $p \in [1, \infty)$, *it holds that*

$$\sup_{n \in \mathbb{N}} \max_{1 \le i,j \le \underline{d}} E\left[\left(\mathfrak{V}_n(\nu)^{ij} \right)^{-p} \right] < \infty, \tag{4.11}$$

    where $\mathfrak{V}_n(\nu)$ *is defined analogously to* $\mathfrak{V}_n$ *with replacing* Σ *by* $\Sigma(\nu)$.

Then we have

$$\sup_{A \in \mathcal{A}^{\mathrm{re}}(\underline{d}^2)} \left| P\left(\mathrm{vec}\left(\tilde{T}_n \right) \in A \right) - P\left(\boldsymbol{X}_n \mathfrak{C}_n^{1/2} \zeta_n \in A \right) \right| \to 0$$

and

$$\sup_{A \in \mathcal{A}^{\mathrm{re}}(\underline{d}^2)} \left| P\left(\mathrm{vec}\left(T_{n,*} \right) \in A | \mathcal{F} \right) - P\left(\boldsymbol{X}_n \mathfrak{C}_n^{1/2} \zeta_n \in A | \mathcal{F} \right) \right| \to^p 0$$

as $n \to \infty$, *provided that* $d = O(n^{\mathfrak{c}})$ *as* $n \to \infty$ *for some* $\mathfrak{c} > 0$.

Now we return to the problem of testing (4.10) simultaneously for $(i,j) \in \Lambda_n$. Here, we consider a more general setting described in the following for the purposes of application (cf. Section 5.2). We suppose that the set Λ_n is decomposed into non-empty disjoint sets $\Lambda_n^1, \ldots, \Lambda_n^{\mathsf{L}}$ as $\Lambda_n = \bigcup_{\ell=1}^{\mathsf{L}} \Lambda_n^\ell$. We consider the problem of testing

$$\bigwedge_{\lambda \in \Lambda_n^\ell} H_0^\lambda \qquad \mathrm{vs} \qquad \bigvee_{\lambda \in \Lambda_n^\ell} H_1^\lambda \tag{4.12}$$

simultaneously for $\ell = 1, \ldots, \mathsf{L}$. Here, for a subset $\mathcal{L}$ of Λ_n, $\bigwedge_{\lambda \in \mathcal{L}} H_0^\lambda$ (resp. $\bigvee_{\lambda \in \mathcal{L}} H_1^\lambda$) denotes the hypothesis that H_0^λ is true for all $\lambda \in \mathcal{L}$ (resp. H_1^λ is true for some $\lambda \in \mathcal{L}$). For simplicity of notation, we set $\mathsf{H}_0^\ell := \bigwedge_{\lambda \in \Lambda_n^\ell} H_0^\lambda$ and $\mathsf{H}_1^\ell := \bigvee_{\lambda \in \Lambda_n^\ell} H_1^\lambda$. If we let L be the number of elements in Λ_n and write $\Lambda_n = \{\lambda_1, \ldots, \lambda_{\mathsf{L}}\}$ and set $\Lambda_n^\ell = \{\lambda_\ell\}$ for $\ell = 1, \ldots, \mathsf{L}$, we recover the original problem of testing (4.10) simultaneously for $(i,j) \in \Lambda_n$.

Our aim is the strong control of the family-wise error rate (FWER) in this problem. More formally, let Θ_n be a set of pairs (μ, σ) of coefficient processes, which is considered as the set of all data generating processes we are interested in (note that the data generating process may vary with n mainly because the dimensions d and r may depend on n). For each $\theta \in \Theta_n$, we denote by $\mathcal{L}_n(\theta)$ the set of all indices $\ell \in \{1, \ldots, \mathsf{L}\}$ for which the hypothesis H_0^ℓ holds true when θ is the true data generating process. Then, the FWER for $\theta \in \Theta_n$, which is denoted by $\mathbf{FWER}(\theta)$, is defined as the probability that H_0^ℓ for some $\ell \in \mathcal{L}_n(\theta)$

is rejected when θ is the true data generating process. Given the significance level $\alpha \in (0, 1)$, we aim at constructing multiple testing procedures such that

$$\limsup_{n \to \infty} \mathbf{FWER}(\theta_n) \leq \alpha \tag{4.13}$$

for any sequence $\theta_n \in \Theta_n$ $(n = 1, 2, \dots)$ of data generating processes. To accomplish this, we employ the stepdown procedure of Romano & Wolf [61] which we describe in the following. First, given a fixed index ℓ, we shall use the test statistic $\mathsf{T}_n^\ell := \max_{\lambda \in \Lambda_n^\ell} |T_n^\lambda|$ for the problem (4.12). Next, we sort the observed test statistics in descending order and denote them by

$$\mathsf{T}_n^{\ell_1} \geq \cdots \geq \mathsf{T}_n^{\ell_\mathsf{L}}.$$

Also, for every subset $\mathcal{L} \subset \{1, \dots, \mathsf{L}\}$, suppose that we have a critical value $c_n^{\mathcal{L}}(1 - \alpha)$ to test the null $\bigwedge_{\lambda \in \mathcal{L}} H_0^\lambda$ against the alternative $\bigvee_{\lambda \in \mathcal{L}} H_1^\lambda$. Those critical values can be random variables and will be specified later. Then the stepdown procedure reads as follows:

1. Let $\mathcal{L}_1 := \{1, \dots, \mathsf{L}\}$. If $\mathsf{T}_n^{\ell_1} \leq c_n^{\mathcal{L}_1}(1 - \alpha)$, then accept all the hypotheses and stop; otherwise, reject $\mathsf{H}_0^{\ell_1}$ and continue.
2. Let $\mathcal{L}_2 := \mathcal{L}_1 \setminus \{\ell_1\}$. If $\mathsf{T}_n^{\ell_2} \leq c_n^{\mathcal{L}_2}(1 - \alpha)$, then accept all the hypotheses H_0^ℓ for $\ell \in \mathcal{L}_2$ and stop; otherwise, reject $\mathsf{H}_0^{\ell_2}$ and continue.

   $\vdots$

k. Let $\mathcal{L}_k := \mathcal{L}_{k-1} \setminus \{\ell_{k-1}\}$. If $\mathsf{T}_n^{\ell_k} \leq c_n^{\mathcal{L}_k}(1 - \alpha)$, then accept all the hypotheses H_0^ℓ for $\ell \in \mathcal{L}_k$ and stop; otherwise, reject $\mathsf{H}_0^{\ell_k}$ and continue.

   $\vdots$

L. If $\mathsf{T}_n^{\lambda_\mathsf{L}} \leq c_n^{\{\ell_\mathsf{L}\}}(1 - \alpha)$, then accept $\mathsf{H}_0^{\ell_\mathsf{L}}$; otherwise, reject $\mathsf{H}_0^{\ell_\mathsf{L}}$.

According to Theorem 3 of [61], the above stepdown procedure satisfies (4.13) if the critical values $c_n^{\mathcal{L}}(1 - \alpha)$, $\mathcal{L} \subset \{1, \dots, \mathsf{L}\}$, satisfy the following conditions:

(i) $c_n^{\mathcal{L}}(1 - \alpha) \leq c_n^{\mathcal{L}'}(1 - \alpha)$ whenever $\mathcal{L} \subset \mathcal{L}' \subset \{1, \dots, \mathsf{L}\}$.

(ii) For any sequence $\theta_n \in \Theta_n$ $(n = 1, 2, \dots)$, it holds that

$$\limsup_{n \to \infty} P\left(\max_{\ell \in \mathcal{L}_n(\theta_n)} \mathsf{T}_n^\ell > c_n^{\mathcal{L}_n(\theta_n)}(1 - \alpha) \right) \leq \alpha$$

whenever θ_n is the true data generating process for every n.

The first method to construct the desired critical values is the well-known *Bonferroni-Holm method*. Namely, we set $c_n^{\mathcal{L}}(1 - \alpha) := q_{N(0,1)}(1 - \alpha / (2\#[\bigcup_{\ell \in \mathcal{L}} \Lambda_n^\ell]))$ for every $\mathcal{L} \subset \Lambda$, where $q_{N(0,1)}$ denotes the quantile function of the standard normal distribution and $\#[\bigcup_{\ell \in \mathcal{L}} \Lambda_n^\ell]$ is the number of elements in $\bigcup_{\ell \in \mathcal{L}} \Lambda_n^\ell$. The second method is to use the $(1 - \alpha)$-quantile of $\max_{\ell \in \mathcal{L}} \mathsf{T}_n^\ell$. Of course, we cannot analytically compute the quantiles of $\max_{\ell \in \mathcal{L}} \mathsf{T}_n^\ell$ in general, so we approximate them by resampling as in [61, 14]. Formally, setting $\mathsf{T}_{n,*}^\ell := \max_{\lambda \in \Lambda_n^\ell} |T_{n,*}^\lambda|$, we use the $\mathcal{F}$-conditional $(1 - \alpha)$-quantile of $\max_{\ell \in \mathcal{L}} \mathsf{T}_{n,*}^\ell$ as $c_n^{\mathcal{L}}(1 - \alpha)$, which can be evaluated by simulation. We refer to this method as the *Romano-Wolf method* in the following.

Corollary 4.2. *Suppose that the assumptions of Proposition 4.2 are satisfied for any sequence $(\mu, \sigma) = (\mu^{(n)}, \sigma^{(n)}) \in \Theta_n$ $(n = 1, 2, \dots)$ of data generating processes whenever $(\mu^{(n)}, \sigma^{(n)})$ is the true data generating process for every n. Then, both the Bonferroni-Holm and Romano-Wolf methods satisfy conditions (i)–(ii), so (4.13) holds true.*

Remark 4.5. The Romano-Wolf method takes account of the dependence structure of the test statistics while the Bonferroni-Holm method ignores it, so the former is generally more powerful than the latter, especially when the test statistics are strongly dependent on each other. Meanwhile, we need no resampling to implement the Bonferroni-Holm method, so it is computationally more attractive than the Romano-Wolf method.

Remark 4.6 (Application to threshold selection in covariance estimation)**.** Another possible application of Theorem 4.1 would be selection of the thresholds in high-dimensional quadratic covariation estimation from high-frequency data (see e.g. Wang & Zou [68] for such an estimation method): We refer to Section 4.1 of Chen [11] for details on such an application in the case of i.i.d. observations.

5. Simulation study and an empirical illustration

In this section we present a small Monte Carlo study to assess the finite sample performance of the multiple testing procedures proposed in Section 4.1. We also demonstrate how the proposed methodology works in a real world using high-frequency data from the components of the S&P 100 index.

5.1. Simulations

We focus on the problem of testing the hypotheses (4.10) simultaneously for $(i, j) \in \Lambda_n$. The simulation design is basically adopted from [28], but we include only the first factor representing the market factor in our model. Specifically, we simulate model (4.9) with the following specification[2]:

$$dY_t^d = \mu dt + \sqrt{v_t} dB_t^d, \qquad dR_t^j = \gamma_j^\top d\underline{B}_t \quad (j = 1, \dots, \underline{d})$$

and

$$dv_t = \kappa(\theta - v_t)dt + \eta\sqrt{v_t}\left(\rho dB_t^d + \sqrt{1 - \rho^2} dB_t^{d+1}\right). \tag{5.1}$$

Here, $\mu, \kappa, \theta, \eta$ and ρ are constants, $\underline{B}_t = (B_t^1, \dots, B_t^{\underline{d}})$, and $\gamma_1, \dots, \gamma_{\underline{d}}$ are $\underline{d}$-dimensional random vectors independent of B. The values of $\beta^1, \dots, \beta^{\underline{d}}$ are independently drawn from the uniform distribution on $[0.25, 2.25]$. We set $\mu =$

---

[2] One can show that the volatility process σ generated by (5.1) *locally* satisfy the condition (4.5) for any $p \in [1, \infty)$ as long as the Feller condition $2\kappa\theta > \eta^2$ is satisfied. In fact, one can show this by setting $\Omega_n(\nu) := \{\inf_{t \in [0,1]} \sigma_t \geq \nu^{-1}\}$ and taking smoothed versions of σ_t analogous to the one considered in [2] as $\sigma(\nu)$'s for $\nu = 1, 2, \dots$.

0.05, $\kappa = 3$, $\theta = 0.09$, $\eta = 0.3$ and $\rho = -0.6$. The initial value v_0 is drawn from the stationary distribution of the process $(v_t)_{t \in [0,1]}$, i.e. the gamma distribution with shape $2\kappa\theta/\eta^2$ and rate $2\kappa/\eta^2$. We assume that $\Gamma := (\gamma_i^\top \gamma_j)_{1 \le i,j \le \underline{d}}$ is a block diagonal matrix with 10 blocks of size $(\underline{d}/10) \times (\underline{d}/10)$ whose diagonals are uniformly generated from $[0.2, 0.5]$ and the corresponding correlation matrices have the constant correlation of ρ_γ. We set $\underline{d} = 100$ and vary ρ_γ as $\rho_\gamma \in \{0.25, 0.5, 0.75\}$.

For each scenario, we compute the FWERs and the average powers (i.e. the average probabilities of rejecting the false null hypotheses) of the Bonferroni-Holm and Romano-Wolf methods at the 5% level based on 10,000 Monte Carlo iterations respectively. Here, we generate 999 bootstrap resamples for the Romano-Wolf method.

Tables 1 and 2 report the results. We see from Table 1 that both the methods succeed in controlling the FWERs under the nominal level 5%, although both are

TABLE 1
Family-wise error rates at the 5% level

rn	$n = 26$	$n = 39$	$n = 78$	$n = 130$	$n = 195$	$n = 390$
$\rho_\gamma = 0.25$						
Holm	0.010	0.004	0.002	0.003	0.008	0.018
RW	0.022	0.007	0.003	0.004	0.009	0.018
$\rho_\gamma = 0.50$						
Holm	0.009	0.004	0.002	0.003	0.007	0.017
RW	0.023	0.008	0.004	0.005	0.009	0.019
$\rho_\gamma = 0.75$						
Holm	0.008	0.003	0.002	0.003	0.006	0.010
RW	0.026	0.011	0.006	0.008	0.014	0.023

Note. This table reports the family-wise error rates at the 5% level of multiple testing for the hypotheses (4.10) by the Bonferroni-Holm (BH) and Romano-Wolf (RW) methods, respectively. The reported values are based on 10,000 Monte Carlo iterations. 999 bootstrap resamples are generated to implement the RW method.

TABLE 2
Average powers at the 5% level

rn	$n = 26$	$n = 39$	$n = 78$	$n = 130$	$n = 195$	$n = 390$
$\rho_\gamma = 0.25$						
Holm	0.000	0.000	0.000	0.005	0.046	0.563
RW	0.000	0.000	0.000	0.006	0.048	0.567
$\rho_\gamma = 0.50$						
Holm	0.000	0.001	0.028	0.421	0.950	1.000
RW	0.000	0.001	0.037	0.458	0.956	1.000
$\rho_\gamma = 0.75$						
Holm	0.001	0.007	0.262	0.953	1.000	1.000
RW	0.004	0.017	0.393	0.977	1.000	1.000

Note. This table reports the average powers at the 5% level of multiple testing for the hypotheses (4.10) by the Bonferroni-Holm (BH) and Romano-Wolf (RW) methods, respectively. The reported values are based on 10,000 Monte Carlo iterations. 999 bootstrap resamples are generated to implement the RW method.

rather conservative. Table 2 shows that the average powers in both the methods tend to 1 as n and ρ_γ increase. The table also reveals that the Romano-Wolf method is more powerful than the Bonferroni-Holm method. As expected, the difference of the average powers between two methods becomes pronounced as the correlation ρ_γ of the residual processes increases.

5.2. Empirical illustration

We apply our methodology to high-frequency returns of the components of the S&P 100 index while taking the SPDR S&P 500 ETF (SPY) as the observable factor process. The sample period is the one month, March 2018, and we regard this period as the interval $[0, 1]$ (over-night returns are ignored). The data are provided by Bloomberg. Following Fan *et al.* [28], we use 15 minute returns to avoid notable market microstructure effects. To illuminate the block diagonal structure reported in [28], we sort the assets by their Global Industry Classification Standard (GICS) sectors while we construct the log-price processes Y^j, $j = 1, \ldots, \underline{d}$.

We begin by examining the sparsity of the quadratic covariation matrix of the assets without taking account of the factor process. The top panel of Figure 1 shows the corresponding realized correlation matrix. Here, we perform multiple testing for the hypotheses (1.1) using the Romano-Wolf method with 999 bootstrap resamples and change the entries for which the null hypotheses are not rejected at the 5% level to blanks. The violet squares indicate GICS sector classifications. Namely, all assets in the same square belong to the same sector. We clearly find that the raw realized correlation matrix is far from sparse, i.e. most the entries are not blank. In fact, our test suggests that about 90.9% pairs would have significant correlations at the 5% level. Meanwhile, the bottom panel of Figure 1 shows the realized correlation matrix of the residual processes of the assets regressed on SPY. Again, we perform multiple testing for the hypotheses (4.10) as above to change the entries with insignificant correlations to blanks. The violet squares have the same meaning as above. In contrast to the first case, the realized correlation matrix exhibits the remarkable diagonal structure inherited from the assets' sectors. In this case only about 4.3% pairs are significantly correlated at the 5% level.

To investigate this block diagonal structure more deeply, we conduct another multiple testing for the absence of covariations within and between sectors after regressing assets on SPY. Formally, let $G_1, \ldots, G_N$ be all the sectors, then we set $I_k := \{i \in \{1, \ldots, \underline{d}\} : \text{the } i\text{-th asset } Y^i \text{ belongs to the sector } G_k\}$ for every $k = 1, \ldots, N$ and $\Lambda_n^{(k,l)} := \Lambda_n \cap (I_k \times I_l)$ for all $k, l = 1, \ldots, N$. We test the null hypothesis $\bigwedge_{\lambda \in \Lambda_n^{(k,l)}} H_0^\lambda$ against the alternative $\bigvee_{\lambda \in \Lambda_n^{(k,l)}} H_1^\lambda$ simultaneously for all $1 \le k \le l \le N$ using the Romano-Wolf method with 999 bootstrap resamples. In our analysis there are totally $N = 11$ sectors: Consumer Discretionary, Consumer Staples, Financials, Health Care, Industrials, Information Technology, Materials, Real Estate, Telecommunication Services, and Utilities. Since Materials and Real Estate contain only one asset respectively, we exclude the
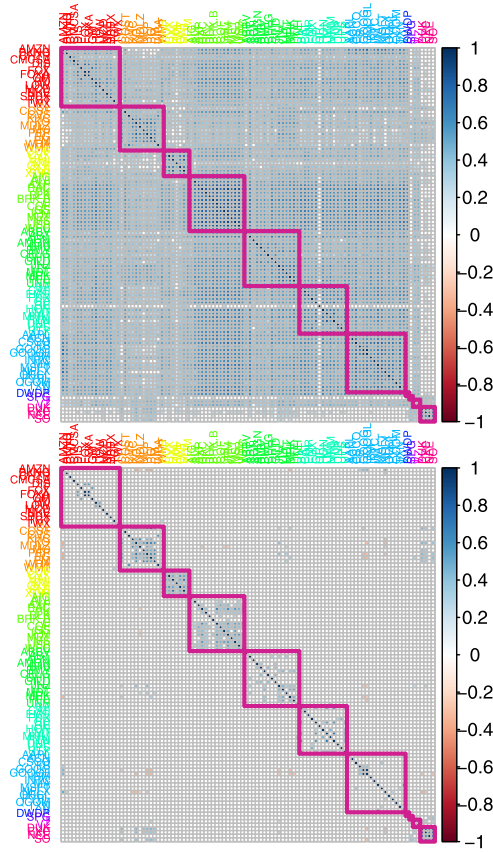
FIG 1. *Realized correlation matrices of the S&P 100 assets (top) and their residual processes regressed on SPY (bottom). They are computed from 15 minute returns in March 2018, where we ignore over-night returns. We perform multiple testing for whether each the entry is zero or not using the Romano-Wolf method with 999 bootstrap resamples, then the entries which are not significantly away from zero at the 5% level are made blank. The violet squares indicate sector blocks. The figure was depicted using the R function* `corrplot` *from the* `corrplot` *package.*

case $k = l$ from the above hypotheses when G_k is Materials or Real Estate. The results are reported in Table 3. As expected, the p-values for the absence of within-sector covariations are very small across all the sectors, which suggests within-sector covariations should exist for all the sectors. In contrast, we find that between-sector covariations can be insignificant for several pairs. For example, assets belonging to Materials (M) are not significantly correlated with assets belonging to the other sectors at the 5% level. The table also reveals a similar between-sector covariation pattern to the one observed in [28]. Namely, they report that the correlation between Energy (E) and Financials (F) disappears but Consumer Staples (CS) and Utilities (U) remain strongly correlated after 2010, which is consistent with the p-values reported in Table 3.

TABLE 3
p-values of multiple testing for the absence of within- and between-secor covariations (the null hypotheses are the absence of covariations).

	CD	CS	E	F	HC	I	IT	M	RE	TS	U
CD	0.001	0.002	0.119	0.246	0.003	0.246	0.001	0.076	0.246	0.226	0.045
CS		0.001	0.044	0.007	0.001	0.413	0.001	0.891	0.025	0.017	0.001
E			0.001	0.502	0.446	0.211	0.076	0.932	0.098	0.662	0.076
F				0.001	0.246	0.246	0.076	0.846	0.246	0.662	0.246
HC					0.001	0.001	0.003	0.932	0.308	0.008	0.024
I						0.001	0.246	0.502	0.846	0.662	0.224
IT							0.001	0.446	0.246	0.072	0.004
M								–	0.932	0.909	0.932
RE									–	0.256	0.004
TS										0.001	0.001
U											0.001

Note. The p-values are computed using the Romano-Wolf method with 999 bootstrap resamples. The sector names are abbreviated as follows: CD: Consumer Discretionary; CS: Consumer Staples; F: Financials; HC: Health Care; I: Industrials; IT: Information Technology; M: Materials; RE: Real Estate; TS: Telecommunication Services; U: Utilities.

Overall, our methodology partially provides a statistically formal support of the findings by [28], although the scope of our analysis is quite limited and thus more comprehensive empirical studies will be necessary.

Appendix A: Proofs for Section 3

A.1. Additional notation

This subsection introduces some additional notation related to multi-way arrays and derivatives, which are necessary for the subsequent proofs.

As in Section 2.2, $\mathbb{K}$ denotes the real field $\mathbb{R}$ or the complex field $\mathbb{C}$. We consider a vector space V over $\mathbb{K}$. Let $N_1, \ldots, N_q$ be positive integers. For $T \in V^{N_1 \times \cdots \times N_q}$ and $x \in \mathbb{K}^{N_1 \times \cdots \times N_q}$, we set

$$T[x] := \sum_{(i_1,\ldots,i_q) \in \prod_{k=1}^q [N_k]} T^{i_1,\ldots,i_q} x^{i_1,\ldots,i_q} \in V.$$

In particular, for $x_j \in \mathbb{K}^{N_j}$ $(j = 1, \ldots, q)$ we have

$$T[x_1 \otimes \cdots \otimes x_q] = \sum_{(i_1,\ldots,i_q) \in \prod_{k=1}^q [N_k]} T^{i_1,\ldots,i_q} x_1^{i_1} \cdots x_q^{i_q}.$$

Here, note that we identify $\mathbb{K}^{N_1} \otimes \cdots \otimes \mathbb{K}^{N_q}$ with $\mathbb{K}^{N_1 \times \cdots \times N_q}$ in the canonical way (see Section 2.2). Moreover, we evidently have

$$|T[x]| \leq \|T\|_{\ell_\infty} \|x\|_{\ell_1}. \tag{A.1}$$

Now suppose that $\mathbb{K} = \mathbb{R}$ and V is a real Hilbert space. Then we have

$$\langle T[x], v \rangle_V = \sum_{(i_1,\ldots,i_q) \in \prod_{k=1}^q [N_k]} \langle T^{i_1,\ldots,i_q}, v \rangle_V x^{i_1,\ldots,i_q} = \langle T, v \rangle_V[x] \qquad \text{(A.2)}$$

for any $v \in V$ (recall (2.1)). Let V_0 be another real Hilbert space and $N_1', \ldots, N_p' \in \mathbb{N}$. Then, for any $S \in V_0^{N_1' \times \cdots \times N_p'}$ and $y \in \mathbb{R}^{N_1' \times \cdots \times N_p'}$, it holds that

$$T[x] \otimes S[y] = (T \otimes S)[x \otimes y]. \qquad \text{(A.3)}$$

In fact, we have

$$T[x] \otimes S[y]$$
$$= \sum_{(i_1,\ldots,i_q) \in \prod_{k=1}^q [N_k]} \sum_{(j_1,\ldots,j_p) \in \prod_{k=1}^p [N_k']} (T^{i_1,\ldots,i_q} x^{i_1,\ldots,i_q}) \otimes (S^{j_1,\ldots,j_p} y^{j_1,\ldots,j_p})$$
$$= \sum_{(i_1,\ldots,i_q) \in \prod_{k=1}^q [N_k]} \sum_{(j_1,\ldots,j_p) \in \prod_{k=1}^p [N_k']} (T^{i_1,\ldots,i_q} \otimes S^{j_1,\ldots,j_p}) x^{i_1,\ldots,i_q} y^{j_1,\ldots,j_p}$$
$$= (T \otimes S)[x \otimes y].$$

Let $\phi = (\phi(y))_{y \in \mathbb{R}^N}$ be a real-valued function. If ϕ is a C^∞ function, we define the $\mathbb{R}$-valued N-dimensional q-way array $\partial_y^{\otimes q} \phi(y)$ by

$$\partial_y^{\otimes q} \phi(y) = (\partial_{y^{i_1} \ldots y^{i_q}} \phi(y))_{1 \leq i_1,\ldots,i_q \leq N} \in \mathbb{R}^{N \times \cdots \times N}$$

for any $y \in \mathbb{R}^N$ and $q \in \mathbb{N}$, where $\partial_{y^{i_1} \ldots y^{i_q}} := \partial^q / \partial y^{i_1} \cdots \partial y^{i_q}$. We set $\partial_y^{\otimes 0} \phi(y) := \phi(y)$ by convention. In general, we say that ϕ is *rapidly decreasing* if ϕ is a C^∞ function and

$$\sup_{y \in \mathbb{R}^N} (1 + \|y\|_{\ell_2})^A \|\partial_y^{\otimes q} \phi(y)\|_{\ell_\infty} < \infty$$

for any $A > 0$ and $q \in \mathbb{Z}_+$. When ϕ is rapidly decreasing, we define its Fourier transform $\hat{\phi} : \mathbb{R}^N \to \mathbb{C}$ by

$$\hat{\phi}(\mathsf{y}) = \int_{\mathbb{R}^N} \phi(y) e^{-y[\mathsf{i}\mathsf{y}]} dy, \qquad \mathsf{y} \in \mathbb{R}^N.$$

Here, i denotes the imaginary unit. By Theorem 7.4(c) from [62], one has

$$T[(\mathsf{i}\mathsf{y})^{\otimes q}] \hat{\phi}(\mathsf{y}) = T[\widehat{\partial_y^{\otimes q} \phi}(\mathsf{y})] \qquad \text{(A.4)}$$

for any $\mathsf{y} \in \mathbb{R}^N$, $q \in \mathbb{N}$ and $\mathbb{C}$-valued N-dimensional q-way array $T \in \mathbb{C}^{N \times \cdots \times N}$.

If $i_1 = \cdots = i_q = i$, we will write $\partial_{y^{i_1} \ldots y^{i_q}}$ as $\partial_{y^i}^q$. We set $\partial_{y^i}^0 \varphi(y) = \varphi(y)$ by convention. For a multi-index $\alpha = (\alpha_1, \ldots, \alpha_N) \in \mathbb{Z}_+^N$, we write $\partial_y^\alpha := \partial_{y^1}^{\alpha_1} \cdots \partial_{y^N}^{\alpha_N}$ as usual. Given a subset $\mathcal{A} = \{a_1, \ldots, a_k\}$ of $\{1, \ldots, s\}$, we will write $\prod_{a \in \mathcal{A}} \partial_{y^{i_a}} := \partial_{y^{i_{a_1}} \ldots y^{i_{a_k}}}$. We set $(\prod_{a \in \emptyset} \partial_{y^{i_a}}) \phi(y) := \phi(y)$ by convention.

A.2. Proof of Theorem 3.1

We begin by noting that it is enough to prove the theorem for the special case that all the rows of the matrix $\boldsymbol{X}_n$ are identical:

Lemma A.1. *Suppose that the claim of Theorem 3.1 holds true if $\boldsymbol{X}_n^{1\cdot} = \cdots = \boldsymbol{X}_n^{m\cdot}$ for every $n \in \mathbb{N}$. Then the claim of Theorem 3.1 holds true for the general case as well.*

Proof. Define the $m \times md$ matrix $\overline{\Upsilon}_n$ by

$$
\overline{\Upsilon}_n = \begin{pmatrix}
(\Upsilon_n^{1\cdot})^\top & 0 & \cdots & \cdots & 0 \\
0 & (\Upsilon_n^{2\cdot})^\top & 0 & \cdots & 0 \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
0 & \ddots & \ddots & \ddots & 0 \\
0 & \cdots & \cdots & 0 & (\Upsilon_n^{m\cdot})^\top
\end{pmatrix}.
$$

We also define the $m \times md$ random matrix $\overline{\boldsymbol{X}}_n$ so that all the rows are identical to the md-dimensional random vector given by

$$
((\boldsymbol{X}_n^{1\cdot})^\top, \ldots, (\boldsymbol{X}_n^{m\cdot})^\top).
$$

In addition, we define the md-dimensional random vector $\overline{Z}_n$ so that

$$
\overline{Z}_n^\top = (\underbrace{Z_n^\top, \ldots, Z_n^\top}_{m})^\top.
$$

By assumption we can apply Theorem 3.1 with taking $\overline{\Upsilon}_n$, $\overline{\boldsymbol{X}}_n$ and $\overline{Z}_n$ as Υ_n, $\boldsymbol{X}_n$ and Z_n respectively, which yields the desired result. $\square$

Taking account of Lemma A.1, we focus only on the case that $\boldsymbol{X}_n^{1\cdot} = \cdots = \boldsymbol{X}_n^{m\cdot} =: X_n$ for every $n \in \mathbb{N}$.

Next we recall the following anti-concentration inequality called Nazarov's inequality in [19]:

Proposition A.1 (Nazarov's inequality). *Let ξ be an m-dimensional centered Gaussian vector such that $\|\xi^j\|_2 \geq a$ for all $j = 1, \ldots, m$ and some constant $a > 0$. Then for any $y \in \mathbb{R}^m$ and $\varepsilon > 0$,*

$$
P(\xi \leq y + \varepsilon) - P(\xi \leq y) \leq \frac{\varepsilon}{a} \left(\sqrt{2 \log m} + 2 \right).
$$

The above form of Nazarov's inequality is found in [20]. An application of the above result immediately yields the following anti-concentration inequality for a mixed-normal random vector:

Lemma A.2. *Let ξ be an m-dimensional standard Gaussian vector. Also, let Γ be an $m \times m$ symmetric positive-semidefinite random matrix independent of ξ. Then for any $y \in \mathbb{R}^m$ and $b, \varepsilon > 0$,*

$$
P(\Gamma^{1/2}\xi \leq y + \varepsilon) - P(\Gamma^{1/2}\xi \leq y) \leq \frac{\varepsilon}{\sqrt{b}} \left(\sqrt{2 \log m} + 2 \right) + P \left(\min \operatorname{diag}(\Gamma) < b \right).
$$

Now we turn to the main body of the proof. As is pointed out in the Introduction, the key part of the proof is to derive reasonable estimates for the quantities

$$E[f(Z_n, X_n)] - E[f(\mathfrak{Z}_n, X_n)] \tag{A.5}$$

for smooth functions $f : \mathbb{R}^{2d} \to \mathbb{R}$. In fact, the remaining part of the proof is essentially the same as the one for the high-dimensional central limit theorem of [19]. To get a reasonable estimate for (A.5), we derive an interpolation formula for it, borrowing an idea from [57]. Namely, we use the duality between iterated Malliavin derivatives and multiple Skorohod integrals combined with the interpolation method in the frequency domain introduced in [57] (see also [66]).

Following [57], we set

$$\lambda_n(\theta; \mathsf{z}, \mathsf{x}) = \theta M_n[\mathsf{iz}] + 2^{-1}(1 - \theta^2)\mathfrak{C}_n[(\mathsf{iz})^{\otimes 2}] + W_n[\mathsf{iz}] + X_n[\mathsf{ix}]$$

and

$$\varphi_n(\theta; \mathsf{z}, \mathsf{x}) = E[e^{\lambda_n(\theta; \mathsf{z}, \mathsf{x})}]$$

for $\theta \in [0, 1]$ and $\mathsf{z}, \mathsf{x} \in \mathbb{R}^d$. We first derive a representation formula for the derivative of $\varphi_n(\theta; \mathsf{z}, \mathsf{x})$ with respect to θ. For this purpose we need the following Malliavin derivative version of the (generalized) Faà di Bruno formula for the iterated derivative of a composition of functions:

Lemma A.3. *Let q, r be positive integers and $g = (g(x))_{x \in \mathbb{R}^r}$ be a real-valued C^q function all of whose partial derivatives up to order q are of polynomial growth. Then, for any $F \in \mathbb{D}_{q,\infty}(\mathbb{R}^r)$, we have $g(F) \in \mathbb{D}_{q,\infty}$ and*

$$D^q g(F) = \sum_{\alpha \in \mathcal{A}(q)} \sum_{\nu \in \mathcal{N}_r(\alpha)} \mathsf{C}(\alpha, \nu) \partial_{x_1}^{|\nu_{\cdot 1}|} \cdots \partial_{x_r}^{|\nu_{\cdot r}|} g(F) \, \mathrm{Sym} \left(\bigotimes_{i=1}^{q} \bigotimes_{j=1}^{r} (D^i F^j)^{\otimes \nu_{ij}} \right),$$

where

$$\mathsf{C}(\alpha, \nu) = \frac{q!}{\prod_{i=1}^{q} (i!)^{\alpha_i} \prod_{j=1}^{r} \nu_{ij}!}.$$

Noting that Malliavin derivatives can be characterized by directional derivatives along Cameron-Martin shifts (cf. Chapter 15 of [37]), we can derive Lemma A.3 from the usual Faà di Bruno formula (found in e.g. [51]). Alternatively, we can prove Lemma A.3 in a parallel way to the usual Faà di Bruno formula using the chain rule for Malliavin derivatives (see e.g. Theorem 15.78 of [37]) instead of that for standard ones.

Lemma A.4. *Under the assumptions of Theorem 3.1, the partial derivative $\partial_\theta \varphi_n(\theta; \mathsf{z}, \mathsf{x})$ exists and it is given by*

$$\partial_\theta \varphi_n(\theta; \mathsf{z}, \mathsf{x}) = \theta \sum_{i,j=1}^{d} E \left[e^{\lambda(\theta; \mathsf{z}, \mathsf{x})} \left(\left\langle D^{q_j} M_n^i, u_n^j \right\rangle_{H^{\otimes q_j}} - \mathfrak{C}_n^{ij} \right) (\mathsf{iz}^i)(\mathsf{iz}^j) \right]$$

$$+ \sum_{j=1}^{d} \sum_{\alpha \in \mathcal{A}(q_j)} \sum_{\nu \in \mathcal{N}_4^*(\alpha)} \mathsf{C}(\alpha, \nu) \theta^{|\nu \cdot 1|} \{2^{-1}(1 - \theta^2)\}^{|\nu \cdot 2|} (\mathsf{i}\mathsf{z}^j)$$

$$\times E\left[e^{\lambda(\theta;\mathsf{z},\mathsf{x})} \Delta_{n,j}(\nu)[(\mathsf{i}\mathsf{z})^{\otimes(\nu_{i1} + 2\nu_{i2} + \nu_{i3})} \otimes (\mathsf{i}\mathsf{x})^{\otimes \nu_{i4}}] \right].$$

Proof. By assumption the function $\theta \mapsto \varphi_n(\theta; \mathsf{z}, \mathsf{x})$ is evidently differentiable and we have

$$\partial_\theta \varphi_n(\theta; \mathsf{z}, \mathsf{x}) = E[e^{\lambda_n(\theta;\mathsf{z},\mathsf{x})}(M_n[\mathsf{i}\mathsf{z}] - \theta \mathfrak{C}_n[(\mathsf{i}\mathsf{z})^{\otimes 2}])]. \qquad (A.6)$$

By duality we obtain

$$E[e^{\lambda_n(\theta;\mathsf{z},\mathsf{x})} M_n^j] = E[\langle D^{q_j} \Re[e^{\lambda_n(\theta;\mathsf{z},\mathsf{x})}], u_n^j \rangle_{H^{\otimes q_j}}] + \mathsf{i} E[\langle D^{q_j} \Im[e^{\lambda_n(\theta;\mathsf{z},\mathsf{x})}], u_n^j \rangle_{H^{\otimes q_j}}]$$

for every j. Therefore, Lemma A.3 yields

$$E[e^{\lambda_n(\theta;\mathsf{z},\mathsf{x})} M_n^j]$$
$$= \sum_{\alpha \in \mathcal{A}(q_j)} \sum_{\nu \in \mathcal{N}_4(\alpha)} \mathsf{C}(\alpha, \nu) \theta^{|\nu \cdot 1|} \{2^{-1}(1 - \theta^2)\}^{|\nu \cdot 2|} \cdot \mathsf{i}^{|\nu \cdot 1| + 2|\nu \cdot 2| + |\nu \cdot 3| + |\nu \cdot 4|}$$
$$\times E\left[e^{\lambda(\theta;\mathsf{z},\mathsf{x})} \left\langle \bigotimes_{i=1}^{q_j} (D^i M_n[\mathsf{z}])^{\otimes \nu_{i1}} \otimes (D^i \mathfrak{C}_n[\mathsf{z}^{\otimes 2}])^{\otimes \nu_{i2}} \right. \right.$$
$$\left. \left. \otimes (D^i W_n[\mathsf{z}])^{\otimes \nu_{i3}} \otimes (D^i X_n[\mathsf{x}])^{\otimes \nu_{i4}}, u_n^j \right\rangle_{H^{\otimes q_j}} \right],$$

where we also use the identity

$$\langle \mathrm{Sym}(f), g \rangle_{H^{\otimes q}} = \langle f, g \rangle_{H^{\otimes q}} \qquad (A.7)$$

holding for any $f, g \in H^{\otimes q}$ such that g is symmetric. Now, by (A.3) we have

$$\bigotimes_{i=1}^{q_j} (D^i M_n[\mathsf{z}])^{\otimes \nu_{i1}} \otimes (D^i \mathfrak{C}_n[\mathsf{z}^{\otimes 2}])^{\otimes \nu_{i2}} \otimes (D^i W_n[\mathsf{z}])^{\otimes \nu_{i3}} \otimes (D^i X_n[\mathsf{x}])^{\otimes \nu_{i4}}$$
$$= \bigotimes_{i=1}^{q_j} (D^i M_n)^{\otimes \nu_{i1}}[\mathsf{z}^{\otimes \nu_{i1}}] \otimes (D^i \mathfrak{C}_n)^{\otimes \nu_{i2}}[\mathsf{z}^{\otimes 2\nu_{i2}}] \otimes (D^i W_n)^{\otimes \nu_{i3}}[\mathsf{z}^{\otimes \nu_{i3}}]$$
$$\otimes (D^i X_n)^{\otimes \nu_{i4}}[\mathsf{x}^{\otimes \nu_{i4}}]$$
$$= \left(\bigotimes_{i=1}^{q_j} (D^i M_n)^{\otimes \nu_{i1}} \otimes (D^i \mathfrak{C}_n)^{\otimes \nu_{i2}} \otimes (D^i W_n)^{\otimes \nu_{i3}} \otimes (D^i X_n)^{\otimes \nu_{i4}} \right)$$
$$\times [\mathsf{z}^{\otimes(\nu_{i1} + 2\nu_{i2} + \nu_{i3})} \otimes \mathsf{x}^{\otimes \nu_{i4}}],$$

so using (A.2) we obtain

$$E[e^{\lambda_n(\theta;\mathsf{z},\mathsf{x})} M_n^j]$$
$$= \sum_{\alpha \in \mathcal{A}(q_j)} \sum_{\nu \in \mathcal{N}_4(\alpha)} \mathsf{C}(\alpha, \nu) \theta^{|\nu \cdot 1|} \{2^{-1}(1 - \theta^2)\}^{|\nu \cdot 2|} \cdot \mathsf{i}^{|\nu \cdot 1| + 2|\nu \cdot 2| + |\nu \cdot 3| + |\nu \cdot 4|}$$

$$\times E\left[e^{\lambda(\theta;\mathsf{z},\mathsf{x})}\Delta_{n,j}(\nu)[\mathsf{z}^{\otimes(\nu_{i1}+2\nu_{i2}+\nu_{i3})}\otimes\mathsf{x}^{\otimes\nu_{i4}}]\right]$$

$$= \theta E\left[e^{\lambda(\theta;\mathsf{z},\mathsf{x})}\left\langle D^{q_j}M_n, u_n^j\right\rangle_{H^{\otimes q_j}}[\mathsf{iz}]\right]$$

$$+ \sum_{\alpha\in\mathcal{A}(q_j)}\sum_{\nu\in\mathcal{N}_4^*(\alpha)}\mathsf{C}(\alpha,\nu)\theta^{|\nu\cdot 1|}\{2^{-1}(1-\theta^2)\}^{|\nu\cdot 2|}$$

$$\times E\left[e^{\lambda(\theta;\mathsf{z},\mathsf{x})}\Delta_{n,j}(\nu)[(\mathsf{iz})^{\otimes(\nu_{i1}+2\nu_{i2}+\nu_{i3})}\otimes(\mathsf{ix})^{\otimes\nu_{i4}}]\right].$$

Combining this identity with (A.6), we obtain the desired result. □

The following lemma is presumably a standard result. We prove it for the shake of completeness.

Lemma A.5. *Let $f = (f(y))_{y\in\mathbb{R}^N}$ be a real-valued C^∞ function all of whose partial derivatives are of polynomial growth. Then there is a sequence $(f_j)_{j=1}^\infty$ of compactly supported real-valued C^∞ functions on $\mathbb{R}^N$ such that*

$$E\left[\xi_0\partial_y^\alpha f_j(\xi_1,\ldots,\xi_N)\right] \to E\left[\xi_0\partial_y^\alpha f(\xi_1,\ldots,\xi_N)\right] \tag{A.8}$$

as $j\to\infty$ for any $\xi_0,\xi_1,\ldots,\xi_N\in L^{\infty-}$ and $\alpha\in\mathbb{Z}_+^N$.

Proof. Take a C^∞ function $\phi:\mathbb{R}^N\to[0,\infty)$ having compact support and satisfying $\phi(0) = 1$. For every $j = 1,2,\ldots$, we define the function $\phi_j:\mathbb{R}^N\to[0,\infty)$ by $\phi_j(y) = \phi(j^{-1}y)$, $y\in\mathbb{R}^N$. Then we define $f_j := f\phi_j$ for $j = 1,2,\ldots$. f_j is evidently a C^∞ function with compact support. Moreover, we have $\partial_y^\alpha f_j(y) \to \partial_y^\alpha f(y)$ as $j\to\infty$ for any $y\in\mathbb{R}^N$ and $\alpha\in\mathbb{Z}_+^N$. In addition, for any $s\in\mathbb{N}$, there is a constant $C > 0$ which depends only on ϕ and s such that $|f_j(y)| \le C(|f(y)| + \sum_{k=1}^s\|\partial_y^{\otimes k}f(y)\|_{\ell_1})$ for any $y\in\mathbb{R}^N$; we can easily prove these facts by directly differentiating f_j with the help of the Leibniz formula and the chain rule. Consequently, we have $\sup_{j\in\mathbb{N}}\|\xi_0\partial_y^\alpha f_j(\xi_1,\ldots,\xi_N)\|_2 < \infty$ for any $\alpha\in\mathbb{Z}_+^N$ because $\xi_0,\xi_1,\ldots,\xi_N\in L^{\infty-}$ and all the partial derivatives of f have polynomial growth. Therefore, $(\xi_0\partial_y^\alpha f_j(\xi_1,\ldots,\xi_N))_{j\in\mathbb{N}}$ is uniformly integrable, so the Vitali convergence theorem yields (A.8). This completes the proof. □

Now we get the following interpolation formula for (A.5):

Lemma A.6. *Let $f:\mathbb{R}^{2d}\to\mathbb{R}$ be a C^∞ function all of whose partial derivatives are of polynomial growth. Under the assumptions of Theorem 3.1, we have*

$$E[f(Z_n, X_n)] - E[f(\mathfrak{Z}_n, X_n)]$$

$$= \sum_{i,j=1}^d\int_0^1\theta E\left[\left(\langle D^{q_j}M_n^i, u_n^j\rangle_{H^{\otimes q_j}} - \mathfrak{C}_n^{ij}\right)\partial_{z_i}\partial_{z_j}f(\theta Z_n + \sqrt{1-\theta^2}\mathfrak{Z}_n, X_n)\right]d\theta$$

$$+ \sum_{j=1}^d\sum_{\alpha\in\mathcal{A}(q_j)}\sum_{\nu\in\mathcal{N}_4^*(\alpha)}\mathsf{C}(\alpha,\nu)\int_0^1\theta^{|\nu\cdot 1|}(2^{-1}(1-\theta^2))^{|\nu\cdot 2|}$$

$$\times E\left[\Delta_{n,j}(\nu)[\partial_z^{\otimes|\nu|*}\partial_x^{\otimes|\nu\cdot 4|}\partial_{z_j}f(\theta Z_n + \sqrt{1-\theta^2}\mathfrak{Z}_n, X_n)]\right]d\theta.$$

Proof. Thanks to Lemma A.5, it is enough to prove the lemma when f is rapidly decreasing. In this case the Fourier inversion formula and the Fubini theorem yield

$$E[f(Z_n, X_n)] - E[f(\mathfrak{Z}_n, X_n)] = (2\pi)^{-2d} \int_{\mathbb{R}^{2d}} \hat{f}(\mathsf{z}, \mathsf{x})\{\varphi(1; \mathsf{z}, \mathsf{x}) - \varphi(0; \mathsf{z}, \mathsf{x})\}d\mathsf{z}d\mathsf{x}$$

$$= (2\pi)^{-2d} \int_0^1 d\theta \int_{\mathbb{R}^{2d}} \hat{f}(\mathsf{z}, \mathsf{x})\partial_\theta \varphi(\theta; \mathsf{z}, \mathsf{x})d\mathsf{z}d\mathsf{x}.$$

Hence the desired result follows from Lemma A.4, (A.4) and the Fourier inversion formula. □

We will use the following elementary result in the proof:

Lemma A.7. *Let k, l be two positive integers. Then we have*

$$\partial_{z^{i_1} \dots z^{i_k}} \left(z^{j_1} \cdots z^{j_l}\right) = \sum_{\substack{c_1, \dots, c_k = 1 \\ c_s \neq c_t}}^{l} \prod_{s=1}^{k} 1_{\{j_{c_s} = i_s\}} \prod_{b \neq c_1, \dots, c_k} z^{j_b}$$

for any $i_1, \dots, i_k, j_1, \dots, j_l \in \{1, \dots, d\}$.

One can easily prove the above lemma by induction on k and application of the Leibniz rule, so we omit its proof.

Finally, as in the original CCK theory, a special approximation of the maximum function (called the "smooth max function") will play a crucial role in our proof. The following lemma summarizes the key properties of this smooth max function used in the proof:

Lemma A.8. *Let $\varepsilon > 0$ and set $\beta = \varepsilon^{-1} \log m$. Define the function $\Phi_\beta : \mathbb{R}^m \to \mathbb{R}$ by*

$$\Phi_\beta(w) = \beta^{-1} \log \left(\sum_{j=1}^{m} \exp(\beta w^j)\right), \qquad w \in \mathbb{R}^m. \tag{A.9}$$

Then we have
$$0 \leq \Phi_\beta(w) - \max_{1 \leq j \leq m} w^j \leq \beta^{-1} \log m = \varepsilon \tag{A.10}$$

for every $w \in \mathbb{R}^m$. Moreover, for any C^∞ function $g : \mathbb{R} \to \mathbb{R}$, $s \in \mathbb{N}$, $\varepsilon > 0$ and $w \in \mathbb{R}^m$, it holds that

$$\left\|\partial_w^{\otimes s} g(\varepsilon^{-1}\Phi_\beta(w))\right\|_{\ell_1} \leq C_{g,s} \max\{\varepsilon^{-s}, \varepsilon^{-1}\beta^{s-1}\} = C_{g,s}\varepsilon^{-s}(\log m)^{s-1}, \tag{A.11}$$

where $C_{g,s} > 0$ depends only on g and s.

Proof. First, note that Φ_β is usually denoted by F_β in the literature on the CCK theory. Now, (A.10) is stated in e.g. Eq.(1) of [17]. On the other hand, (A.11) is obtained by applying Lemma 5 in [24] with $h = g$, $n = 1$, $m = s$ and $b = \varepsilon^{-1}$ in their notation. □

Proof of Theorem 3.1. First, as is already noted in the above, for the proof it is enough to focus only on the case that $\boldsymbol{X}_n^{1\cdot} = \cdots = \boldsymbol{X}_n^{m\cdot} =: X_n$ for every $n \in \mathbb{N}$ due to Lemma A.1. Note that in this case we have $\Xi_n z = \Upsilon_n(z \circ X_n)$ for every $z \in \mathbb{R}^d$.

We turn to the main body of the proof. Take a number $\varepsilon > 0$ arbitrarily, and set $\beta = \varepsilon^{-1} \log m$. We define the function $\Phi_\beta : \mathbb{R}^m \to \mathbb{R}$ by (A.9). We also take a C^∞ function $g : \mathbb{R} \to [0, 1]$ such that all the derivatives of g are bounded and $g(t) = 1$ for $t \leq 0$ and $g(t) = 0$ for $t \geq 1$.

Now let us fix a vector $y \in \mathbb{R}^m$ arbitrarily, and define the functions $\varphi : \mathbb{R}^m \to \mathbb{R}$, $\psi : \mathbb{R}^d \to \mathbb{R}$ and $f : \mathbb{R}^{2d} \to \mathbb{R}$ by

$$\varphi(w) = g(\varepsilon^{-1} \Phi_\beta(w - y - \varepsilon)), \qquad w \in \mathbb{R}^m,$$
$$\psi(v) = \varphi(\Upsilon_n v), \qquad v \in \mathbb{R}^d,$$
$$f(z, x) = \psi(z \circ x), \qquad z, x \in \mathbb{R}^d.$$

For any $k, l \in \mathbb{Z}_+$ and any $z, x \in \mathbb{R}^d$, we have

$$\left\| \partial_z^{\otimes k} \partial_x^{\otimes l} f(z, x) \right\|_{\ell_1} = \sum_{i_1, \ldots, i_k, j_1, \ldots, j_l = 1}^{d} \left| \partial_{z^{i_1} \ldots z^{i_k}} \left(z^{j_1} \cdots z^{j_l} \partial_{v^{j_1} \ldots v^{j_l}} \psi(z \circ x) \right) \right|.$$

Applying the Leibniz rule repeatedly (cf. Proposition 5 of [33]), we deduce

$$\left\| \partial_z^{\otimes k} \partial_x^{\otimes l} f(z, x) \right\|_{\ell_1}$$
$$= \sum_{i_1, \ldots, i_k, j_1, \ldots, j_l = 1}^{d} \left| \sum_{\mathcal{A} \subset \{1, \ldots, k\}} \left(\prod_{a \in \mathcal{A}} \partial_{z^{i_a}} \right) (z^{j_1} \cdots z^{j_l}) \left(\prod_{a \notin \mathcal{A}} x^{i_a} \right) \right.$$
$$\left. \times \left(\prod_{a \notin \mathcal{A}} \partial_{v^{i_a}} \right) \partial_{v^{j_1} \ldots v^{j_l}} \psi(z \circ x) \right|$$
$$\leq \sum_{j_1, \ldots, j_l = 1}^{d} \sum_{\mathcal{A} \subset \{1, \ldots, k\}} \sum_{i_1, \ldots, i_k = 1}^{d} \left| \left(\prod_{a \in \mathcal{A}} \partial_{z^{i_a}} \right) (z^{j_1} \cdots z^{j_l}) \right.$$
$$\left. \times \left(\prod_{a \notin \mathcal{A}} x^{i_a} \right) \left(\prod_{a \notin \mathcal{A}} \partial_{v^{i_a}} \right) \partial_{v^{j_1} \ldots v^{j_l}} \psi(z \circ x) \right|.$$

Now let us fix a subset $\mathcal{A}$ of $\{1, \ldots, k\}$. Let r be the number of elements of $\mathcal{A}$ and we write $\mathcal{A} = \{a_1, \ldots, a_r\}$ and $\{1, \ldots, k\} \setminus \mathcal{A} = \{b_1, \ldots, b_{k-r}\}$. Assume $1 \leq r \leq l$. Then, by Lemma A.7 we obtain

$$\sum_{i_1, \ldots, i_k = 1}^{d} \left| \left(\prod_{a \in \mathcal{A}} \partial_{z^{i_a}} \right) (z^{j_1} \cdots z^{j_l}) \left(\prod_{a \notin \mathcal{A}} x^{i_a} \right) \left(\prod_{a \notin \mathcal{A}} \partial_{v^{i_a}} \right) \partial_{v^{j_1} \ldots v^{j_l}} \psi(z \circ x) \right|$$
$$= \sum_{i_1, \ldots, i_k = 1}^{d} \left| \left(\sum_{\substack{c_1, \ldots, c_r = 1 \\ c_s \neq c_t}}^{l} \prod_{s=1}^{r} 1_{\{j_{c_s} = i_{a_s}\}} \prod_{b \neq c_1, \ldots, c_r} z^{j_b} \right) \right.$$

$$\times \left(\prod_{t=1}^{k-r} x^{i_{b_t}}\right) \partial_{v^{i_{b_1}}\ldots v^{i_{b_{k-r}}} v^{j_1}\ldots v^{j_l}} \psi(z\circ x)\Bigg|$$

$$= \sum_{i_{a_1},\ldots,i_{a_r}=1}^{d} \left| \sum_{\substack{c_1,\ldots,c_r=1\\c_s\neq c_t}}^{l} \prod_{s=1}^{r} 1_{\{j_{c_s}=i_{a_s}\}} \prod_{b\neq c_1,\ldots,c_r} z^{j_b} \right|$$

$$\times \sum_{i_{b_1},\ldots,i_{b_{k-r}}=1}^{d} \left| \left(\prod_{t=1}^{k-r} x^{i_{b_t}}\right) \partial_{v^{i_{b_1}}\ldots v^{i_{b_{k-r}}} v^{j_1}\ldots v^{j_l}} \psi(z\circ x) \right|$$

$$\leq \frac{l!}{(l-r)!} \|z\|_{\ell_\infty}^{l-r} \|x\|_{\ell_\infty}^{k-r} \left\| \partial_v^{\otimes(k-r)} \partial_{v^{j_1}\ldots v^{j_l}} \psi(z\circ x) \right\|_{\ell_1}.$$

Note that the above inequality evidently holds true if $\mathcal{A}=\emptyset$. Moreover, we obviously have $\left(\prod_{a\in\mathcal{A}} \partial_{z^{i_a}}\right)\left(z^{j_1}\cdots z^{j_l}\right) = 0$ if $r > l$. Consequently, we infer that

$$\left\|\partial_z^{\otimes k}\partial_x^{\otimes l} f(z,x)\right\|_{\ell_1} \leq \sum_{r=0}^{k\wedge l} r! \binom{k}{r}\binom{l}{r} \|z\|_{\ell_\infty}^{l-r} \|x\|_{\ell_\infty}^{k-r} \left\|\partial_v^{\otimes(k+l-r)}\psi(z\circ x)\right\|_{\ell_1}.$$

Meanwhile, we can easily verify that

$$\partial_{v^{i_1}\ldots v^{i_s}}\psi(v) = \sum_{j_1,\ldots,j_s=1}^{m} \partial_{w^{j_1}\ldots w^{j_s}}\varphi(\Upsilon_n v)\Upsilon_n^{j_1 i_1}\cdots\Upsilon_n^{j_s i_s}$$

for any $s\in\mathbb{N}$ and $i_1,\ldots,i_s\in\{1,\ldots,d\}$. Hence we have

$$\|\partial_v^{\otimes s}\psi(v)\|_{\ell_1} \leq \left(\max_{1\leq j\leq m}\sum_{i=1}^{d}|\Upsilon_n^{ji}|\right)^s \sum_{j_1,\ldots,j_s=1}^{m} |\partial_{w^{j_1}\ldots w^{j_s}}\varphi(\Upsilon_n v)|$$

$$= |\!|\!|\Upsilon_n|\!|\!|_\infty^s \|\partial_w^{\otimes s}\varphi(\Upsilon_n v)\|_{\ell_1}.$$

Now, by (A.11) it holds that

$$\|\partial_w^{\otimes s}\varphi(w)\|_{\ell_1} \leq C_{g,s}\varepsilon^{-s}(\log m)^{s-1}$$

for all $w\in\mathbb{R}^m$, where $C_{g,s} > 0$ is a constant which depends only on g and s. Therefore, we obtain

$$\left\|\partial_z^{\otimes k}\partial_x^{\otimes l} f(z,x)\right\|_{\ell_1}$$
$$\leq c_{g,k,l}\sum_{r=0}^{k\wedge l} r!\binom{k}{r}\binom{l}{r}\|z\|_{\ell_\infty}^{l-r}\|x\|_{\ell_\infty}^{k-r}|\!|\!|\Upsilon_n|\!|\!|_\infty^{k+l-r}\varepsilon^{-(k+l-r)}(\log m)^{k+l-r-1}$$
$$\leq c'_{g,k,l}\left(\|z\|_{\ell_\infty}\vee 1\right)^l\left(\|x\|_{\ell_\infty}\vee 1\right)^k|\!|\!|\Upsilon_n|\!|\!|_\infty^{k+l}\varepsilon_1^{-(k+l)}(\log m)^{k+l-1},$$

where $\varepsilon_1 := \varepsilon\wedge 1$ and $c_{g,k,l}, c'_{g,k,l} > 0$ are constants which depend only on g and k, l (recall $|\!|\!|\Upsilon_n|\!|\!|_\infty \geq 1$ by assumption). We especially infer that all the partial

derivatives of f are of polynomial growth. Therefore, noting that $\Delta_{n,j}(\nu) = 0$ when $\nu \notin \bigcup_{\alpha \in \mathcal{A}(q_j)} \mathcal{N}_4^*(\alpha)$, by (A.1) and Lemma A.6 we obtain

$$\eta_n(\varepsilon) := |E[f(Z_n, X_n)] - E[f(\mathfrak{Z}_n, X_n)]|$$

$$\leq \int_0^1 E\left[\|\Delta_n\|_{\ell_\infty} \sum_{i,j=1}^d \left|\partial_{z_i}\partial_{z_j} f(\theta Z_n + \sqrt{1-\theta^2}\mathfrak{Z}_n, X_n)\right|\right] d\theta$$

$$+ K_{\overline{q}} \sum_{j=1}^d \sum_{\alpha \in \mathcal{A}(q_j)} \sum_{\nu \in \mathcal{N}_4^*(\alpha)} \int_0^1 E\left[\|\Delta_{n,j}(\nu)\|_{\ell_\infty}\right.$$

$$\left. \times \|\partial_z^{\otimes|\nu|_*}\partial_x^{\otimes|\nu_{\cdot 4}|}\partial_{z_j} f(\theta Z_n + \sqrt{1-\theta^2}\mathfrak{Z}_n, X_n)\|_{\ell_1}\right] d\theta$$

$$\leq \int_0^1 E\left[\|X_n\|_{\ell_\infty}^2\||\Upsilon_n\||_\infty^2\|\Delta_n\|_{\ell_\infty} \sum_{i,j=1}^m \left|\partial_{w_i}\partial_{w_j}\varphi(X_n \circ (\theta Z_n + \sqrt{1-\theta^2}\mathfrak{Z}_n))\right|\right] d\theta$$

$$+ K_{\overline{q}} \sum_{\alpha \in \overline{\mathcal{A}}(\overline{q})} \sum_{\nu \in \mathcal{N}_4^*(\alpha)} \int_0^1 E\left[\max_{1 \leq j \leq d}\|\Delta_{n,j}(\nu)\|_{\ell_\infty}\|\partial_z^{\otimes(|\nu|_*+1)}\right.$$

$$\left. \times \partial_x^{\otimes|\nu_{\cdot 4}|} f(\theta Z_n + \sqrt{1-\theta^2}\mathfrak{Z}_n, X_n)\|_{\ell_1}\right] d\theta$$

$$\leq c''_{g,\overline{q}}\varepsilon_1^{-2}(\log m)\||\Upsilon_n\||_\infty^2 E\left[\|X_n\|_{\ell_\infty}^2\|\Delta_n\|_{\ell_\infty}\right]$$

$$+ c''_{g,\overline{q}} \sum_{\alpha \in \overline{\mathcal{A}}(\overline{q})} \sum_{\nu \in \mathcal{N}_4^*(\alpha)} \varepsilon_1^{-|\nu|_{**}-1}(\log m)^{|\nu|_{**}}\||\Upsilon_n\||_\infty^{|\nu|_{**}+1}$$

$$\times E\left[\left(1 + \|X_n\|_{\ell_\infty}^{|\nu|_*+1}\right)\left(1 + \|Z_n\|_{\ell_\infty}^{|\nu_{\cdot 4}|} + \|\mathfrak{Z}_n\|_{\ell_\infty}^{|\nu_{\cdot 4}|}\right) \max_{1 \leq j \leq d}\|\Delta_{n,j}(\nu)\|_{\ell_\infty}\right],$$

where $K_{\overline{q}} > 0$ depends only on $\overline{q}$ and $c''_{g,\overline{q}} > 0$ depends only on g and $\overline{q}$. Now we have

$$P(\Xi_n Z_n \leq y) \leq P(\Phi_\beta(\Upsilon_n(Z_n \circ X_n) - y - \varepsilon) \leq 0) \ (\because \text{Eq.(A.10)})$$
$$\leq E[f(Z_n, X_n)] \leq E[f(\mathfrak{Z}_n, X_n)] + \eta_n(\varepsilon)$$
$$\leq P(\Phi_\beta(\Upsilon_n(\mathfrak{Z}_n \circ X_n) - y - \varepsilon) < \varepsilon) + \eta_n(\varepsilon) \ (\because \text{the definition of } g)$$
$$\leq P(\Xi_n\mathfrak{Z}_n \leq y + 2\varepsilon) + \eta_n(\varepsilon) \ (\because \text{Eq.(A.10)}).$$

Set $\Gamma_n := \Xi_n \mathfrak{C}_n \Xi_n^\top$. Then, by Lemma A.2 we obtain

$$P(\Xi_n Z_n \leq y)$$
$$\leq P(\min \text{diag}(\Gamma_n) < b) + P(\Xi_n\mathfrak{Z}_n \leq y) + \frac{2\varepsilon}{\sqrt{b}}(\sqrt{2\log m} + 2) + \eta_n(\varepsilon)$$

for every $b > 0$. By an analogous argument we also obtain

$$P(\Xi_n Z_n \leq y)$$
$$\geq P(\min \text{diag}(\Gamma_n) < b) - P(\Xi_n\mathfrak{Z}_n \leq y) - \frac{2\varepsilon}{\sqrt{b}}(\sqrt{2\log m} + 2) - \eta_n(\varepsilon).$$

Therefore, we conclude that

$$\sup_{y \in \mathbb{R}^m} |P(\Xi_n Z_n \le y) - P(\Xi_n \mathfrak{Z}_n \le y)|$$

$$\le P(\min \operatorname{diag}(\Gamma_n) < b) + \frac{2\varepsilon}{\sqrt{b}}(\sqrt{2\log m} + 2) + \eta_n(\varepsilon).$$

Taking

$$\varepsilon = \left(\sqrt{\log m} \||\Upsilon_n|\|_\infty^2 E\left[\|X_n\|_{\ell_\infty}^2 \|\Delta_n\|_{\ell_\infty} \right] \right)^{1/3} \vee$$

$$\max_{\nu \in \overline{\mathcal{N}}_4^*(\overline{q})} \left\{ (\log m)^{|\nu|_{**} - \frac{1}{2}} \||\Upsilon_n|\|_\infty^{|\nu|_* + 1} E\left[(1 + \|X_n\|_{\ell_\infty}^{|\nu|_* + 1}) \right. \right.$$

$$\left. \left. \times \left(1 + \|Z_n\|_{\ell_\infty}^{|\nu.4|} + \|\mathfrak{Z}_n\|_{\ell_\infty}^{|\nu.4|} \right) \max_{1 \le j \le d} \|\Delta_{n,j}(\nu)\|_{\ell_\infty} \right] \right\}^{\frac{1}{|\nu|_{**} + 2}},$$

we obtain

$$\limsup_{n \to \infty} \sup_{y \in \mathbb{R}^m} |P(\Xi_n Z_n \le y) - P(\Xi_n \mathfrak{Z}_n \le y)| \le \limsup_{n \to \infty} P(\min \operatorname{diag}(\Gamma_n) < b)$$

by assumption. Letting $b \to 0$, we complete the proof. $\square$

A.3. Proof of Lemma 3.1

Take a number $\varepsilon > 0$ arbitrarily. For any $y \in \mathbb{R}^m$, we have

$$P(Y_n \le y) \le P(\sqrt{\log m} \|Y_n - \Xi_n Z_n\|_{\ell_\infty} > \varepsilon) + P(\Xi_n Z_n \le y + \varepsilon/\sqrt{\log m})$$

$$\le P(\sqrt{\log m} \|Y_n - \Xi_n Z_n\|_{\ell_\infty} > \varepsilon) + P(\Xi_n \mathfrak{Z}_n \le y + \varepsilon/\sqrt{\log m}) + \rho_n,$$

where

$$\rho_n := \sup_{y \in \mathbb{R}^m} |P(\Xi_n Z_n \le y) - P(\Xi_n \mathfrak{Z}_n \le y)|.$$

Therefore, Lemma A.2 yields

$$P(Y_n \le y) \le P(\sqrt{\log m} \|Y_n - \Xi_n Z_n\|_{\ell_\infty} > \varepsilon) + P(\Xi_n \mathfrak{Z}_n \le y)$$

$$+ \frac{2\varepsilon}{\sqrt{b \log m}}(\sqrt{2 \log m} + 2) + P(\min \operatorname{diag}(\Gamma_n) < b) + \rho_n$$

for any $b > 0$. An analogous argument yields

$$P(Y_n \le y) \ge -P(\sqrt{\log m} \|Y_n - \Xi_n Z_n\|_{\ell_\infty} > \varepsilon) + P(\Xi_n \mathfrak{Z}_n \le y)$$

$$- \frac{2\varepsilon}{\sqrt{b \log m}}(\sqrt{2 \log m} + 2) + P(\min \operatorname{diag}(\Gamma_n) < b) - \rho_n,$$

so we conclude that

$$\sup_{y\in\mathbb{R}^m} |P(Y_n \le y) - P(\Xi_n \mathfrak{Z}_n \le y)| \le P(\sqrt{\log m}\|Y_n - \Xi_n Z_n\|_{\ell_\infty} > \varepsilon)$$

$$+ \frac{2\varepsilon}{\sqrt{b\log m}}(\sqrt{2\log m} + 2) + P(\min \operatorname{diag}(\Gamma_n) < b) + \rho_n.$$

Now, by assumption we obtain

$$\limsup_{n\to\infty} \sup_{y\in\mathbb{R}^m} |P(Y_n \le y) - P(\Xi_n \mathfrak{Z}_n \le y)|$$

$$\le \frac{2\varepsilon}{\sqrt{b}}\left(\sqrt{2} + \frac{2}{\sqrt{\log 2}}\right) + \limsup_{n\to\infty} P(\min \operatorname{diag}(\Gamma_n) < b).$$

We first let $\varepsilon \to 0$. After that, we let $b \to 0$. Then we conclude that

$$\limsup_{n\to\infty} \sup_{y\in\mathbb{R}^m} |P(Y_n \le y) - P(\Xi_n \mathfrak{Z}_n \le y)| = 0.$$

This completes the proof. □

A.4. Proof of Proposition 3.1

The proof is analogous to that of Theorem 2 from [17] (see also the proof of Theorem 4.1 from [19]). Setting

$$\Gamma_n := \Xi_n \mathfrak{C}_n \Xi_n^\top, \qquad \widehat{\Gamma}_n := \widehat{\Xi}_n \widehat{\mathfrak{C}}_n \widehat{\Xi}_n^\top, \qquad \mu_n := \Xi_n W_n, \qquad \widehat{\mu}_n := \widehat{\Xi}_n \widehat{W}_n,$$

we have

$$P(\Xi_n \mathfrak{Z}_n \le y|\mathcal{F}) = P(\Gamma_n^{1/2}\xi_n + \mu_n \le y|\mathcal{F})$$

and

$$P(\widehat{\Xi}_n \widehat{\mathfrak{Z}}_n \le y|\mathcal{F}) = P(\widehat{\Gamma}_n^{1/2}\xi_n' + \widehat{\mu}_n \le y|\mathcal{F})$$

for all $y \in \mathbb{R}^m$, where ξ_n and ξ_n' are two independent m-dimensional standard Gaussian vectors jointly independent of $\mathcal{F}$. Therefore, it is enough to prove

$$\sup_{y\in\mathbb{R}^m} |P(\widehat{\Gamma}_n^{1/2}\xi_n' + \widehat{\mu}_n \le y|\mathcal{F}) - P(\Gamma_n^{1/2}\xi_n + \mu_n \le y|\mathcal{F})| \to^p 0$$

as $n \to \infty$. In addition, thanks to the condition (3.7), it suffices to prove the above convergence on the set $\Omega_b := \{\operatorname{diag}(\Gamma_n) \ge b\}$ for an arbitrarily fixed $b > 0$. More precisely, it is enough to prove

$$P\left(\Omega_b \cap \left\{\sup_{y\in\mathbb{R}^m} |P(\widehat{\Gamma}_n^{1/2}\xi_n' + \widehat{\mu}_n \le y|\mathcal{F}) - P(\Gamma_n^{1/2}\xi_n + \mu_n \le y|\mathcal{F})| > \eta\right\}\right) \to 0$$

as $n \to \infty$ for any $\eta > 0$.

We first prove

$$P\left(\Omega_b \cap \left\{\sup_{y\in\mathbb{R}^m} |P(\Gamma_n^{1/2}\xi_n + \widehat{\mu}_n \le y|\mathcal{F}) - P(\Gamma_n^{1/2}\xi_n + \mu_n \le y|\mathcal{F})| > \eta\right\}\right) \to 0$$

$$(A.12)$$

as $n \to \infty$ for any $\eta > 0$. By Nazarov's inequality we have

$$|P(\Gamma_n^{1/2}\xi_n + \widehat{\mu}_n \le y|\mathcal{F}) - P(\Gamma_n^{1/2}\xi_n + \mu_n \le y|\mathcal{F})| \le \frac{\|\widehat{\mu}_n - \mu_n\|_{\ell_\infty}}{\sqrt{b}}\left(\sqrt{2\log m} + 2\right)$$

a.s. on the set Ω_b for every $y \in \mathbb{R}^m$. Since the function $y \mapsto |P(\Gamma_n^{1/2}\xi_n + \widehat{\mu}_n \le y|\mathcal{F}) - P(\Gamma_n^{1/2}\xi_n + \mu_n \le y|\mathcal{F})|$ is a.s. right-continuous, the above result yields

$$\sup_{y\in\mathbb{R}^m} |P(\Gamma_n^{1/2}\xi_n + \widehat{\mu}_n \le y|\mathcal{F}) - P(\Gamma_n^{1/2}\xi_n + \mu_n \le y|\mathcal{F})|$$

$$\le \frac{\|\widehat{\mu}_n - \mu_n\|_{\ell_\infty}}{\sqrt{b}}\left(\sqrt{2\log m} + 2\right)$$

a.s. on the set Ω_b. Hence (A.12) follows from the assumption (3.18).

Thanks to (A.12), it suffices to prove

$$P\left(\Omega_b \cap \left\{ \sup_{y\in\mathbb{R}^m} |P(\widehat{\Gamma}_n^{1/2}\xi_n' + \widehat{\mu}_n \le y|\mathcal{F}) - P(\Gamma_n^{1/2}\xi_n + \widehat{\mu}_n \le y|\mathcal{F})| > \eta \right\}\right) \to 0$$

as $n \to \infty$ for any $\eta > 0$. However, since we have

$$\sup_{y\in\mathbb{R}^m} |P(\widehat{\Gamma}_n^{1/2}\xi_n' + \widehat{\mu}_n \le y|\mathcal{F}) - P(\Gamma_n^{1/2}\xi_n + \widehat{\mu}_n \le y|\mathcal{F})|$$

$$= \sup_{y\in\mathbb{R}^m} |P(\widehat{\Gamma}_n^{1/2}\xi_n' \le y|\mathcal{F}) - P(\Gamma_n^{1/2}\xi_n \le y|\mathcal{F})|,$$

this amounts to proving

$$P\left(\Omega_b \cap \left\{ \sup_{y\in\mathbb{R}^m} |P(\widehat{\Gamma}_n^{1/2}\xi_n' \le y|\mathcal{F}) - P(\Gamma_n^{1/2}\xi_n \le y|\mathcal{F})| > \eta \right\}\right) \to 0 \quad \text{(A.13)}$$

as $n \to \infty$ for any $\eta > 0$. To prove this claim, we take a number $\varepsilon > 0$ arbitrarily and set $\beta = \varepsilon^{-1}\log m$ as in the proof of Theorem 3.1. Then we define the function $\Phi_\beta : \mathbb{R}^m \to \mathbb{R}$ by (A.9). We also take a C^∞ function $g : \mathbb{R} \to [0,1]$ such that all the derivatives of g is bounded and $g(t) = 1$ for $t \le 0$ and $g(t) = 0$ for $t \ge 1$.

Fix a vector $y \in \mathbb{R}^m$ arbitrarily and define the function $\varphi : \mathbb{R}^m \to \mathbb{R}$ by

$$\varphi(w) = g(\varepsilon^{-1}\Phi_\beta(w - y - \varepsilon)), \qquad w \in \mathbb{R}^m.$$

Then we define the stochastic process $\Psi = (\Psi(t))_{t\in[0,1]}$ by

$$\Psi(t) = E\left[\varphi\left(\sqrt{t}\widehat{\Gamma}_n^{1/2}\xi_n' + \sqrt{1-t}\Gamma_n^{1/2}\xi_n\right)|\mathcal{F}\right], \qquad t \in [0,1].$$

We evidently have

$$\frac{d\Psi(t)}{dt} = E\left[\partial_w^{\otimes 1}\varphi\left(\sqrt{t}\widehat{\Gamma}_n^{1/2}\xi_n' + \sqrt{1-t}\Gamma_n^{1/2}\xi_n\right)\left[\frac{\widehat{\Gamma}_n^{1/2}\xi_n'}{\sqrt{t}} - \frac{\Gamma_n^{1/2}\xi_n}{\sqrt{1-t}}\right]|\mathcal{F}\right]$$

$$\text{for all } t \in (0,1)$$

with probability one. Then, Stein's identity yields

$$\frac{d\Psi(t)}{dt} = E\left[\partial_w^{\otimes 2}\varphi\left(\sqrt{t}\widehat{\Gamma}_n^{1/2}\xi_n' + \sqrt{1-t}\Gamma_n^{1/2}\xi_n\right)\left[\widehat{\Gamma}_n - \Gamma_n\right]|\mathcal{F}\right]$$
$$\text{for all } t \in (0,1)$$

with probability one. Consequently, we obtain

$$\left|E\left[\varphi\left(\widehat{\Gamma}_n^{1/2}\xi_n'\right)|\mathcal{F}\right] - E\left[\varphi\left(\Gamma_n^{1/2}\xi_n\right)|\mathcal{F}\right]\right| \le \int_0^1 \left|\frac{d\Psi(t)}{dt}\right| dt$$
$$\le C\varepsilon^{-2}(\log m)\|\widehat{\Gamma}_n - \Gamma_n\|_{\ell_\infty}$$

by Lemmas 3–4 of [17], where $C > 0$ is a constant which depends only on g. Now we have

$$P\left(\widehat{\Gamma}_n^{1/2}\xi_n' \le y|\mathcal{F}\right)$$
$$\le P\left(\Phi_\beta(\widehat{\Gamma}_n^{1/2}\xi_n' - y - \varepsilon) \le 0|\mathcal{F}\right) \ (\because \text{Eq.(A.10)})$$
$$\le E\left[\varphi\left(\widehat{\Gamma}_n^{1/2}\xi_n'\right)|\mathcal{F}\right] \le E\left[\varphi\left(\Gamma_n^{1/2}\xi_n\right)|\mathcal{F}\right] + C\varepsilon^{-2}(\log m)\|\widehat{\Gamma}_n - \Gamma_n\|_{\ell_\infty}$$
$$\le P\left(\Phi_\beta(\Gamma_n^{1/2}\xi_n - y - \varepsilon) < \varepsilon|\mathcal{F}\right) + C\varepsilon^{-2}(\log m)\|\widehat{\Gamma}_n - \Gamma_n\|_{\ell_\infty}$$
$$(\because \text{the definition of } g)$$
$$\le P\left(\Gamma_n^{1/2}\xi_n \le y + 2\varepsilon\right) + C\varepsilon^{-2}(\log m)\|\widehat{\Gamma}_n - \Gamma_n\|_{\ell_\infty} \ (\because \text{Eq.(A.10)}).$$

Since we have on the set Ω_b

$$P(\Gamma_n^{1/2}\xi_n \le y + 2\varepsilon|\mathcal{F}) \le P(\Gamma_n^{1/2}\xi_n \le y|\mathcal{F}) + \frac{2\varepsilon}{\sqrt{b}}(\sqrt{2\log m} + 2)$$

by the Nazarov inequality, we obtain

$$P\left(\widehat{\Gamma}_n^{1/2}\xi_n' \le y|\mathcal{F}\right)$$
$$\le P(\Gamma_n^{1/2}\xi_n \le y|\mathcal{F}) + \frac{2\varepsilon}{\sqrt{b}}(\sqrt{2\log m} + 2) + C\varepsilon^{-2}(\log m)\|\widehat{\Gamma}_n - \Gamma_n\|_{\ell_\infty}$$

a.s. on the set Ω_b. By an analogous argument we also obtain

$$P\left(\widehat{\Gamma}_n^{1/2}\xi_n' \le y|\mathcal{F}\right)$$
$$\ge P(\Gamma_n^{1/2}\xi_n \le y|\mathcal{F}) - \frac{2\varepsilon}{\sqrt{b}}(\sqrt{2\log m} + 2) - C\varepsilon^{-2}(\log m)\|\widehat{\Gamma}_n - \Gamma_n\|_{\ell_\infty}$$

a.s. on the set Ω_b. Therefore, we conclude that

$$\left|P\left(\widehat{\Gamma}_n^{1/2}\xi_n' \le y|\mathcal{F}\right) - P(\Gamma_n^{1/2}\xi_n \le y|\mathcal{F})\right|$$

$$\leq \frac{2\varepsilon}{b}(\sqrt{2\log m} + 2) + C\varepsilon^{-2}(\log m)\|\widehat{\Gamma}_n - \Gamma_n\|_{\ell_\infty}$$

a.s. on the set Ω_b. Since the function $y \mapsto \left|P\left(\widehat{\Gamma}_n^{1/2}\xi'_n \leq y|\mathcal{F}\right) - P(\Gamma_n^{1/2}\xi_n \leq y|\mathcal{F})\right|$ is a.s. right-continuous, the above result implies that

$$\sup_{y \in \mathbb{R}^m} \left|P\left(\widehat{\Gamma}_n^{1/2}\xi'_n \leq y|\mathcal{F}\right) - P(\Gamma_n^{1/2}\xi_n \leq y|\mathcal{F})\right|$$

$$\leq \frac{2\varepsilon}{b}(\sqrt{2\log m} + 2) + C\varepsilon^{-2}(\log m)\|\widehat{\Gamma}_n - \Gamma_n\|_{\ell_\infty}$$

a.s. on the set Ω_b. Hence we deduce

$$P\left(\Omega_b \cap \left\{\sup_{y \in \mathbb{R}^m} |P(\widehat{\Gamma}_n^{1/2}\xi'_n \leq y|\mathcal{F}) - P(\Gamma_n^{1/2}\xi_n \leq y|\mathcal{F})| > \eta\right\}\right)$$

$$\leq P\left(\frac{2\varepsilon}{b}(\sqrt{2\log m} + 2) + C\varepsilon^{-2}(\log m)\|\widehat{\Gamma}_n - \Gamma_n\|_{\ell_\infty} > \eta\right)$$

for all $n \in \mathbb{N}$. Now, take a number $a > 0$ such that $\frac{2a}{b}(\sqrt{2} + 2/\sqrt{\log 2}) \leq \frac{\eta}{2}$ and set $\varepsilon = a/\sqrt{\log m}$. Then the above inequality yields

$$P\left(\Omega_b \cap \left\{\sup_{y \in \mathbb{R}^m} |P(\widehat{\Gamma}_n^{1/2}\xi'_n \leq y|\mathcal{F}) - P(\Gamma_n^{1/2}\xi_n \leq y|\mathcal{F})| > \eta\right\}\right)$$

$$\leq P\left(\frac{C}{a}(\log m)^2\|\widehat{\Gamma}_n - \Gamma_n\|_{\ell_\infty} > \frac{\eta}{2}\right).$$

Therefore, (A.13) follows from the assumption (3.18), which yields the desired result. □

A.5. Proof of Proposition 3.2

We follow Step 3 in the proof of Theorem 2 from [41]. First, by assumption and Theorem 9.2.2 of [26] there is a sequence ε_n of positive numbers tending to 0 such that

$$P\left(\mathcal{E}_n^c\right) \leq \varepsilon_n, \qquad \sup_{x \in \mathbb{R}}\left|P\left(T_n \leq x\right) - P\left(T_n^\dagger \leq x\right)\right| \leq \varepsilon_n$$

for all $n \in \mathbb{N}$, where

$$\mathcal{E}_n = \left\{\sup_{x \in \mathbb{R}}\left|P\left(T_n^\dagger \leq x|\mathcal{F}\right) - P\left(T_n^* \leq x|\mathcal{F}\right)\right| \leq \varepsilon_n\right\}.$$

Next, let us denote by $q_n^\dagger$ the $\mathcal{F}$-conditional quantile function of $T_n^\dagger$. Then, on the set $\mathcal{E}_n \cap E_n$ we have

$$P\left(T_n^* \leq q_n^\dagger(\alpha + \varepsilon_n)|\mathcal{F}\right) \geq P\left(T_n^\dagger \leq q_n^\dagger(\alpha + \varepsilon_n)|\mathcal{F}\right) - \varepsilon_n = \alpha.$$

Hence, on $\mathcal{E}_n \cap E_n$ it holds that $q_n^*(\alpha) \leq q_n^\dagger(\alpha + \varepsilon_n)$. Therefore, we obtain

$$
\begin{aligned}
P\left(T_n \leq q_n^*(\alpha)\right) &\leq P\left(T_n \leq q_n^\dagger(\alpha + \varepsilon_n)\right) + P(\mathcal{E}_n^c) + P(E_n^c) \\
&\leq P\left(T_n^\dagger \leq q_n^\dagger(\alpha + \varepsilon_n)\right) + 2\varepsilon_n + P(E_n^c) = \alpha + 3\varepsilon_n + P(E_n^c).
\end{aligned}
$$

Meanwhile, for any $\omega \in \mathcal{E}_n \cap E_n$ and any $z \in \mathbb{R}$ such that $P(T_n^* \leq z | \mathcal{F})(\omega) \geq \alpha$, we have

$$
\begin{aligned}
&P\left(T_n^\dagger \leq q_n^\dagger(\alpha - \varepsilon_n)(\omega) | \mathcal{F}\right)(\omega) \\
&= \alpha - \varepsilon_n \leq P(T_n^* \leq z | \mathcal{F})(\omega) - \varepsilon_n \leq P\left(T_n^\dagger \leq z | \mathcal{F}\right)(\omega).
\end{aligned}
$$

Hence it holds that $q_n^\dagger(\alpha - \varepsilon_n)(\omega) \leq z$. This implies that $q_n^*(\alpha) \geq q_n^\dagger(\alpha - \varepsilon_n)$ on $\mathcal{E}_n \cap E_n$. Therefore, we obtain

$$
\begin{aligned}
P\left(T_n < q_n^*(\alpha)\right) &\geq P\left(T_n < q_n^\dagger(\alpha - \varepsilon_n)\right) - P(\mathcal{E}_n^c) - P(E_n^c) \\
&\geq P\left(T_n^\dagger < q_n^\dagger(\alpha - \varepsilon_n)\right) - 2\varepsilon_n - P(E_n^c) = \alpha - 3\varepsilon_n - P(E_n^c).
\end{aligned}
$$

Consequently, we obtain $P\left(T_n \leq q_n^*(\alpha)\right) \to \alpha$ as $n \to \infty$. $\qquad\square$

Appendix B: Proofs for Section 4

B.1. Proof of Theorem 4.1

We first introduce some notation. For two sequences $(x_n), (y_n)$ of numbers, the notation $x_n \lesssim y_n$ means that there is a *universal* constant $C > 0$ such that $x_n \leq C y_n$ for all n. Here, the value of the constant C will change from line to line. We define the d-dimensional processes $\mathsf{A} = (\mathsf{A}_t)_{t \in [0,1]}$ and $\mathsf{M} = (\mathsf{M}_t)_{t \in [0,1]}$ by

$$
\mathsf{A}_t = \int_0^t \mu_s ds, \qquad \mathsf{M}_t = \int_0^t \sigma_s dB_s
$$

for every $t \in [0,1]$. If $\phi = (\phi_t)_{t \in [0,1]}$ is an r-dimensional $(\mathcal{F}_t)$-progressively measurable process such that $\int_0^1 \|\phi_t\|_{\ell_2}^2 dt < \infty$ a.s., we define

$$
\int_0^t \phi_s \cdot dB_s := \sum_{a=1}^r \int_0^t \phi_s^a dB_s^a
$$

for all $t \in [0,1]$.

For every $n \in \mathbb{N}$, we set $I_h = I_h^n := (t_{h-1}, t_h]$ for every $h = 1, \ldots, n$ and define the filtration $(\mathcal{G}_t^n)_{t \in [0,1]}$ by $\mathcal{G}_0^n := \mathcal{F}_0$ and

$$
\mathcal{G}_t^n := \mathcal{F}_{t_{h-1}}
$$

when $t \in I_h$ for some $h = 1, \ldots, n$. Then we define the process $(\varsigma_t)_{t \in [0,1]}$ by

$$
\varsigma_t = E[\sigma_t | \mathcal{G}_t^n], \qquad t \in [0,1]
$$

(we subtract the index n from ς_t although it depends on n). For all $i, j = 1, \ldots, d$, we define the symmetric $H^{\otimes 2}$-valued random variable u_n^{ij} by

$$u_n^{ij} := \sqrt{n} \sum_{h=1}^{n} f_n^{ij} 1_{I_h \times I_h},$$

where $f_n^{ij} = \mathrm{Sym}\left(\varsigma^{i\cdot} \otimes \varsigma^{j\cdot}\right)$. We note the following result:

Lemma B.1. *Given an index $n \in \mathbb{N}$, let $\xi = (\xi_t)_{t \in [0,1]}$ and $\eta = (\eta_t)_{t \in [0,1]}$ be $(\mathcal{G}_t^n)$-adapted r-dimensional processes such that $\sup_{t \in [0,1]} E[\|\xi_t\|_{\ell_2}^4 + \|\eta_t\|_{\ell_2}^4] < \infty$. Then $\xi \otimes \eta 1_{I_h \times I_h} \in \mathrm{Dom}(\delta^2)$ and*

$$\delta^2(\xi \otimes \eta 1_{I_h \times I_h}) = \int_{t_{h-1}}^{t_h} \left(\int_{t_{h-1}}^{t} \xi_s \cdot dB_s \right) \eta_t \cdot dB_t + \int_{t_{h-1}}^{t_h} \left(\int_{t_{h-1}}^{t} \xi_s \cdot dB_s \right) \eta_t \cdot dB_t$$

for every $h = 1, \ldots, n$.

Proof. Set $S := \{(s, t) \in [0,1]^2 : s \leq t\}$. For any $t \in [0,1]$, the process

$$(\xi_s \eta_t 1_{(I_h \times I_h) \cap S}(s, t))_{s \in [0,1]}$$

is evidently **F**-predictable and H-valued, so it belongs to $\mathrm{Dom}(\delta)$ and

$$\delta(\xi \eta_t 1_{(I_h \times I_h) \cap S}(\cdot, t)) = \eta_t 1_{I_h}(t) \int_0^t \xi_s 1_{I_h}(s) \cdot dB_s$$

by Proposition 1.3.11 of [55]. Moreover, from the above expression the process

$$(\delta(\xi \eta_t 1_{(I_h \times I_h) \cap S}(\cdot, t)))_{t \in [0,1]}$$

is evidently **F**-predictable and H-valued. Therefore, Proposition 1.3.11 of [55] and Proposition 2.6 of [58] imply that $\xi \otimes \eta 1_{(I_h \times I_h) \cap S}$ belongs to $\mathrm{Dom}(\delta^2)$ and

$$\delta^2(\xi \otimes \eta 1_{(I_h \times I_h) \cap S}) = \int_{t_{h-1}}^{t_h} \left(\int_{t_{h-1}}^{t} \xi_s \cdot dB_s \right) \eta_t \cdot dB_t.$$

Similarly, we can show that $\xi \otimes \eta 1_{(I_h \times I_h) \cap S^c} \in \mathrm{Dom}(\delta^2)$ and

$$\delta^2(\xi \otimes \eta 1_{(I_h \times I_h) \cap S^c}) = \int_{t_{h-1}}^{t_h} \left(\int_{t_{h-1}}^{t} \eta_s \cdot dB_s \right) \xi_t \cdot dB_t.$$

This completes the proof. □

Thanks to Lemma B.1, we have $u_n^{ij} \in \mathrm{Dom}(\delta^2)$, so we can define the variable M_n^{ij} by

$$M_n^{ij} = \delta^2(u_n^{ij}).$$

Next we prove some auxiliary results. We begin by noting some elementary facts which are frequently used throughout the proof. First, for any random variable ξ and any $p, q \in (0, \infty)$, it holds that

$$\||\xi|^q\|_p = \|\xi\|_{pq}^q.$$

Second, for two random variables ξ, η and numbers $p \in (0, \infty)$, $q \in (1, \infty)$, we have

$$\|\xi\eta\|_p \leq \|\xi\|_{qp}\|\eta\|_{\frac{q}{q-1}p}.$$

This is a consequence of the Hölder inequality. These facts will be used without reference in the following. We also refer to two inequalities which are repeatedly used throughout the proof. The first one is the following integral version of the Minkowski inequality:

Proposition B.1. *Let* $(\mathcal{X}, \mathcal{A}, \mathfrak{m})$ *be a* σ*-finite measure space and* $f : \mathcal{X} \times \Omega \to [0, \infty]$ *be an* $\mathcal{A} \otimes \mathcal{F}$*-measurable function. Then we have*

$$\left\| \int_{\mathcal{X}} f(x)\mathfrak{m}(dx) \right\|_p \leq \int_{\mathcal{X}} \|f(x)\|_p \mathfrak{m}(dx)$$

for all $p \in [1, \infty]$.

Proposition B.1 is an easy consequence of the standard Minkowski inequality via approximating the function f by simple functions (see also Proposition C.4 of [37]).

The second one is the following Burkholder-Davis-Gundy inequality with a sharp constant:

Proposition B.2 (Barlow & Yor [3], Proposition 4.2)**.** *There is a universal constant* $c > 0$ *such that*

$$\left\| \sup_{0 \leq t \leq T} |M_t| \right\|_p \leq c\sqrt{p} \left\| \langle M \rangle_T^{1/2} \right\|_p$$

for any $p \in [2, \infty)$ *and any continuous martingale* $M = (M_t)_{t \in [0,T]}$ *with* $M_0 = 0$.

We then prove some auxiliary estimates.

Lemma B.2. *There is a universal constant* $C > 0$ *such that*

$$\left\| \sum_{h=h_0+1}^{h_1} \int_{t_{h-1}}^{t_h} \left(\int_{t_{h-1}}^t \eta_s \cdot dB_s \right) \xi_t \cdot dB_t \right\|_p$$
$$\leq C \frac{p\sqrt{h_1 - h_0}}{n} \sup_{t_{h_0} < t < t_{h_1}} \|\xi_t\|_{qp, \ell_2} \sup_{t_{h_0} < s < t_{h_1}} \|\eta_s\|_{\frac{q}{q-1}p, \ell_2}$$

for any $p \in [2, \infty)$, $q \in (1, \infty)$, $n \in \mathbb{N}$, $h_0, h_1 = 0, 1, \ldots, n$ *such that* $h_0 < h_1$ *and any* r*-dimensional* $(\mathcal{F}_t)$*-progressively measurable processes* ξ *and* η *such that* $\sup_{t \in [0,1]}(\|\xi_t\|_p + \|\eta_t\|_p) < \infty$ *for all* $p \in [1, \infty)$.

Proof. Set $q' = q/(1-q)$. By Propositions B.1–B.2 we have

$$\left\| \sum_{h=h_0+1}^{h_1} \int_{t_{h-1}}^{t_h} \left(\int_{t_{h-1}}^{t} \eta_s \cdot dB_s \right) \xi_t \cdot dB_t \right\|_p$$

$$\lesssim \sqrt{p} \left\| \sqrt{ \sum_{h=h_0+1}^{h_1} \int_{t_{h-1}}^{t_h} \|\xi_t\|_{\ell_2}^2 \left(\int_{t_{h-1}}^{t} \eta_s \cdot dB_s \right)^2 dt } \right\|_p$$

$$\leq \sqrt{ p \sum_{h=h_0+1}^{h_1} \int_{t_{h-1}}^{t_h} \|\xi_t\|_{qp,\ell_2}^2 \left\| \int_{t_{h-1}}^{t} \eta_s \cdot dB_s \right\|_{q'p}^2 dt }$$

$$\lesssim p \sqrt{ \sum_{h=h_0+1}^{h_1} \int_{t_{h-1}}^{t_h} \|\xi_t\|_{qp,\ell_2}^2 \left\| \sqrt{ \int_{t_{h-1}}^{t} \|\eta_s\|_{\ell_2}^2 ds } \right\|_{q'p}^2 dt }$$

$$\leq \frac{p\sqrt{h_1-h_0}}{n} \sup_{t_{h_0} < t < t_{h_1}} \|\xi_t\|_{qp,\ell_2} \sup_{t_{h_0} < s < t_{h_1}} \|\eta_s\|_{q'p,\ell_2}.$$

Hence we obtain the desired result. $\qquad\qquad\qquad\qquad\qquad\qquad\quad\square$

Lemma B.3. *There is a universal constant $C > 0$ such that*

$$\max_{1 \leq h \leq n-1} \sup_{t \in [t_{h-1}, t_{h+1}]} \left\| \xi_t - E[\xi_t | \mathcal{F}_{t_{h-1}}] \right\|_{2p,\ell_2} \leq C \sqrt{\frac{p}{n}} \sup_{0 \leq u \leq v \leq 1} \|D_u \xi_v\|_{2p,\ell_2}$$

for all $n \in \mathbb{N}$, $p \in [2, \infty)$, $\mathsf{r} \in \mathbb{N}$ and any r-dimensional $(\mathcal{F}_t)$-progressively measurable process ξ such that $\xi_t \in \mathbb{D}_{1,\infty}(\mathbb{R}^{\mathsf{r}})$ for all $t \in [0,1]$.

Proof. By the Clark-Ocone formula (Proposition A.1 of [56]) we have

$$\xi_t^a = E[\xi_t^a | \mathcal{F}_{t_{h-1}}] + \int_{t_{h-1}}^{t} E[D_s \xi_t^a | \mathcal{F}_s] \cdot dB_s \qquad \text{a.s.}$$

for all $t \in [t_{h-1}, t_{h+1}]$ and $a = 1, \ldots, \mathsf{r}$. Therefore, it suffices to show that there is a universal constant $C' > 0$ such that

$$\max_{1 \leq h \leq n-1} \sup_{t \in [t_{h-1}, t_{h+1}]} \left\| \sum_{a=1}^{\mathsf{r}} \left(\int_{t_{h-1}}^{t} E[D_s \xi_t^a | \mathcal{F}_s] \cdot dB_s \right)^2 \right\|_p$$

$$\leq C' \frac{p}{n} \sup_{0 \leq u \leq v \leq 1} \|D_u \xi_v\|_{2p,\ell_2}^2$$

for all $n \in \mathbb{N}$ and $p \in [2, \infty)$.

Fix $h = 1, \ldots, n-1$ and $t \in [t_{h-1}, t_{h+1}]$ arbitrarily. By Itô's formula we have

$$\sum_{a=1}^{\mathsf{r}} \left(\int_{t_{h-1}}^{\tau} E[D_s \xi_t^a | \mathcal{F}_s] \cdot dB_s \right)^2$$

$$= 2 \int_{t_{h-1}}^{\tau} \sum_{a=1}^{\mathsf{r}} \left(\int_{t_{h-1}}^{s} E[D_u \xi_t^a | \mathcal{F}_u] \cdot dB_u \right) E[D_s \xi_t^a | \mathcal{F}_s] \cdot dB_s$$

$$+ \int_{t_{h-1}}^{\tau} \sum_{a=1}^{\mathsf{r}} \| E[D_s \xi_t^a | \mathcal{F}_s] \|_{\ell_2}^2 \, ds =: \mathbf{I}_\tau + \mathbf{II}_\tau$$

for every $\tau \in [t_{h-1}, t]$. The Lyapunov inequality and Proposition B.1 yield

$$\| \mathbf{II}_\tau \|_p \le \frac{1}{n} \sup_{0 \le u \le v \le 1} \| D_u \xi_v \|_{2p,\ell_2}^2.$$

Meanwhile, we have

$$\| \mathbf{I}_\tau \|_p \lesssim \sqrt{p} \left\| \sqrt{ \sum_{b=1}^{r} \int_{t_{h-1}}^{\tau} \left| \sum_{a=1}^{\mathsf{r}} \left(\int_{t_{h-1}}^{s} E[D_u \xi_t^a | \mathcal{F}_u] \cdot dB_u \right) E[D_s^{(b)} \xi_t^a | \mathcal{F}_s] \right|^2 ds } \right\|_p$$

$$(\because \text{Proposition B.2})$$

$$\le \sqrt{p} \left\| \int_{t_{h-1}}^{\tau} \sum_{a=1}^{\mathsf{r}} \left(\int_{t_{h-1}}^{s} E[D_u \xi_t^a | \mathcal{F}_u] \cdot dB_u \right)^2 \sum_{a=1}^{r} \sum_{b=1}^{r} E[D_s^{(b)} \xi_t^a | \mathcal{F}_s]^2 ds \right\|_{p/2}^{1/2}$$

$$(\because \text{Schwarz})$$

$$\le \sqrt{ p \int_{t_{h-1}}^{\tau} \left\| \sum_{a=1}^{\mathsf{r}} \left(\int_{t_{h-1}}^{s} E[D_u \xi_t^a | \mathcal{F}_u] dB_u \right)^2 E \left[\| D_s \xi_t \|_{\ell_2}^2 | \mathcal{F}_s \right] \right\|_{p/2} ds }$$

$$(\because \text{Lyapunov, Proposition B.1})$$

$$\le \sqrt{ p \int_{t_{h-1}}^{\tau} \left\| \sum_{a=1}^{\mathsf{r}} \left(\int_{t_{h-1}}^{s} E[D_u \xi_t^a | \mathcal{F}_u] \cdot dB_u \right)^2 \right\|_p \left\| E \left[\| D_s \xi_t \|_{\ell_2}^2 | \mathcal{F}_s \right] \right\|_p ds }$$

$$(\because \text{Schwarz})$$

$$\le \sqrt{ p \sup_{0 \le u \le v \le 1} \| D_u \xi_v \|_{2p,\ell_2}^2 \int_{t_{h-1}}^{\tau} \left\| \sum_{a=1}^{\mathsf{r}} \left(\int_{t_{h-1}}^{s} E[D_u \xi_t^a | \mathcal{F}_u] \cdot dB_u \right)^2 \right\|_p ds }$$

$$(\because \text{Lyapunov}).$$

Therefore, defining the function $g : [t_{h-1}, t] \to [0, \infty)$ by

$$g(\tau) = \left\| \sum_{a=1}^{\mathsf{r}} \left(\int_{t_{h-1}}^{\tau} E[D_s \xi_t^a | \mathcal{F}_s] \cdot dB_s \right)^2 \right\|_p^2, \qquad \tau \in [t_{h-1}, t],$$

we obtain

$$g(\tau) \le \frac{2}{n^2} \sup_{0 \le u \le v \le 1} \| D_u \xi_v \|_{2p,\ell_2}^4 + C_0 p \sup_{0 \le u \le v \le 1} \| D_u \xi_v \|_{2p,\ell_2}^2 \int_{t_{h-1}}^{\tau} \sqrt{g(s)} ds$$

for any $\tau \in [t_{h-1}, t]$ with some universal constant $C_0 > 0$. Hence the Bihari inequality (cf. Section 3 of [8]) yields

$$\sqrt{g(t)} \leq \frac{\sqrt{2}}{n} \sup_{0 \leq u \leq v \leq 1} \|D_u \xi_v\|_{2p,\ell_2}^2 + \frac{C_0 p}{2n} \sup_{0 \leq u \leq v \leq 1} \|D_u \xi_v\|_{2p,\ell_2}^2.$$

This implies that the desired result holds true with the constant $C' = 1/\sqrt{2} + C_0/2$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Lemma B.4. *Under the assumptions of Theorem 4.1, it holds that $\|\varsigma_t^{i\cdot}\|_{2p,\ell_2} \leq \|\sigma_t^{i\cdot}\|_{2p,\ell_2} = \|\Sigma_t^{ii}\|_p^{1/2}$ for any $t \in [0,1]$, $i = 1, \ldots, d$ and $p \geq 1$.*

Proof. The last equality is evident from the identity $\|\sigma_t^{i\cdot}\|_{\ell_2}^2 = \Sigma_t^{ii}$. Meanwhile, the Lyapunov inequality yields

$$\|\varsigma_t^{i\cdot}\|_{\ell_2}^2 = \sum_{a=1}^r \left(E\left[\sigma_t^{ia}|\mathcal{G}_t^n\right] \right)^2 \leq \sum_{a=1}^r E\left[\left(\sigma_t^{ia}\right)^2 |\mathcal{G}_t^n \right] = E\left[\|\sigma_t^{i\cdot}\|_{\ell_2}^2 |\mathcal{G}_t^n \right].$$

Therefore, the Lyapunov inequality again yields

$$E\left[\|\varsigma_t^{i\cdot}\|_{\ell_2}^{2p} \right] \leq E\left[\left(E\left[\|\sigma_t^{i\cdot}\|_{\ell_2}^2 |\mathcal{G}_t^n \right] \right)^p \right] \leq E\left[\|\sigma_t^{i\cdot}\|_{\ell_2}^{2p} \right].$$

This means $\|\varsigma_t^{i\cdot}\|_{2p,\ell_2} \leq \|\sigma_t^{i\cdot}\|_{2p,\ell_2}$. $\qquad\qquad\qquad\qquad\qquad\qquad\square$

Lemma B.5. *Under the assumptions of Theorem 4.1, for all $i = 1, \ldots, d$ and $u \in [0,1]$, $\varsigma_u^{i\cdot} \in \mathbb{D}_{2,\infty}(\mathbb{R}^r)$ and $D_s \varsigma_u^{i\cdot}, D_{s,t} \varsigma_u^{i\cdot}$ are $\mathcal{G}_u^n$-measurable for any $s, t \in [0,1]$. Moreover, the following estimates hold true for any $p \in [1, \infty)$ and $s, t \in [0,1]$:*

$$\|D_t \varsigma_u^{i\cdot}\|_{2p,\ell_2} \leq \|D_t \sigma_u^{i\cdot}\|_{2p,\ell_2}, \tag{B.1}$$

$$\|D_{s,t} \varsigma_u^{i\cdot}\|_{2p,\ell_2} \leq \|D_{s,t} \sigma_u^{i\cdot}\|_{2p,\ell_2}, \tag{B.2}$$

$$\max_{1 \leq k \leq d} \left\| \sum_{a=1}^r \varsigma_s^{ka} D_s^{(a)} \varsigma_u^{i\cdot} \right\|_{p,\ell_2} \leq \max_{1 \leq k \leq d} \|\Sigma_u^{kk}\|_p^{1/2} \|D_s \sigma_u^{i\cdot}\|_{2p,\ell_2}, \tag{B.3}$$

$$\max_{1 \leq k,l \leq d} \left\| \sum_{a,b=1}^r \varsigma_s^{ka} \varsigma_t^{lb} D_{s,t}^{(a,b)} \varsigma_u^{i\cdot} \right\|_{p,\ell_2} \leq \max_{1 \leq k \leq d} \|\Sigma_s^{kk}\|_{\frac{3}{2}p} \|D_{s,t} \sigma_u^{i\cdot}\|_{3p,\ell_2}. \tag{B.4}$$

Proof. First, by Proposition 3.1 of [38] $\varsigma_u^{i\cdot} \in \mathbb{D}_{2,2}(\mathbb{R}^r)$ and we have

$$D_s^{(a)} \varsigma_u^{i\cdot} = E\left[D_s^{(a)} \sigma_u^{i\cdot} |\mathcal{G}_u^n \right] 1_{[0,(\lceil nu \rceil - 1)/n]}(s)$$

and

$$D_{s,t}^{(a,b)} \varsigma_u^{i\cdot} = E\left[D_{s,t}^{(a,b)} \sigma_u^{i\cdot} |\mathcal{G}_u^n \right] 1_{[0,(\lceil nu \rceil - 1)/n]^2}(s,t)$$

for any $s, t \in [0,1]$ and $a, b = 1, \ldots, r$. In particular, $D_s \varsigma_u^{i\cdot}, D_{s,t} \varsigma_u^{i\cdot}$ are $\mathcal{G}_u^n$-measurable. Moreover, (B.1)–(B.2) can be shown in an analogous way to the proof of Lemma B.4, which also implies that $\varsigma_u^{i\cdot} \in \mathbb{D}_{2,\infty}(\mathbb{R}^r)$.

Next, the Schwarz inequality, Lemma B.4 and (B.1) yield

$$
\left\| \sum_{a=1}^{r} \varsigma_s^{ka} D_s^{(a)} \varsigma_u^{i\cdot} \right\|_{p,\ell_2} = \left\| \sqrt{ \sum_{b=1}^{r} \left(\sum_{a=1}^{r} \varsigma_s^{ka} D_s^{(a)} \varsigma_u^{ib} \right)^2 } \right\|_p
$$
$$
\leq \left\| \|\varsigma_s^{k\cdot}\|_{\ell_2} \left\| D_s \varsigma_u^{i\cdot} \right\|_{\ell_2} \right\|_p \leq \left\| \Sigma_u^{kk} \right\|_p^{1/2} \left\| D_s \sigma_u^{i\cdot} \right\|_{2p,\ell_2},
$$

and thus we obtain (B.3).

Finally, the Schwarz inequality, Lemma B.4 and (B.2) yield

$$
\left\| \sum_{a,b=1}^{r} \varsigma_s^{ka} \varsigma_t^{lb} D_{s,t}^{(a,b)} \varsigma_u^{i\cdot} \right\|_{p,\ell_2} = \left\| \sqrt{ \sum_{c=1}^{r} \left(\sum_{a,b=1}^{r} \varsigma_s^{ka} \varsigma_t^{lb} D_{s,t}^{(a,b)} \varsigma_u^{ic} \right)^2 } \right\|_p
$$
$$
\leq \left\| \|\varsigma_s^{k\cdot}\|_{\ell_2} \|\varsigma_t^{l\cdot}\|_{\ell_2} \|D_{s,t}^2 \varsigma_u^{i\cdot}\|_{\ell_2} \right\|_p \leq \left\| \|\varsigma_s^{k\cdot}\|_{\ell_2} \|\varsigma_t^{l\cdot}\|_{\ell_2} \right\|_{\frac{3}{2}p} \|D_{s,t}^2 \varsigma_u^{i\cdot}\|_{3p,\ell_2}
$$
$$
\leq \left\| \Sigma_s^{kk} \right\|_{\frac{3}{2}p}^{1/2} \left\| \Sigma_t^{ll} \right\|_{\frac{3}{2}p}^{1/2} \|D_{s,t}^2 \sigma_u^{i\cdot}\|_{3p,\ell_2},
$$

so we obtain (B.4) and thus complete the proof. □

Now we turn to the main body of the proof. We begin by evaluating the approximation error between $\sqrt{n}(\widehat{[Y^i, Y^j]}_1^n - [Y^i, Y^j]_1)$ and M_n^{ij}.

Lemma B.6. *Under the assumptions of Theorem 4.1, it holds that*

$$
\max_{1 \leq i,j \leq d} \left\| \sqrt{n} \widehat{[A^i, A^j]}_1^n \right\|_p \leq \frac{1}{\sqrt{n}} \max_{1 \leq i \leq d} \sup_{0 \leq s \leq 1} \|\mu_s^i\|_{2p}^2
$$

for any $p \in [1, \infty)$ and $n \in \mathbb{N}$. Moreover, there is a universal constant $C > 0$ such that

$$
\left\| \sqrt{n} \widehat{[M^i, A^j]}_1^n \right\|_p
$$
$$
\leq \frac{C}{\sqrt{n}} \sup_{0 \leq s \leq 1} \|\sigma_s^{i\cdot}\|_{2p,\ell_2} \left(\sqrt{p} \sup_{0 \leq t \leq 1} \left\| \mu_t^j \right\|_{2p} + p \sup_{0 \leq u \leq v \leq 1} \|D_u \mu_v^j\|_{2p,\ell_2} \right)
$$

for any $p \in [2, \infty)$, $n \in \mathbb{N}$ and $i, j = 1, \ldots, d$.

Proof. The first claim is an immediate consequence of the Hölder inequality and Proposition B.1.

To prove the second claim, by Itô's formula we decompose $\sqrt{n} \widehat{[M^i, A^j]}_1^n$ as

$$
\sqrt{n} \widehat{[M^i, A^j]}_1^n = \sqrt{n} \sum_{h=1}^{n} \left\{ \int_{t_{h-1}}^{t_h} \left(\int_{t_{h-1}}^{t} \mu_s^j ds \right) \sigma_t^{i\cdot} \cdot dB_t \right.
$$
$$
\left. + \int_{t_{h-1}}^{t_h} \left(\int_{t_{h-1}}^{t} \sigma_s^{i\cdot} \cdot dB_s \right) \mu_t^j dt \right\}
$$

$$=: \mathbf{I}_n^{ij} + \mathbf{II}_n^{ij}.$$

By Propositions B.1–B.2 we have

$$\|\mathbf{I}_n^{ij}\|_p \lesssim \sqrt{np} \left\| \sqrt{\sum_{h=1}^n \int_{t_{h-1}}^{t_h} \left(\int_{t_{h-1}}^t \mu_s^j ds \right)^2 \|\sigma_t^{i\cdot}\|_{\ell_2}^2 dt} \right\|_p$$

$$\leq \sqrt{\frac{p}{n}} \sup_{0 \leq s \leq 1} \left\| \mu_s^j \right\|_{2p} \sup_{0 \leq t \leq 1} \|\sigma_t^{i\cdot}\|_{2p,\ell_2}.$$

In the meantime, we further decompose $\mathbf{II}_n^{ij}$ as

$$\mathbf{II}_n^{ij} = \sqrt{n} \sum_{h=1}^n \left\{ \int_{t_{h-1}}^{t_h} \left(\int_{t_{h-1}}^t \sigma_s^{i\cdot} \cdot dB_s \right) E\left[\mu_t^j | \mathcal{G}_t^n\right] dt \right.$$

$$\left. + \int_{t_{h-1}}^{t_h} \left(\int_{t_{h-1}}^t \sigma_s^{i\cdot} \cdot dB_s \right) \left(\mu_t^j - E\left[\mu_t^j | \mathcal{G}_t^n\right] \right) dt \right\}$$

$$=: \mathbf{II}_n^{ij}(1) + \mathbf{II}_n^{ij}(2).$$

Since $E\left[\mu_t^j | \mathcal{G}_t^n\right]$ is $\mathcal{F}_{t_{h-1}}$-measurable for $t \in I_h$, we have

$$\mathbf{II}_n^{ij}(1) = \sqrt{n} \sum_{h=1}^n \int_{t_{h-1}}^{t_h} \left(\int_{t_{h-1}}^{t_h} 1_{(t_{h-1},t]}(s) E\left[\mu_t^j | \mathcal{G}_t^n\right] \sigma_s^{i\cdot} \cdot dB_s \right) dt.$$

Therefore, the stochastic Fubini theorem (e.g. Corollary 5.28 of [50]) yields

$$\mathbf{II}_n^{ij}(1) = \sqrt{n} \sum_{h=1}^n \int_{t_{h-1}}^{t_h} \left(\int_{t_{h-1}}^{t_h} 1_{(t_{h-1},t]}(s) E\left[\mu_t^j | \mathcal{G}_t^n\right] dt \right) \sigma_s^{i\cdot} \cdot dB_s.$$

Hence, Propositions B.1–B.2 and the Lyapunov inequality imply that

$$\left\| \mathbf{II}_n^{ij}(1) \right\|_p \lesssim \sqrt{np} \left\| \sqrt{\sum_{h=1}^n \int_{t_{h-1}}^{t_h} \left(\int_{t_{h-1}}^{t_h} 1_{(t_{h-1},t]}(s) E\left[\mu_t^j | \mathcal{G}_t^n\right] dt \right)^2 \|\sigma_s^{i\cdot}\|_{\ell_2}^2 ds} \right\|_p$$

$$\leq \sqrt{\frac{p}{n}} \sup_{0 \leq t \leq 1} \|\mu_t^j\|_{2p} \sup_{0 \leq s \leq 1} \|\sigma_s^{i\cdot}\|_{2p,\ell_2}.$$

Meanwhile, Propositions B.1–B.2 and Lemma B.3 yield

$$\left\| \mathbf{II}_n^{ij}(2) \right\|_p \leq \sqrt{n} \sum_{h=1}^n \int_{t_{h-1}}^{t_h} \left\| \int_{t_{h-1}}^t \sigma_s^{i\cdot} \cdot dB_s \right\|_{2p} \left\| \mu_t^j - E\left[\mu_t^j | \mathcal{G}_t^n\right] \right\|_{2p} dt$$

$$\lesssim p \sum_{h=1}^n \int_{t_{h-1}}^{t_h} \left\| \sqrt{\int_{t_{h-1}}^t \|\sigma_s^{i\cdot}\|_{\ell_2}^2 ds} \right\|_{2p} dt \sup_{0 \leq u \leq v \leq 1} \|D_u \mu_v^j\|_{2p,\ell_2}$$

$$\leq \frac{p}{\sqrt{n}} \sup_{0 \leq s \leq 1} \|\sigma_s^{i\cdot}\|_{2p,\ell_2} \sup_{0 \leq u \leq v \leq 1} \|D_u \mu_v^j\|_{2p,\ell_2}.$$

Combining these estimates, we complete the proof. $\qquad\square$

Lemma B.7. *Under the assumptions of Theorem 4.1, there is a universal constant $C > 0$ such that*

$$\max_{1 \leq i,j \leq d} \left\| \sqrt{n} \left(\left[\widehat{\mathsf{M}^i, \mathsf{M}^j} \right]_1^n - [Y^i, Y^j]_1 \right) - M_n^{ij} \right\|_p$$

$$\leq C \frac{p^{3/2}}{\sqrt{n}} \max_{1 \leq i \leq d} \sup_{0 \leq s \leq 1} \left\| \Sigma_s^{ii} \right\|_p^{1/2} \max_{1 \leq j \leq d} \sup_{0 \leq u \leq v \leq 1} \| D_u \sigma_v^{j\cdot} \|_{2p,\ell_2}$$

for every $n \in \mathbb{N}$ and $p \in [2, \infty)$.

Proof. By Itô's formula we deduce the following decomposition:

$$\sqrt{n} \left(\left[\widehat{\mathsf{M}^i, \mathsf{M}^j} \right]_1^n - [Y^i, Y^j]_1 \right)$$

$$= \sqrt{n} \sum_{h=1}^n \left\{ \int_{t_{h-1}}^{t_h} \left(\int_{t_{h-1}}^t \sigma_s^{j\cdot} \cdot dB_s \right) \sigma_t^{i\cdot} \cdot dB_t + \int_{t_{h-1}}^{t_h} \left(\int_{t_{h-1}}^t \sigma_s^{i\cdot} \cdot dB_s \right) \sigma_t^{j\cdot} \cdot dB_t \right\}$$

$$=: \mathbf{I}_n^{ij} + \mathbf{II}_n^{ij}.$$

Propositions B.1–B.2 and Lemmas B.3–B.4 yield

$$\left\| \sqrt{n} \sum_{h=1}^n \int_{t_{h-1}}^{t_h} \left(\int_{t_{h-1}}^t \sigma_s^{j\cdot} \cdot dB_s \right) \left(\sigma_t^{i\cdot} - \varsigma_t^{i\cdot} \right) \cdot dB_t \right\|_p$$

$$\lesssim \sqrt{np} \left\| \sqrt{\sum_{h=1}^n \int_{t_{h-1}}^{t_h} \left(\int_{t_{h-1}}^t \sigma_s^{j\cdot} \cdot dB_s \right)^2 \left\| \sigma_t^{i\cdot} - \varsigma_t^{i\cdot} \right\|_{\ell_2}^2 dt} \right\|_p$$

$$\leq \sqrt{np} \sqrt{\sum_{h=1}^n \int_{t_{h-1}}^{t_h} \left\| \int_{t_{h-1}}^t \sigma_s^{j\cdot} \cdot dB_s \right\|_{2p}^2 \left\| \sigma_t^{i\cdot} - \varsigma_t^{i\cdot} \right\|_{2p,\ell_2}^2 dt}$$

$$\lesssim p \sqrt{\sup_{0 \leq s \leq 1} \left\| \Sigma_s^{jj} \right\|_p \sup_{0 \leq t \leq 1} \left\| \sigma_t^{i\cdot} - \varsigma_t^{i\cdot} \right\|_{2p,\ell_2}^2}$$

$$\lesssim \frac{p^{3/2}}{\sqrt{n}} \sup_{0 \leq s \leq 1} \left\| \Sigma_s^{jj} \right\|_p^{1/2} \sup_{0 \leq u \leq v \leq 1} \| D_u \sigma_v^{i\cdot} \|_{2p,\ell_2}$$

and

$$\left\| \sqrt{n} \sum_{h=1}^n \int_{t_{h-1}}^{t_h} \left(\int_{t_{h-1}}^t \left(\sigma_s^{j\cdot} - \varsigma_s^{j\cdot} \right) \cdot dB_s \right) \varsigma_t^{i\cdot} \cdot dB_t \right\|_p$$

$$\lesssim \sqrt{np} \left\| \sqrt{\sum_{h=1}^n \int_{t_{h-1}}^{t_h} \left(\int_{t_{h-1}}^t \left(\sigma_s^{j\cdot} - \varsigma_s^{j\cdot} \right) \cdot dB_s \right)^2 \| \varsigma_t^{i\cdot} \|_{\ell_2}^2 dt} \right\|_p$$

$$\lesssim \sqrt{n}p \sqrt{\sum_{h=1}^{n} \int_{t_{h-1}}^{t_h} \left\| \sqrt{\int_{t_{h-1}}^{t} \left\| \sigma_s^{j\cdot} - \varsigma_s^{j\cdot} \right\|_{\ell_2}^2 ds} \right\|_{2p}^2 \|\varsigma_t^{i\cdot}\|_{2p,\ell_2}^2 dt}$$

$$\lesssim \frac{p^{3/2}}{\sqrt{n}} \sup_{0 \le u \le v \le 1} \|D_u \sigma_v^{j\cdot}\|_{2p,\ell_2} \sup_{0 \le s \le 1} \left\| \Sigma_s^{ii} \right\|_p^{1/2}.$$

Hence we obtain

$$\left\| \mathbf{I}_n^{ij} - \sqrt{n} \sum_{h=1}^{n} \int_{t_{h-1}}^{t_h} \left(\int_{t_{h-1}}^{t} \varsigma_s^{j\cdot} \cdot dB_s \right) \varsigma_t^{i\cdot} \cdot dB_t \right\|_p$$

$$\lesssim \frac{p^{3/2}}{\sqrt{n}} \max_{1 \le i,j \le d} \sup_{0 \le u \le v \le 1} \|D_u \sigma_v^{j\cdot}\|_{2p,\ell_2} \sup_{0 \le s \le 1} \left\| \Sigma_s^{ii} \right\|_p^{1/2}.$$

Analogously we can prove

$$\left\| \mathbf{II}_n^{ij} - \sqrt{n} \sum_{h=1}^{n} \int_{t_{h-1}}^{t_h} \left(\int_{t_{h-1}}^{t} \varsigma_s^{i\cdot} \cdot dB_s \right) \varsigma_t^{j\cdot} \cdot dB_t \right\|_p$$

$$\lesssim \frac{p^{3/2}}{\sqrt{n}} \max_{1 \le i,j \le d} \sup_{0 \le u \le v \le 1} \|D_u \sigma_v^{j\cdot}\|_{2p,\ell_2} \sup_{0 \le s \le 1} \left\| \Sigma_s^{ii} \right\|_p^{1/2}.$$

Consequently, the desired result follows from Lemma B.1. □

Next we establish some properties of M_n^{ij} which are necessary for the application of our main theorem. The first result gives the moment bounds.

Lemma B.8. *Under the assumptions of Theorem 4.1, there is a universal constant $C > 0$ such that*

$$\max_{1 \le i,j \le d} \|M_n^{ij}\|_p \le Cp \max_{1 \le i \le d} \sup_{0 \le t \le 1} \|\Sigma_t^{ii}\|_p$$

for all $n \in \mathbb{N}$ and $p \in [2, \infty)$.

Proof. This is an immediate consequence of Lemmas B.1–B.2 and B.4. □

Second, we prove the Malliavin differentiability of M_n^{ij} and compute its Malliavin derivatives. For this purpose we prove an auxiliary result. Recall that we have $\mathbb{D}_{1,2}(H) \subset \mathrm{Dom}(\delta)$ by Proposition 1.3.1 of [55].

Lemma B.9. *Let $k \in \mathbb{N}$.*

(a) *Suppose that $u \in \mathbb{D}_{k,2}(H)$ satisfies $D_{t_1,\ldots,t_j}^{(a_1,\ldots,a_j)} u \in \mathrm{Dom}(\delta)$ for all $j = 1, \ldots, k$, $a_1, \ldots, a_j \in \{1, \ldots, r\}$ and $t_1, \ldots, t_j \in [0,1]$ and*

$$\sum_{j=1}^{k} E \left[\int_{[0,1]^j} \left\| \delta \left(D_{t_1,\ldots,t_j} u \right) \right\|_{\ell_2}^2 dt_1 \cdots dt_j \right] < \infty. \qquad (B.5)$$

Then we have $\delta(u) \in \mathbb{D}_{k,2}$ and

$$D_{t_1,\ldots,t_k}\delta(u) = \delta\left(D_{t_1,\ldots,t_k}u\right) + k\operatorname{Sym}\left(D^{k-1}u\right)(t_1,\ldots,t_k) \tag{B.6}$$

for all $t_1,\ldots,t_k \in [0,1]$.

(b) *If $u \in \mathbb{D}_{k,2}(H)$ is $\mathbf{F}$-adapted, then $\delta(u) \in \mathbb{D}_{k,2}$ and (B.6) holds true for all $t_1,\ldots,t_k \in [0,1]$.*

Proof. (a) We prove the claim by induction on k. When $k = 1$, the claim follows from Proposition 1.3.8 of [55]. Next, supposing that the claim holds true for $k = K \in \mathbb{N}$, we prove the claim for $k = K + 1$. From (B.5) we have

$$E\left[\int_0^1 \|\delta\left(D_t\left(D_{t_1,\ldots,t_K}u\right)\right)\|_{\ell_2}^2 \, dt\right] < \infty \tag{B.7}$$

for all $t_1,\ldots,t_K \in [0,1]$. Moreover, by the assumption of the induction (B.6) holds true for all $t_1,\ldots,t_K \in [0,1]$. Now let us take $t_1,\ldots,t_K \in [0,1]$ arbitrarily, and set $v := D_{t_1,\ldots,t_K}u$. Then, by assumptions and Proposition 1.3.8 of [55], $\delta(v) \in \mathbb{D}_{1,2}$ and $D_t\delta(v) = \delta(D_t v) + v(t)$ for all $t \in [0,1]$. Therefore, by (B.6) we have $D_{t_1,\ldots,t_K}\delta(u) \in \mathbb{D}_{1,2}$ and

$$
\begin{aligned}
D_t\left(D_{t_1,\ldots,t_K}\delta(u)\right) &= \delta(D_t v) + v(t) + K D_t \operatorname{Sym}\left(D^{K-1}u\right)(t_1,\ldots,t_K) \\
&= \delta\left(D_{t_1,\ldots,t_K,t}u\right) + (K+1)\operatorname{Sym}\left(D^K u\right)(t_1,\ldots,t_K,t)
\end{aligned}
$$

for all $t \in [0,1]$. This implies that the claim also holds true for $k = K + 1$ and thus completes the proof.

(b) This claim is an immediate consequence of claim (a) and Propositions 1.2.8 and 1.3.11 of [55]. $\qquad\square$

We then obtain the following result.

Lemma B.10. *Under the assumptions of Theorem 4.1, the following statements hold true for any $n \in \mathbb{N}$ and $i, j = 1,\ldots,d$:*

(a) *$f_n^{ij} \in \mathbb{D}_{2,\infty}(H^{\otimes 2})$ and it holds that*

$$D_s^{(a)} f_n^{ij}(u,v) = \frac{1}{2}(D_s^{(a)}\varsigma_u^{i\cdot} \otimes \varsigma_v^{j\cdot} + \varsigma_u^{i\cdot} \otimes D_s^{(a)}\varsigma_v^{j\cdot} + D_s^{(a)}\varsigma_u^{j\cdot} \otimes \varsigma_v^{i\cdot} + \varsigma_u^{j\cdot} \otimes D_s^{(a)}\varsigma_v^{i\cdot}) \tag{B.8}$$

and

$$
\begin{aligned}
&D_{s,t}^{(a,b)} f_n^{ij}(u,v) \\
&= \frac{1}{2}\left(D_{s,t}^{(a,b)}\varsigma_u^{i\cdot} \otimes \varsigma_v^{j\cdot} + D_s^{(a)}\varsigma_u^{i\cdot} \otimes D_t^{(b)}\varsigma_v^{j\cdot} + D_t^{(b)}\varsigma_u^{i\cdot} \otimes D_s^{(a)}\varsigma_v^{j\cdot} + \varsigma_u^{i\cdot} \otimes D_{s,t}^{(a,b)}\varsigma_v^{j\cdot} \right. \\
&\quad\left. + D_{s,t}^{(a,b)}\varsigma_u^{j\cdot} \otimes \varsigma_v^{i\cdot} + D_s^{(a)}\varsigma_u^{j\cdot} \otimes D_t^{(b)}\varsigma_v^{i\cdot} + D_t^{(b)}\varsigma_u^{j\cdot} \otimes D_s^{(a)}\varsigma_v^{i\cdot} + \varsigma_u^{j\cdot} \otimes D_{s,t}^{(a,b)}\varsigma_v^{i\cdot} \right)
\end{aligned}
\tag{B.9}
$$

for any $a, b = 1,\ldots,r$ and $s, t, u, v \in [0,1]$.

(b) $u_n^{ij} \in \mathbb{D}_{2,\infty}(H^{\otimes 2})$ *and* $u_n^{ij}(s,\cdot)^{a\cdot}, D_s^{(a)} u_n^{ij}(t,\cdot)^{b\cdot} \in \mathrm{Dom}(\delta),\ D_{s,t}^{(a,b)} u_n^{ij} \in$ $\mathrm{Dom}(\delta^2)$ *for any* $a,b = 1,\dots,r$ *and* $s,t \in [0,1]$.

(c) $M_n^{ij} \in \mathbb{D}_{2,\infty}$ *and we have*

$$D_s^{(a)} M_n^{ij} = \delta^2(D_s^{(a)} u_n^{ij}) + 2\delta(u_n^{ij}(s,\cdot)^{a\cdot})$$

and

$$D_{s,t}^{(a,b)} M_n^{ij} = \delta^2\left(D_{s,t}^{(a,b)} u_n^{ij}\right) + 2\delta\left(D_s^{(a)} u_n^{ij}(t,\cdot)^{b\cdot}\right)$$
$$+ 2\delta\left(D_t^{(b)} u_n^{ij}(s,\cdot)^{a\cdot}\right) + 2u_n^{ij}(s,t)^{ab}$$

for any $a,b = 1,\dots,r$ *and* $s,t \in [0,1]$.

Proof. Claim (a) follows from Lemma B.5 as well as Lemma 15.82 and Theorem 15.83 of [37]. Claim (b) is a consequence of claim (a), Proposition 1.3.11 of [55] and Lemmas B.1 and B.5.

Now we prove claim (c). By Lemma B.1 we can rewrite M_n^{ij} as

$$M_n^{ij} = 2\sqrt{n} \sum_{h=1}^{n} \int_{I_h} \delta(f_n^{ij}(\cdot,t)1_{I_h \cap [0,t]}) \cdot dB_t. \tag{B.10}$$

From claim (a), $f_n^{ij}(\cdot,t) \in \mathbb{D}_{2,\infty}(H)$ for all $t \in [0,1]$, so Lemma B.9(b) implies that $\delta(f_n^{ij}(\cdot,t)1_{I_h \cap [0,t]}) \in \mathbb{D}_{2,2}(\mathbb{R}^r)$ and

$$D_u^{(a)}\delta(f_n^{ij}(\cdot,t)1_{I_h \cap [0,t]}) = \delta(D_u^{(a)} f_n^{ij}(\cdot,t)1_{I_h \cap [0,t]}) + f_n^{ij}(u,t)^{a\cdot}1_{I_h \cap [0,t]}(u),$$
$$D_{u,v}^{(a,b)}\delta(f_n^{ij}(\cdot,t)1_{I_h \cap [0,t]}) =$$
$$\delta(D_{u,v}^{(a,b)} f_n^{ij}(\cdot,t)1_{I_h \cap [0,t]}) + D_v^{(b)} f_n^{ij}(u,t)^{a\cdot}1_{I_h \cap [0,t]}(u) + D_u^{(a)} f_n^{ij}(v,t)^{b\cdot}1_{I_h \cap [0,t]}(v)$$

for all $a,b = 1,\dots,r$ and $u,v \in [0,1]$. These formulae imply that the process $(\delta(f_n^{ij}(\cdot,t)))_{t \in [0,1]}$ belongs to $\mathbb{D}_{2,2}(H)$, so Lemma B.9(b) again implies that $\int_{I_h} \delta(f_n^{ij}(\cdot,t)1_{I_h \cap [0,t]}) \cdot dB_t \in \mathbb{D}_{2,2}$ and

$$D_u^{(a)}\left(\int_{I_h} \delta(f_n^{ij}(\cdot,t)1_{I_h \cap [0,t]}) \cdot dB_t\right)$$
$$= \int_{I_h} D_u^{(a)}\delta(f_n^{ij}(\cdot,t)1_{I_h \cap [0,t]}) \cdot dB_t + \delta(f_n^{ij}(\cdot,u)^{a\cdot}1_{I_h \cap [0,u]})1_{I_h}(u)$$
$$= \int_{I_h} \delta(D_u^{(a)} f_n^{ij}(\cdot,t)1_{I_h \cap [0,t]}) \cdot dB_t + \int_{I_h} f_n^{ij}(u,t)^{a\cdot}1_{I_h \cap [0,t]}(u) \cdot dB_t$$
$$+ \delta(f_n^{ij}(\cdot,u)^{a\cdot}1_{I_h \cap [0,u]})1_{I_h}(u)$$
$$= \int_{I_h}\left(\int_{t_{h-1}}^{t} D_u^{(a)} f_n^{ij}(s,t)dB_s\right) \cdot dB_t + \delta(f_n^{ij}(u,\cdot)^{a\cdot}1_{I_h \times I_h}(u,\cdot)) \tag{B.11}$$

and

$$D_{u,v}^{(a,b)}\left(\int_{I_h} \delta(f_n^{ij}(\cdot,t)1_{I_h \cap [0,t]}) \cdot dB_t\right)$$

$$
\begin{aligned}
&= \int_{I_h} D_{u,v}^{(a,b)} \delta(f_n^{ij}(\cdot,t) 1_{I_h \cap [0,t]}) \cdot dB_t + D_v^{(b)} \delta(f_n^{ij}(\cdot,u)^{\cdot a} 1_{I_h \cap [0,u]}) 1_{I_h}(u) \\
&\quad + D_u^{(a)} \delta(f_n^{ij}(\cdot,v)^{\cdot b} 1_{I_h \cap [0,v]}) 1_{I_h}(v) \\
&= \int_{I_h} \Big\{ \delta(D_{u,v}^{(a,b)} f_n^{ij}(\cdot,t) 1_{I_h \cap [0,t]}) + D_v^{(b)} f_n^{ij}(u,t)^{a \cdot} 1_{I_h \cap [0,t]}(u) \\
&\qquad + D_u^{(a)} f_n^{ij}(v,t)^{b \cdot} 1_{I_h \cap [0,t]}(v) \Big\} \cdot dB_t \\
&\quad + \Big\{ \delta(D_v^{(b)} f_n^{ij}(\cdot,u)^{\cdot a} 1_{I_h \cap [0,u]}) + f_n^{ij}(v,u)^{ba} 1_{I_h \cap [0,u]}(v) \Big\} 1_{I_h}(u) \\
&\quad + \Big\{ \delta(D_u^{(a)} f_n^{ij}(\cdot,v)^{\cdot b} 1_{I_h \cap [0,v]}) + f_n^{ij}(u,v)^{ab} 1_{I_h \cap [0,v]}(u) \Big\} 1_{I_h}(v) \\
&= \int_{I_h} \Bigg\{ \int_{t_{h-1}}^t D_{u,v}^{(a,b)} f_n^{ij}(s,t) dB_s + D_v^{(b)} f_n^{ij}(u,t)^{a \cdot} 1_{I_h \cap [0,t]}(u) \\
&\qquad + D_u^{(a)} f_n^{ij}(v,t)^{b \cdot} 1_{I_h \cap [0,t]}(v) \Bigg\} \cdot dB_t \\
&\quad + \delta(D_v^{(b)} f_n^{ij}(\cdot,u)^{\cdot a} 1_{I_h \cap [0,u]}) 1_{I_h}(u) + \delta(D_u^{(a)} f_n^{ij}(\cdot,v)^{\cdot b} 1_{I_h \cap [0,v]}) 1_{I_h}(v) \\
&\quad + f_n^{ij}(u,v)^{ab} 1_{I_h \times I_h}(u,v) \\
&= \int_{I_h} \Bigg\{ \int_{t_{h-1}}^t D_{u,v}^{(a,b)} f_n^{ij}(s,t) dB_s \Bigg\} \cdot dB_t + \delta(D_v^{(b)} f_n^{ij}(u,\cdot)^{a \cdot} 1_{I_h \times I_h}(u,\cdot)) \\
&\quad + \delta(D_u^{(a)} f_n^{ij}(v,\cdot)^{b \cdot} 1_{I_h \times I_h}(v,\cdot)) + f_n^{ij}(u,v)^{ab} 1_{I_h \times I_h}(u,v) \qquad \text{(B.12)}
\end{aligned}
$$

for all $a,b = 1,\ldots,r$ and $u,v \in [0,1]$. Now, noting that formulae (B.8)–(B.9) can be rewritten as

$$
D_s^{(a)} f_n^{ij}(u,v) = \mathrm{Sym}\left(D_s^{(a)} \varsigma^{i \cdot} \otimes \varsigma^{j \cdot} \right)(u,v) + \mathrm{Sym}\left(\varsigma^{i \cdot} \otimes D_s^{(a)} \varsigma^{j \cdot} \right)(u,v)
$$

and

$$
\begin{aligned}
D_{s,t}^{(a,b)} f_n^{ij}(u,v) &= \mathrm{Sym}\left(D_{s,t}^{(a,b)} \varsigma^{i \cdot} \otimes \varsigma^{j \cdot} \right)(u,v) + \mathrm{Sym}\left(D_s^{(a)} \varsigma^{i \cdot} \otimes D_t^{(b)} \varsigma^{j \cdot} \right)(u,v) \\
&\quad + \mathrm{Sym}\left(D_t^{(b)} \varsigma^{i \cdot} \otimes D_s^{(a)} \varsigma^{j \cdot} \right)(u,v) + \mathrm{Sym}\left(\varsigma_u^{i \cdot} \otimes D_{s,t}^{(a,b)} \varsigma_v^{j \cdot} \right)(u,v),
\end{aligned}
$$

by Lemma B.1 we obtain

$$
2 \int_{I_h} \left(\int_{t_{h-1}}^t D_u^{(a)} f_n^{ij}(s,t) dB_s \right) \cdot dB_t = \delta^2 \left(D_u^{(a)} f_n^{ij} \right), \qquad \text{(B.13)}
$$

$$
2 \int_{I_h} \left\{ \int_{t_{h-1}}^t D_{u,v}^{(a,b)} f_n^{ij}(s,t) dB_s \right\} \cdot dB_t = \delta^2 \left(D_{u,v}^{(a,b)} f_n^{ij} \right). \qquad \text{(B.14)}
$$

Now claim (c) follows from (B.10)–(B.14) and the assumptions of the lemma. $\quad\square$

Third, we prove the Malliavin differentiability of Σ and $\mathfrak{C}_n$ as well as establish moment estimates for their Malliavin derivatives.

Lemma B.11. *Under the assumptions of Theorem 4.1, for any $i, j = 1, \ldots, d$ and $t \in [0,1]$, $\Sigma_t^{ij} \in \mathbb{D}_{2,\infty}$ and*

$$\|D_u \Sigma_t^{ij}\|_{p,\ell_2} \leq 2 \max_{1 \leq i,j \leq d} \left\|\Sigma_t^{ii}\right\|_p^{1/2} \|D_u \sigma_t^{j\cdot}\|_{2p,\ell_2},$$

$$\|D_{u,v} \Sigma_t^{ij}\|_{p,\ell_2} \leq 2 \Bigg(\max_{1 \leq i,j \leq d} \left\|\Sigma_t^{ii}\right\|_p^{1/2} \|D_{u,v} \sigma_t^{j\cdot}\|_{2p,\ell_2}$$

$$+ \max_{1 \leq i,j \leq d} \|D_u \sigma_t^{i\cdot}\|_{2p,\ell_2} \|D_v \sigma_t^{j\cdot}\|_{2p,\ell_2} \Bigg)$$

for any $p \in [1,\infty)$ and $u, v \in [0,1]$.

Proof. Since $\Sigma_t^{ij} = \sum_{a=1}^r \sigma_t^{ia} \sigma_t^{ja}$, Theorem 15.78 of [37] implies that $\Sigma_t^{ij} \in \mathbb{D}_{2,\infty}$ for all $t \in [0,1]$ and

$$D\Sigma_t^{ij} = \sum_{a=1}^r \left(\sigma_t^{ja} D\sigma_t^{ia} + \sigma_t^{ia} D\sigma_t^{ja} \right)$$

and

$$D^2 \Sigma_t^{ij} = \sum_{a=1}^r \left(\sigma_t^{ja} D^2 \sigma_t^{ia} + D\sigma_t^{ja} \otimes D\sigma_t^{ia} + D\sigma_t^{ia} \otimes D\sigma_t^{ja} + \sigma_t^{ia} D^2 \sigma_t^{ja} \right).$$

In particular, we have

$$\|D_u \Sigma_t^{ij}\|_{\ell_2} \leq \sum_{a=1}^r \left(|\sigma_t^{ja}| \|D\sigma_t^{ia}\|_{\ell_2} + |\sigma_t^{ia}| \|D\sigma_t^{ja}\|_{\ell_2} \right)$$

$$\leq \sqrt{\Sigma_t^{jj}} \|D\sigma_t^{i\cdot}\|_{\ell_2} + \sqrt{\Sigma_t^{ii}} \|D\sigma_t^{j\cdot}\|_{\ell_2}$$

and

$$\|D_{u,v} \Sigma_t^{ij}\|_{\ell_2}$$
$$\leq \sum_{a=1}^r \left(|\sigma_t^{ja}| \|D_{u,v} \sigma_t^{ia}\|_{\ell_2} + \|D_u \sigma_t^{ja}\|_{\ell_2} \|D_v \sigma_t^{ia}\|_{\ell_2} + \|D_u \sigma_t^{ia}\|_{\ell_2} \|D_v \sigma_t^{ja}\|_{\ell_2} \right.$$
$$\left. + |\sigma_t^{ia}| \|D_{u,v} \sigma_t^{ja}\|_{\ell_2} \right)$$
$$\leq \sqrt{\Sigma_t^{jj}} \|D_{u,v} \sigma_t^{i\cdot}\|_{\ell_2} + \|D_u \sigma_t^{j\cdot}\|_{\ell_2} \|D_v \sigma_t^{i\cdot}\|_{\ell_2} + \|D_u \sigma_t^{i\cdot}\|_{\ell_2} \|D_v \sigma_t^{j\cdot}\|_{\ell_2}$$
$$+ \sqrt{\Sigma_t^{ii}} \|D_{u,v} \sigma_t^{j\cdot}\|_{\ell_2}$$

by the triangular and Schwarz inequalities. Hence we complete the proof by the Hölder inequality. □

Lemma B.12. *Under the assumptions of Theorem 4.1, $\mathfrak{C}_n \in \mathbb{D}_{2,\infty}(\mathbb{R}^{d^2 \times d^2})$ and*

$$\|D_u \mathfrak{C}_n^{(i-1)d+j,(k-1)d+l}\|_{p,\ell_2} \leq 8 \max_{1 \leq i,j \leq d} \sup_{0 \leq s,t \leq 1} \|\Sigma_s^{ii}\|_{2p}^{3/2} \|D_u \sigma_t^{j\cdot}\|_{4p,\ell_2},$$

$$\left\| D_{u,v}\mathfrak{C}_n^{(i-1)d+j,(k-1)d+l} \right\|_{p,\ell_2}$$

$$\leq 8 \max_{1\leq i,j,k\leq d} \sup_{0\leq s,t\leq 1} \|\Sigma_s^{ii}\|_{2p}^{3/2} \|D_{u,v}\sigma_t^{j\cdot}\|_{4p,\ell_2}$$

$$+ 24 \max_{1\leq i,j,k\leq d} \sup_{0\leq s,t,\tau\leq 1} \|\Sigma_s^{ii}\|_{2p} \|D_u\sigma_t^{j\cdot}\|_{4p,\ell_2} \|D_v\sigma_\tau^{k\cdot}\|_{4p,\ell_2}$$

for any $p\in[1,\infty)$, $n\in\mathbb{N}$, $i,j,k,l=1,\ldots,d$ *and* $u,v\in[0,1]$.

Proof. By Remark 15.87 of [37], $\int_{I_h}\Sigma_t^{ij}dt\in\mathbb{D}_{2,\infty}$ and

$$D\left(\int_{I_h}\Sigma_t^{ij}dt\right)=\int_{I_h}D\Sigma_t^{ij}dt, \qquad D^2\left(\int_{I_h}\Sigma_t^{ij}dt\right)=\int_{I_h}D^2\Sigma_t^{ij}dt$$

for any $i,j=1,\ldots,d$ and $h=1,\ldots,n$. Therefore, by Theorem 15.78 of [37], the Schwarz inequality and Proposition B.1 we obtain $\mathfrak{C}_n^{(i-1)d+j,(k-1)d+l}\in\mathbb{D}_{2,\infty}$ and

$$\|D_u\mathfrak{C}_n^{(i-1)d+j,(k-1)d+l}\|_{p,\ell_2}\leq 4\max_{1\leq i,j,k,l\leq d}\sup_{0\leq s,t\leq 1}\|\Sigma_s^{ik}\|_{2p}\|D_u\Sigma_t^{kl}\|_{2p,\ell_2},$$

$$\|D_{u,v}\mathfrak{C}_n^{(i-1)d+j,(k-1)d+l}\|_{p,\ell_2}$$

$$\leq 4\max_{1\leq i,j,k,l\leq d}\sup_{0\leq s,t\leq 1}\left(\|\Sigma_s^{ik}\|_{2p}\|D_{u,v}^2\Sigma_t^{kl}\|_{2p,\ell_2}+\|D_u\Sigma_s^{ik}\|_{2p,\ell_2}\|D_v\Sigma_t^{jl}\|_{2p,\ell_2}\right)$$

for any $p\geq 1$, $i,j,k,l=1,\ldots,d$ and $u,v\in[0,1]$. Now the desired result follows from the Schwarz inequality and Lemma B.11. $\qquad\square$

Now we proceed to checking the conditions of Theorem 3.1 in the current setting.

Lemma B.13. *Under the assumptions of Theorem 4.1, there is a universal constant* $C>0$ *such that*

$$\max_{1\leq i,j,k,l\leq d}\left\|\left\langle D^2 M_n^{ij},u_n^{kl}\right\rangle_{H^{\otimes 2}}-\mathfrak{C}_n^{(i-1)d+j,(k-1)d+l}\right\|_p$$

$$\leq C\frac{p}{\sqrt{n}}\left(\max_{1\leq i,j\leq d}\sup_{0\leq s\leq 1}\left\|\Sigma_s^{jj}\right\|_{2p}^{3/2}\sup_{0\leq s,t,u\leq 1}\left\|D_{s,t}\sigma_u^{i\cdot}\right\|_{4p,\ell_2}\right.$$

$$\left.+\max_{1\leq i,k\leq d}\sup_{0\leq s,u\leq 1}\left\|\Sigma_u^{kk}\right\|_{2p}\left\|D_s\sigma_u^{i\cdot}\right\|_{4p,\ell_2}^2\right)$$

$$+C\sqrt{\frac{p}{n}}\max_{1\leq i,j\leq d}\sup_{0\leq s,t,u\leq 1}\left\|D_s\sigma_u^{i\cdot}\right\|_{4p,\ell_2}\left\|\Sigma_t^{jj}\right\|_{2p}^{3/2}$$

for any $n\in\mathbb{N}$ *and* $p\in[2,\infty)$.

Proof. By Lemma B.10 the desired result follows once we verify the following statements for all $p\in[2,\infty)$ (note (A.7)):

$$\left\|\langle\delta^2(D^2 u_n^{ij}),u_n^{kl}\rangle_{H^{\otimes 2}}\right\|_p$$

$$\lesssim \frac{p}{\sqrt{n}} \left(\max_{1 \le i,j \le d} \sup_{0 \le s \le 1} \left\| \Sigma_s^{jj} \right\|_{2p}^{3/2} \sup_{0 \le s,t,u \le 1} \left\| D_{s,t} \sigma_u^{i\cdot} \right\|_{4p,\ell_2} \right.$$

$$\left. + \max_{1 \le i,k \le d} \sup_{0 \le s,u \le 1} \left\| \Sigma_u^{kk} \right\|_{2p} \left\| D_s \sigma_u^{i\cdot} \right\|_{4p,\ell_2}^2 \right), \quad \text{(B.15)}$$

$$\left\| \langle \delta(D u_n^{ij}), u_n^{kl} \rangle_{H^{\otimes 2}} \right\|_p \lesssim \sqrt{\frac{p}{n}} \max_{1 \le i,j \le d} \sup_{0 \le s,t,u \le 1} \left\| D_s \sigma_u^{i\cdot} \right\|_{4p,\ell_2} \| \Sigma_t^{jj} \|_{2p}^{3/2}, \quad \text{(B.16)}$$

$$\left\| 2 \langle u_n^{ij}, u_n^{kl} \rangle_{H^{\otimes 2}} - \mathfrak{C}_n^{(i-1)d+j,(k-1)d+l} \right\|_p$$

$$\lesssim \sqrt{\frac{p}{n}} \max_{1 \le i \le d} \sup_{0 \le u \le v \le 1} \| D_u \sigma_v^{i\cdot} \|_{4p,\ell_2} \max_{1 \le j \le d} \sup_{0 \le t \le 1} \| \Sigma_t^{jj} \|_{2p}^{3/2}. \quad \text{(B.17)}$$

We first verify (B.15). We can rewrite $\langle \delta^2(D^2 u_n^{ij}), u_n^{kl} \rangle_{H^{\otimes 2}}$ as

$$\langle \delta^2(D^2 u_n^{ij}), u_n^{kl} \rangle_{H^{\otimes 2}} = n \sum_{h,h'=1}^{n} \sum_{a,b=1}^{r} \int_{I_h \times I_h} \delta^2(D_{s,t}^{(a,b)} f_n^{ij} 1_{I_{h'}} \times 1_{I_{h'}}) f_n^{kl}(s,t)^{ab} ds dt$$

$$= n \sum_{h,h'=1}^{n} \sum_{a,b=1}^{r} \int_{I_h \times I_h} \delta^2 \left(D_{s,t}^{(a,b)} f_n^{ij} 1_{I_{h'}} \times 1_{I_{h'}} \right) \varsigma_s^{ka} \varsigma_t^{lb} ds dt,$$

where the last identity holds true because $\langle \varphi, \mathrm{Sym}(\psi) \rangle_{H^{\otimes 2}} = \langle \varphi, \psi \rangle_{H^{\otimes 2}}$ for any $\varphi, \psi \in H^{\otimes 2}$ if φ is symmetric. Then, noting that $f_n^{ij}(u,v)$ is $\mathcal{F}_{t_{h'-1}}$-measurable when $u, v \in I_{h'}$ by construction, Corollary 1.2.1 of [55] yields

$$\langle \delta^2(D^2 u_n^{ij}), u_n^{kl} \rangle_{H^{\otimes 2}} = n \sum_{h=1}^{n} \sum_{h':h'>h} \sum_{a,b=1}^{r} \int_{I_h \times I_h} \delta^2(D_{s,t}^{(a,b)} f_n^{ij} 1_{I_{h'}} \times 1_{I_{h'}}) \varsigma_s^{ka} \varsigma_t^{lb} ds dt.$$

Moreover, since ς_u is $\mathcal{F}_{t_{h-1}}$-measurable when $u \in I_h$, using Lemma B.1 and Exercise 2.30 in Chapter 3 of [39] repeatedly, we obtain

$$\langle \delta^2(D^2 u_n^{ij}), u_n^{kl} \rangle_{H^{\otimes 2}}$$

$$= n \sum_{h=1}^{n} \int_{I_h \times I_h} \sum_{h':h'>h} \delta^2 \left(\sum_{a,b=1}^{r} \varsigma_s^{ka} \varsigma_t^{lb} D_{s,t}^{(a,b)} f_n^{ij} 1_{I_{h'}} \times 1_{I_{h'}} \right) ds dt.$$

Hence Proposition B.1 yields

$$\left\| \langle \delta^2(D^2 u_n^{ij}), u_n^{kl} \rangle_{H^{\otimes 2}} \right\|_p$$

$$\le n \sum_{h=1}^{n} \int_{I_h \times I_h} \left\| \sum_{h':h'>h} \delta^2 \left(\sum_{a,b=1}^{r} \varsigma_s^{ka} \varsigma_t^{lb} D_{s,t}^{(a,b)} f_n^{ij} 1_{I_{h'}} \times 1_{I_{h'}} \right) \right\|_p ds dt.$$

Now, from (B.9) we infer that

$$\sum_{a,b=1}^{r} \varsigma_s^{ka} \varsigma_t^{lb} D_{s,t}^{(a,b)} f_n^{ij}(u,v)$$

$$
= \frac{1}{2} \left\{ \left(\sum_{a,b=1}^{r} \varsigma_s^{ka} \varsigma_t^{lb} D_{s,t}^{(a,b)} \varsigma_u^{i\cdot} \right) \otimes \varsigma_v^{j\cdot} + \left(\sum_{a=1}^{r} \varsigma_s^{ka} D_s^{(a)} \varsigma_u^{i\cdot} \right) \otimes \left(\sum_{b=1}^{r} \varsigma_t^{lb} D_t^{(b)} \varsigma_v^{j\cdot} \right) \right.
$$

$$
+ \left(\sum_{b=1}^{r} \varsigma_t^{lb} D_t^{(b)} \varsigma_u^{i\cdot} \right) \otimes \left(\sum_{a=1}^{r} \varsigma_s^{ka} D_s^{(a)} \varsigma_v^{j\cdot} \right) + \varsigma_u^{i\cdot} \otimes \left(\sum_{a,b=1}^{r} \varsigma_s^{ka} \varsigma_t^{lb} D_{s,t}^{(a,b)} \varsigma_v^{j\cdot} \right)
$$

$$
+ \left(\sum_{a,b=1}^{r} \varsigma_s^{ka} \varsigma_t^{lb} D_{s,t}^{(a,b)} \varsigma_u^{j\cdot} \right) \otimes \varsigma_v^{i\cdot} + \left(\sum_{a=1}^{r} \varsigma_s^{ka} D_s^{(a)} \varsigma_u^{j\cdot} \right) \otimes \left(\sum_{b=1}^{r} \varsigma_t^{lb} D_t^{(b)} \varsigma_v^{i\cdot} \right)
$$

$$
+ \left. \left(\sum_{b=1}^{r} \varsigma_t^{lb} D_t^{(b)} \varsigma_u^{j\cdot} \right) \otimes \left(\sum_{a=1}^{r} \varsigma_s^{ka} D_s^{(a)} \varsigma_v^{i\cdot} \right) + \varsigma_u^{j\cdot} \otimes \left(\sum_{a,b=1}^{r} \varsigma_s^{ka} \varsigma_t^{lb} D_{s,t}^{(a,b)} \varsigma_v^{i\cdot} \right) \right\} .
$$

Hence Lemmas B.2 and B.4–B.5 yield

$$
\sup_{0 \le s,t \le 1} \left\| \sum_{h':h'>h} \delta^2 \left(\sum_{a,b=1}^{r} \varsigma_s^{ka} \varsigma_t^{lb} D_{s,t}^{(a,b)} f_n^{ij} 1_{I_{h'}} \times 1_{I_{h'}} \right) \right\|_p
$$

$$
\lesssim \frac{p}{\sqrt{n}} \left(\sup_{0 \le s,t,u,v \le 1} \left\| \sum_{a,b=1}^{r} \varsigma_s^{ka} \varsigma_t^{lb} D_{s,t}^{(a,b)} \varsigma_u^{i\cdot} \right\|_{\frac{4}{3}p, \ell_2} \left\| \varsigma_v^{j\cdot} \right\|_{4p, \ell_2} \right.
$$

$$
+ \sup_{0 \le s,t,u,v \le 1} \left\| \sum_{a=1}^{r} \varsigma_s^{ka} D_s^{(a)} \varsigma_u^{i\cdot} \right\|_{2p, \ell_2} \left\| \sum_{b=1}^{r} \varsigma_t^{lb} D_t^{(b)} \varsigma_v^{j\cdot} \right\|_{2p, \ell_2}
$$

$$
+ \left. \sup_{0 \le s,t,u,v \le 1} \left\| \varsigma_u^{i\cdot} \right\|_{4p, \ell_2} \left\| \sum_{a,b=1}^{r} \varsigma_s^{ka} \varsigma_t^{lb} D_{s,t}^{(a,b)} \varsigma_v^{j\cdot} \right\|_{\frac{4}{3}p, \ell_2} \right)
$$

$$
\lesssim \frac{p}{\sqrt{n}} \left(\max_{1 \le i,j \le d} \sup_{0 \le s \le 1} \left\| \Sigma_s^{jj} \right\|_{2p}^{3/2} \sup_{0 \le s,t,u \le 1} \left\| D_{s,t} \sigma_u^{i\cdot} \right\|_{4p, \ell_2} \right.
$$

$$
+ \left. \max_{1 \le i,k \le d} \sup_{0 \le s,u \le 1} \left\| \Sigma_u^{kk} \right\|_{2p} \left\| D_s \sigma_u^{i\cdot} \right\|_{4p, \ell_2}^2 \right) .
$$

Therefore, we obtain (B.15).

Next we verify (B.16). We have

$$
\langle \delta(D u_n^{ij}), u_n^{kl} \rangle_{H^{\otimes 2}} = n \sum_{h=1}^{n} \sum_{a=1}^{r} \int_{I_h \times I_h} \delta(D_s^{(a)} f_n^{ij}(\cdot, t) 1_{I_h}(\cdot)) \cdot f_n^{kl}(s,t)^{a\cdot} ds dt .
$$

Since $f_n^{kl}(s,t)$ is $\mathcal{F}_{t_{h-1}}$-measurable when $s,t \in I_h$, by Proposition 1.3.11 of [55] and Exercise 2.30 in Chapter 3 of [39] we obtain

$$
\langle \delta(D u_n^{ij}), u_n^{kl} \rangle_{H^{\otimes 2}}
$$

$$= n \sum_{h=1}^{n} \int_{I_h \times I_h} \left(\int_{I_h} \left(\sum_{a,b=1}^{r} f_n^{kl}(s,t)^{ab} D_s^{(a)} f_n^{ij}(u,t)^{b\cdot} \right) \cdot dB_u \right) ds dt.$$

Therefore, Propositions B.1–B.2 and the Schwarz inequality yield

$$\left\| \langle \delta(D u_n^{ij}), u_n^{kl} \rangle_{H^{\otimes 2}} \right\|_p$$

$$\lesssim \sqrt{p} n \sum_{h=1}^{n} \int_{I_h \times I_h} \left\| \sqrt{\int_{I_h} \sum_{c=1}^{r} \left(\sum_{a,b=1}^{r} f_n^{kl}(s,t)^{ab} D_s^{(a)} f_n^{ij}(u,t)^{bc} \right)^2 du} \right\|_p ds dt$$

$$\leq \sqrt{p} n \sum_{h=1}^{n} \int_{I_h \times I_h} \left\| \int_{I_h} \| f_n^{kl}(s,t) \|_{\ell_2}^2 \| D_s f_n^{ij}(u,t) \|_{\ell_2}^2 du \right\|_{p/2}^{1/2} ds dt$$

$$\leq \sqrt{p} n \sum_{h=1}^{n} \int_{I_h \times I_h} \sqrt{\int_{I_h} \| f_n^{kl}(s,t) \|_{2p,\ell_2}^2 \| D_s f_n^{ij}(u,t) \|_{2p,\ell_2}^2 du} \, ds dt$$

$$\leq \sqrt{\frac{p}{n}} \sup_{0 \leq s,t,u \leq 1} \| f_n^{kl}(s,t) \|_{2p,\ell_2} \| D_s f_n^{ij}(u,t) \|_{2p,\ell_2}.$$

Since we have $\| f_n^{kl}(s,t) \|_{2p,\ell_2} \leq \sqrt{\| \Sigma_s^{kk} \|_{2p} \| \Sigma_t^{ll} \|_{2p}}$ and

$$\| D_s f_n^{ij}(u,t) \|_{2p,\ell_2} \leq 2 \max_{1 \leq i,j \leq d} \sup_{0 \leq s,t,u \leq 1} \| D_s \sigma_u^{i\cdot} \|_{4p,\ell_2} \| \Sigma_t^{jj} \|_{2p}^{1/2}$$

by the Schwarz inequality and Lemmas B.4–B.5, we obtain (B.16).

Finally we verify (B.17). We can rewrite $2 \langle u_n^{ij}, u_n^{kl} \rangle_{H^{\otimes 2}}$ as

$$2 \langle u_n^{ij}, u_n^{kl} \rangle_{H^{\otimes 2}}$$

$$= n \sum_{h=1}^{n} \sum_{a,b=1}^{r} \int_{I_h \times I_h} \left(\varsigma_s^{ia} \varsigma_t^{jb} \varsigma_s^{ka} \varsigma_t^{lb} + \varsigma_s^{ia} \varsigma_t^{jb} \varsigma_t^{kb} \varsigma_s^{la} \right) ds dt$$

$$= n \sum_{h=1}^{n} \left\{ \left(\int_{I_h} \varsigma_s^{i\cdot} \cdot \varsigma_s^{k\cdot} ds \right) \left(\int_{I_h} \varsigma_s^{j\cdot} \cdot \varsigma_s^{l\cdot} ds \right) + \left(\int_{I_h} \varsigma_s^{i\cdot} \cdot \varsigma_s^{l\cdot} ds \right) \left(\int_{I_h} \varsigma_s^{j\cdot} \cdot \varsigma_s^{k\cdot} ds \right) \right\}.$$

Note that we have

$$\left\| n \sum_{h=1}^{n} \left(\int_{I_h} (\varsigma_s^{i\cdot} - \sigma_s^{i\cdot}) \cdot \varsigma_s^{k\cdot} ds \right) \left(\int_{I_h} \varsigma_s^{j\cdot} \cdot \varsigma_s^{l\cdot} ds \right) \right\|_p$$

$$\leq n \sum_{h=1}^{n} \left(\int_{I_h} \| \varsigma_s^{i\cdot} - \sigma_s^{i\cdot} \|_{4p,\ell_2} \| \varsigma_s^{k\cdot} \|_{4p,\ell_2} ds \right) \left(\int_{I_h} \| \varsigma_s^{j\cdot} \|_{4p,\ell_2} \| \varsigma_s^{l\cdot} \|_{4p,\ell_2} ds \right)$$

$$\lesssim \sqrt{\frac{p}{n}} \sup_{0 \leq u \leq v \leq 1} \| D_u \sigma_v^{i\cdot} \|_{4p,\ell_2} \max_{1 \leq j \leq d} \sup_{0 \leq t \leq 1} \| \Sigma_t^{jj} \|_{2p}^{3/2}$$

for any $i,j,k,l = 1, \ldots, d$ by Proposition B.1 and Lemmas B.3 and B.4. Therefore, we obtain (B.17) and complete the proof of the lemma. $\qquad\square$

Lemma B.14. *Under the assumptions of Theorem 4.1, we have*

$$\left\|\langle D^2 F, u_{ij}^n\rangle_{H^{\otimes 2}}\right\|_p \le \frac{1}{\sqrt{n}} \sup_{0 \le s,t \le 1} \|D_{s,t}F\|_{2p,\ell_2} \max_{1 \le i \le d} \sup_{0 \le s \le 1} \left\|\Sigma_s^{ii}\right\|_{2p}$$

for any $p \in [1,\infty)$, $n \in \mathbb{N}$ and $F \in \mathbb{D}_{2,\infty}$.

Proof. Since $D^2 F$ is symmetric, (A.7) yields

$$\langle D^2 F, u_{ij}^n\rangle_{H^{\otimes 2}} = \sqrt{n} \sum_{h=1}^n \langle D^2 F, \varsigma_s^{i\cdot} \otimes \varsigma_t^{j\cdot} 1_{I_h \times I_h}\rangle_{H^{\otimes 2}}$$

$$= \sqrt{n} \sum_{h=1}^n \int_{I_h \times I_h} D_{s,t}F \cdot \varsigma_s^{i\cdot} \otimes \varsigma_t^{j\cdot} \, ds\, dt.$$

Therefore, by Proposition B.1, the Schwarz inequality and Lemma B.4 we have

$$\left\|\langle D^2 F, u_{ij}^n\rangle_{H^{\otimes 2}}\right\|_p \le \sqrt{n} \sum_{h=1}^n \int_{I_h \times I_h} \left\| D_{s,t}F \cdot \varsigma_s^{i\cdot} \otimes \varsigma_t^{j\cdot} \right\|_p ds\, dt$$

$$\le \sqrt{n} \sum_{h=1}^n \int_{I_h \times I_h} \|D_{s,t}F\|_{2p,\ell_2} \left\|\varsigma_s^{i\cdot}\right\|_{4p,\ell_2} \left\|\varsigma_t^{j\cdot}\right\|_{4p,\ell_2} ds\, dt$$

$$= \frac{1}{\sqrt{n}} \sup_{0 \le s,t \le 1} \|D_{s,t}F\|_{2p,\ell_2} \max_{1 \le i \le d} \sup_{0 \le s \le 1} \left\|\Sigma_s^{ii}\right\|_{2p}.$$

This completes the proof. $\qquad\square$

Lemma B.15. *Under the assumptions of Theorem 4.1, there is a universal constant $C > 0$ such that*

$$\max_{1 \le i,j,k \le d} \|D_s M_n^{ij} \cdot \varsigma_s^{k\cdot}\|_p \le C\left(p \max_{1 \le i \le d} \sup_{0 \le s,t \le 1} \left\|D_s \sigma_t^{i\cdot}\right\|_{3p,\ell_2} \max_{1 \le j \le d} \sup_{0 \le t \le 1} \|\Sigma_t^{jj}\|_{\frac{3}{2}p} \right.$$
$$\left. + \sqrt{p} \max_{1 \le i \le d} \sup_{0 \le t \le 1} \left\|\Sigma_t^{ii}\right\|_{\frac{3}{2}p}^{3/2} \right)$$

for any $p \in [2,\infty)$, $n \in \mathbb{N}$ and $s \in (0,1]$.

Proof. By Lemma B.10 the desired result follows once we verify the following statements for all $p \in [2,\infty)$:

$$\|\delta^2(D_s u_n^{ij}) \cdot \varsigma_s^{k\cdot}\|_p \lesssim p \max_{1 \le i \le d} \sup_{0 \le s,t \le 1} \left\|D_s \sigma_t^{i\cdot}\right\|_{3p,\ell_2} \max_{1 \le j \le d} \sup_{0 \le t \le 1} \|\Sigma_t^{jj}\|_{\frac{3}{2}p}, \quad \text{(B.18)}$$

$$\left\|\delta(u_n^{ij}(s,\cdot)) \cdot \varsigma_s^{k\cdot}\right\|_p \lesssim \sqrt{p} \max_{1 \le i \le d} \sup_{0 \le t \le 1} \left\|\Sigma_t^{ii}\right\|_{\frac{3}{2}p}^{3/2}. \quad \text{(B.19)}$$

Let h be the unique integer such that $s \in I_h$. First we verify (B.18). Since $f_n^{ij}(u,v)$ is $\mathcal{F}_{t_{h'-1}}$-measurable when $u,v \in I_{h'}$, by Corollary 1.2.1 of [55] we have

$$\delta^2(D_s^{(a)} u_n^{ij}) = \sqrt{n} \sum_{h'=1}^n \delta^2(D_s^{(a)} f_n^{ij} 1_{I_{h'} \times I_{h'}}) = \sqrt{n} \sum_{h':h'>h} \delta^2(D_s^{(a)} f_n^{ij} 1_{I_{h'} \times I_{h'}})$$

for any $a = 1, \ldots, r$. Hence Lemmas B.1 and B.5 as well as Exercise 2.30 in Chapter 3 of [39] yield

$$\delta^2(D_s u_n^{ij}) \cdot \varsigma_s^{k\cdot} = \sqrt{n} \sum_{h':h'>h} \delta^2 \left(\sum_{a=1}^r \varsigma_s^{ka} D_s^{(a)} f_n^{ij} 1_{I_{h'} \times I_{h'}} \right).$$

Since (B.8) implies that

$$\sum_{a=1}^r \varsigma_s^{ka} D_s^{(a)} f_n^{ij}$$
$$= \frac{1}{2} \left\{ \left(\sum_{a=1}^r \varsigma_s^{ka} D_s^{(a)} \varsigma^{i\cdot} \right) \otimes \varsigma^{j\cdot} + \varsigma^{i\cdot} \otimes \left(\sum_{a=1}^r \varsigma_s^{ka} D_s^{(a)} \varsigma^{j\cdot} \right) \right.$$
$$\left. + \left(\sum_{a=1}^r \varsigma_s^{ka} D_s^{(a)} \varsigma^{j\cdot} \right) \otimes \varsigma^{i\cdot} + \varsigma^{j\cdot} \otimes \left(\sum_{a=1}^r \varsigma_s^{ka} D_s^{(a)} \varsigma^{i\cdot} \right) \right\},$$

Lemmas B.1–B.2 and B.4–B.5 imply that

$$\|\delta^2(D_s u_n^{ij}) \cdot \varsigma_s^{k\cdot}\|_p$$
$$\lesssim p \max_{1 \leq i,k \leq d} \sup_{0 \leq t \leq 1} \left\| \sum_{a=1}^r \varsigma_s^{ka} D_s^{(a)} \varsigma_t^{i\cdot} \right\|_{\frac{3}{2}p,\ell_2} \max_{1 \leq j \leq d} \sup_{0 \leq t \leq 1} \|\varsigma_t^{j\cdot}\|_{3p,\ell_2}$$
$$\leq p \max_{1 \leq i,k \leq d} \sup_{0 \leq t \leq 1} \|\Sigma_s^{kk}\|_{\frac{3}{2}p}^{1/2} \|D_s \sigma_t^{i\cdot}\|_{3p,\ell_2} \max_{1 \leq j \leq d} \sup_{0 \leq t \leq 1} \|\Sigma_t^{jj}\|_{\frac{3}{2}p,\ell_2}^{1/2}$$
$$\leq p \max_{1 \leq i \leq d} \sup_{0 \leq s,t \leq 1} \|D_s \sigma_t^{i\cdot}\|_{3p,\ell_2} \max_{1 \leq j \leq d} \sup_{0 \leq t \leq 1} \|\Sigma_t^{jj}\|_{\frac{3}{2}p}.$$

Next we verify (B.19). Proposition 1.3.11 of [55] yields

$$\delta(u_n^{ij}(s,\cdot)) = \frac{\sqrt{n}}{2} \sum_{h'=1}^n 1_{I_{h'}}(s) \int_{I_{h'}} \left(\varsigma_s^{i\cdot} \otimes \varsigma_t^{j\cdot} + \varsigma_s^{j\cdot} \otimes \varsigma_t^{i\cdot} \right) dB_t$$
$$= \frac{\sqrt{n}}{2} \int_{I_h} \left(\varsigma_s^{i\cdot} \otimes \varsigma_t^{j\cdot} + \varsigma_s^{j\cdot} \otimes \varsigma_t^{i\cdot} \right) dB_t,$$

so we obtain

$$\delta(u_n^{ij}(s,\cdot)) \cdot \varsigma_s^{k\cdot} = \frac{\sqrt{n}}{2} \int_{I_h} \left\{ (\varsigma_s^{k\cdot} \cdot \varsigma_s^{i\cdot}) \varsigma_t^{j\cdot} + (\varsigma_s^{k\cdot} \cdot \varsigma_s^{j\cdot}) \varsigma_t^{i\cdot} \right\} dB_t.$$

Therefore, Propositions B.1–B.2, the Schwarz inequality and Lemma B.4 yield

$$\left\| \delta(u_n^{ij}(s,\cdot)) \cdot \varsigma_s^{k\cdot} \right\|_p \lesssim \sqrt{np} \left\| \sqrt{\int_{I_h} \left\| (\varsigma_s^{k\cdot} \cdot \varsigma_s^{i\cdot}) \varsigma_t^{j\cdot} + \left(\varsigma_s^{k\cdot} \cdot \varsigma_s^{j\cdot} \right) \varsigma_t^{i\cdot} \right\|_{\ell_2}^2 dt} \right\|_p$$
$$\leq \sqrt{np \int_{I_h} \left\| (\varsigma_s^{k\cdot} \cdot \varsigma_s^{i\cdot}) \varsigma_t^{j\cdot} + \left(\varsigma_s^{k\cdot} \cdot \varsigma_s^{j\cdot} \right) \varsigma_t^{i\cdot} \right\|_{p,\ell_2}^2 dt}$$

$$\leq 2 \max_{1 \leq i,j,k \leq d} \sqrt{np \int_{I_h} \|\varsigma_s^{k\cdot} \cdot \varsigma_s^{i\cdot}\|_{\frac{3}{2}p}^2 \left\|\varsigma_t^{j\cdot}\right\|_{3p,\ell_2}^2 dt}$$

$$\leq 2\sqrt{p} \max_{1 \leq i \leq d} \sup_{0 \leq t \leq 1} \left\|\varsigma_t^{i\cdot}\right\|_{3p,\ell_2}^3 \leq 2\sqrt{p} \max_{1 \leq i \leq d} \sup_{0 \leq t \leq 1} \left\|\Sigma_t^{ii}\right\|_{\frac{3}{2}p}^{3/2}.$$

This completes the proof. $\qquad\square$

Lemma B.16. *Under the assumptions of Theorem 4.1, there is a universal constant $C > 0$ such that*

$$\left\|\langle DM_n^{ij} \otimes DM_n^{i'j'}, u_n^{kl}\rangle_{H^{\otimes}}\right\|_p$$

$$\leq \frac{C}{\sqrt{n}} \left(p \max_{1 \leq i \leq d} \sup_{0 \leq s,t \leq 1} \left\|D_s \sigma_t^{i\cdot}\right\|_{6p,\ell_2} \max_{1 \leq j \leq d} \sup_{0 \leq t \leq 1} \|\Sigma_t^{jj}\|_{3p} \right.$$

$$\left. + \sqrt{p} \max_{1 \leq i \leq d} \sup_{0 \leq t \leq 1} \left\|\Sigma_t^{ii}\right\|_{3p}^{3/2} \right)^2,$$

$$\left\|\langle DM_n^{ij} \otimes DF, u_n^{kl}\rangle_{H^{\otimes}}\right\|_p$$

$$\leq \frac{C}{\sqrt{n}} \left(p \max_{1 \leq i \leq d} \sup_{0 \leq s,t \leq 1} \left\|D_s \sigma_t^{i\cdot}\right\|_{6p,\ell_2} \max_{1 \leq j \leq d} \sup_{0 \leq t \leq 1} \|\Sigma_t^{jj}\|_{3p}^{3/2} \right.$$

$$\left. + \sqrt{p} \max_{1 \leq i \leq d} \sup_{0 \leq t \leq 1} \left\|\Sigma_t^{ii}\right\|_{3p}^2 \right) \sup_{0 \leq s \leq 1} \|D_s F\|_{3p,\ell_2},$$

$$\left\|\langle DF \otimes DG, u_n^{kl}\rangle_{H^{\otimes}}\right\|_p$$

$$\leq \frac{C}{\sqrt{n}} \max_{1 \leq i \leq d} \sup_{0 \leq t \leq 1} \left\|\Sigma_t^{ii}\right\|_{3p} \sup_{0 \leq s \leq 1} \|D_s F\|_{3p,\ell_2} \sup_{0 \leq s \leq 1} \|D_s G\|_{3p,\ell_2}$$

for all $n \in \mathbb{N}$, $i,j,i',j',k,l = 1, \ldots, d$, $F, G \in \mathbb{D}_{1,\infty}$ and $p \in [2,\infty)$.

Proof. For any H-valued random variables ξ, η, we have

$$\langle \xi \otimes \eta, u_n^{ij}\rangle_{H^{\otimes 2}}$$

$$= \frac{\sqrt{n}}{2} \sum_{h=1}^n \left\{ \left(\int_{I_h} \xi_s \cdot \varsigma_s^{i\cdot} ds \right) \left(\int_{I_h} \eta_s \cdot \varsigma_s^{j\cdot} ds \right) + \left(\int_{I_h} \xi_s \cdot \varsigma_s^{j\cdot} ds \right) \left(\int_{I_h} \eta_s \cdot \varsigma_s^{i\cdot} ds \right) \right\}.$$

Hence, Proposition B.1 yields

$$\left\|\langle \xi \otimes \eta, u_n^{ij}\rangle_{H^{\otimes 2}}\right\|_p \leq \sqrt{n} \max_{1 \leq i,j \leq d} \sum_{h=1}^n \left(\int_{I_h} \|\xi_s \cdot \varsigma_s^{i\cdot}\|_{2p} ds \right) \left(\int_{I_h} \|\eta_s \cdot \varsigma_s^{j\cdot}\|_{2p} ds \right).$$

Now the desired result follows from the Schwarz inequality and Lemmas B.4 and B.15. $\qquad\square$

Proof of Theorem 4.1. Set $\mathfrak{S}_n := \mathfrak{C}_n^{1/2}\zeta_n$. By the hypercontractivity of Gaussian variables, we have

$$E\left[|\mathfrak{S}_n^k|^p \mid \mathcal{F}\right] \leq \left(\sqrt{(p-1)\mathfrak{C}_n^{kk}} \right)^p$$

for any $k = 1, \ldots, d^2$ and $p \in [2, \infty)$. Hence we obtain

$$\max_{1 \le k \le d^2} \left\| \mathfrak{S}_n^k \right\|_p \le \sqrt{p-1} \left\| \mathfrak{C}_n^{kk} \right\|_{p/2}^{1/2} \le 2\sqrt{p-1} \max_{1 \le i \le d} \sup_{0 \le t \le 1} \|\Sigma_t^{ii}\|_p \qquad (\text{B.20})$$

for any $p \in [2, \infty)$ by the Hölder inequality and Proposition B.1.

   We turn to the main body. First we prove claim (a). Note that we have $\| \|\xi\|_{\ell_\infty} \|_p \le k^{1/p} \max_{1 \le i \le k} \|\xi^i\|_p$ for any $p \ge 1$, $k \in \mathbb{N}$ and k-dimensional random vector ξ. Then, thanks to Lemmas 3.1, B.6 and B.7, it suffices to prove

$$\lim_{n \to \infty} \sup_{y \in \mathbb{R}^m} |P\left(\Xi_n \left(M_n + W_n\right) \le y\right) - P(\Xi_n(\mathfrak{S}_n + W_n) \le y)| = 0.$$

To prove this equation, we apply Theorem 3.1. For this purpose we need to verify conditions (3.7)–(3.17). (3.7) follows from (4.1). (3.8) follows from Lemma B.13. (3.9) follows from Lemmas B.12 and B.14. (3.10) follows from Lemma B.14. (3.11) follows from Lemmas B.8, B.14 and (B.20). (3.12)–(3.13) follow from Lemmas B.12 and B.16. (3.14) follows from Lemmas B.8, B.16 and (B.20). (3.15) follows from Lemmas B.12 and B.16. (3.16)–(3.17) follow from Lemmas B.8, B.12, B.16 and (B.20). So we complete the proof of claim (a).

   Next, if a k-dimensional random vector ξ satisfy $\max_{1 \le i \le k} \|\xi^i\|_p \le Ap^{r/2}$ for any $p \in \mathbb{N}$ with some constants $A > 0$ and $r \in \mathbb{N}$, then Lemma A.7 and Proposition A.1 of [45] imply that $\| \|\xi\|_{\ell_\infty} \|_p \le A \log^{r/2}(2k - 1 + e^{pr/2-1})$ for any $p > 0$ with $pr \ge 2$. Using this fact, we can prove claim (b) in the same way as the proof of claim (a). $\square$

B.2. Proof of Theorem 4.2

For every $\nu \in \mathbb{N}$, define the process $Y(\nu) = (Y(\nu)_t)_{t \in [0,1]}$ by

$$Y(\nu)_t = Y_0 + \int_0^t \mu(\nu)_s ds + \int_0^t \sigma(\nu)_s dB_s, \qquad t \in [0,1].$$

By the local property of Itô integrals (cf. pages 17–18 of [55]) we have $Y_t = Y(\nu)_t$ on $\Omega_n(\nu)$ for all $t \in [0,1]$. Therefore, setting

$$S_n(\nu) := \text{vec} \left[\sqrt{n} \left([\widehat{Y(\nu), Y(\nu)}]_1^n - [Y(\nu), Y(\nu)]_1 \right) \right],$$

we obtain

$$\rho_n(\nu) :=$$
$$\sup_{y \in \mathbb{R}^m} \left| P\left(\Xi_n(\nu)\left(S_n(\nu) + W_n(\nu)\right) \le y\right) - P(\Xi_n(\nu)(\mathfrak{C}_n(\nu)^{1/2}\zeta_n + W_n(\nu)) \le y) \right| \to 0$$

as $n \to \infty$ by Theorem 4.1. Now, for every $y \in \mathbb{R}^m$, we have

$$P\left(\Xi_n \left(S_n + W_n\right) \le y\right) \le P\left(\Xi_n(\nu)\left(S_n(\nu) + W_n(\nu)\right) \le y\right) + P(\Omega_n(\nu)^c)$$

$$\leq P(\Xi_n(\nu)(\mathfrak{C}_n(\nu)^{1/2}\zeta_n+W_n(\nu)) \leq y)+\rho_n(\nu)+P(\Omega_n(\nu)^c)$$
$$\leq P(\Xi_n(\mathfrak{C}_n^{1/2}\zeta_n+W_n) \leq y)+\rho_n(\nu)+2P(\Omega_n(\nu)^c).$$

By an analogous argument we also have

$$P\left(\Xi_n\left(S_n+W_n\right) \leq y\right) \geq P(\Xi_n(\mathfrak{C}_n^{1/2}\zeta_n+W_n) \leq y)-\rho_n(\nu)-2P(\Omega_n(\nu)^c).$$

Consequently, we obtain

$$\limsup_{n\to\infty} \sup_{y\in\mathbb{R}^m} \left|P\left(\Xi_n\left(S_n+W_n\right) \leq y\right) - P(\Xi_n(\mathfrak{C}_n^{1/2}\zeta_n+W_n) \leq y)\right|$$
$$\leq 2\limsup_{n\to\infty} P(\Omega_n(\nu)^c).$$

Letting $\nu \to \infty$, we complete the proof. $\square$

B.3. Proof of Proposition 4.1

We introduce some notation. Given a process $\xi = (\xi_t)_{t\in[0,1]}$ and an interval $I = (S,T] \subset [0,1]$, we set

$$\xi(I) := \xi_T - \xi_S, \qquad \xi(I)_t := \xi_{t\wedge T} - \xi_{t\wedge S} \quad (t \in [0,1]).$$

Also, we define

$$L(h)^{ij} := \mathsf{M}^i(I_h)\mathsf{M}^j(I_h) - [\mathsf{M}^i,\mathsf{M}^j](I_h)$$

and

$$L(h)_t^{ij} := \mathsf{M}^i(I_h)_t\mathsf{M}^j(I_h)_t - [\mathsf{M}^i,\mathsf{M}^j](I_h)_t \quad (t \in [0,1])$$

for $i,j = 1,\ldots,d$ and $h = 1,\ldots,n$.

Next we remark that a localization procedure allows us to reduce the situation of the proposition to the case that $\mu = \mu(\nu)$ and $\sigma = \sigma(\nu)$ for all $n,\nu \in \mathbb{N}$:

Lemma B.17. *Suppose that the statement of Proposition 4.1 holds true when we additionally assume $\mu = \mu(\nu)$ and $\sigma = \sigma(\nu)$ for all $n,\nu \in \mathbb{N}$. Then the original statement of Proposition 4.1 holds true as well.*

The proof of Lemma B.17 is analogous to the one of Theorem 4.2, so we omit it.

Lemma B.18. *There is a universal constant $C > 0$ such that*

$$\max_{1\leq i,j,k,l\leq d} \sup_{1\leq h\leq n} n^2 \left\|\mathsf{A}^i(I_h)\mathsf{X}^j(I_h)\mathsf{Y}^k(I_h)\mathsf{Z}^l(I_h)\right\|_p$$
$$\leq \frac{C}{\sqrt{n}} \max_{1\leq i,j\leq d} \sup_{0\leq,ts\leq 1} \left\|\mu_s^{ii}\right\|_{4p} \left(\left\|\mu_t^j\right\|_{4p}^3 + p^{3/2} \left\|\Sigma_t^{jj}\right\|_{2p}^{3/2}\right)$$

for any $\mathsf{X},\mathsf{Y},\mathsf{Z} \in \{\mathsf{A},\mathsf{M}\}$, $p \in [2,\infty)$ and $n \in \mathbb{N}$.

Proof. This is an immediate consequence of the Hölder inequality and Propositions B.1–B.2. $\square$

Lemma B.19. *There is a universal constant $C > 0$ such that*

$$\max_{1 \le i,j,k,l \le d} \left\| n \sum_{h=1}^{n} \int_{t_{h-1}}^{t_h} \mathsf{M}^j(I_h)_s \mathsf{M}^k(I_h)_s \mathsf{M}^l(I_h)_s d\mathsf{M}^i_s \right\|_p \le C \frac{p^2}{\sqrt{n}} \max_{1 \le i \le d} \sup_{0 \le s \le 1} \left\| \Sigma^{ii}_s \right\|^2_{2p}$$

for all $p \in [2, \infty)$ and $n \in \mathbb{N}$.

Proof. Propositions B.1–B.2 yield

$$\left\| n \sum_{h=1}^{n} \int_{t_{h-1}}^{t_h} \mathsf{M}^j(I_h)_s \mathsf{M}^k(I_h)_s \mathsf{M}^l(I_h)_s d\mathsf{M}^i_s \right\|_p$$

$$\lesssim n\sqrt{p} \left\| \sqrt{\sum_{h=1}^{n} \int_{t_{h-1}}^{t_h} \mathsf{M}^j(I_h)^2_s \mathsf{M}^k(I_h)^2_s \mathsf{M}^l(I_h)^2_s d[\mathsf{M}^i, \mathsf{M}^i]_s} \right\|_p$$

$$\le n \sqrt{p \sum_{h=1}^{n} \int_{t_{h-1}}^{t_h} \left\| \mathsf{M}^j(I_h)_s \right\|^2_{4p} \left\| \mathsf{M}^k(I_h)_s \right\|^2_{4p} \left\| \mathsf{M}^l(I_h)_s \right\|^2_{4p} \left\| \Sigma^{ii}_s \right\|_{2p} ds}$$

$$\lesssim np^2 \sqrt{\sum_{h=1}^{n} \int_{t_{h-1}}^{t_h} \left\| \sqrt{\int_{t_{h-1}}^{s} \Sigma^{jj}_u du} \right\|^2_{4p} \left\| \sqrt{\int_{t_{h-1}}^{s} \Sigma^{kk}_u du} \right\|^2_{4p} \left\| \sqrt{\int_{t_{h-1}}^{s} \Sigma^{ll}_u du} \right\|^2_{4p} \left\| \Sigma^{ii}_s \right\|_{2p} ds}$$

$$\le \frac{p^2}{\sqrt{n}} \max_{1 \le i \le d} \sup_{0 \le s \le 1} \left\| \Sigma^{ii}_s \right\|^2_{2p}.$$

This completes the proof. $\qquad\qquad\square$

Lemma B.20. *Suppose that the assumptions of of Proposition 4.1 hold true under the additional assumption that $\mu = \mu(\nu)$ and $\sigma = \sigma(\nu)$ for all $n, \nu \in \mathbb{N}$. Then there is a universal constant $C > 0$ such that*

$$\max_{1 \le i,j,k,l \le d} \left\| n \sum_{h=1}^{n-\nu} L(h)^{ij} [\mathsf{M}^k, \mathsf{M}^l](I_{h+\nu}) \right\|_p$$

$$\le \frac{C}{\sqrt{n}} \left(p \max_{1 \le i \le d} \sup_{0 \le t \le 1} \| \Sigma^{ii}_t \|^2_{2p} \right.$$

$$\left. + p^{3/2} \max_{1 \le i \le d} \sup_{0 \le t \le 1} \| \Sigma^{ii}_t \|^{3/2}_{2p} \max_{1 \le k \le d} \sup_{0 \le u \le v \le 1} \| D_u \sigma^{k\cdot}_v \|_{4p, \ell_2} \right)$$

for all $p \in [2, \infty)$, $n \in \mathbb{N}$ and $\nu \in \{0, 1\}$.

Proof. We decompose the target quantity as

$$n \sum_{h=1}^{n-\nu} L(h)^{ij} [\mathsf{M}^k, \mathsf{M}^l](I_{h+\nu})$$

$$= n \sum_{h=1}^{n-\nu} L(h)^{ij} \int_{I_{h+\nu}} E\left[\Sigma^{kl}_t | \mathcal{F}_{t_{h-1}} \right] dt + n \sum_{h=1}^{n-\nu} L(h)^{ij} \int_{I_{h+\nu}} \left(\Sigma^{kl}_t - E\left[\Sigma^{kl}_t | \mathcal{F}_{t_{h-1}} \right] \right) dt$$

$$=: \mathbf{I}_n + \mathbf{II}_n.$$

First we consider $\mathbf{I}_n$. Set

$$\phi_h := \int_{I_{h+\nu}} E\left[\Sigma_t^{kl}|\mathcal{F}_{t_{h-1}}\right] dt, \qquad h = 1, \ldots, n - \nu.$$

Then we can rewrite $\mathbf{I}_n$ as

$$\mathbf{I}_n = n \sum_{h=1}^{n-\nu} \left\{ \int_{I_h} \left(\int_{t_{h-1}}^t \sigma_s^{i\cdot} \cdot dB_s \right) \phi_h \sigma^{j\cdot} \cdot dB_t + \int_{I_h} \left(\int_{t_{h-1}}^t \phi_h \sigma_s^{j\cdot} \cdot dB_s \right) \sigma^{i\cdot} \cdot dB_t \right\}.$$

Therefore, Lemmas B.4–B.2 and the Hölder inequality imply that

$$\|\mathbf{I}_n\|_p \lesssim p\sqrt{n} \sup_{0 \le s \le 1} \|\sigma_s^{i\cdot}\|_{4p,\ell_2} \sup_{0 \le s < 1-\nu/n} \|\phi_{\lceil ns\rceil}\sigma_s^{j\cdot}\|_{\frac{4}{3}p,\ell_2}$$

$$\le p\sqrt{n} \sup_{0 \le s \le 1} \|\Sigma_s^{ii}\|_{2p}^{1/2} \sup_{0 \le s < 1-\nu/n} \|\phi_{\lceil ns\rceil}\|_{2p}\|\Sigma_s^{jj}\|_{2p}^{1/2}.$$

Now, Proposition B.1 and the Lyapunov and Schwarz inequalities yield

$$\sup_{0 \le s < 1-\nu/n} \|\phi_{\lceil ns\rceil}\|_{2p} \le \frac{1}{n} \sup_{0 \le t \le 1} \|\Sigma_t^{kl}\|_{2p} \le \frac{1}{n} \max_{1 \le k \le d} \sup_{0 \le t \le 1} \|\Sigma_t^{kk}\|_{2p}.$$

Consequently, we obtain

$$\|\mathbf{I}_n\|_p \lesssim \frac{p}{\sqrt{n}} \max_{1 \le i \le d} \sup_{0 \le t \le 1} \|\Sigma_t^{ii}\|_{2p}^2.$$

Next we consider $\mathbf{II}_n$. By Proposition B.1 we have

$$\|\mathbf{II}_n\|_p \le n \sum_{h=1}^{n-\nu} \|L(h)^{ij}\|_{2p} \int_{I_{h+\nu}} \left\| \Sigma_t^{kl} - E\left[\Sigma_t^{kl}|\mathcal{F}_{t_{h-1}}\right] \right\|_{2p} dt.$$

Itô's formula, Lemmas B.2 and B.4 yield

$$\|L(h)^{ij}\|_{2p} \lesssim \frac{p}{n} \max_{1 \le i \le d} \sup_{0 \le t \le 1} \|\Sigma_t^{ii}\|_{2p}.$$

Meanwhile, Lemmas B.3 and B.11 yield

$$\left\| \Sigma_t^{kl} - E\left[\Sigma_t^{kl}|\mathcal{F}_{t_{h-1}}\right] \right\|_{2p} \lesssim \sqrt{\frac{p}{n}} \sup_{0 \le u \le v \le 1} \|D_u\Sigma_v^{kl}\|_{2p,\ell_2}$$

$$\lesssim \sqrt{\frac{p}{n}} \max_{1 \le k,l \le d} \sup_{0 \le u \le v \le 1} \|\Sigma_v^{ll}\|_{2p}^{1/2}\|D_u\sigma_v^{k\cdot}\|_{4p,\ell_2}.$$

Consequently, we obtain

$$\|\mathbf{II}_n\|_p \lesssim \sqrt{\frac{p^3}{n}} \max_{1 \le i \le d} \sup_{0 \le t \le 1} \|\Sigma_t^{ii}\|_{2p}^{3/2} \max_{1 \le k \le d} \sup_{0 \le u \le v \le 1} \|D_u\sigma_v^{k\cdot}\|_{4p,\ell_2}.$$

We thus complete the proof. □

Lemma B.21. *Under the assumptions of of Lemma B.20, there is a universal constant $C > 0$ such that*

$$\max_{1 \leq i,j,k,l \leq d} \left\| n \sum_{h=1}^{n} \int_{t_{h-1}}^{t_h} L(h)_t^{kl} d[\mathsf{M}^i, \mathsf{M}^j]_t \right\|_p$$

$$\leq \frac{C}{\sqrt{n}} \left(p \max_{1 \leq i \leq d} \sup_{0 \leq t \leq 1} \|\Sigma_t^{ii}\|_{2p}^2 \right.$$

$$\left. + p^{3/2} \max_{1 \leq i \leq d} \sup_{0 \leq t \leq 1} \|\Sigma_t^{ii}\|_{2p}^{3/2} \max_{1 \leq k \leq d} \sup_{0 \leq u \leq v \leq 1} \|D_u \sigma_v^{k\cdot}\|_{4p,\ell_2} \right)$$

for all $p \in [2, \infty)$ and $n \in \mathbb{N}$.

Proof. By Itô's formula we can rewrite the target quantity as

$$n \sum_{h=1}^{n} \int_{t_{h-1}}^{t_h} L(h)_t^{kl} d[\mathsf{M}^i, \mathsf{M}^j]_t$$

$$= n \sum_{h=1}^{n} L(h)^{kl} [\mathsf{M}^i, \mathsf{M}^j](I_h) + n \sum_{h=1}^{n} \int_{t_{h-1}}^{t_h} [\mathsf{M}^i, \mathsf{M}^j](I_h)_t dL(h)_t^{kl}.$$

Since we have

$$n \sum_{h=1}^{n} \int_{t_{h-1}}^{t_h} [\mathsf{M}^i, \mathsf{M}^j](I_h)_t dL(h)_t^{kl}$$

$$= n \sum_{h=1}^{n} \left\{ \int_{t_{h-1}}^{t_h} \left(\int_{t_{h-1}}^{t} \sigma_s^{k\cdot} \cdot dB_s \right) [\mathsf{M}^i, \mathsf{M}^j](I_h)_t \sigma_t^{l\cdot} \cdot dB_t \right.$$

$$\left. + \int_{t_{h-1}}^{t_h} \left(\int_{t_{h-1}}^{t} \sigma_s^{l\cdot} \cdot dB_s \right) [\mathsf{M}^i, \mathsf{M}^j](I_h)_t \sigma_t^{k\cdot} \cdot dB_t \right\}$$

by Itô's formula, Lemmas B.2 and B.4 yield

$$\left\| n \sum_{h=1}^{n} \int_{t_{h-1}}^{t_h} [\mathsf{M}^i, \mathsf{M}^j](I_h)_s dL(h)_s^{kl} \right\|_p$$

$$\lesssim \sqrt{n} p \max_{1 \leq k,l \leq d} \sup_{0 \leq s \leq 1} \|\sigma_s^{k\cdot}\|_{4p,\ell_2} \sup_{0 \leq s \leq 1} \|[\mathsf{M}^i, \mathsf{M}^j](I_h)_s \sigma_s^{l\cdot}\|_{\frac{4}{3}p,\ell_2}$$

$$\leq \sqrt{n} p \max_{1 \leq k,l \leq d} \sup_{0 \leq s \leq 1} \|\Sigma_s^{kk}\|_{2p}^{1/2} \sup_{0 \leq s \leq 1} \|[\mathsf{M}^i, \mathsf{M}^j](I_h)_s\|_{2p} \|\Sigma_s^{ll}\|_{2p}^{1/2}.$$

Since Proposition B.1 and the Schwarz inequality imply that

$$\|[\mathsf{M}^i, \mathsf{M}^j](I_h)_s\|_{2p} \leq \frac{1}{n} \sup_{0 \leq s \leq 1} \|\Sigma_s^{ij}\|_{2p} \leq \frac{1}{n} \max_{1 \leq i \leq d} \sup_{0 \leq s \leq 1} \|\Sigma_s^{ii}\|_{2p},$$

we obtain

$$\left\| n \sum_{h=1}^{n} \int_{t_{h-1}}^{t_h} [\mathsf{M}^i, \mathsf{M}^j](I_h)_s dL(h)_s^{kl} \right\|_p \lesssim \frac{p}{\sqrt{n}} \max_{1 \leq i \leq d} \sup_{0 \leq s \leq 1} \|\Sigma_s^{ii}\|_{2p}^2.$$

Combining this estimate with Lemma B.20, we obtain the desired result. $\quad\square$

Lemma B.22. *Under the assumptions of of Lemma B.20, there is a universal constant $C > 0$ such that*

$$
\max_{1 \leq i,j,k,l \leq d} \left\| n \sum_{h=1}^{n} \mathsf{M}^i(I_h)\mathsf{M}^j(I_h)\mathsf{M}^k(I_h)\mathsf{M}^l(I_h) \right.
$$

$$
- n \sum_{h=1}^{n} \{ [\mathsf{M}^i,\mathsf{M}^j](I_h)[\mathsf{M}^k,\mathsf{M}^l](I_h) + [\mathsf{M}^i,\mathsf{M}^k](I_h)[\mathsf{M}^j,\mathsf{M}^l](I_h)
$$

$$
\left. + [\mathsf{M}^j,\mathsf{M}^k](I_h)[\mathsf{M}^i,\mathsf{M}^l](I_h) \} \right\|_p
$$

$$
\leq \frac{C}{\sqrt{n}} \left(p^2 \max_{1 \leq i \leq d} \sup_{0 \leq t \leq 1} \|\Sigma_t^{ii}\|_{2p}^2 \right.
$$

$$
\left. + p^{3/2} \max_{1 \leq i \leq d} \sup_{0 \leq t \leq 1} \|\Sigma_t^{ii}\|_{2p}^{3/2} \max_{1 \leq k \leq d} \sup_{0 \leq u \leq v \leq 1} \|D_u \sigma_v^{k\cdot}\|_{4p,\ell_2} \right)
$$

for all $p \in [2,\infty)$ and $n \in \mathbb{N}$.

Proof. Using Itô's formula repeatedly, we have

$$
\mathsf{M}^i(I_h)\mathsf{M}^j(I_h)\mathsf{M}^k(I_h)\mathsf{M}^l(I_h)
$$

$$
= \int_{t_{h-1}}^{t_h} \mathsf{M}^j(I_h)_s \mathsf{M}^k(I_h)_s \mathsf{M}^l(I_h)_s d\mathsf{M}^i_s + \int_{t_{h-1}}^{t_h} \mathsf{M}^i(I_h)_s \mathsf{M}^k(I_h)_s \mathsf{M}^l(I_h)_s d\mathsf{M}^j_s
$$

$$
+ \int_{t_{h-1}}^{t_h} \mathsf{M}^i(I_h)_s \mathsf{M}^j(I_h)_s \mathsf{M}^l(I_h)_s d\mathsf{M}^k_s + \int_{t_{h-1}}^{t_h} \mathsf{M}^i(I_h)_s \mathsf{M}^j(I_h)_s \mathsf{M}^k(I_h)_s d\mathsf{M}^l_s
$$

$$
+ \int_{t_{h-1}}^{t_h} L(h)_s^{kl} d[\mathsf{M}^i,\mathsf{M}^j]_s + \int_{t_{h-1}}^{t_h} L(h)_s^{jl} d[\mathsf{M}^i,\mathsf{M}^k]_s + \int_{t_{h-1}}^{t_h} L(h)_s^{jk} d[\mathsf{M}^i,\mathsf{M}^l]_s
$$

$$
+ \int_{t_{h-1}}^{t_h} L(h)_s^{il} d[\mathsf{M}^j,\mathsf{M}^k]_s + \int_{t_{h-1}}^{t_h} L(h)_s^{ik} d[\mathsf{M}^j,\mathsf{M}^l]_s + \int_{t_{h-1}}^{t_h} L(h)_s^{ij} d[\mathsf{M}^k,\mathsf{M}^l]_s
$$

$$
+ [\mathsf{M}^i,\mathsf{M}^j](I_h)[\mathsf{M}^k,\mathsf{M}^l](I_h) + [\mathsf{M}^i,\mathsf{M}^k](I_h)[\mathsf{M}^j,\mathsf{M}^l](I_h)
$$

$$
+ [\mathsf{M}^j,\mathsf{M}^k](I_h)[\mathsf{M}^i,\mathsf{M}^l](I_h)
$$

for every h. Therefore, the desired result follows from Lemmas B.19 and B.21. $\quad\square$

Lemma B.23. *Under the assumptions of of Lemma B.20, there is a universal constant $C > 0$ such that*

$$
\max_{1 \leq i,j,k,l \leq d} \left\| n \sum_{h=1}^{n-1} \mathsf{M}^i(I_h)\mathsf{M}^j(I_h) \int_{t_h}^{t_{h+1}} \mathsf{M}^l(I_{h+1})_s d\mathsf{M}^k_s \right\|_p \leq C \frac{p^2}{\sqrt{n}} \max_{1 \leq i \leq d} \sup_{0 \leq s \leq 1} \|\Sigma_s^{ii}\|_{2p}^2
$$

for all $p \in [2,\infty)$ and $n \in \mathbb{N}$.

Proof. Proposition B.1–B.2 yield

$$
\left\| n \sum_{h=1}^{n-1} \mathsf{M}^i(I_h)\mathsf{M}^j(I_h) \int_{t_h}^{t_{h+1}} \mathsf{M}^l(I_{h+1})_s d\mathsf{M}^k_s \right\|_p
$$

$$
\lesssim n\sqrt{p} \left\| \sqrt{\sum_{h=1}^{n-1} \mathsf{M}^i(I_h)^2 \mathsf{M}^j(I_h)^2 \int_{t_h}^{t_{h+1}} \mathsf{M}^l(I_{h+1})_s^2 d[\mathsf{M}^k, \mathsf{M}^k]_s } \right\|_p
$$

$$
\leq n \sqrt{ p \sum_{h=1}^{n-1} \left\| \mathsf{M}^i(I_h) \right\|_{4p}^2 \left\| \mathsf{M}^j(I_h) \right\|_{4p}^2 \int_{t_h}^{t_{h+1}} \left\| \mathsf{M}^l(I_{h+1})_s \right\|_{4p}^2 \left\| \Sigma_s^{kk} \right\|_{2p} ds }
$$

$$
\lesssim np^2 \sqrt{ \sum_{h=1}^{n-1} \left\| \sqrt{\int_{I_h} \Sigma_s^{ii} ds} \right\|_{4p}^2 \left\| \sqrt{\int_{I_h} \Sigma_s^{jj} ds} \right\|_{4p}^2 \int_{t_h}^{t_{h+1}} \left\| \sqrt{\int_{t_h}^{s} \Sigma_s^{ll} ds} \right\|_{4p}^2 \left\| \Sigma_s^{kk} \right\|_{2p} ds }
$$

$$
\leq \frac{p^2}{\sqrt{n}} \max_{1 \leq i \leq d} \sup_{0 \leq s \leq 1} \left\| \Sigma_s^{ii} \right\|_{2p}^2.
$$

This completes the proof. $\qquad\square$

Lemma B.24. *Under the assumptions of of Lemma B.20, there is a universal constant $C > 0$ such that*

$$
\max_{1 \leq i,j,k,l \leq d} \left\| n \sum_{h=1}^{n-1} \mathsf{M}^i(I_h)\mathsf{M}^j(I_h)\mathsf{M}^k(I_{h+1})\mathsf{M}^l(I_{h+1}) - n \sum_{h=1}^{n-1} [\mathsf{M}^i, \mathsf{M}^j](I_h)[\mathsf{M}^k, \mathsf{M}^l](I_{h+1}) \right\|_p
$$

$$
\leq \frac{C}{\sqrt{n}} \left(p^2 \max_{1 \leq i \leq d} \sup_{0 \leq t \leq 1} \left\| \Sigma_t^{ii} \right\|_{2p}^2 + p^{3/2} \max_{1 \leq i \leq d} \sup_{0 \leq t \leq 1} \left\| \Sigma_t^{ii} \right\|_{2p}^{3/2} \max_{1 \leq k \leq d} \sup_{0 \leq u \leq v \leq 1} \left\| D_u \sigma_v^{k\cdot} \right\|_{4p, \ell_2} \right)
$$

for all $p \in [2, \infty)$ and $n \in \mathbb{N}$.

Proof. By Itô's formula we have

$$
\mathsf{M}^i(I_h)\mathsf{M}^j(I_h)\mathsf{M}^k(I_{h+1})\mathsf{M}^l(I_{h+1})
$$

$$
= \mathsf{M}^i(I_h)\mathsf{M}^j(I_h) \int_{t_h}^{t_{h+1}} \mathsf{M}^l(I_{h+1})_s d\mathsf{M}^k_s + \mathsf{M}^i(I_h)\mathsf{M}^j(I_h) \int_{t_h}^{t_{h+1}} \mathsf{M}^k(I_{h+1})_s d\mathsf{M}^l_s
$$

$$
+ L(h)^{ij} [\mathsf{M}^k, \mathsf{M}^l](I_{h+1}) + [\mathsf{M}^i, \mathsf{M}^j](I_h)[\mathsf{M}^k, \mathsf{M}^l](I_{h+1})
$$

for every h. Therefore, the desired result follows from Lemmas B.20 and B.23. $\qquad\square$

Proof of Proposition 4.1. Thanks to Lemma B.17, throughout the proof we may assume $\mu = \mu(\nu)$ and $\sigma = \sigma(\nu)$ for all $n, \nu \in \mathbb{N}$.

(a) According to Lemmas B.18, B.22 and B.24, it suffices to show that

$$
E \left[\max_{1 \leq i,j,k,l \leq d} \left| n[\mathsf{M}^i, \mathsf{M}^j](I_n)[\mathsf{M}^k, \mathsf{M}^l](I_n) \right| \right] = O(n^{-\varpi}), \tag{B.21}
$$

$$E\left[\max_{1\leq i,j,k,l\leq d}\left|n\sum_{h=1}^{n-1}[\mathsf{M}^i,\mathsf{M}^j](I_h)\left\{[\mathsf{M}^k,\mathsf{M}^l](I_{h+1})-[\mathsf{M}^k,\mathsf{M}^l](I_h)\right\}\right|\right]=O(n^{-\varpi}).$$
(B.22)

(B.21) is evident from assumptions. In the meantime, the Schwarz inequality and Proposition B.1 yield

$$E\left[\max_{1\leq i,j,k,l\leq d}\left|n\sum_{h=1}^{n-1}[\mathsf{M}^i,\mathsf{M}^j](I_h)\left\{[\mathsf{M}^k,\mathsf{M}^l](I_{h+1})-[\mathsf{M}^k,\mathsf{M}^l](I_h)\right\}\right|\right]$$

$$\leq n\sum_{h=1}^{n-1}E\left[\max_{1\leq i,j\leq d}\left|[\mathsf{M}^i,\mathsf{M}^j](I_h)\right|\max_{1\leq k,l\leq d}\left|[\mathsf{M}^k,\mathsf{M}^l](I_{h+1})-[\mathsf{M}^k,\mathsf{M}^l](I_h)\right|\right]$$

$$\leq \sup_{0\leq t\leq 1}\left\|\max_{1\leq i,j\leq d}\left|\Sigma_t^{ij}\right|\right\|_2 \sup_{0<t\leq 1-\frac{1}{n}}\left\|\max_{1\leq k,l\leq d}\left|\Sigma_{t+\frac{1}{n}}^{kl}-\Sigma_t^{kl}\right|\right\|_2,$$
(B.23)

so (B.22) also follows from assumptions. This completes the proof.

(b) By assumptions we have $E[\max_{1\leq i,j,k,l\leq d}\left|n[\mathsf{M}^i,\mathsf{M}^j](I_n)[\mathsf{M}^k,\mathsf{M}^l](I_n)\right|]=O(n^{-1})$. Moreover, from (B.23) and assumptions, we also have

$$E\left[\max_{1\leq i,j,k,l\leq d}\left|n\sum_{h=1}^{n-1}[\mathsf{M}^i,\mathsf{M}^j](I_h)\left\{[\mathsf{M}^k,\mathsf{M}^l](I_{h+1})-[\mathsf{M}^k,\mathsf{M}^l](I_h)\right\}\right|\right]=O(n^{-\gamma}).$$

Therefore, the desired result follows from Lemmas B.18, B.22 and B.24 as well as Lemma A.7 and Proposition A.1 of [45]. $\square$

B.4. Proof of Proposition 4.2

An analogous argument to the proof of Theorem 4.2 allows us to assume $\mu = \mu(\nu)$ and $\Sigma = \Sigma(\nu)$ for all $n, \nu \in \mathbb{N}$.

Define the $\underline{d}^2 \times d^2$ random matrix $\hat{\boldsymbol{X}}_n$ by

$$\hat{\boldsymbol{X}}_n^{(i-1)\underline{d}+j,(k-1)d+l}=\begin{cases}\widehat{[Y^j,Y^d]}_1^n/\sqrt{\widehat{\mathfrak{V}_n^{ij}}} & \text{if } k=i,\ l=d,\\ [Y^i,Y^d]_1/\sqrt{\widehat{\mathfrak{V}_n^{ij}}} & \text{if } k=j,\ l=d,\\ -\widehat{[Y^d,Y^d]}_1^n/\sqrt{\widehat{\mathfrak{V}_n^{ij}}} & \text{if } k=l=d\\ -[Y^i,Y^j]_1/\sqrt{\widehat{\mathfrak{V}_n^{ij}}} & \text{if } k=i,l=j\\ 0 & \text{otherwise.}\end{cases}$$

for $i,j=1,\ldots,\underline{d}$ and $k,l=1,\ldots,d$. We also define the $\underline{d}^2 \times d^2$ matrix $\underline{\Upsilon}_n$ by

$$\underline{\Upsilon}_n^{(i-1)\underline{d}+j,(k-1)d+l}=\begin{cases}1 & \text{if } k\in\{i,j\},l=d,\\ -1 & \text{if } k=i,l=j \text{ or } k=l=d,\\ 0 & \text{otherwise}\end{cases}$$

for $i,j=1,\ldots,\underline{d}$ and $k,l=1,\ldots,d$. Then we set

$$\Xi_n=\begin{pmatrix}\boldsymbol{X}_n\\ -\boldsymbol{X}_n\end{pmatrix},\qquad \hat{\Xi}_n=\begin{pmatrix}\hat{\boldsymbol{X}}_n\\ -\hat{\boldsymbol{X}}_n\end{pmatrix},\qquad \Upsilon_n=\begin{pmatrix}\underline{\Upsilon}_n\\ -\underline{\Upsilon}_n\end{pmatrix}.$$

Since we have

$$\hat{\Xi}_n S_n = \begin{pmatrix} \mathrm{vec}(T_n) \\ -\mathrm{vec}(T_n) \end{pmatrix}, \qquad \hat{\Xi}_n S_n^* = \begin{pmatrix} \mathrm{vec}(T_n^*) \\ -\mathrm{vec}(T_n^*) \end{pmatrix}$$

as well as all the diagonal entries of $\Xi_n \mathfrak{C}_n \Xi_n^\top$ are equal to 1 by the definition of $\boldsymbol{X}_n$, Lemma 3.1 and Proposition 3.1 imply that it suffices to prove the following equations:

$$\sup_{y \in (\infty, \infty]^{2\underline{d}^2}} |P(\Xi_n S_n \le y) - P(\Xi_n \mathfrak{C}_n^{1/2} \zeta_n \le y)| \to 0, \qquad (\mathrm{B.24})$$

$$\sqrt{\log d} \|\hat{\Xi}_n S_n - \Xi_n S_n\|_{\ell_\infty} \to^p 0, \qquad (\mathrm{B.25})$$

$$(\log d)^2 \|\hat{\Xi}_n \hat{\mathfrak{C}}_n \hat{\Xi}_n^\top - \Xi_n \mathfrak{C}_n \Xi_n^\top\|_{\ell_\infty} \to^p 0 \qquad (\mathrm{B.26})$$

as $n \to \infty$.

We begin by proving (B.24). Since $\boldsymbol{X}_n = \boldsymbol{X}_n \circ \underline{\Upsilon}_n$ and $\||\underline{\Upsilon}_n\||_\infty = 4$, an application of Theorem 4.1 implies that the desired result follows once we show that $\boldsymbol{X}_n \in \mathbb{D}_{2,\infty}(\mathbb{R}^{\underline{d}^2} \otimes \mathbb{R}^{d^2})$ and

$$\sup_{n \in \mathbb{N}} \max_{1 \le i \le \underline{d}^2, 1 \le j \le d^2} \left(\|\boldsymbol{X}_n^{ij}\|_p + \sup_{0 \le t \le 1} \|D_t \boldsymbol{X}_n^{ij}\|_{p,\ell_2} + \sup_{0 \le s,t \le 1} \|D_{s,t} \boldsymbol{X}_n^{ij}\|_{p,\ell_2} \right) < \infty$$

for all $p \in [2, \infty)$. By Remark 15.87 of [37], we have $[Y^i, Y^j]_1 \in \mathbb{D}_{2,\infty}$ and $D^q[Y^i, Y^j]_1 = \int_0^1 D^q \Sigma_t^{ij} dt$ for any $i, j = 1, \dots, d$ and $q = 1, 2$. Therefore, by Corollary 15.80 of [37], the desired result follows once we show that $1/\sqrt{\mathfrak{V}_n^{ij}} \in \mathbb{D}_{2,\infty}$ for any $i, j = 1, \dots, \underline{d}$ and

$$\sup_{n \in \mathbb{N}} \max_{1 \le i,j \le \underline{d}} \left(\left\| \frac{1}{\sqrt{\mathfrak{V}_n^{ij}}} \right\|_p + \sup_{0 \le t \le 1} \left\| D_t \left(\frac{1}{\sqrt{\mathfrak{V}_n^{ij}}} \right) \right\|_{p,\ell_2} \right.$$

$$\left. + \sup_{0 \le s,t \le 1} \left\| D_{s,t} \left(\frac{1}{\sqrt{\mathfrak{V}_n^{ij}}} \right) \right\|_{p,\ell_2} \right) < \infty \qquad (\mathrm{B.27})$$

for all $p \in [2, \infty)$. Note that we have

$$\|D_t[Y^i, Y^j]_1\|_{p,\ell_2} \le \sup_{0 \le u \le 1} \|D_t \Sigma_u^{ij}\|_{p,\ell_2},$$

$$\left\| D_{s,t}[Y^i, Y^j]_1 \right\|_{p,\ell_2} \le \sup_{0 \le u \le 1} \left\| D_{s,t} \Sigma_u^{ij} \right\|_{p,\ell_2}$$

for all $i, j = 1, \dots, d$, $p \in [2, \infty)$ and $s, t \in [0, 1]$ by Proposition B.1. Therefore, Lemmas B.11–B.12 and Corollary 15.80 of [37] imply that $\mathfrak{V}_n^{ij} \in \mathbb{D}_{2,\infty}$ for any $i, j = 1, \dots, \underline{d}$ and

$$\sup_{n \in \mathbb{N}} \max_{1 \le i,j \le \underline{d}} \left(\left\| \mathfrak{V}_n^{ij} \right\|_p + \sup_{0 \le t \le 1} \left\| D_t \mathfrak{V}_n^{ij} \right\|_{p,\ell_2} + \sup_{0 \le s,t \le 1} \left\| D_{s,t} \mathfrak{V}_n^{ij} \right\|_{p,\ell_2} \right) < \infty$$

for all $p \in [2, \infty)$. Now, since we can write $1/\sqrt{\mathfrak{V}_n^{ij}} = (\mathfrak{V}_n^{ij})^{5/2}(\mathfrak{V}_n^{ij})^{-3}$, Theorem 15.78 and Lemma 15.152 of [37] as well as (4.11) imply that $1/\sqrt{\mathfrak{V}_n^{ij}} \in \mathbb{D}_{2,\infty}$ for any $i, j = 1, \ldots, \underline{d}$ and (B.27) holds true for all $p \in [2, \infty)$. Hence we complete the proof of (B.24).

Next we prove (B.25)–(B.26). First, note that we have $\|S_n\|_{\ell_\infty} = O_p(n^\eta)$ as $n \to \infty$ for any $\eta > 0$ by Corollary 4.1 and (B.20). Since $\|[Y, Y]_1\|_{\ell_\infty} = O_p(n^\eta)$ as $n \to \infty$ for any $\eta > 0$ by assumptions, this especially yields $\|\widehat{[Y, Y]}_1^n\|_{\ell_\infty} = O_p(n^\eta)$ as $n \to \infty$ for any $\eta > 0$. Next we verify

$$\max_{1 \le i,j \le \underline{d}} |\hat{\mathfrak{V}}_n^{ij} - \mathfrak{V}_n^{ij}| = O_p(n^{-\varpi})$$

as $n \to \infty$ for any $\varpi \in (0, \gamma)$. In fact, by definition we have

$$\max_{1 \le i,j \le \underline{d}} |\hat{\mathfrak{V}}_n^{ij} - \mathfrak{V}_n^{ij}|$$
$$\lesssim \left(\left\|\widehat{[Y,Y]}_1^n\right\|_{\ell_\infty} + \|[Y,Y]_1\|_{\ell_\infty} \right) \left(\left\|\hat{\mathfrak{C}}_n\right\|_{\ell_\infty} + \|\mathfrak{C}_n\|_{\ell_\infty} \right)$$
$$\times \left(\left\|\widehat{[Y,Y]}_1^n - [Y,Y]_1\right\|_{\ell_\infty} + \left\|\hat{\mathfrak{C}}_n - \mathfrak{C}_n\right\|_{\ell_\infty} \right).$$

Since $\|\mathfrak{C}_n\|_{\ell_\infty} = O_p(n^\eta)$ for any $\eta > 0$ by assumptions, the desired result follows from Proposition 4.1 and the results noted above. In particular, it holds that $\max_{1 \le i,j \le \underline{d}} |1/\sqrt{\hat{\mathfrak{V}}_n^{ij}}| = O_p(n^\eta)$ for any $\eta > 0$ because $\max_{1 \le i,j \le \underline{d}} |1/\sqrt{\mathfrak{V}_n^{ij}}| = O_p(n^\eta)$ for any $\eta > 0$ by assumptions. Moreover, we have

$$\|\hat{\boldsymbol{X}}_n - \boldsymbol{X}_n\|_{\ell_\infty}$$
$$\le \max_{1 \le i,j \le \underline{d}} \left| \frac{1}{\sqrt{\hat{\mathfrak{V}}_n^{ij}}} \right| \left\|\widehat{[Y,Y]}_1^n - [Y,Y]_1\right\|_{\ell_\infty} + \max_{1 \le i,j \le \underline{d}} \left| \frac{1}{\sqrt{\hat{\mathfrak{V}}_n^{ij}}} - \frac{1}{\sqrt{\mathfrak{V}_n^{ij}}} \right| \|[Y,Y]_1\|_{\ell_\infty},$$

and thus it holds that $\|\hat{\boldsymbol{X}}_n - \boldsymbol{X}_n\|_{\ell_\infty} = O_p(n^{-\varpi})$ as $n \to \infty$ for any $\varpi \in (0, \gamma)$. Noting that we have $\|\boldsymbol{X}_n\|_{\ell_\infty} = O_p(n^\eta)$ for any $\eta > 0$ by assumptions, this particularly implies that $\|\hat{\boldsymbol{X}}_n\|_{\ell_\infty} = O_p(n^\eta)$ for any $\eta > 0$.

Now since we have

$$\|\hat{\Xi}_n S_n - \Xi_n S_n\|_{\ell_\infty} \le 4\|\hat{\boldsymbol{X}}_n - \boldsymbol{X}_n\|_{\ell_\infty} \|S_n\|_{\ell_\infty}$$

and

$$\|\hat{\Xi}_n \hat{\mathfrak{C}}_n \hat{\Xi}_n^\top - \Xi_n \mathfrak{C}_n \Xi_n^\top\|_{\ell_\infty}$$
$$\le 16 \left\{ \|\hat{\boldsymbol{X}}_n\|_{\ell_\infty}^2 \|\hat{\mathfrak{C}}_n - \mathfrak{C}_n\|_{\ell_\infty} + \left(\|\hat{\boldsymbol{X}}_n\|_{\ell_\infty} + \|\boldsymbol{X}_n\|_{\ell_\infty} \right) \|\mathfrak{C}_n\|_{\ell_\infty} \|\hat{\boldsymbol{X}}_n - \boldsymbol{X}_n\|_{\ell_\infty} \right\},$$

(B.25)–(B.26) follow from the results remarked above. Thus we complete the proof. $\square$

B.5. Proof of Corollary 4.2

By construction both the Bonferroni-Holm and Romano-Wolf methods evidently satisfy condition (i). So it remains to check that they also satisfy (ii). Since it holds that $\max_{\ell \in \mathcal{L}_n(\theta_n)} \mathsf{T}_n^\ell = \max_{\ell \in \mathcal{L}_n(\theta_n)} \max_{\lambda \in \Lambda_n^\ell} |\check{\tilde{T}}_n^\lambda|$, Proposition 4.2 yields

$$P\left(\max_{\ell \in \mathcal{L}_n(\theta_n)} \mathsf{T}_n^\ell > c_n^{\mathcal{L}_n(\theta_n)}(1-\alpha)\right) - P\left(\max_{\ell \in \mathcal{L}_n(\theta_n)} \max_{k \in \mathcal{K}_n^\ell} \left|\tilde{\zeta}_n^k\right| > c_n^{\mathcal{L}_n(\theta_n)}(1-\alpha)\right) \to 0$$

as $n \to \infty$, where $\mathcal{K}_n^\ell := \{(i-1)\underline{d} + j : (i,j) \in \Lambda_n^\ell\}$ and $\tilde{\zeta}_n := \boldsymbol{X}_n \mathfrak{C}_n^{1/2} \zeta_n$. Now if we use the Bonferroni-Holm method, we have

$$P\left(\max_{\ell \in \mathcal{L}_n(\theta_n)} \max_{k \in \mathcal{K}_n^\ell} \left|\tilde{\zeta}_n^k\right| > c_n^{\mathcal{L}_n(\theta_n)}(1-\alpha)\right)$$

$$\leq \sum_{\ell \in \mathcal{L}_n(\theta_n)} \sum_{k \in \mathcal{K}_n^\ell} P\left(\left|\tilde{\zeta}_n^k\right| > q_{N(0,1)}\left(1 - \frac{\alpha}{2\#[\bigcup_{\ell \in \mathcal{L}_n(\theta_n)} \mathcal{K}_n^\ell]}\right)\right) = \alpha,$$

so condition (ii) is satisfied. Meanwhile, if we use the Romano-Wolf method, Propositions 3.2 and 4.2 yield

$$P\left(\max_{\ell \in \mathcal{L}_n(\theta_n)} \max_{k \in \mathcal{K}_n^\ell} \left|\tilde{\zeta}_n^k\right| > c_n^{\mathcal{L}_n(\theta_n)}(1-\alpha)\right) \to \alpha$$

as $n \to \infty$, so condition (ii) is satisfied. Thus we complete the proof. $\square$

Acknowledgements

References

[1] Aït-Sahalia, Y. & Xiu, D. (2017). Using principal component analysis to estimate a high dimensional factor model with high-frequency data. *J. Econometrics* **201**, 384–399. MR3717571

[2] Alòs, E. & Ewald, C.-O. (2008). Malliavin differentiability of the Heston volatility and applications to option pricing. *Adv. in Appl. Probab.* **40**, 144–162. MR2411818

[3] Barlow, M. T. & Yor, M. (1982). Semi-martingale inequalities via the Garsia-Rodemich-Rumsey lemma and application to local times. *J. Funct. Anal.* **49**, 198–229. MR0680660

[4] Barndorff-Nielsen, O. E. & Shephard, N. (2004). Econometric analysis of realized covariation: High frequency based covariance, regression, and correlation in financial economics. *Econometrica* **72**, 885–925. MR2051439

[5] Bibinger, M. & Mykland, P. A. (2016). Inference for multi-dimensional high-frequency data with an application to conditional independence testing. *Scand. J. Stat.* **43**, 1078–1102. MR3573677

[6] Bickel, P. J. & Levina, E. (2008). Covariance regularization by thresholding. *Ann. Statist.* **36**, 2577–2604. MR2485008

[7] Bickel, P. J. & Levina, E. (2008). Regularized estimation of large covariance matrices. *Ann. Statist.* **36**, 199–227. MR2387969

[8] Bihari, I. (1956). A generalization of a lemma of Bellman and its application to uniqueness problems of differential equations. *Acta Math. Hungar.* **7**, 81–94. MR0079154

[9] Brownlees, C., Nualart, E. & Sun, Y. (2018). Realized networks. *J. Appl. Econometrics* **33**, 986–1006. MR3887275

[10] Chang, J., Qiu, Y., Yao, Q. & Zou, T. (2018). Confidence regions for entries of a large precision matrix. *J. Econometrics* **206**, 57–82. MR3840783

[11] Chen, X. (2018). Gaussian and bootstrap approximations for high-dimensional U-statistics and their applications. *Ann. Statist.* **46**, 642–678. MR3782380

[12] Chen, X. & Kato, K. (2017). Jackknife multiplier bootstrap: finite sample approximations to the U-process supremum with applications. Working paper. arXiv: 1708.02705. MR2635313

[13] Chen, X. & Kato, K. (2019). Randomized incomplete U-statistics in high dimensions. *Ann. Statist. (forthcoming)*.

[14] Chernozhukov, V., Chetverikov, D. & Kato, K. (2013). Gaussian approximations and multiplier bootstrap for maxima of sums of high-dimensional random vectors. *Ann. Statist.* **41**, 2786–2819. MR3161448

[15] Chernozhukov, V., Chetverikov, D. & Kato, K. (2014). Gaussian approximation of suprema of empirical processes. *Ann. Statist.* **42**, 1564–1597. MR3262461

[16] Chernozhukov, V., Chetverikov, D. & Kato, K. (2014). Testing many moment inequalities. Working paper. arXiv: 1312.7614v4.

[17] Chernozhukov, V., Chetverikov, D. & Kato, K. (2015). Comparison and anti-concentration bounds for maxima of Gaussian random vectors. *Probab. Theory Related Fields* **162**, 47–70. MR3350040

[18] Chernozhukov, V., Chetverikov, D. & Kato, K. (2016). Empirical and multiplier bootstraps for suprema of empirical processes of increasing complexity, and related Gaussian couplings. *Stochastic Process. Appl.* **126**, 3632–3651. MR3565470

[19] Chernozhukov, V., Chetverikov, D. & Kato, K. (2017). Central limit theorems and bootstrap in high dimensions. *Ann. Probab.* **45**, 2309–2353. MR3693963

[20] Chernozhukov, V., Chetverikov, D. & Kato, K. (2017). Detailed proof of Nazarov's inequality. Unpublished paper. arXiv: 1711.10696.

[21] Christensen, K., Podolskij, M., Thamrongrat, N. & Veliyev, B. (2017). Inference from high-frequency data: A subsampling approach. *J. Econometrics* **197**, 245–272. MR3614200

[22] Clément, E. & Gloter, A. (2011). Limit theorems in the Fourier transform

method for the estimation of multivariate volatility. *Stochastic Process. Appl.* **121**, 1097–1124. MR2775108

[23] Dai, C., Lu, K. & Xiu, D. (2017). Knowing factors or factor loadings, or neither? Evaluating estimators of large covariance matrices with noisy and asynchronous data. Tech. Rep. 17-02, The University of Chicago Booth School of Business.

[24] Deng, H. & Zhang, C.-H. (2017). Beyond Gaussian approximation: Bootstrap for maxima of sums of independent random vectors. Working paper. arXiv: 1705.09528. MR3616522

[25] Dovonon, P., Gonçalves, S. & Meddahi, N. (2013). Bootstrapping realized multivariate volatility measures. *J. Econometrics* **172**, 49–65. MR2997129

[26] Dudley, R. M. (2002). *Real analysis and probability.* Cambridge University Press. MR1932358

[27] Fama, E. F. & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* **33**, 3–56.

[28] Fan, J., Furger, A. & Xiu, D. (2016). Incorporating global industrial classification standard into portfolio allocation: A simple factor-based large covariance matrix estimator with high-frequency data. *J. Bus. Econom. Statist.* **34**, 489–503. MR3547991

[29] Fan, J. & Kim, D. (2018). Robust high-dimensional volatility matrix estimation for high-frequency factor model. *J. Amer. Statist. Assoc.* **113**, 1268–1283. MR3862356

[30] Fan, J., Li, Y. & Yu, K. (2012). Vast volatility matrix estimation using high-frequency data for portfolio selection. *J. Amer. Statist. Assoc.* **107**, 412–428. MR2949370

[31] Fan, J., Liao, Y. & Mincheva, M. (2011). High-dimensional covariance matrix estimation in approximate factor models. *Ann. Statist.* **39**, 3320–3356. MR3012410

[32] Fukasawa, M. (2011). Asymptotic analysis for stochastic volatility: martingale expansion. *Finance Stoch.* **15**, 635–654. MR2863637

[33] Hardy, M. (2006). Combinatorics of partial derivatives. *Electron. J. Combin.* **13**, 1–12. MR2200529

[34] Hounyo, U. (2017). Bootstrapping integrated covariance matrix estimators in noisy jump-diffusion models with non-synchronous trading. *J. Econometrics* **197**, 130–152. MR3598650

[35] Jacod, J. & Protter, P. (2012). *Discretization of processes.* Springer. MR2859096

[36] Jacod, J. & Todorov, V. (2009). Testing for common arrivals of jumps for discretely observed multidimensional processes. *Ann. Statist.* **37**, 1792–1838. MR2533472

[37] Janson, S. (1997). *Gaussian Hilbert space.* Cambridge University Press. MR1474726

[38] Jolis, M. & Sanz, M. (1990). On generalized multiple stochastic integrals and multiparameter anticipative calculus. In H. Korezlioglu & A. S. Ustunel, eds., *Stochastic analysis and related topics II*, vol. 1444 of *Lecture Notes in Math.* Springer, pp. 142–182. MR1276930

[39] Karatzas, I. & Shreve, S. E. (1998). *Brownian motion and stochastic calculus.* Springer, 2nd edn. MR1121940

[40] Kato, K. & Kurisu, D. (2017). Bootstrap confidence bands for spectral estimation of Lévy densities under high-frequency observations. Working paper. arXiv: 1705.00586.

[41] Kato, K. & Sasaki, Y. (2018). Uniform confidence bands in deconvolution with unknown error distribution. *J. Econometrics* **207**, 129–161. MR3856764

[42] Kim, D., Kong, X.-B., Li, C.-X. & Wang, Y. (2018). Adaptive thresholding for large volatility matrix estimation based on high-frequency financial data. *J. Econometrics* **203**, 69–79. MR3758328

[43] Kim, D., Liu, Y. & Wang, Y. (2018). Large volatility matrix estimation with factor-based diffusion model for high-frequency financial data. *Bernoulli* **24**, 3657–3682. MR3788185

[44] Kim, D., Wang, Y. & Zou, J. (2016). Asymptotic theory for large volatility matrix estimation based on high-frequency financial data. *Stochastic Process. Appl.* **126**, 3527–3577. MR3549717

[45] Koike, Y. (2019). Gaussian approximation of maxima of Wiener functionals and its application to high-frequency data. *Ann. Statist.* **47**, 1663–1687. MR3911126

[46] Kurose, Y. & Omori, Y. (2016). Dynamic equicorrelation stochastic volatility. *Comput. Statist. Data Anal.* **100**, 795–813. MR3505833

[47] Kurose, Y. & Omori, Y. (2018). Multiple-block dynamic equicorrelations with realized measures, leverage and endogeneity. *Econom. Stat. (in press).*

[48] Li, J., Todorov, V. & Tauchen, G. (2017). Jump regressions. *Econometrica* **85**, 173–195. MR3611769

[49] Lunde, A., Shephard, N. & Sheppard, K. (2016). Econometric analysis of vast covariance matrices using composite realized kernels and their application to portfolio choice. *J. Bus. Econom. Statist.* **34**, 504–518. MR3547992

[50] Medvegyev, P. (2007). *Stochastic integration theory.* Oxford University Press. MR2345169

[51] Mishkov, R. L. (2000). Generalization of the formula of Faa di Bruno for a composite function with a vector argument. *International Journal of Mathematics and Mathematical Sciences* **24**, 481–491. MR1781515

[52] Naganuma, N. (2013). Smoothness of densities of generalized locally nondegenerate Wiener functionals. *Stoch. Anal. Appl.* **31**, 609–631. MR3175788

[53] Nourdin, I., Nualart, D. & Peccati, G. (2016). Quantitative stable limit theorems on the Wiener space. *Ann. Probab.* **44**, 1–41. MR3456331

[54] Nourdin, I. & Peccati, G. (2012). *Normal approximations with Malliavin calculus: From Stein's method to universality.* Cambridge University Press. MR2962301

[55] Nualart, D. (2006). *The Malliavin calculus and related topics.* Springer, 2nd edn. MR2200233

[56] Nualart, D. & Pardoux, E. (1988). Stochastic calculus with anticipating integrands. *Probab. Theory Related Fields* **78**, 535–581. MR0950346

[57] Nualart, D. & Yoshida, N. (2017). Asymptotic expansion of Skorohod

integrals. Working paper. Available at arXiv: https://arxiv.org/abs/1801.00120.

[58] Nualart, D. & Zakai, M. (1988). Generalized multiple stochastic integrals and the representation of Wiener functionals. *Stochastics* **23**, 311–330. MR0959117

[59] Pelger, M. (2019). Large-dimensional factor modeling based on high-frequency observations. *J. Econometrics* **208**, 23–42. MR3906960

[60] Podolskij, M. & Vetter, M. (2010). Understanding limit theorems for semi-martingales: a short survey. *Stat. Neerl.* **64**, 329–351. MR2683464

[61] Romano, J. P. & Wolf, M. (2005). Exact and approximate stepdown methods for multiple hypothesis testing. *J. Amer. Statist. Assoc.* **100**, 94–108. MR2156821

[62] Rudin, W. (1991). *Functional analysis*. McGraw-Hill, 2nd edn. MR1157815

[63] Sakamoto, Y. & Yoshida, N. (2004). Asymptotic expansion formulas for functionals of ϵ-Markov processes with a mixing property. *Ann. Inst. Statist. Math.* **56**, 545–597. MR2095018

[64] Tao, M., Wang, Y., Yao, Q. & Zou, J. (2011). Large volatility matrix inference via combining low-frequency and high-frequency approaches. *J. Amer. Statist. Assoc.* **106**, 1025–1040. MR2894761

[65] Tao, M., Wang, Y. & Zhou, H. (2013). Optimal sparse volatility matrix estimation for high-dimensional Itô processes with measurement errors. *Ann. Statist.* **41**, 1816–1864. MR3127850

[66] Tudor, C. A. & Yoshida, N. (2018). Asymptotic expansion for vector-valued sequences of random variables with focus on Wiener chaos. *Stochastic Process. Appl. (in press)*.

[67] Ubukata, M. (2010). Large-scale portfolios using realized covariance matrix: evidence from the Japanese stock market. *Economics Bulletin* **30**, 2906–2919.

[68] Wang, Y. & Zou, J. (2010). Vast volatility matrix estimation for high-frequency financial data. *Ann. Statist.* **38**, 943–978. MR2604708

[69] Yoshida, N. (1997). Malliavin calculus and asymptotic expansion for martingales. *Probab. Theory Related Fields* **109**, 301–342. MR1481124

[70] Zhang, D. & Wu, W. B. (2017). Gaussian approximation for high dimensional time series. *Ann. Statist.* **45**, 1895–1919. MR3718156

[71] Zhang, X. & Cheng, G. (2018). Gaussian approximation for high dimensional vector under physical dependence. *Bernoulli* **24**, 2640–2675. MR3779697