Stein's method for multivariate Brownian approximations of sums under dependence

Mikołaj J. Kasprzak

University of Luxembourg, Department of Mathematics, , Maison du Nombre, 6 Avenue de la Fonte, L-4364 Esch-sur-Alzette, Luxembourg

Abstract

We use Stein's method to obtain a bound on the distance between scaled p-dimensional random walks and a p-dimensional (correlated) Brownian motion. We consider dependence schemes including those in which the summands in scaled sums are weakly dependent and their p components are strongly correlated. As an example application, we prove a functional limit theorem for exceedances in an m-scans process, together with a bound on the rate of convergence. We also find a bound on the rate of convergence of scaled U-statistics to Brownian motion, representing an example of a sum of strongly dependent terms.
ⓒ 2020 Elsevier B.V. All rights reserved.

1. Introduction

In the seminal paper [1], Barbour addressed the problem of providing bounds on the rate of convergence in functional limit results (or invariance principles as they are often called in the literature). He observed that the celebrated Stein's method, first introduced in [38] as a tool for proving the Central Limit Theorem, may also be used in the setup of the *Functional* Central Limit Theorem. This theorem, whose early versions are attributed to Donsker [13], says that for a sequence of i.i.d. real random variables $(X_n)_{n=1}^{\infty}$ with mean zero and unit variance, the random process

$$\mathbf{Y}_n(t) = n^{-1/2} \sum_{i=1}^{\lfloor nt \rfloor} X_i, \quad t \in [0, 1] \tag{1.1}$$

converges in distribution to the standard Brownian motion with respect to the Skorokhod topology.

Through a careful and technical adaptation of Stein's method to the framework of Brownian-motion approximation and a subsequent repetitive use of Taylor's theorem, Barbour [1] proved a powerful estimate on a distance between the law of $\mathbf{Y}_n$ in (1.1) and the Wiener measure. Specifically, he considered test functions g acting on the Skorokhod space $D([0, 1], \mathbb{R})$ of càdlàg real-valued maps on [0, 1], such that g takes values in the reals, does not grow faster than a cubic, is twice Fréchet differentiable and its second derivative is Lipschitz. Denoting by $\mathbf{Z}$ the Brownian motion on [0, 1] and adopting the notation of (1.1), his result says that

$$|\mathbb{E}g(\mathbf{Y}_n) - \mathbb{E}g(\mathbf{Z})| \le C_g \frac{\mathbb{E}|X_1|^3 + \sqrt{\log n}}{\sqrt{n}},$$

where C_g is a constant, independent of n, yet depending on the (carefully defined) *smoothness properties* of g. Among the applications and extensions considered by Barbour are an analysis of the empirical distribution function of i.i.d. random variables and the Wald–Wolfowitz theorem often used to construct tests in non-parametric statistics [41].

Our aim in this paper is to extend the results of [1] to approximations of scaled sums of univariate and *multivariate* random variables with different *dependence structures* by univariate and *multivariate* Wiener processes.

1.1. Motivation

Functional limit results play an important role in applied fields. Researchers often choose to model discrete phenomena with continuous processes arising as scaling limits of discrete ones. The reason is that those scaling limits may be studied using stochastic analysis and are more robust to changes in local details. Questions about the rate of convergence in functional limit results are equivalent to ones about the error those researchers make when doing so. Obtaining bounds on a certain distance between the scaled discrete and the limiting continuous processes provides a way of quantifying this error.

Our motivation in this paper comes from the desire to fill in a gap in the theory but we are also motivated by examples related to applications.

One of those, studied in the example in Section 4 of this paper, considers *exceedances of the m-scans process*. For a sequence of i.i.d. random variables $X_1, X_2, \ldots$, the one-dimensional m-scans process is given by $R_i = \sum_{k=0}^{m-1} X_{i+k}$. The number of its exceedances of a real number a is given by

$$Y = \sum_{i=1}^{n} \mathbb{1}[R_i > a].$$

As noted in [6, Example 9.2], this statistic has been studied by many authors, including [17] and [29]. It is of high importance in many areas of applied statistics and has been used, for instance, to evaluate the significance of observed inhomogeneities in the distribution of markers along the length of long DNA sequences (see [10,23]). Y may be normalised and centralised and then shown to converge in distribution to the standard normal law. Berry–Esseen bounds on the rate of this convergence have been found in [11, Theorem 4.1] and [6, Example 9.2]. We are interested in studying the functional convergence of a multidimensional version of Y.

Another example concerns *bivariate U-statistics* and is treated in Theorem 3.9 of this paper. Bivariate U-statistics are defined to be random variables of the form:

$$S_n^2(h) = \sum_{1 \leq i_1 < i_2 \leq n} h(X_{i_1}, X_{i_2}), \quad n \geq 1$$

for a symmetric real (or complex) function h on $\mathcal{S}^2$ (where $\mathcal{S}$ is some measurable space) and a sequence of i.i.d. random variables $(X_i)_{i \geq 1}$ taking values in $\mathcal{S}$. Because of their appealing properties, they are central objects in the field of Mathematical Statistics, as described in [27] and many commonly used statistics can be expressed in terms of certain U-statistics or approximated by them. They also appear in decompositions of more general statistics into sums of terms of a simpler form (see, e.g. [36, Chapter 6] or [35] and [40]) and play an important role in the study of random fields (see, e.g. [8, Chapter 4]). The appealing properties of *non-degenerate bivariate U-statistics*, i.e. those such that, for

$$w(x) = \mathbb{E}h(x, X_1),$$

$0 < \mathrm{Var}[w(X_1)] < \infty$, include their asymptotic behaviour. It can be described by a Strong Law of Large Numbers [20], a central limit theorem [19] or the functional central limit theorem (e.g. [21, Chapter XI]), which will be studied in this paper. Other interesting results include those connected to large deviations for U-statistics (see [14]), Berry–Esseen-type bounds (see [7]) and other bounds on the speed of convergence in the U-statistic CLT (see [33]). *Degenerate U-statistics* have also received much attention in the recent years with [12] providing bounds on the speed of convergence in de Jong's theorem [22] and proving its multidimensional version.

Our theoretical motivation is expressed in Proposition 3.5 of this paper. It seems natural to ask whether techniques similar to those of [1] may be used to study a process of the form

$$t \mapsto n^{-1/2} \sum_{i=1}^{\lfloor nt \rfloor} X_i, \quad t \in [0, 1] \tag{1.2}$$

where $\{X_i : i = 1, \ldots, n\}$ is a collection of i.i.d. random vectors in $\mathbb{R}^p$ for $p > 1$ with a given covariance matrix Σ. Interesting questions arising include those about the rate of convergence of the process in (1.2) to the correlated p-dimensional Brownian motion created from a standard Brownian motion $\mathbf{B}$ by premultiplying it by $\Sigma^{1/2}$. In this context, the role played by Σ in the quality of this approximation seems worth paying attention to.

1.2. Contribution of the paper

The main achievements of the paper are the following:

(a) A very general result providing a bound on the distance between a process of the form

$$\mathbf{Y}_n(t) = \left(\sum_{i=1}^{\lambda_1} X_{i,1} J_{i,1}(t), \ldots, \sum_{i=1}^{\lambda_p} X_{i,p} J_{i,p}(t) \right), \quad t \in [0, 1],$$

where:

- the numbers λ_j are such that $\lambda_j \leq n$;
- p is a fixed positive integer;

- the collection of vectors $X_i = (X_{i,1}, \ldots, X_{i,p})$ for $i = 1, \ldots, n$ is allowed to be *dependent* and those vectors themselves are allowed to have non-identity covariance matrices;
- the collection of (possibly random) functions

$$\left\{ J_{i,k} \in D([0,1], \mathbb{R}) : i = 1, \ldots n, k = 1, \ldots, p \right\}$$

  is independent of the collection of vectors $(X_i)_{i=1}^{n}$ from the previous point;

and a correlated p-dimensional Brownian motion. The bound is presented in Theorem 3.1 and provides a substantial extension of the result of [1], which bounds the rate of convergence in the classical, one-dimensional Donsker's invariance principle.
(b) A novel functional central limit theorem involving the number of exceedances in the multidimensional m-scans process, together with bounds on the rate of convergence, presented in the example in Section 4.
(c) A novel bound on the rate of convergence in the functional central limit theorem for non-degenerate, bivariate U-statistics (for a classical proof of the theorem see, for instance, [18]), which is presented in Theorem 3.9.
(d) A technical result, presented in Proposition 2.3, showing that our bounds' converging to zero implies weak convergence of the underlying processes with respect to the Skorokhod and uniform topologies. This result is a direct extension of [3, Proposition 3.1] to the multidimensional setting.

We provide explicit values for all the constants appearing in our bounds. To our best knowledge, none of the authors who have considered functional approximations with Stein's method so far have done so. We do it as we hope that this will make our results more powerful when used in applications.

The technique which is central in obtaining all the bounds is Stein's method.

1.3. Stein's method for distributional approximation

In [38] it is observed that a random variable Z has standard normal law if and only if $\mathbb{E}Zf(Z) = \mathbb{E}f'(Z)$ for all smooth functions f. Therefore, if, for a random variable W with mean zero and unit variance, $\mathbb{E}f'(W) - \mathbb{E}Wf(W)$ is close to zero for a large class of functions f, then the law of W should be approximately Gaussian. This leads to a method of bounding the speed of convergence to the normal distribution. Instead of evaluating $|\mathbb{E}h(W) - \mathbb{E}h(Z)|$ directly for a given function h, one can first find an $f = f_h$ solving the following *Stein equation*:

$$f'(w) - wf(w) = h(w) - \mathbb{E}h(Z)$$

and then find a bound on $|\mathbb{E}f'(W) - \mathbb{E}Wf(W)|$. This approach, called *Stein's method*, often turns out to be surprisingly easy and has also proved to be useful for approximations by distributions other than normal.

The aim of the generalised version of Stein's method is to find a bound for the quantity $|\mathbb{E}_{\nu_n}h - \mathbb{E}_\mu h|$, where μ is the target (known) distribution, ν_n is the approximating law and h is chosen from a suitable class of real-valued test functions $\mathcal{H}$. The procedure can be described in terms of three steps. First, an operator $\mathcal{A}$ acting on a class of real-valued functions is sought, such that

$$(\forall f \in \text{Domain}(\mathcal{A}) \quad \mathbb{E}_\nu \mathcal{A}f = 0) \quad \Longleftrightarrow \quad \nu = \mu,$$

where μ is the target distribution. Then, for a given function $h \in \mathcal{H}$, the Stein equation

$$\mathcal{A}f = h - \mathbb{E}_\mu h$$

has to be solved. Finally, using properties of the solution and various mathematical tools (among which the most popular are Taylor's expansions in the continuous case, Malliavin calculus, as described in [30], and coupling methods), an explicit bound is sought for the quantity $|\mathbb{E}_{\nu_n} \mathcal{A}f_h|$.

An accessible account of the method can be found, for example, in the surveys [28] and [34] as well as the books [2] and [6], which treat the cases of Poisson and normal approximation, respectively, in detail. The reference [39] is a database of information and publications connected to Stein's method.

Approximations by laws of diffusion processes have not been covered in the Stein's method literature very widely, with the notable exceptions of [1,3,9,37] and recently [4,24,25]. Our aim in this paper is to develop it in a direction not previously explored by other authors while completely natural given the direction in which the finite-dimensional Stein's method literature has evolved.

1.4. Structure of the paper

In Section 2 we define the spaces of test functions we will be working with and the corresponding norms which will appear in the bounds. We also present Proposition 2.3 giving circumstances under which the bounds obtained later in the paper converging to zero imply weak convergence of the considered probability distributions. Section 3 gives statements of the main results of the paper, mentioned above. Section 4 presents the example concerning exceedances of an m-scans process. Section 5 contains all the proofs preceded by finding the Stein equation for approximation by the law of interest, solving it and examining properties of the solutions. In the appendix we present the proof of the aforementioned Proposition 2.3.

2. Notation and spaces M, M^1, M^2 and M^0

The following notation is used throughout the paper. For a function w defined on the interval $[0, 1]$ and taking values in a Euclidean space, we define

$$\|w\| = \sup_{t \in [0,1]} |w(t)|,$$

where $|\cdot|$ denotes the Euclidean norm. We also let p be an integer such that $p \geq 1$ and $D^p = D([0, 1], \mathbb{R}^p)$ be the Skorokhod space of all càdlàg functions on $[0, 1]$ taking values in $\mathbb{R}^p$. In the literature, this space is usually equipped with the *Skorokhod topology* generated by the *Skorokhod metric* σ given by

$$\sigma(w, v) = \inf_{\lambda \in \Lambda} \max\{\|\lambda - I\|, \|w - v \circ \lambda\|\},$$

where I is the identity function and Λ is the set of all strictly increasing continuous bijections on $[0, 1]$. We will most often consider the topology generated by the supremum norm, though.

In the sequel, for $i = 1, \ldots, p$, e_i will denote the ith unit vector of the canonical basis of $\mathbb{R}^p$ and the ith component of $x \in \mathbb{R}^p$ will be represented by $x^{(i)}$, i.e. $x = \left(x^{(1)}, \ldots, x^{(p)}\right)$.

Let $p \in \mathbb{N}$. Let us define:

$$\|f\|_L := \sup_{w \in D^p} \frac{|f(w)|}{1 + \|w\|^3},$$

and let L be the Banach space of continuous functions $f : D^p \to \mathbb{R}$ such that $\|f\|_L < \infty$. Following [1], we now define $M \subset L$ to be the set of the twice Fréchet differentiable functions f, such that:

$$\|D^2 f(w + h) - D^2 f(w)\| \le k_f \|h\|, \tag{2.1}$$

for some constant k_f, uniformly in $w, h \in D^p$. By $D^k f$ we mean the kth Fréchet derivative of f and the norm of k-linear form B on L is defined to be

$$\|B\| = \sup_{\|h_i\| \le 1 \,\forall i=1,\dots,k} |B[h_1, \dots, h_k]|,$$

where

$$B[h_1, \dots, h_k]$$

denotes B applied to arguments $h_1, \dots, h_k$. Note the following lemma, which can be proved in an analogous way to that used to show (2.6) and (2.7) of [1]. We omit the proof here.

Lemma 2.1. *For every $g \in M$, let:*

$$\|g\|_M := \sup_{w \in D^p} \frac{|g(w)|}{1 + \|w\|^3} + \sup_{w \in D^p} \frac{\|Dg(w)\|}{1 + \|w\|^2} + \sup_{w \in D^p} \frac{\|D^2 g(w)\|}{1 + \|w\|}$$
$$+ \sup_{w,h \in D^p} \frac{\|D^2 g(w + h) - D^2 g(w)\|}{\|h\|}.$$

Then, for all $g \in M$, we have $\|g\|_M < \infty$.

For future reference, we let $M^1 \subset M$ be the class of functionals $g \in M$ such that:

$$\|g\|_{M^1} := \sup_{w \in D^p} \frac{|g(w)|}{1 + \|w\|^3} + \sup_{w \in D^p} \|Dg(w)\| + \sup_{w \in D^p} \|D^2 g(w)\|$$
$$+ \sup_{w,h \in D^p} \frac{\|D^2 g(w + h) - D^2 g(w)\|}{\|h\|} < \infty. \tag{2.2}$$

and $M^2 \subset M$ be the class of functionals $g \in M$ such that:

$$\|g\|_{M^2} := \sup_{w \in D^p} \frac{|g(w)|}{1 + \|w\|^3} + \sup_{w \in D^p} \frac{\|Dg(w)\|}{1 + \|w\|} + \sup_{w \in D^p} \frac{\|D^2 g(w)\|}{1 + \|w\|}$$
$$+ \sup_{w,h \in D^p} \frac{\|D^2 g(w + h) - D^2 g(w)\|}{\|h\|} < \infty. \tag{2.3}$$

We also let M^0 be the class of functionals $g \in M$ such that:

$$\|g\|_{M^0} := \sup_{w \in D^p} |g(w)| + \sup_{w \in D^p} \|Dg(w)\| + \sup_{w \in D^p} \|D^2 g(w)\|$$
$$+ \sup_{w,h \in D^p} \frac{\|D^2 g(w + h) - D^2 g(w)\|}{\|h\|} < \infty.$$

We note that $M^0 \subset M^1 \subset M^2 \subset M$. We shall refer to those different classes of functions in the results presented in the remainder of this paper. In each case we aim to obtain our bounds for the largest possible class, yet it is not always possible to do so for class M or even M^2. Hence, the introduction of the above presented restrictions of M is necessary for a recovery of the full strength of our results.

The next proposition is a p-dimensional version of [3, Proposition 3.1] and shows conditions, under which convergence of the sequence of expectations of a functional g under the approximating measures to the expectation of g under the target measure for all $g \in M^0$ implies weak convergence of the measures of interest. The proposition will be later used to conclude weak convergence from bounds derived in the theorems of the next section. Its proof can be found in the Appendix.

Definition 2.2. $Y \in D([0, 1], \mathbb{R}^p)$ is called piecewise constant if $[0, 1]$ can be divided into intervals of constancy $[a_k, a_{k+1})$ such that $(Y(t_1) - Y(t_2)) = 0$ for all $t_1, t_2 \in [a_k, a_{k+1})$.

Proposition 2.3.   *Suppose that, for each $n \geq 1$, the random element $\mathbf{Y}_n$ of D^p is piecewise constant and let $r_n > 0$ be such that the intervals of constancy are of length at least r_n. Let $(\mathbf{Z}_n)_{n \geq 1}$ be random elements of D^p converging in distribution in D^p, with respect to the Skorokhod topology, to a random element $\mathbf{Z} \in C([0, 1], \mathbb{R}^p) \subset D^p$. If there exists a sequence $(\kappa_n)_{n \geq 1}$ such that $\kappa_n \log^2(1/r_n) \xrightarrow{n \to \infty} 0$ and*

$$|\mathbb{E}g(\mathbf{Y}_n) - \mathbb{E}g(\mathbf{Z}_n)| \leq C\kappa_n \|g\|_{M^0} \tag{2.4}$$

for each $g \in M^0$ then $\mathbf{Y}_n \Rightarrow \mathbf{Z}$ (converges weakly) in D^p, in both the uniform and the Skorokhod topology.

Remark 2.4. The formulation of Proposition 2.3 is almost identical to that of [3, Proposition 3.1] with the only difference being that $\mathbf{Y}_n$ and $\mathbf{Z}_n$ are allowed to be p-dimensional for $p > 1$. For completeness, the appendix contains a more detailed proof than the one presented in [3], which may be used by the reader to derive extensions or other versions of the result.

3. Main results

3.1. Scaled sum of dependent vectors with dependent components

Theorem 3.1 studies a scaled sum of locally dependent terms whose components are (strongly) dependent. It bounds the error on its approximation by a correlated Brownian motion for test functions in M^1.

Theorem 3.1 (*Dependent Components and Locally Dependent Summands*)**.** *Let n and p be positive integers. Consider an array of mean-zero random variables*

$$\{X_{i,j} : i = 1, \ldots, n, j = 1, \ldots, p\},$$

with a positive definite covariance matrix $\tilde{\Sigma}_n$. Let

(a) *$\lambda_j \leq n$, for $j = 1, \ldots p$, be deterministic positive integers;*
(b) *$\mathbb{A}_i \subset \{1, 2, \ldots, n\}$, for $i = 1, \ldots, n$ be a set such that $X_i = (X_{i,1}, \ldots, X_{i,p})$ is independent of $\{X_j : j \in \mathbb{A}_i^c\}$;*
(c) *$\mathbb{A}_{ij} \subset \{1, \ldots, n\}$, for $i, j = 1, \ldots n$ be a set such that (X_i, X_j) and $\{X_k : k \in A_{ij}^c\}$ are independent.*
(d) *$J_{i,k} \in D([0, 1], \mathbb{R})$ for $i = 1, \ldots, n$ and $k = 1, \ldots, p$, be (possibly random) functions, independent of the family $\{X_{i,k} : i = 1, \ldots, n, k = 1, \ldots, p\}$.*

Assume that:

$$\sup_{\substack{i_1,i_2,i_3\in\{1,\ldots,n\} \\ k_1,k_2,k_3\in\{1,\ldots,p\}}} \mathbb{E}\left[\|J_{i_1,k_1}\|\|J_{i_2,k_2}\|\|J_{i_3,k_3}\|\right] < \infty.$$

Let

$$\mathbf{Y}_n(t) = \left(\sum_{i=1}^{\lambda_1} X_{i,1}J_{i,1}(t), \ldots, \sum_{i=1}^{\lambda_p} X_{i,p}J_{i,p}(t)\right), \quad t\in[0,1].$$

Furthermore, for a standard p-dimensional Brownian motion **B** *and a positive definite covariance matrix* $\Sigma\in\mathbb{R}^{p\times p}$, *let* $\mathbf{Z}=\Sigma^{1/2}\mathbf{B}$. *Then, for any* $g\in M^1$, *as defined by* (2.2):

$$|\mathbb{E}g(\mathbf{Y}_n) - \mathbb{E}g(\mathbf{Z})| \le \|g\|_{M^1}\sum_{i=1}^{7}\epsilon_i,$$

where:

$$\epsilon_1 = \frac{1}{6}\sum_{i=1}^{n}\mathbb{E}\left\{\left(\sum_{k,l,m=1}^{p}\left[(X_{i,k})^2\|J_{i,k}\|^2\mathbb{1}_{[1,\lambda_k]}(i)\left(\sum_{j\in\mathbb{A}_i}X_{j,l}\|J_{j,l}\|\mathbb{1}_{[1,\lambda_l]}(j)\right)^2\right.\right.\right.$$
$$\left.\left.\left.\cdot\left(\sum_{j\in\mathbb{A}_i}X_{j,m}\|J_{j,m}\|\mathbb{1}_{[1,\lambda_m]}(j)\right)^2\right]\right)^{1/2}\right\};$$

$$\epsilon_2 = \frac{1}{3}\sum_{i=1}^{n}\sum_{j\in\mathbb{A}_i}\sum_{k,l=1}^{p}\mathbb{E}\left\{\left[\sum_{m=1}^{p}\left(X_{i,k}\|J_{i,k}\|X_{j,l}\|J_{j,l}\|\mathbb{1}_{[1,\lambda_k]}(i)\mathbb{1}_{[1,\lambda_l]}(j)\right.\right.\right.$$
$$\left.\left.\left.\cdot\sum_{r\in\mathbb{A}_{ij}\cap\mathbb{A}_i^c}X_{r,m}\|J_{r,m}\|\mathbb{1}_{[1,\lambda_m]}(r)\right)^2\right]^{1/2}\right\};$$

$$\epsilon_3 = \frac{1}{3}\sum_{i=1}^{n}\sum_{j\in\mathbb{A}_i}\sum_{k,l=1}^{p}\left\{\left|\mathbb{E}\left[X_{i,k}X_{j,l}\right]\right|\mathbb{1}_{[1,\lambda_k]}(i)\mathbb{1}_{[1,\lambda_l]}(j)\right.$$
$$\left.\cdot\mathbb{E}\left[\|J_{i,k}\|\|J_{j,l}\|\sqrt{\sum_{m=1}^{p}\left(\sum_{r\in\mathbb{A}_{ij}}X_{r,m}\|J_{r,m}\|\mathbb{1}_{[1,\lambda_m]}(r)\right)^2}\right]\right\};$$

$$\epsilon_4 = \frac{1}{2}\sum_{k,l=1}^{p}\sum_{i=1}^{\lambda_k\wedge\lambda_l}\left|\frac{\Sigma_{k,l}}{\sqrt{\lambda_k\lambda_l}} - \mathbb{E}[X_{i,k}X_{i,l}]\right|;$$

$$\epsilon_5 = \frac{1}{2}\sum_{k,l=1}^{p}\sum_{i=1}^{\lambda_k}\sum_{j\in\mathbb{A}_i\setminus\{i\}}\left|\mathbb{E}[X_{i,k}X_{j,l}]\right|;$$

$$\epsilon_6 = \frac{6\sqrt{5}}{\sqrt{2\log 2}}\left(\sum_{i=1}^{p}\frac{\log(2\lambda_i)}{\lambda_i}\right)^{1/2}\left(\sum_{i=1}^{p}\Sigma_{i,i}\right)^{1/2};$$

$$\epsilon_7 = \sum_{k=1}^{p} \sum_{i=1}^{\lambda_k} \sqrt{\mathbb{E}\left[(X_{i,k})^2\right]} \mathbb{E}\left\|J_{i,k} - \mathbb{1}_{[i/\lambda_k,1]}\right\|.$$

Remark 3.2 (*Relevance of Terms in the Bound*).

(a) Terms $\epsilon_1, \epsilon_2, \epsilon_3$ correspond to a Berry–Esseen-type bound involving third moments of the summands, and also account for local dependence between the summands;

(b) Terms ϵ_4 and ϵ_5 involve a variance estimation with the latter corresponding to the off-diagonal terms of the covariance matrix of the summands, accounting for the dependence;

(c) Term ϵ_6 comes from estimates on the moments of the Brownian modulus of continuity and accounts for the transition from the Skorokhod space to the Wiener space of continuous functions;

(d) Term ϵ_7 describes the randomness of the functions $J_{i,k}$ and their distance from indicators $\mathbb{1}_{[i/\lambda_k,1]}$.

Remark 3.3 (*Convergence of the Bound and Process Weak Convergence*). By Proposition 2.3, if, in Theorem 3.1, $J_{i,k} = \mathbb{1}_{[i/\lambda_k,1]}$ for all $i = 1, \ldots, n$ and $k = 1, \ldots, p$ and the bound $\sum_{i=1}^{7} \epsilon_i$ converges to 0 faster than $\frac{1}{\log^2(\max(\lambda_1,\ldots,\lambda_p))}$, then $\mathbf{Y}_n$ converges to $\mathbf{Z}$ in distribution with respect to the uniform topology. We note that, in practice, one might expect that ϵ_4 and ϵ_5 will be the slowest vanishing terms.

Remark 3.4 (*Independent Summands*). If the summands are independent in Theorem 3.1, i.e. $\mathbb{A}_i = \{i\}$ for all i, then ϵ_2 and ϵ_5 disappear from the bound and ϵ_1 and ϵ_3 become simpler. The new bound takes the following form

$$|\mathbb{E}g(\mathbf{Y}_n) - \mathbb{E}g(\mathbf{Z})| \leq \|g\|_{M^1} \left(\epsilon_1 + \epsilon_3 + \epsilon_4 + \epsilon_6 + \epsilon_7\right),$$

where:

$$\epsilon_1 = \frac{1}{6} \sum_{i=1}^{n} \mathbb{E}\left\{\left[\sum_{k,l,m=1}^{p} \left(X_{i,k} X_{i,l} X_{i,m} \|J_{i,k}\| \|J_{i,l}\| \|J_{i,m}\| \mathbb{1}_{[1,\lambda_k]\cap[1,\lambda_l]\cap[1,\lambda_m]}(i)\right)^2\right]^{1/2}\right\};$$

$$\epsilon_3 = \frac{1}{3} \sum_{k,l=1}^{p} \sum_{i=1}^{\min(\lambda_k,\lambda_l)} \left\{\left|\mathbb{E}\left[X_{i,k} X_{i,l}\right]\right| \mathbb{E}\left[\|J_{i,k}\| \|J_{i,l}\| \sqrt{\sum_{m=1}^{p} \left(X_{i,m} \|J_{i,m}\| \mathbb{1}_{[1,\lambda_m]}(i)\right)^2}\right]\right\};$$

$$\epsilon_4 = \frac{1}{2} \sum_{k,l=1}^{p} \sum_{i=1}^{\min(\lambda_k,\lambda_l)} \left|\frac{\Sigma_{k,l}}{\sqrt{\lambda_k \lambda_l}} - \mathbb{E}[X_{i,k} X_{i,l}]\right|;$$

$$\epsilon_6 = \frac{6\sqrt{5}}{\sqrt{2\log 2}} \left(\sum_{i=1}^{p} \frac{\log(2\lambda_i)}{\lambda_i}\right)^{1/2} \left(\sum_{i=1}^{p} \Sigma_{i,i}\right)^{1/2};$$

$$\epsilon_7 = \sum_{k=1}^{p} \sum_{i=1}^{\lambda_k} \sqrt{\mathbb{E}\left[(X_{i,k})^2\right]} \mathbb{E}\left\|J_{i,k} - \mathbb{1}_{[i/\lambda_k,1]}\right\|.$$

In this case, it is also possible to derive a bound for the larger class of test functions M (see Section 2). A bound for such test functions, in the case of independent summands, is obtained in Proposition 3.5.

3.2. Scaled sum of independent vectors with dependent components

The next result treats quantitatively the case of independent p-dimensional terms with dependent components, whose scaled sum can be compared to a correlated p-dimensional Brownian motion:

Proposition 3.5 (*Independent Summands with Dependent Components*). *Suppose that $X_1, \ldots,$ X_n, where $X_i = \left(X_i^{(1)}, \ldots, X_i^{(p)}\right)$ for $i = 1, \ldots, n$, are i.i.d. random vectors in $\mathbb{R}^p$. Suppose that each has a positive definite symmetric covariance matrix $\Sigma \in \mathbb{R}^{p \times p}$ and mean zero. Let:*

$$\mathbf{Y}_n(t) = n^{-1/2} \sum_{i=1}^{\lfloor nt \rfloor} X_i, \quad t \in [0, 1]$$

and for $\mathbf{B}$, a standard p-dimensional Brownian motion, let $\mathbf{Z} = \Sigma^{1/2}\mathbf{B}$. Then, for any $g \in M$:

$$|\mathbb{E}g(\mathbf{Y}_n) - \mathbb{E}g(\mathbf{Z})|$$

$$\leq \|g\|_M n^{-1/2} \left\{ \sqrt{\log 2n} \left[\frac{6\sqrt{5}}{\sqrt{\pi \log 2}} \left(\sum_{i=1}^p \Sigma_{i,i} \right)^{1/2} + \frac{93 p^{1/2}}{\sqrt{2 \log 2}} \sum_{i=1}^p |\Sigma_{i,i}|^{3/2} \right] \right.$$

$$+ \frac{1}{6} \left(p^{1/2} \sum_{m=1}^p \mathbb{E} \left| X_1^{(m)} \right|^3 + 2 \sum_{k,l=1}^p |\Sigma_{k,l}| \left(\sum_{m=1}^p \mathbb{E} \left| X_1^{(m)} \right|^2 \right)^{1/2} \right)$$

$$\left. + n^{-1} (\log 2n)^{3/2} p^{1/2} \frac{2160}{\sqrt{\pi} (\log 2)^{3/2}} \sum_{i=1}^p |\Sigma_{i,i}|^{3/2} \right\}.$$

Remark 3.6. The bound in Proposition 3.5 is of order $\frac{\sqrt{\log n}}{\sqrt{n}}$. We are not aware of any reference providing a bound in a similar setup (i.e. in a multidimensional version of Donsker's theorem) but we note that our bound is of the same order as the bound derived in [1] for one-dimensional Donsker's theorem.

Remark 3.7. If the components are uncorrelated and scaled in Proposition 3.5, i.e. $\Sigma = I_{p \times p}$, then the bound simplifies in the following way:

$$|\mathbb{E}g(\mathbf{Y}_n) - \mathbb{E}g(\mathbf{Z})|$$

$$\leq \|g\|_M n^{-1/2} \left\{ \sqrt{\log 2n} \left[\frac{6\sqrt{5} p^{1/2}}{\sqrt{2 \log 2}} + \frac{93 p^{3/2}}{\sqrt{\pi \log 2}} \right] \right.$$

$$\left. + \frac{1}{6} \left(p^{1/2} \sum_{m=1}^p \mathbb{E} \left| X_1^{(m)} \right|^3 + 2 p^{3/2} \right) + n^{-1} (\log 2n)^{3/2} p^{3/2} \frac{2160}{\sqrt{\pi} (\log 2)^{3/2}} \right\}.$$

Remark 3.8. For fixed p, by Proposition 2.3, Proposition 3.5 implies that $\mathbf{Y}_n$ converges in distribution to $\mathbf{Z}$ in the uniform topology as the bound is of order $\frac{\sqrt{\log n}}{\sqrt{n}}$. If one made p depend on n the bound would also converge to zero as $n \to \infty$ as long as $p = o\left(n^{1/5}\right)$.

3.3. Non-degenerate bivariate U-statistics

The next result will be proved using ideas similar to those used to prove Theorem 3.1. It treats non-degenerate bivariate U-statistics. Those, as observed for instance in [18, Corollary 1], after proper rescaling, represent a process created out of globally dependent summands and converge to standard Brownian motion in distribution under certain conditions. We find a bound for the rate of this convergence.

We note that bivariate U-statistics are defined to be random variables of the form:

$$S_n^2(h) = \sum_{1 \le i_1 < i_2 \le n} h(X_{i_1}, X_{i_2}), \quad n \ge 1$$

for a symmetric real (or complex) function h on $\mathcal{S}^2$ (where $\mathcal{S}$ is some measurable space) and a sequence of i.i.d. random variables $(X_i)_{i \ge 1}$ taking values in $\mathcal{S}$. Here, we only consider non-degenerate U-statistics, i.e. those with $0 < \sigma_w^2 = \mathrm{Var}(w(X_1)) < \infty$, where $w(x) = \mathbb{E}[h(X_1, x)]$. The reason is that in the case of degenerate ones (i.e. those satisfying $\mathrm{Var}(w(X_1)) = 0$) the limit in the invariance principle is non-Gaussian (see [18, Corollary 1]), which is beyond the scope of this paper.

Theorem 3.9 (*Non-degenerate Bivariate U-statistics*). *Let $X_1, X_2, \ldots$ be i.i.d. random variables taking values in some measurable space $\mathcal{S}$ and let $h : \mathcal{S}^2 \to \mathbb{R}$ be a symmetric function such that $\mathbb{E}[h(X_1, X_2)] = 0$, $\mathbb{E}[h^2(X_1, X_2)] = \sigma_h^2 < \infty$. Also, suppose that, for the function $w(x) = \mathbb{E}[h(X_1, x)]$, we have that: $0 < \sigma_w^2 = \mathrm{Var}(w(X_1))$ and $\mathbb{E}|w(X_1)|^3 < \infty$. Let:*

$$\mathbf{Y}_n(t) = \frac{n^{-3/2}}{\sigma_w t} \sum_{1 \le i_1 < i_2 \le \lfloor nt \rfloor} h(X_{i_1}, X_{i_2}), \quad t \in [0, 1]$$

and let $\mathbf{Z}$ be a standard Brownian motion. Then, for any $g \in M^2$, as defined by (2.3):

$$|\mathbb{E}g(\mathbf{Y}_n) - \mathbb{E}g(\mathbf{Z})| \le \|g\|_{M^2} n^{-1/2} \left[\left(141 + 16 \frac{\sigma_h^2}{\sigma_w^2} + 12 \left(\frac{\sigma_h^2}{\sigma_w^2} - 2 \right)^{1/2} \right) \sqrt{\log 3n} \right.$$
$$\left. + 43 + \frac{\mathbb{E}|w(X_1)|^3 + 2\sigma_w^2 \mathbb{E}|w(X_1)|}{6\sigma_w^3} \right].$$

Remark 3.10 (*Discussion of the Bound*). The term $\frac{\mathbb{E}|w(X_1)|^3 + 2\sigma_w^2 \mathbb{E}|w(X_1)|}{6\sigma_w^3}$ appearing in the bound comes from the comparison of the process given by

$$\tilde{\mathbf{Y}}_n(t) = \frac{n^{-3/2}}{\sigma_w t} \sum_{1 \le i_1 < i_2 \le \lfloor nt \rfloor} \left(w(X_{i_1}) + w(X_{i_2}) \right), \quad t \in [0, 1]$$

and a piecewise constant Gaussian process. It involves a Berry–Esseen-type third absolute moment component. The remaining terms come from the comparison of $\mathbf{Y}_n$ and $\tilde{\mathbf{Y}}_n$ and from the comparison of the piecewise constant Gaussian process and Brownian motion, for which the Brownian modulus of continuity is used.

The bound is of order $\frac{\sqrt{\log n}}{\sqrt{n}}$. We are not aware of any reference providing a bound on the rate of *functional* convergence of non-degenerate U-statistics but we note that our bound

is of the same order as the bound obtained in [1] for the rate of convergence in the classical Donsker's theorem.

Remark 3.11. By Proposition 2.3, Theorem 3.9 implies that $\mathbf{Y}_n$ converges in distribution to $\mathbf{Z}$ in the uniform (and Skorokhod) topology.

Remark 3.12. The constants in Theorems 3.1, 3.9 and Proposition 3.5 are not optimal as they are often estimated in a crude manner in the proofs presented in the section below. The constants are, however, expressed explicitly, which is often not the case in related pieces of literature. We also have no information about the optimality of the orders of the obtained bounds.

4. Example: Exceedances of the m-scans process

Consider an extension of the one-dimensional result [6, Example 9.2, p. 254] to the multidimensional and functional setting. For $j = 1, 2, \ldots$, let $V_j = \left(V_{j,1}, \ldots, V_{j,p}\right)$ be i.i.d. random vectors in $\mathbb{R}^p$. For $k = 1, \ldots, p$ and $i = 1, 2, \ldots$ let $R_{i,k} = \sum_{l=0}^{m-1} V_{i+l,k}$ be an m-scans process. Let $a = (a_1, \ldots, a_p) \in \mathbb{R}^p$ and suppose that $n > m$.

For $k = 1, \ldots, p$, let $\pi_k = \mathbb{P}(R_{1,k} \leq a_k)$ and for $i = 1, \ldots, n$ and $k = 1, \ldots, p$, let

$$X_{i,k} = \frac{1}{n} \left(\sum_{j=1}^{n} \mathbb{1}[R_{n(i-1)+j,k} \leq a_k] \right) - \pi_k.$$

Extending [11, (4.1)], we have that, for $k, l = 1, \ldots, p$ and for $\psi_{k,l}(d) = \mathbb{P}\left[R_{d+1,k} \leq a_k, R_{1,l} \leq a_l\right] - \pi_k \pi_l$,

$$\mathbb{E}\left[X_{i,k} X_{i,l}\right] = \frac{1}{n} \left(\psi_{k,l}(0) + \sum_{d=1}^{m-1} \left(1 - \frac{d}{n}\right) \left(\psi_{l,k}(d) + \psi_{k,l}(d)\right) \right). \tag{4.1}$$

Let $X_i = (X_{i,1}, \ldots, X_{i,p})$ for $i = 1, \ldots, n$. Note that $\mathbb{A}_i = \{i - 1, i, i + 1\}$ satisfies the requirement that X_i is independent of $\{X_j : j \in \mathbb{A}_i^c\}$ and that we can take $\mathbb{A}_{ij} = \mathbb{A}_i \cup \mathbb{A}_j$. Furthermore, for all $k, l \in \{1, \ldots, p\}$,

$$\mathbb{E}\left[X_{i,k} X_{i+1,l}\right] = \frac{1}{n^2} \sum_{d=1}^{m-1} d\psi_{k,l}(d). \tag{4.2}$$

Consider

$$\mathbf{Y}_n(t) = \sum_{i=1}^{\lfloor nt \rfloor} \left(X_{i,1}, \ldots, X_{i,p}\right) \quad t \in [0, 1].$$

Let $\Sigma \in \mathbb{R}^{p \times p}$ be given by

$$\Sigma_{k,l} = \psi_{k,l}(0) + \sum_{d=1}^{m-1} \left(\psi_{l,k}(d) + \psi_{k,l}(d)\right). \tag{4.3}$$

We will bound the distance between $\mathbf{Y}_n$ and $\mathbf{Z} = \Sigma^{1/2}\mathbf{B}$, where $\mathbf{B}$ is a standard p-dimensional Brownian motion. Using the notation of Theorem 3.1, note that for all $k \in \{1, \ldots, p\}$, $\lambda_k = n$, for all $i \in \{1, \ldots, n\}$, $J_{i,k} = \mathbb{1}_{[i/n,1]}$ and

(1) By Cauchy–Schwarz and Jensen inequalities and (4.1),

$$
\epsilon_1 \leq \frac{3}{2n^{1/2}} \sum_{k,l,r=1}^{p} \left\{ \left(\psi_{k,k}(0) + 2\sum_{d=1}^{m-1} \left(1 - \frac{d}{n}\right) (\psi_{k,k}(d)) \right)^{1/2} \right.
$$
$$
\cdot \left(\psi_{l,l}(0) + 2\sum_{d=1}^{m-1} \left(1 - \frac{d}{n}\right) (\psi_{l,l}(d)) \right)^{1/2}
$$
$$
\left. \cdot \left(\psi_{r,r}(0) + 2\sum_{d=1}^{m-1} \left(1 - \frac{d}{n}\right) (\psi_{r,r}(d)) \right)^{1/2} \right\};
$$

(2) By Cauchy–Schwarz and Jensen inequalities and (4.1),

$$
\epsilon_2 \leq \frac{2}{3n^{1/2}} \sum_{k,l,r=1}^{p} \left\{ \left(\psi_{k,k}(0) + 2\sum_{d=1}^{m-1} \left(1 - \frac{d}{n}\right) (\psi_{k,k}(d)) \right)^{1/2} \right.
$$
$$
\cdot \left(\psi_{l,l}(0) + 2\sum_{d=1}^{m-1} \left(1 - \frac{d}{n}\right) (\psi_{l,l}(d)) \right)^{1/2}
$$
$$
\left. \cdot \left(\psi_{r,r}(0) + 2\sum_{d=1}^{m-1} \left(1 - \frac{d}{n}\right) (\psi_{r,r}(d)) \right)^{1/2} \right\};
$$

(3) By Cauchy–Schwarz and Jensen inequalities and (4.1),

$$
\epsilon_3 \leq \frac{2}{n^{1/2}} \sum_{k,l,r=1}^{p} \left\{ \left(\psi_{k,k}(0) + 2\sum_{d=1}^{m-1} \left(1 - \frac{d}{n}\right) (\psi_{k,k}(d)) \right)^{1/2} \right.
$$
$$
\cdot \left(\psi_{l,l}(0) + 2\sum_{d=1}^{m-1} \left(1 - \frac{d}{n}\right) (\psi_{l,l}(d)) \right)^{1/2}
$$
$$
\left. \cdot \left(\psi_{r,r}(0) + 2\sum_{d=1}^{m-1} \left(1 - \frac{d}{n}\right) (\psi_{r,r}(d)) \right)^{1/2} \right\};
$$

(4) By (4.1) and (4.3),

$$
\epsilon_4 = \frac{1}{2n} \sum_{k,l=1}^{p} \left| \sum_{d=1}^{m-1} d(\psi_{l,k}(d) + \psi_{k,l}(d)) \right|;
$$

(5) By (4.2),

$$
\epsilon_5 \leq \frac{1}{n} \sum_{l,k=1}^{p} \sum_{d=1}^{m-1} d\psi_{k,l}(d);
$$

(6) By (4.3),

$$
\epsilon_6 = \frac{6\sqrt{5}p^{1/2}}{\sqrt{2\log 2}} \frac{\sqrt{\log(2n)}}{\sqrt{n}} \left[\sum_{k=1}^{p} \left(\psi_{k,k}(0) + 2\sum_{d=1}^{m-1} \psi_{k,k}(d) \right) \right]^{1/2};
$$

(7) Since for all $k \in \{1, \ldots, p\}$ and $i \in \{1, \ldots, n\}$, $J_{i,k} = \mathbb{1}_{[i/n,1]}$,

$$\epsilon_7 = 0.$$

By Theorem 3.1, for any $g \in M^1$, as defined in (2.2),

$$|\mathbb{E}g(\mathbf{Y}_n) - \mathbb{E}g(\mathbf{Z})| \leq \|g\|_{M^1} \sum_{i=1}^{7} \epsilon_i,$$

which gives the desired bound. The bound clearly approaches zero faster than $\log^{-2}(n)$, as $n \to \infty$. Indeed, terms $\epsilon_1, \epsilon_2, \epsilon_3$ converge to zero at rate $n^{-1/2}$, ϵ_4 and ϵ_5 do so at rate n^{-1}, ϵ_6 at rate $\frac{\sqrt{\log n}}{\sqrt{n}}$ and $\epsilon_7 = 0$. This, by Proposition 2.3, implies that $\mathbf{Y}_n$ converges in distribution to $\mathbf{Z}$ with respect to the uniform topology.

5. Proofs of the main results

The main tool used in the proofs of Theorems 3.1, 3.9 and Proposition 3.5 is Stein's method. It can be used in a surprisingly easy way to find a distance of the processes of interest from certain scaled sums of Gaussian random variables, which approximate the limiting continuous Gaussian process.

First, we set up Stein's method for distributions of certain D^p-valued random objects expressed as scaled sums of Gaussian random variables. Using a collection of Ornstein–Uhlenbeck processes with a Gaussian stationary law, we will construct a process whose stationary law is that of our target distribution. Then, we will find the infinitesimal generator $\mathcal{A}$ of that process and deduce that $\mathcal{A}g = g - \mathbb{E}_\mu g$ can be used as our Stein equation, where μ is the target law. This follows from the fact that $\mathbb{E}_\mu \mathcal{A}g = 0$ for all g in the domain of $\mathcal{A}$. We will then solve the Stein equation for all $g \in M$, using the analysis of [26], and use some appealing properties of the Ornstein–Uhlenbeck semigroup to prove bounds on the derivatives of the solution.

5.1. Setting up Stein's method

Let $n, p \in \mathbb{N}_+$ and let $\tilde{Z}_{i,k}$'s be centred Gaussian random variables for $i = 1, \ldots, n$, $k = 1, \ldots, p$. Suppose that

(a) the covariance matrix of $\left(\tilde{Z}_{1,1}, \ldots, \tilde{Z}_{1,p}, \tilde{Z}_{2,1}, \ldots, \tilde{Z}_{2,p}, \ldots, \tilde{Z}_{n,1}, \ldots, \tilde{Z}_{n,p} \right)$ is given by $\Sigma_n \in \mathbb{R}^{(np) \times (np)}$;

(b) $\{J_{i,k} \in D([0,1], \mathbb{R}) : i = 1, \ldots, n, \ k = 1, \ldots, p\}$ is a collection of functions independent of $\{\tilde{Z}_{i,k} : i = 1, \ldots, n, \ k = 1, \ldots, p\}$;

(c) $\lambda_k \leq n$, for all $k = 1, \ldots, p$.

Let

$$\mathbf{D}_n(t) = \left(\sum_{i=1}^{\lambda_1} \tilde{Z}_{i,1} J_{i,1}(t), \ldots, \sum_{i=1}^{\lambda_p} \tilde{Z}_{i,p} J_{i,p}(t) \right), \quad t \in [0,1], \tag{5.1}$$

Now let $\{(\mathscr{X}_{i,j}(u), u \geq 0) : i = 1, \ldots, n, \ j = 1, \ldots, p\}$ be an array of i.i.d. Ornstein–Uhlenbeck processes with stationary law $\mathcal{N}(0,1)$, i.e. independent processes such that each weakly solves the following stochastic differential equation

$$dx_t = -x_t \, dt + \sqrt{2} \, dB(t), \quad x_0 \sim \mathcal{N}(0,1),$$

for $B(t)$, $t \geq 0$ denoting the standard Wiener process. Suppose that the collection $\{(\mathscr{X}_{i,j}(u), u \geq 0) : i = 1, \ldots, n, j = 1, \ldots, p\}$ is independent of the collection $\{J_{i,k} : i = 1, \ldots, n, k = 1, \ldots, p\}$. Consider:

$$\tilde{\mathscr{U}}(u) = (\Sigma_n)^{1/2} \left(\mathscr{X}_{1,1}(u), \ldots, \mathscr{X}_{1,p}(u), \mathscr{X}_{2,1}(u), \ldots, \mathscr{X}_{2,p}(u), \ldots, \right.$$
$$\left. \mathscr{X}_{n,1}(u), \ldots \mathscr{X}_{n,p}(u) \right)^T$$

for $u \geq 0$ and write $\mathscr{U}_{i,k}(u) = \left(\tilde{\mathscr{U}}(u) \right)_{p(i-1)+k}$ for $i = 1, \ldots, n$ and $k = 1, \ldots, p$. This notation is introduced for convenience, in order to define the following process:

$$\mathbf{W}_n(t, u) = \left(\sum_{i=1}^{\lambda_1} \mathscr{U}_{i,1}(u) J_{i,1}(t), \ldots, \sum_{i=1}^{\lambda_p} \mathscr{U}_{i,p}(u) J_{i,p}(t) \right), \qquad t \in [0, 1], \quad u \geq 0.$$

The stationary law of the process $(\mathbf{W}_n(\cdot, u))_{u \geq 0}$ is exactly the law of $\mathbf{D}_n$. We claim that:

Proposition 5.1. *The infinitesimal generator $\mathcal{A}_n$ of the process $(\mathbf{W}_n(\cdot, u))_{u \geq 0}$ acts on any $f \in M$ in the following way:*

$$\mathcal{A}_n f(w) = -Df(w)[w] + \mathbb{E} D^2 f(w) [\mathbf{D}_n, \mathbf{D}_n].$$

Remark 5.2. By definition, the first Fréchet derivative of a function, at a certain point, is a linear map, while the second Fréchet derivative of a function, at a certain point, is a bilinear map. In Proposition 5.1, and throughout this paper, $Df(w)[w]$ denotes the first Fréchet derivative of f, at w, applied to w and $D^2 f(w) [\mathbf{D}_n, \mathbf{D}_n]$ is the second Fréchet derivative of f, at w, applied to $\mathbf{D}_n$ and $\mathbf{D}_n$.

Remark 5.3. The generator in Proposition 5.1 can also be written in the following way:

$$\mathcal{A}_n f(w) = -Df(w)[w] + \sum_{k,l=1}^{p} \sum_{i=1}^{\lambda_k} \sum_{j=1}^{\lambda_l} (\Sigma_n)_{p(i-1)+k, p(j-1)+l} \, \mathbb{E} D^2 f(w) \left[e_k J_{i,k}, e_l J_{j,l} \right].$$

Let us prove a lemma that will be used in the proof of Proposition 5.1.

Lemma 5.4. *We have, for $u \geq 0$, $v \geq 0$:*

$$\mathbf{W}_n(\cdot, u + v) - e^{-v} \mathbf{W}_n(\cdot, u) \overset{\mathcal{D}}{=} \sigma(v) \mathbf{D}_n(\cdot)$$

for $\sigma^2(v) = 1 - e^{-2v}$.

Proof. We can construct i.i.d. standard Brownian motions $\mathscr{B}_{i,j}$ such that $\left(\mathscr{X}_{i,j}(u), u \geq 0 \right) = \left(e^{-u} \mathscr{B}_{i,j}(e^{2u}), u \geq 0 \right)$ (see, for instance [31, Subsection 4.4.3]). Then, writing $\mathbf{W}_n = \left(\mathbf{W}_n^{(1)}, \ldots, \mathbf{W}_n^{(p)} \right)$ and $\mathbf{D}_n = \left(\mathbf{D}_n^{(1)}, \ldots, \mathbf{D}_n^{(k)} \right)$ we obtain for all $k = 1, \ldots, p$:

$$\mathbf{W}_n^{(k)}(\cdot, u + v) - e^{-v} \mathbf{W}_n^{(k)}(\cdot, u)$$
$$= \sum_{i=1}^{\lambda_k} \left[\mathscr{U}_{i,k}(u + v) - e^{-v} \mathscr{U}_{i,k}(u) \right] J_{i,k}(\cdot)$$
$$= \sum_{i=1}^{\lambda_k} \left[\left(\tilde{\mathscr{U}}(u + v) \right)_{p(i-1)+k} - e^{-v} \left(\tilde{\mathscr{U}}(u) \right)_{p(i-1)+k} \right] J_{i,k}(\cdot)$$

$$\overset{(*)}{=} \sum_{j=1}^{n} \sum_{l=1}^{p} \sum_{i=1}^{\lambda_k} \left(\Sigma_n^{1/2}\right)_{p(i-1)+k,\, p(j-1)+l} \left[\mathscr{X}_{j,l}(u+v) - e^{-v}\mathscr{X}_{j,l}(u)\right] J_{i,k}(\cdot)$$

$$\overset{\mathcal{D}}{=} e^{-(u+v)} \sum_{j=1}^{n} \sum_{l=1}^{p} \sum_{i=1}^{\lambda_k} \left(\Sigma_n^{1/2}\right)_{p(i-1)+k,\, p(j-1)+l} \left[\mathscr{B}_{j,l}\left(e^{2(u+v)}\right) - \mathscr{B}_{j,l}\left(e^{2u}\right)\right] J_{i,k}(\cdot)$$

$$\overset{\mathcal{D}}{=} \sigma(v)\mathbf{D}_n^{(k)}(\cdot),$$

as $\mathscr{B}_{j,l}\left(e^{2(u+v)}\right) - \mathscr{B}_{j,l}\left(e^{2u}\right) \sim \mathcal{N}\left(0, e^{2(u+v)} - e^{2u}\right)$. In the above formula, the equality $(*)$ represents the matrix multiplication formula. $\quad\square$

Proof of Proposition 5.1. Note that the semigroup of $(\mathbf{W}_n(\cdot, u))_{u \geq 0}$, acting on L of Section 2 is defined by:

$$(T_{n,u}f)(w) := \mathbb{E}\left[f\left(\mathbf{W}_n(\cdot, u)\right) \mid \mathbf{W}_n(\cdot, 0) = w\right] = \mathbb{E}\left[f\left(we^{-u} + \sigma(u)\mathbf{D}_n(\cdot)\right)\right], \tag{5.2}$$

where the last equality follows from Lemma 5.4. By (5.2) and Lemma 2.1 we have that, for every $f \in M$:

$$\left| (T_{n,u}f)(w) - f(w) - \mathbb{E}Df(w)[\sigma(u)\mathbf{D}_n - w(1 - e^{-u})] \right.$$

$$\left. - \frac{1}{2}\mathbb{E}D^2 f(w)\left[\sigma(u)\mathbf{D}_n - w(1 - e^{-u}), \sigma(u)\mathbf{D}_n - w(1 - e^{-u})\right] \right|$$

$$\leq \|f\|_M \mathbb{E}\|\sigma(u)\mathbf{D}_n - w(1 - e^{-u})\|^3$$

$$\leq K_1(1 + \|w\|^3)u^{3/2}$$

for a constant K_1 depending only on f, where the last inequality follows from the fact that for $u \geq 0$, $\sigma^3(u) \leq 3u^{3/2}$ and $(1 - e^{-u})^3 \leq u^{3/2}$. So:

$$\left| (T_{n,u}f - f)(w) + uDf(w)[w] - u\mathbb{E}D^2 f(w)[\mathbf{D}_n, \mathbf{D}_n] \right|$$

$$\leq \left| (T_{n,u}f)(w) - f(w) - \mathbb{E}Df(w)[\sigma(u)\mathbf{D}_n - w(1 - e^{-u})] \right.$$

$$\left. - \frac{1}{2}\mathbb{E}D^2 f(w)[\sigma(u)\mathbf{D}_n - w(1 - e^{-u}), \sigma(u)\mathbf{D}_n - w(1 - e^{-u})] \right| + \left|\sigma(u)\mathbb{E}Df(w)[\mathbf{D}_n]\right|$$

$$+ \left|(u - 1 + e^{-u})Df(w)[w]\right| + \left|\left(\frac{\sigma^2(u)}{2} - u\right)\mathbb{E}D^2 f(w)[\mathbf{D}_n, \mathbf{D}_n]\right|$$

$$+ \left|\frac{(1 - e^{-u})^2}{2}D^2 f(w)[w, w]\right| + \left|\sigma(u)(1 - e^{-u})\mathbb{E}D^2 f(w)[\mathbf{D}_n, w]\right|$$

$$\leq K_2 u^{3/2}\left[(1 + \|w\|^3) + (1 + \|w\|^2)\|w\| + (1 + \|w\|)\mathbb{E}\|\mathbf{D}_n\|^2\right.$$

$$\left. + (1 + \|w\|)\|w\|^2 + (1 + \|w\|)\|w\|\mathbb{E}\|\mathbf{D}_n\|\right] + \left|\sigma(u)\mathbb{E}Df(w)[\mathbf{D}_n]\right|$$

$$\leq K_3(1 + \|w\|^3)u^{3/2}, \tag{5.3}$$

for some constants K_2 and K_3 depending only on f. The last inequality follows from the fact that:

$$\mathbb{E}Df(w)[\mathbf{D}_n] = \sum_{k=1}^{p} \sum_{i=1}^{\lambda_k} \mathbb{E}Df(w)\left[J_{i,k}e_k\right] \mathbb{E}[\tilde{Z}_{i,k}] = 0.$$

Therefore, by (5.3), we obtain that:

$$\mathcal{A}_n f(w) := \lim_{u \searrow 0} \frac{T_{n,u} f(w) - f(w)}{u} = -Df(w)[w] + \mathbb{E}D^2 f(w)[\mathbf{D}_n, \mathbf{D}_n],$$

as required. □

Now we prove the following:

Proposition 5.5. *For any* $g \in M$ *such that* $\mathbb{E}g(\mathbf{D}_n) = 0$, *the Stein equation* $\mathcal{A}_n f_n = g$ *is solved by:*

$$f_n = \phi_n(g) = -\int_0^\infty T_{n,u} g \, du, \tag{5.4}$$

where $(T_{n,u} f)(w) = \mathbb{E}\left[f\left(we^{-u} + \sigma(u)\mathbf{D}_n\right)\right]$ *for* $\sigma^2(v) = 1 - e^{-2v}$. *Furthermore:*

(A) $\|D\phi_n(g)(w)\| \le \|g\|_M \left(1 + \frac{2}{3}\|w\|^2 + \frac{4}{3}\mathbb{E}\|\mathbf{D}_n\|^2\right),$

(B) $\|D^2\phi_n(g)(w)\| \le \|g\|_M \left(\frac{1}{2} + \frac{\|w\|}{3} + \frac{\mathbb{E}\|\mathbf{D}_n\|}{3}\right),$

(C) $\dfrac{\left\|D^2\phi_n(g)(w+h) - D^2\phi_n(g)(w)\right\|}{\|h\|}$

$\le \sup_{w,h \in D^p} \dfrac{\|D^2(g+c)(w+h) - D^2(g+c)(w)\|}{3\|h\|}. \tag{5.5}$

for any constant function $c : D^p \to \mathbb{R}$ *and for all* $w, h \in D^p$.

Remark 5.6. It is worth noting that obtaining a bound for $\mathbb{E}\|\mathbf{D}_n\|$ or $\mathbb{E}\|\mathbf{D}_n\|^2$ that does not blow up with $n \to \infty$ is not easy, unless $\mathbf{D}_n$ is a martingale and Doob's L^2 inequality can be used to show that $\mathbb{E}\|\mathbf{D}_n\|^2 \le \mathbb{E}|\mathbf{D}_n(1)| = \mathbb{E}\sqrt{\sum_{i=1}^p \mathbf{D}_n^{(i)}(1)}$. This is, for instance, the case, if $\tilde{Z}_i = \left(\tilde{Z}_{i,1}, \ldots, \tilde{Z}_{i,p}\right)$'s are independent and $J_{i,k}$'s are independent.

Proof. The first part of the proposition follows by the argument used to prove [26, Proposition 4.4] upon noting that we can readily substitute $\mathbf{D}_n$ in the place of Z therein due to $\mathbb{E}\|\mathbf{D}_n\|^3$ being finite. What follows is a sketch summary of this argument. Using dominated convergence theorem, we note that, for any $f \in M$ and $w \in D([0, 1], \mathbb{R})$,

$$\left(\frac{d}{ds}\right)^+ T_{n,s} f(w) = \lim_{h \searrow 0} T_{n,s}\left[\frac{T_{n,h} - I}{h} f(w)\right] = \lim_{h \searrow 0} \mathbb{E}\left[\frac{T_{n,h} - I}{h} f(we^{-s} + \sigma(s)\mathbf{D}_n)\right]$$

$$= \mathbb{E}\left[\lim_{h \searrow 0} \frac{T_{n,h} - I}{h} f(we^{-s} + \sigma(s)\mathbf{D}_n)\right] = T_{n,s}\mathcal{A}_n f(w).$$

Similarly, for $s > 0$, $\left(\frac{d}{ds}\right)^- T_{n,s} f = T_{n,s}\mathcal{A}_n f$ because:

$$\lim_{h \searrow 0} \frac{1}{-h}\left[T_{n,s-h} f - T_{n,s} f\right](w) - T_{n,s}\mathcal{A}_n f(w)$$

$$= \lim_{h \searrow 0} T_{n,s-h}\left[\left(\frac{T_{n,h} - I}{h} - \mathcal{A}_n\right) f\right](w) + \lim_{h \searrow 0} \left(T_{n,s-h} - T_{n,s}\right)\mathcal{A}_n f(w)$$

$$= \lim_{h \searrow 0} \mathbb{E}\left[\left(\frac{T_{n,h} - I}{h} - \mathcal{A}_n\right) f(we^{-s+h} + \sigma(s-h)\mathbf{D}_n)\right]$$

$$+ \lim_{h \searrow 0} \mathbb{E}\left[\mathcal{A}_n f(we^{-s+h} + \sigma(s-h)\mathbf{D}_n) - \mathcal{A}_n f(we^{-s} + \sigma(s)\mathbf{D}_n)\right]$$

$$= 0$$

again, by dominated convergence and an argument similar to (5.3). Thus, for all $f \in M$ and $s > 0$, we have

$$\frac{d}{ds} T_{n,s} f = T_{n,s} \mathcal{A}_n f$$

and so, by the fundamental theorem of calculus, for any $r > 0$,

$$T_{n,r} f - f = \int_0^r T_{n,s} \mathcal{A}_n f \, ds.$$

Applying this to $f = \int_0^t T_{n,u} g \, du$ (which belongs to M, for instance by [1, (2.23), (2.24)]), for some $t > 0$, we obtain for any $r > 0$ and any $w \in D([0,1], \mathbb{R})$,

$$T_{n,r} \int_0^t T_{n,u} g(w) du - \int_0^t T_{n,u} g(w) du = \int_0^r T_{n,s} \mathcal{A}_n \left(\int_0^t T_{n,u} g(w) du \right) ds. \qquad (5.6)$$

On the other hand, for all $w \in D[0,1]$ and $h > 0$:

$$\frac{1}{h}[T_{n,h} - I] \int_0^t T_{n,u} g(w) du = \frac{1}{h} \int_0^t [T_{n,u+h} g(w) - T_{n,u} g(w)] du$$

$$= \frac{1}{h} \int_t^{t+h} T_{n,u} g(w) du - \frac{1}{h} \int_0^h T_{n,u} g(w) du$$

$$\stackrel{(5.2)}{=} \frac{1}{h} \int_t^{t+h} \mathbb{E}[g(we^{-u} + \sigma(u)\mathbf{D}_n)] du - \frac{1}{h} \int_0^h \mathbb{E}[g(we^{-u} + \sigma(u)\mathbf{D}_n)] du. \qquad (5.7)$$

Taking $h \to 0$ in (5.7) and noting that

$$\lim_{h \searrow 0} \left[\frac{1}{h} \int_0^h \mathbb{E} g\left(we^{-s} + \sigma(s)\mathbf{D}_n \right) ds \right] = g(w),$$

as proved in [26, (4.6)], yields

$$\mathcal{A}_n \left(\int_0^t T_{n,u} g \, du \right) = T_{n,t} g - g. \qquad (5.8)$$

Now, taking $t \to \infty$ in (5.6) and applying dominated convergence, we obtain

$$T_{n,r} \int_0^\infty T_{n,u} g(w) - \int_0^\infty T_{n,u} g(w) du = \int_0^r T_{n,s} \lim_{t \to \infty} \mathcal{A}_n \left(\int_0^t T_{n,u} g(w) du \right) ds$$

$$\stackrel{(5.8)}{=} - \int_0^r T_{n,s} g(w) ds. \qquad (5.9)$$

Furthermore, by [26, Lemma 4.1], $\int_0^\infty T_{n,u} g \, du$ is in the domain of $\mathcal{A}_n$. Therefore, dividing both sides of (5.9) by r and taking $r \searrow 0$ gives

$$\mathcal{A}_n \left(\int_0^\infty T_{n,u} g(w) du \right) \stackrel{(5.9)}{=} - \lim_{r \searrow 0} \frac{1}{r} \int_0^r T_{n,s} g(w) ds$$

$$= - \lim_{r \searrow 0} \left[\frac{1}{r} \int_0^r \mathbb{E} g\left(we^{-s} + \sigma(s)\mathbf{D}_n \right) ds \right]$$

$$= -g(w),$$

where the last equality follows from [26, (4.6)]. This lets us conclude that the Stein equation $\mathcal{A}_n f_n = g$ is indeed solved by:

$$f_n = \phi_n(g) = -\int_0^\infty T_{n,u} g \, du.$$

Now, note that for ϕ_n defined in (5.4) we get:

$$\phi_n(g)(w+h) - \phi_n(g)(w)$$
$$\overset{(5.2)}{=} -\mathbb{E}\int_0^\infty \left[g\left((w+h)e^{-u} + \sigma(u)\mathbf{D}_n\right) - g\left(we^{-u} + \sigma(u)\mathbf{D}_n\right)\right] du$$

and so dominated convergence (which can be applied because of [26, (4.2)]) gives:

$$D^k \phi_n(g)(w) = -\mathbb{E}\int_0^\infty e^{-ku} D^k g(we^{-u} + \sigma(u)\mathbf{D}_n) du, \quad k = 1, 2. \tag{5.10}$$

Now, using (5.10) observe that:

(A) $\|D\phi_n(g)(w)\|$
$$\leq \int_0^\infty e^{-u}\mathbb{E}\|Dg(we^{-u} + \sigma(u)\mathbf{D}_n)\| du$$
$$\leq \|g\|_M \int_0^\infty \left(e^{-u} + 2\|w\|^2 e^{-3u} + 2\mathbb{E}\|\mathbf{D}_n\|^2(e^{-u} - e^{-3u})\right) du$$
$$\leq \|g\|_M \left(1 + \frac{2}{3}\|w\|^2 + \frac{4}{3}\mathbb{E}\|\mathbf{D}_n\|^2\right),$$

(B) $\|D^2 \phi_n(g)(w)\|$
$$\leq \int_0^\infty e^{-2u}\mathbb{E}\left\|D^2 g(we^{-u} + \sigma(u)\mathbf{D}_n)\right\| du$$
$$\leq \|g\|_M \int_0^\infty e^{-2u}(1 + \mathbb{E}\|we^{-u} + \sigma(u)\mathbf{D}_n\|) du$$
$$\leq \|g\|_M \left(\frac{1}{2} + \frac{\|w\|}{3} + \frac{\mathbb{E}\|\mathbf{D}_n\|}{3}\right),$$

(C) $\dfrac{\|D^2\phi_n(g)(w+h) - D^2\phi_n(g)(w)\|}{\|h\|}$
$$\leq \|h\|^{-1}\left\| \mathbb{E}\int_0^\infty e^{-2u} D^2 g((w+h)e^{-u} + \sigma(u)\mathbf{D}_n) - e^{-2u} D^2 g(we^{-u} + \sigma(u)\mathbf{D}_n) du\right\|$$
$$\leq \sup_{w,h \in D^p} \frac{\|D^2 g(w+h) - D^2 g(w)\|}{\|h\|}\int_0^\infty e^{-2u} e^{-u} du$$
$$= \sup_{w,h \in D^p} \frac{\|D^2(g+c)(w+h) - D^2(g+c)(w)\|}{3\|h\|},$$

uniformly in $g \in M$, for any constant c, which proves (5.5).   □

5.2. An auxiliary result

We now move to proving the main results of the paper. We start with an auxiliary lemma in which we use Stein's method combined with Taylor expansions to bound the distance between

$\mathbf{Y}_n$, as defined in Theorem 3.1 and $\mathbf{D}_n$, as defined in (5.1). This result is of independent interest and will be used in all the proofs in this Section.

Lemma 5.7. *Consider the setup of Theorem 3.1. Let $\mathbf{D}_n$ be defined as in (5.1) for the covariance matrix Σ_n equal to the covariance matrix of $\left(X_{1,1}, \ldots, X_{1,p}, \ldots, X_{n,1}, \ldots, X_{n,p}\right)$. Assume that the two collections $\{\tilde{Z}_{i,k} \,:\, i = 1, \ldots, n,\, k = 1, \ldots, p\}$ and $\{X_{i,k} \,:\, i = 1, \ldots, n,\, k = 1, \ldots, p\}$ are independent. Let $g \in M$, as defined in Section 2. Then:*

$$
\left|\mathbb{E}g(\mathbf{Y}_n) - \mathbb{E}g(\mathbf{D}_n)\right|
$$

$$
\leq \frac{\|g\|_M}{6} \sum_{i=1}^n \mathbb{E}\left\{\left(\sum_{k,l,m=1}^p \left[(X_{i,k})^2 \|J_{i,k}\|^2 \mathbb{1}_{[1,\lambda_k]}(i) \left(\sum_{j\in\mathbb{A}_i} X_{j,l}\|J_{j,l}\|\mathbb{1}_{[1,\lambda_l]}(j)\right)^2\right.\right.\right.
$$

$$
\left.\left.\left. \cdot \left(\sum_{j\in\mathbb{A}_i} X_{j,m}\|J_{j,m}\|\mathbb{1}_{[1,\lambda_m]}(j)\right)^2\right]\right)^{1/2}\right\}
$$

$$
+ \frac{\|g\|_M}{3} \sum_{i=1}^n \sum_{j\in\mathbb{A}_i} \sum_{k,l=1}^p \mathbb{E}\left\{\left[\sum_{m=1}^p \left(X_{i,k}\|J_{i,k}\|\,X_{j,l}\|J_{j,l}\|\,\mathbb{1}_{[1,\lambda_k]}(i)\mathbb{1}_{[1,\lambda_l]}(j)\right.\right.\right.
$$

$$
\left.\left.\left. \cdot \sum_{r\in\mathbb{A}_{ij}\cap\mathbb{A}_i^c} X_{r,m}\|J_{r,m}\|\mathbb{1}_{[1,\lambda_m]}(r)\right)^2\right]^{1/2}\right\}
$$

$$
+ \frac{\|g\|_M}{3} \sum_{i=1}^n \sum_{j\in\mathbb{A}_i} \sum_{k,l=1}^p \left\{\left|\mathbb{E}\left[X_{i,k}X_{j,l}\right]\right|\mathbb{1}_{[1,\lambda_k]}(i)\mathbb{1}_{[1,\lambda_l]}(j)\right.
$$

$$
\left. \cdot \mathbb{E}\left[\|J_{i,k}\|\,\|J_{j,l}\|\sqrt{\sum_{m=1}^p \left(\sum_{r\in\mathbb{A}_{ij}} X_{r,m}\|J_{r,m}\|\mathbb{1}_{[1,\lambda_m]}(r)\right)^2}\right]\right\}.
$$

The proof of Lemma 5.7 is based on manipulating the Stein operator, given in Proposition 5.1, using Taylor's theorem.

Proof of Lemma 5.7. Let $g_n = g - \mathbb{E}g(\mathbf{D}_n)$ and $f_n = \phi_n(g_n)$, as defined in (5.4). From Proposition 5.1 we know that:

$$
\left|\mathbb{E}g(\mathbf{Y}_n) - \mathbb{E}g(\mathbf{D}_n)\right| = \left|\mathbb{E}\left[Df_n(\mathbf{Y}_n)[\mathbf{Y}_n] - D^2 f_n(\mathbf{Y}_n)[\mathbf{D}_n, \mathbf{D}_n]\right]\right|.
$$

Let

$$
\mathbf{Y}_n^j = \sum_{k\in\mathbb{A}_j^c} \left(X_{k,1}\mathbb{1}_{[1,\lambda_1]}(k)J_{k,1}, \ldots, X_{k,p}\mathbb{1}_{[1,\lambda_p]}(k)J_{k,p}\right)
$$

and

$$
\mathbf{Y}_n^{ij} = \sum_{k\in\mathbb{A}_{ij}^c} \left(X_{k,1}\mathbb{1}_{[1,\lambda_1]}(k)J_{k,1}, \ldots, X_{k,p}\mathbb{1}_{[1,\lambda_p]}(k)J_{k,p}\right).
$$

Hence, $\mathbf{Y}_n^j$ is independent of X_j for all j and $\mathbf{Y}_n^{ij}$ is independent of (X_i, X_j) for all i, j. Therefore

$$\mathbb{E}Df_n(\mathbf{Y}_n^i)\left[\left(X_{i,1}\mathbb{1}_{[1,\lambda_1]}(i)J_{i,1}, \ldots, X_{i,p}\mathbb{1}_{[1,\lambda_p]}(i)J_{i,p}\right)\right] = 0.$$

For $\{e_k : k = 1, \ldots, p\}$ denoting the elements of the canonical basis of $\mathbb{R}^p$ and for $i \in \{1, \ldots, n\}$, we have the following identities and inequalities (note that inequality $(*)$ follows from Taylor's theorem):

$$
\begin{aligned}
&\left| \mathbb{E}Df_n(\mathbf{Y}_n)\left[\left(X_{i,1}\mathbb{1}_{[1,\lambda_1]}(i)J_{i,1}, \ldots, X_{i,p}\mathbb{1}_{[1,\lambda_p]}(i)J_{i,p}\right)\right] \right. \\
&\quad \left. - \mathbb{E}\left[\sum_{j\in\mathbb{A}_i}\sum_{k,l=1}^{p}\left(X_{i,k}\mathbb{1}_{[1,\lambda_k]}(i)\right)\left(X_{j,l}\mathbb{1}_{[1,\lambda_l]}(j)\right)D^2f_n(\mathbf{Y}_n^i)\left[e_kJ_{i,k}, e_lJ_{j,l}\right]\right] \right| \\
&= \left| \mathbb{E}Df_n(\mathbf{Y}_n)\left[\left(X_{i,1}\mathbb{1}_{[1,\lambda_1]}(i)J_{i,1}, \ldots, X_{i,p}\mathbb{1}_{[1,\lambda_p]}(i)J_{i,p}\right)\right] \right. \\
&\quad - \mathbb{E}Df_n(\mathbf{Y}_n^i)\left[\left(X_{i,1}\mathbb{1}_{[1,\lambda_1]}(i)J_{i,1}, \ldots, X_{i,p}\mathbb{1}_{[1,\lambda_p]}(i)J_{i,p}\right)\right] \\
&\quad - \mathbb{E}D^2f_n(\mathbf{Y}_n^i)\left[\left(X_{i,1}\mathbb{1}_{[1,\lambda_1]}(i)J_{i,1}, \ldots, X_{i,p}\mathbb{1}_{[1,\lambda_p]}(i)J_{i,p}\right), \right. \\
&\qquad\qquad\qquad \left.\left. \sum_{j\in\mathbb{A}_i}\left(X_{j,1}\mathbb{1}_{[1,\lambda_1]}(j)J_{j,1}, \ldots X_{j,p}\mathbb{1}_{[1,\lambda_p]}(j)J_{j,p}\right)\right]\right| \\
&\overset{(*)}{\leq} \frac{1}{2}\sup_{w,h\in D^p}\frac{\left\|D^2f_n(w+h) - D^2f_n(w)\right\|}{\|h\|} \\
&\quad \cdot \mathbb{E}\left[\left\|\left(X_{i,1}\mathbb{1}_{[1,\lambda_1]}(i)J_{i,1}, \ldots, X_{i,p}\mathbb{1}_{[1,\lambda_p]}(i)J_{i,p}\right)\right\| \right. \\
&\qquad \left. \cdot \left\|\sum_{j\in\mathbb{A}_i}\left(X_{j,1}\mathbb{1}_{[1,\lambda_1]}(j)J_{j,1}, \ldots, X_{j,p}\mathbb{1}_{[1,\lambda_p]}(j)J_{j,p}\right)\right\|\|Y_n - Y_n^i\|\right] \\
&\overset{(5.5)(C)}{\leq} \frac{\|g\|_M}{6}\mathbb{E}\left[\left\|\left(X_{i,1}\mathbb{1}_{[1,\lambda_1]}(i)J_{i,1}, \ldots, X_{i,p}\mathbb{1}_{[1,\lambda_p]}(i)J_{i,p}\right)\right\| \right. \\
&\qquad \left. \cdot \left\|\sum_{j\in\mathbb{A}_i}\left(X_{j,1}\mathbb{1}_{[1,\lambda_1]}(j)J_{j,1}, \ldots, X_{j,p}\mathbb{1}_{[1,\lambda_p]}(j)J_{j,p}\right)\right\|\|Y_n - Y_n^i\|\right] \\
&= \frac{\|g\|_M}{6}\mathbb{E}\left[\left\|\left(X_{i,1}\mathbb{1}_{[1,\lambda_1]}(i)J_{i,1}, \ldots, X_{i,p}\mathbb{1}_{[1,\lambda_p]}(i)J_{i,p}\right)\right\| \right.
\end{aligned}
$$

$$\cdot \left\| \sum_{j \in \mathbb{A}_i} \left(X_{j,1} \mathbb{1}_{[1,\lambda_1]}(j) J_{j,1}, \ldots, X_{j,p} \mathbb{1}_{[1,\lambda_p]}(j) J_{j,p} \right) \right\|^2 \right]$$

$$\leq \frac{\|g\|_M}{6} \mathbb{E} \left\{ \left(\sum_{k,l,m=1}^{p} \left[(X_{i,k})^2 \|J_{i,k}\|^2 \mathbb{1}_{[1,\lambda_k]}(i) \left(\sum_{j \in \mathbb{A}_i} X_{j,l} \|J_{j,l}\| \mathbb{1}_{[1,\lambda_l]}(j) \right)^2 \right. \right. \right.$$

$$\left. \left. \left. \cdot \left(\sum_{j \in \mathbb{A}_i} X_{j,m} \|J_{j,m}\| \mathbb{1}_{[1,\lambda_m]}(j) \right)^2 \right] \right)^{1/2} \right\}. \tag{5.11}$$

Furthermore, for all $i, j \in \{1, \ldots, n\}$,

$$\left| \mathbb{E} \left[X_{i,k} \mathbb{1}_{[1,\lambda_k]}(i) X_{j,l} \mathbb{1}_{[1,\lambda_l]}(j) D^2 f_n(\mathbf{Y}_n^i) \left[e_k J_{i,k}, e_l J_{j,l} \right] \right] \right.$$
$$\left. - \mathbb{E} \left[X_{i,k} \mathbb{1}_{[1,\lambda_k]}(i) X_{j,l} \mathbb{1}_{[1,\lambda_l]}(j) D^2 f_n(\mathbf{Y}_n^{i,j}) \left[e_k J_{i,k}, e_l J_{j,l} \right] \right] \right|$$

$$\overset{(5.5)(C)}{\leq} \frac{\|g\|_M}{3} \mathbb{E} \left\{ \left[\sum_{m=1}^{p} \left(X_{i,k} \|J_{i,k}\| X_{j,l} \|J_{j,l}\| \mathbb{1}_{[1,\lambda_k]}(i) \mathbb{1}_{[1,\lambda_l]}(j) \right. \right. \right.$$

$$\left. \left. \left. \cdot \sum_{r \in \mathbb{A}_{ij} \cap \mathbb{A}_i^c} X_{r,m} \|J_{r,m}\| \mathbb{1}_{[1,\lambda_m]}(r) \right)^2 \right]^{1/2} \right\} \tag{5.12}$$

and

$$\left| \mathbb{E} \left[X_{i,k} \mathbb{1}_{[1,\lambda_k]}(i) X_{j,l} \mathbb{1}_{[1,\lambda_l]}(j) D^2 f_n \left(\mathbf{Y}_n^{i,j} \right) \left[e_k J_{i,k}, e_l J_{j,l} \right] \right] \right.$$
$$\left. - \mathbb{E} \left[X_{i,k} X_{j,l} \right] \mathbb{1}_{[1,\lambda_k]}(i) \mathbb{1}_{[1,\lambda_l]}(j) \mathbb{E} \left[D^2 f_n(\mathbf{Y}_n) \left[e_k J_{i,k}, e_l J_{j,l} \right] \right] \right|$$

$$= \left| \mathbb{E} \left[X_{i,k} X_{j,l} \right] \mathbb{1}_{[1,\lambda_k]}(i) \mathbb{1}_{[1,\lambda_l]}(j) \mathbb{E} \left[\left(D^2 f_n(\mathbf{Y}_n) - D^2 f_n \left(\mathbf{Y}_n^{i,j} \right) \right) \left[e_k J_{i,k}, e_l J_{j,l} \right] \right] \right|$$

$$\overset{(5.5)(C)}{\leq} \frac{\|g\|_M}{3} \left| \mathbb{E} \left[X_{i,k} X_{j,l} \right] \right| \mathbb{1}_{[1,\lambda_k]}(i) \mathbb{1}_{[1,\lambda_l]}(j)$$

$$\cdot \mathbb{E} \left[\|J_{i,k}\| \|J_{j,l}\| \sqrt{\sum_{m=1}^{p} \left(\sum_{r \in \mathbb{A}_{ij}} X_{r,m} \|J_{r,m}\| \mathbb{1}_{[1,\lambda_m]}(r) \right)^2} \right]. \tag{5.13}$$

Summing (5.11) over $i = 1, \ldots, n$ and (5.12) and (5.13) over $i = 1, \ldots, n$, $j \in \mathbb{A}_i$ and $k, l = 1, \ldots, p$ will give us a bound on $|\mathbb{E} \mathcal{A}_n g(\mathbf{Y}_n)|$, as defined in Proposition 5.1, i.e. a bound on $|\mathbb{E} g(\mathbf{Y}_n) - \mathbb{E} g(\mathbf{D}_n)|$. $\square$

5.3. Proof of Theorem 3.1

In the proof of Theorem 3.1, we will use auxiliary processes $\tilde{\mathbf{D}}_n$ and $\tilde{\mathbf{A}}_n$. In order to define them, we let $\left(\tilde{Z}_{1,1}, \ldots, \tilde{Z}_{1,p}, \tilde{Z}_{2,1}, \ldots, \tilde{Z}_{2,p}, \ldots, \tilde{Z}_{n,1}, \ldots, \tilde{Z}_{n,p} \right)$ be a centred Gaussian vector with the same covariance as that of $(X_{1,1}, \ldots, X_{1,p}, \ldots, X_{n,1}, \ldots, X_{n,p})$ and independent of $(X_{1,1}, \ldots, X_{1,p}, \ldots, X_{n,1}, \ldots, X_{n,p})$. We also let $\left\{ (Z_{i,1}, \ldots, Z_{i,p}) : i = 1, \ldots, n \right\}$ be a collection of i.i.d. Gaussian vectors with mean zero and covariance Σ, independent of the

collections $\{J_{i,k} : i = 1, \ldots, n, \ k = 1, \ldots, p\}$ and $\{X_{i,k} : i = 1, \ldots, n, \ k = 1, \ldots, p\}$. The auxiliary processes are defined for $t \in [0, 1]$ in the following way:

$$\tilde{\mathbf{D}}_n(t) = \left(\sum_{i=1}^{\lambda_1} \tilde{Z}_{i,1} \mathbb{1}_{[i/\lambda_1, 1]}(t), \ldots, \sum_{i=1}^{\lambda_p} \tilde{Z}_{i,p} \mathbb{1}_{[i/\lambda_p, 1]}(t) \right) ; \tag{5.14}$$

$$\tilde{\mathbf{A}}_n(t) = \left(\frac{1}{\sqrt{\lambda_1}} \sum_{i=1}^{\lambda_1} Z_{i,1} \mathbb{1}_{[i/\lambda_1, 1]}(t), \ldots, \frac{1}{\sqrt{\lambda_p}} \sum_{i=1}^{\lambda_p} Z_{i,p} \mathbb{1}_{[i/\lambda_p, 1]}(t) \right) . \tag{5.15}$$

Step 1 of the proof below makes a straightforward use of the mean value theorem to bound the distance between $\mathbf{D}_n$, as defined by (5.1) and $\tilde{\mathbf{D}}_n$. In **Step 2** the distance between $\tilde{\mathbf{D}}_n$ and $\tilde{\mathbf{A}}_n$ is bounded using bounds on the distance between two multivariate Gaussian distributions ([32, Proposition 2.8]). In **Step 3** we couple $\tilde{\mathbf{A}}_n$ and $\mathbf{Z}$ in order to obtain a bound on $\mathbb{E}\|\tilde{\mathbf{A}}_n - \mathbf{Z}\|$ and then apply the mean value theorem again to bound $|\mathbb{E}g(\tilde{\mathbf{A}}_n) - \mathbb{E}g(\mathbf{Z})|$ for all $g \in M^1$. Those three steps combined with Lemma 5.7 yield the assertion. In short:

$$|\mathbb{E}g(\mathbf{Y}_n) - \mathbb{E}g(\mathbf{Z})| \leq \underbrace{|\mathbb{E}g(\mathbf{Y}_n) - \mathbb{E}g(\mathbf{D}_n)|}_{\text{Lemma 5.7}} + \underbrace{\left|\mathbb{E}g(\mathbf{D}_n) - \mathbb{E}g(\tilde{\mathbf{D}}_n)\right|}_{\text{Step 1}}$$

$$+ \underbrace{\left|\mathbb{E}g(\tilde{\mathbf{D}}_n) - \mathbb{E}g(\tilde{\mathbf{A}}_n)\right|}_{\text{Step 2}} + \underbrace{\left|\mathbb{E}g(\tilde{\mathbf{A}}_n) - \mathbb{E}g(\mathbf{Z})\right|}_{\text{Step 3}} .$$

Proof of Theorem 3.1.

Step 1. Note that, for $\mathbf{D}_n$ of Lemma 5.7 and $\tilde{\mathbf{D}}_n$ of (5.14),

$$\left| \mathbb{E}g(\mathbf{D}_n) - \mathbb{E}g(\tilde{\mathbf{D}}_n) \right|$$

$$\leq \|g\|_{M^1} \mathbb{E}\|\mathbf{D}_n - \tilde{\mathbf{D}}_n\|$$

$$\leq \|g\|_{M^1} \mathbb{E}\left\{ \sup_{t \in [0,1]} \sqrt{\sum_{k=1}^{p} \left[\sum_{i=1}^{\lambda_k} \tilde{Z}_{i,k} \left(J_{i,k}(t) - \mathbb{1}_{[i/\lambda_k, 1]}(t) \right) \right]^2} \right\}$$

$$\leq \|g\|_{M^1} \sum_{k=1}^{p} \sum_{i=1}^{\lambda_k} \mathbb{E}|\tilde{Z}_{i,k}| \mathbb{E} \left\| J_{i,k} - \mathbb{1}_{[i/\lambda_k, 1]} \right\|$$

$$\leq \|g\|_{M^1} \sum_{k=1}^{p} \sum_{i=1}^{\lambda_k} \sqrt{\mathbb{E}\left[(X_{i,k})^2 \right]} \mathbb{E} \left\| J_{i,k} - \mathbb{1}_{[i/\lambda_k, 1]} \right\| , \tag{5.16}$$

giving ϵ_7.

Step 2. Let $\lambda = \sum_{k=1}^{p} \lambda_k$ and consider function $f : \mathbb{R}^\lambda \to D^p[0, 1]$ given by:

$$f\left(x_{1,1}, \ldots, x_{\lambda_1, 1}, \ldots, x_{1,p}, \ldots, x_{\lambda_p, p}\right) = \left(\sum_{i=1}^{\lambda_1} x_{i,1} \mathbb{1}_{[i/\lambda_1, 1]}, \ldots, \sum_{i=1}^{\lambda_p} x_{i,p} \mathbb{1}_{[i/\lambda_p, 1]} \right) .$$

This function is twice differentiable with:

(A) $Df(x)[(h_{1,1}, \ldots, h_{\lambda_1,1}, \ldots, h_{1,p}, \ldots, h_{\lambda_p,p})]$

$$= \left(\sum_{i=1}^{\lambda_1} h_{i,1} \mathbb{1}_{[i/\lambda_1,1]}, \ldots, \sum_{i=1}^{\lambda_p} h_{i,p} \mathbb{1}_{[i/\lambda_p,1]} \right)$$

(B) $D^2 f(x)[h^{(1)}, h^{(2)}] = 0$

for all $x, h = (h_{1,1}, \ldots, h_{\lambda_1,1}, \ldots, h_{1,p}, \ldots, h_{\lambda_p,p}), h^{(1)}, h^{(2)} \in \mathbb{R}^{np}$. We notice that for the canonical basis vectors $e_i, e_j \in \mathbb{R}^{np}$ we have:

$$\left| D^2(g \circ f)(x)[e_i, e_j] \right| = \left| D^2 g(f(x))[Df(x)[e_i], Df(x)[e_j]] \right| \leq \sup_{w \in D} \| D^2 g(w) \|$$

for all $x \in \mathbb{R}^{np}$. This follows from the fact that $|Df(x)[e_i]| = 1$. Therefore, we can apply [32, Proposition 2.8] to the function $g \circ f$ and, recalling the definitions of $\tilde{\mathbf{D}}_n$ in (5.14) and $\tilde{\mathbf{A}}_n$ in (5.15), obtain

$$|\mathbb{E}g(\tilde{\mathbf{A}}_n) - \mathbb{E}g(\tilde{\mathbf{D}}_n)|$$
$$\leq \frac{1}{2} \|g\|_{M^1} \sum_{k,l=1}^{p} \left[\sum_{i=1}^{\lambda_k} \sum_{j \in \mathbb{A}_i \setminus \{i\}} \left| \mathbb{E}\left[X_{i,k} X_{j,l} \right] \right| + \sum_{i=1}^{\lambda_k \wedge \lambda_l} \left| \frac{\Sigma_{k,l}}{\sqrt{\lambda_k \lambda_l}} - \mathbb{E}[X_{i,k} X_{i,l}] \right| \right], \qquad (5.17)$$

giving $\epsilon_4 + \epsilon_5$.

Step 3. We now realise a p-dimensional Brownian motion $\mathbf{B}$ and let $\mathbf{Z} = \Sigma^{1/2}\mathbf{B}$. We also let

$$\tilde{\mathbf{A}}_n^{(j)}(t) = \mathbf{Z}^{(j)}\left(l/\lambda_j \right), \quad \text{for } t \in \left[l/\lambda_j, (l+1)/\lambda_j \right)$$

for every $j = 1, \ldots, p$, which agrees in distribution with our original definition (5.15) of $\tilde{\mathbf{A}}_n = \left(\tilde{\mathbf{A}}_n^{(1)}, \ldots, \tilde{\mathbf{A}}_n^{(p)} \right)$. Now, note that, using Jensen's inequality , we have:

$$\mathbb{E}\|\tilde{\mathbf{A}}_n - \mathbf{Z}\| \leq \left(\sum_{i=1}^{p} \mathbb{E} \left\| \tilde{\mathbf{A}}_n^{(i)} - \mathbf{Z}^{(i)} \right\|^2 \right)^{1/2}$$

$$= \sqrt{ \sum_{i=1}^{p} \mathbb{E} \sup_{t \in [0,1]} \left| \mathbf{Z}^{(i)}(t) - \mathbf{Z}^{(i)}\left(\frac{\lfloor \lambda_i t \rfloor}{\lambda_i} \right) \right|^2 }$$

$$\leq \left(\sum_{i=1}^{p} \Sigma_{i,i} \right)^{1/2} \sqrt{ \sum_{i=1}^{p} \mathbb{E} \sup_{t \in [0,1]} \left| \mathbf{B}^{(i)}(t) - \mathbf{B}^{(i)}\left(\frac{\lfloor \lambda_i t \rfloor}{\lambda_i} \right) \right|^2 }$$

$$\leq \frac{6\sqrt{5}}{\sqrt{2 \log 2}} \left(\sqrt{ \sum_{i=1}^{p} \frac{\log(2\lambda_i)}{\lambda_i} } \right) \left(\sum_{i=1}^{p} \Sigma_{i,i} \right)^{1/2},$$

where the third inequality follows because $\left\| \Sigma^{1/2} \right\|_2 = \sqrt{\lambda_{\max}(\Sigma)} \le \left(\sum_{i=1}^{p} \Sigma_{i,i} \right)^{1/2}$ (for $\lambda_{\max}(\Sigma)$ denoting the largest eigenvalue of Σ) and the last inequality follows by [16, Lemma 3]. Therefore:

$$
|\mathbb{E}g(\tilde{\mathbf{A}}_n) - \mathbb{E}g(\mathbf{Z})| \stackrel{\text{MVT}}{\le} \sup_{w \in D^p} \|Dg(w)\| \mathbb{E}\|\mathbf{Z} - \tilde{\mathbf{A}}_n\|
$$

$$
\le \|g\|_{M^1} \frac{6\sqrt{5}}{\sqrt{2\log 2}} \left(\sqrt{\sum_{i=1}^{p} \frac{\log(2\lambda_i)}{\lambda_i}} \right) \left(\sum_{i=1}^{p} \Sigma_{i,i} \right)^{1/2}, \tag{5.18}
$$

giving ϵ_6.

Now, Lemma 5.7 (which gives $\epsilon_1 + \epsilon_2 + \epsilon_3$), combined with (5.17), (5.16), (5.18), yields the assertion. $\square$

5.4. Proof of Proposition 3.5

The proof of Proposition 3.5 is similar to that of Lemma 5.7 and **Step 3** of the proof of Theorem 3.1. Due to the independence of the summands, the bound on the distance between $\mathbf{Y}_n$ and the pre-limiting Gaussian process has a simpler form than the one appearing in Theorem 3.1. We now work with all $g \in M$, contrary to what is done in the proof of Theorem 3.1. Hence, we need to bound both the first and second moment of the supremum distance between the pre-limiting process and the correlated Brownian motion. This is necessary for the mean value theorem to be applied in the final step.

Proof of Proposition 3.5. Let $\mathbf{D}_n$ be as in (5.1) with Σ_n such that the vectors $\left(\tilde{Z}_i \right)_{i=1}^{n}$ are i.i.d with the same covariance structure as that of $(X_i)_{i=1}^{n}$ and for all $i = 1, \ldots, n$ and $k = 1, \ldots, p$, $J_{i,k} = \mathbb{1}_{[i/n,1]}$. Assume that the two collections $\{\tilde{Z}_{i,k} : i = 1, \ldots, n, k = 1, \ldots, p\}$ and $\{X_{i,k} : i = 1, \ldots, n, k = 1, \ldots, p\}$ are independent. Let $g \in M$, $g_n = g - \mathbb{E}[g(\mathbf{D}_n)]$, $f_n = \phi_n(g_n)$, as in (5.4).

Note that for $\mathbf{Y}_n^j = \mathbf{Y}_n - \frac{1}{\sqrt{n}} X_j \mathbb{1}_{[j/n,1]}$, $j = 1, \ldots, n$, $\mathbf{Y}_n^j$ is independent of X_j and

$$
\left| n^{-1/2} \mathbb{E}Df_n(\mathbf{Y}_n) \left[X_j \mathbb{1}_{[j/n,1]} \right] - n^{-1} \sum_{k,l=1}^{p} \Sigma_{k,l} \mathbb{E}D^2 f_n(\mathbf{Y}_n^j) \left[e_k \mathbb{1}_{[j/n,1]}, e_l \mathbb{1}_{[j/n,1]} \right] \right|
$$

$$
= \left| n^{-1/2} \mathbb{E}Df_n(\mathbf{Y}_n) \left[X_j \mathbb{1}_{[j/n,1]} \right] - n^{-1/2} \mathbb{E}Df_n(\mathbf{Y}_n^j) \left[X_j \mathbb{1}_{[j/n,1]} \right] \right.
$$

$$
\left. - n^{-1} \mathbb{E}D^2 f_n(\mathbf{Y}_n^j) \left[X_j \mathbb{1}_{[j/n,1]}, X_j \mathbb{1}_{[j/n,1]} \right] \right|
$$

$$
\le \frac{n^{-3/2}}{2} \sup_{w,h \in D^p} \frac{\left\| D^2 f_n(w+h) - D^2 f_n(w) \right\|}{\|h\|} \mathbb{E} \left\| X_j \mathbb{1}_{[j/n,1]} \right\|^3
$$

$$
\le n^{-3/2} \frac{\|g\|_M}{6} \mathbb{E} \left\| X_j \mathbb{1}_{[j/n,1]} \right\|^3
$$

$$
= n^{-3/2} \frac{\|g\|_M}{6} \mathbb{E} \left[\left(\left(X_j^{(1)} \right)^2 + \cdots + \left(X_j^{(p)} \right)^2 \right)^{3/2} \right]
$$

$$
\le p^{1/2} n^{-3/2} \frac{\|g\|_M}{6} \sum_{m=1}^{p} \mathbb{E} \left| X_j^{(m)} \right|^3, \tag{5.19}
$$

where the first inequality follows by Taylor's theorem and the second one by (5.5)(C). Also, by (5.5)(C):

$$
\left| n^{-1} \sum_{k,l=1}^{p} \Sigma_{k,l} \mathbb{E} D^2 f_n(\mathbf{Y}_n^j) \left[e_k \mathbb{1}_{[j/n,1]}, e_l \mathbb{1}_{[j/n,1]} \right] \right.
$$
$$
\left. - n^{-1} \sum_{k,l=1}^{p} \Sigma_{k,l} \mathbb{E} D^2 f_n(\mathbf{Y}_n) \left[e_k \mathbb{1}_{[j/n,1]}, e_l \mathbb{1}_{[j/n,1]} \right] \right|
$$
$$
\leq n^{-3/2} \frac{\|g\|_M}{3} \sum_{k,l=1}^{p} |\Sigma_{k,l}| \left(\sum_{m=1}^{p} \mathbb{E} \left| X_j^{(m)} \right|^2 \right)^{1/2}. \tag{5.20}
$$

Let us now realise a p-dimensional Brownian motion $\mathbf{B}$ and let $\mathbf{Z} = \Sigma^{1/2} \mathbf{B}$. We realise it in such a way that $\Sigma^{-1/2} \mathbf{D}_n(j/n) = \mathbf{B}(j/n)$ for every $j = 1, \ldots, n$, which agrees in distribution with our original definition of $\mathbf{D}_n$. Now, note that, by [16, Lemma 3] and Doob's L^3 inequality:

(A) $\quad \mathbb{E}\|\mathbf{Z} - \mathbf{D}_n\| \leq \sqrt{\sum_{i=1}^{p} \mathbb{E} \left\| \mathbf{Z}^{(i)} - \mathbf{D}_n^{(i)} \right\|^2} \leq \frac{6\sqrt{5}}{\sqrt{2\log 2}} n^{-1/2} \sqrt{\log 2n} \left(\sum_{i=1}^{p} |\Sigma_{i,i}| \right)^{1/2};$

(B) $\quad \mathbb{E}\|\mathbf{Z} - \mathbf{D}_n\|^3 \leq p^{1/2} \sum_{i=1}^{p} \mathbb{E}\|\mathbf{Z}^{(i)} - \mathbf{D}_n^{(i)}\|^3$

$$
\leq p^{1/2} \frac{1080}{\sqrt{\pi}(\log 2)^{3/2}} n^{-3/2} (\log 2n)^{3/2} \sum_{i=1}^{p} |\Sigma_{i,i}|^{3/2};
$$

(C) $\quad (\mathbb{E}\|\mathbf{Z}\|^3)^{2/3} \leq \left(p^{1/2} \sum_{i=1}^{p} \mathbb{E}\|\mathbf{Z}^{(i)}\|^3 \right)^{2/3} \leq \frac{9p^{1/3}}{2\pi^{1/3}} \left(\sum_{i=1}^{p} |\Sigma_{i,i}|^{3/2} \right)^{2/3}.$

Therefore:

$$
|\mathbb{E}g(\mathbf{D}_n) - \mathbb{E}g(\mathbf{Z})|
$$
$$
\overset{\text{MVT}}{\leq} \mathbb{E} \left[\sup_{c \in [0,1]} \|Dg(\tilde{\mathbf{Z}} + c(\mathbf{D}_n - \mathbf{Z}))\| \|\mathbf{Z} - \mathbf{D}_n\| \right]
$$
$$
\leq \|g\|_M \mathbb{E} \left[\sup_{c \in [0,1]} (1 + \|\mathbf{Z} + c(\mathbf{D}_n - \mathbf{Z})\|^2) \|\mathbf{Z} - \mathbf{D}_n\| \right]
$$
$$
\leq \|g\|_M \left\{ \mathbb{E}\|\mathbf{Z} - \mathbf{D}_n\| + 2\mathbb{E}\|\mathbf{Z} - \mathbf{D}_n\|^3 + 2(\mathbb{E}\|\mathbf{Z}\|^3)^{2/3} (\mathbb{E}\|\mathbf{D}_n - \mathbf{Z}\|^3)^{1/3} \right\}
$$
$$
\leq \|g\|_M \left\{ n^{-1/2} \sqrt{\log 2n} \left[\frac{6\sqrt{5}}{\sqrt{2\log 2}} \left(\sum_{i=1}^{p} |\Sigma_{i,i}| \right)^{1/2} + \frac{54 \cdot 5^{1/3} p^{1/2}}{\sqrt{\pi \log 2}} \sum_{i=1}^{p} |\Sigma_{i,i}|^{3/2} \right] \right.
$$
$$
\left. + n^{-3/2} (\log 2n)^{3/2} p^{1/2} \frac{2160}{\sqrt{\pi}(\log 2)^{3/2}} \sum_{i=1}^{p} |\Sigma_{i,i}|^{3/2} \right\}. \tag{5.21}
$$

We now sum (5.19) and (5.20) and sum them over j, which, combined with (5.21) yields the result. $\square$

5.5. *Proof of Theorem 3.9*

In **Step 1** of the proof of Theorem 3.9, we consider a scaled sum of i.i.d random variables $w(X_i)$ and apply Lemma 5.7 together with an argument similar to **Step 1** and **Step 3** of the proof of Theorem 3.1 in order to bound the distance between this scaled sum and $\mathbf{Z}$. In **Step 2** we bound the distance between this scaled sum and our original process $\mathbf{Y}_n$ by bounding the second moment of the supremum distance between them and then using the mean value theorem.

Proof of Theorem 3.9. Let $g \in M^2$.

Step 1. As in the proof of the invariance principle for U-statistics of [18], we start by considering the behaviour of the following process $(\tilde{\mathbf{Y}}_n(t), t \geq 0)$:

$$\tilde{\mathbf{Y}}_n(t) = \frac{n^{-3/2}}{\sigma_w t} \sum_{1 \leq i_1 < i_2 \leq \lfloor nt \rfloor} \left(w(X_{i_1}) + w(X_{i_2}) \right) = \frac{1}{\sqrt{n}\sigma_w} \sum_{i=1}^{n} w(X_i) J_{i,n}(t),$$

where $J_{i,n}(t) = \frac{(\lfloor nt \rfloor - 1)\mathbb{1}_{[i/n,1]}(t)}{nt}$. Recall that $w(x) = \mathbb{E}h(X_1, x)$. Let $\mathbf{A}_n(t) = n^{-1/2} \sum_{i=1}^{n} Z_i J_{i,n}(t)$ and $\hat{\mathbf{A}}_n(t) = n^{-1/2} \sum_{i=1}^{\lfloor nt \rfloor} Z_i$, where $Z_i \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, 1)$ and the collections $(Z_i)_{i=1}^{n}$ and $(X_i)_{i=1}^{n}$ are independent.

Note that Lemma 5.7 readily yields that:

$$\left| \mathbb{E}g(\tilde{\mathbf{Y}}_n) - \mathbb{E}g(\mathbf{A}_n) \right| \leq \frac{\|g\|_M}{6\sigma_w^3} n^{-1/2} \left(\mathbb{E}|w(X_1)|^3 + 2\sigma_w^2 \mathbb{E}|w(X_1)| \right), \tag{5.22}$$

as $\|J_{i,n}\| \leq 1$ for all $i, n \in \mathbb{N}$ and $w(X_i)$'s for $i = 1, \ldots, n$ are independent.

We see that, by Doob's L^2 inequality, we have for every m:

$$\mathbb{E}\left[\max_{1 \leq l \leq m} \left| \sum_{i=1}^{l} Z_i \right| \right]^2 \leq 4m = 4\sum_{i=1}^{m} 1.$$

Therefore, using [15, Theorem 1] for inequality $(*)$, we obtain:

$$\text{(A)} \quad \mathbb{E}\|\mathbf{A}_n - \hat{\mathbf{A}}_n\|^2 \leq n^{-1}\mathbb{E}\left[\max_{1 \leq l \leq n} \left| \frac{l-1}{l+1} - 1 \right| \left| \sum_{i=1}^{l} Z_i \right| \right]^2$$

$$\leq n^{-1} 4\mathbb{E}\left[\max_{1 \leq l \leq n} \left| \frac{\sum_{i=1}^{l} Z_i}{l+1} \right| \right]^2$$

$$\overset{(*)}{\leq} 32n^{-1} \sum_{i=1}^{n} \frac{1}{i^2}$$

$$\leq \frac{16\pi^2}{3} n^{-1};$$

$$\text{(B)} \quad \mathbb{E}\|\mathbf{A}_n - \hat{\mathbf{A}}_n\| \leq \sqrt{\mathbb{E}\|\mathbf{A}_n - \hat{\mathbf{A}}_n\|^2} \leq \frac{4\pi}{\sqrt{3}} n^{-1/2}. \tag{5.23}$$

Doob's L^2 inequality readily gives us:

$$\mathbb{E}\|\hat{\mathbf{A}}_n\|^2 = \mathbb{E}\left[\left(\max_{1 \leq m \leq n} n^{-1/2} \left| \sum_{i=1}^{m} Z_i \right| \right)^2 \right] \leq 4. \tag{5.24}$$

It follows that:

$$|\mathbb{E}g(\mathbf{A}_n) - \mathbb{E}g(\hat{\mathbf{A}}_n)|$$

$$\leq \mathbb{E}\left[\sup_{c \in [0,1]} \|Dg((1-c)\hat{\mathbf{A}}_n + c\mathbf{A}_n)\| \|\mathbf{A}_n - \hat{\mathbf{A}}_n\|\right]$$

$$\leq \|g\|_{M^2} \mathbb{E}\left[\sup_{c \in [0,1]} (1 + \|\hat{\mathbf{A}}_n + c(\mathbf{A}_n - \hat{\mathbf{A}}_n)\|) \|\mathbf{A}_n - \hat{\mathbf{A}}_n\|\right]$$

$$\leq \|g\|_{M^2} \left(\mathbb{E}\|\mathbf{A}_n - \hat{\mathbf{A}}_n\| + \mathbb{E}\|\mathbf{A}_n - \hat{\mathbf{A}}_n\|^2 + \sqrt{\mathbb{E}\|\hat{\mathbf{A}}_n\|^2}\sqrt{\mathbb{E}\|\mathbf{A}_n - \hat{\mathbf{A}}_n\|^2}\right)$$

$$\leq \|g\|_{M^2} \left(\frac{12\pi}{\sqrt{3}} n^{-1/2} + \frac{16\pi^2}{3} n^{-1}\right), \tag{5.25}$$

where the first inequality follows from the mean value theorem and the last one follows from (5.23) and (5.24). Also, by [16, Lemma 3] and Doob's L^2 inequality:

(A) $\mathbb{E}\|\hat{\mathbf{A}}_n - \mathbf{Z}\| \leq \dfrac{30}{\sqrt{\pi \log 2}} n^{-1/2}\sqrt{\log 2n}$

(B) $\mathbb{E}\|\hat{\mathbf{A}}_n - \mathbf{Z}\|^2 \leq \dfrac{90}{\log 2} n^{-1} \log 2n$

(C) $\mathbb{E}\|\mathbf{Z}\|^2 \leq 4$

and therefore:

$$|\mathbb{E}g(\hat{\mathbf{A}}_n) - \mathbb{E}g(\mathbf{Z})|$$

$$\leq \|g\|_{M^2} \left(\mathbb{E}\|\hat{\mathbf{A}}_n - \mathbf{Z}\| + \mathbb{E}\|\hat{\mathbf{A}}_n - \mathbf{Z}\|^2 + \sqrt{\mathbb{E}\|\mathbf{Z}\|^2}\sqrt{\mathbb{E}\|\hat{\mathbf{A}}_n - \mathbf{Z}\|^2}\right)$$

$$\leq \|g\|_{M^2} n^{-1/2} \left[\left(\frac{30}{\sqrt{\pi \log 2}} + \frac{12\sqrt{5}}{\sqrt{2 \log 2}}\right)\sqrt{\log 2n} + \frac{90}{\log 2} n^{-1/2} \log 2n\right]. \tag{5.26}$$

Step 2. We now wish to find a bound on $|\mathbb{E}g(\tilde{\mathbf{Y}}_n) - \mathbb{E}g(\mathbf{Y}_n)|$. Note that:

$$\mathbf{Y}_n - \tilde{\mathbf{Y}}_n = \frac{n^{-3/2}}{\sigma_w t} \sum_{1 \leq i_1 < i_2 \leq \lfloor nt \rfloor} \left(h(X_{i_1}, X_{i_2}) - w(X_{i_1}) - w(X_{i_2})\right).$$

Let $\phi_h^2 = \mathbb{E}h^2(X_1, X_2)$. First, note that, if $\mu = \mathcal{L}(X_1)$ (i.e. μ is the law of X_1),

$$\mathbb{E}\left[(h(X_1, X_2) - w(X_1) - w(X_2))(h(X_1, X_3) - w(X_1) - w(X_3))\right]$$

$$= \mathbb{E}\left[h(X_1, X_2)h(X_1, X_3)\right] - 2\mathbb{E}\left[h(X_1, X_2)w(X_1)\right] + \mathbb{E}w^2(X_1)$$

$$= \int \int \int h(x, y)h(x, z)\mu(dx)\mu(dy)\mu(dz)$$

$$\quad - 2\int \int h(x, y)\int h(x, z)\mu(dz)\mu(dx)\mu(dy)$$

$$\quad + \int \int h(x, y)\mu(dy)\int h(x, z)\mu(dz)\mu(dx) = 0,$$

where the first equality follows by the fact that $w(X_2)$ is independent of $h(X_1, X_3)$, $w(X_1)$ and $w(X_3)$, $w(X_3)$ is independent of $h(X_1, X_2)$, $w(X_1)$ and $w(X_2)$, and $\mathbb{E}w(X_2) = \mathbb{E}w(X_3) = 0$.

Therefore:

$$
\mathbb{E}\left[\sum_{1\le i_1<i_2\le m}\left(h(X_{i_1},X_{i_2})-w(X_{i_1})-w(X_{i_2})\right)\right]^2
$$

$$
=\binom{m}{2}\mathbb{E}\left[h(X_1,X_2)-w(X_1)-w(X_2)\right]^2
$$

$$
=\binom{m}{2}\left[\sigma_h^2+2\sigma_w^2-4\int\int h(x,y)\int h(x,z)\mu(dz)\mu(dx)\mu(dy)\right]
$$

$$
=\binom{m}{2}(\sigma_h^2-2\sigma_w^2). \tag{5.27}
$$

Now, $\sum_{1\le i_1<i_2\le m}\left(h(X_{i_1},X_{i_2})-w(X_{i_1})-w(X_{i_2})\right)$ is a martingale with respect to the filtration $(\sigma(X_1,\ldots,X_m))_{m\ge1}$. Indeed:

$$
\mathbb{E}\left[\sum_{1\le i_1<i_2\le m+1}\left(h(X_{i_1},X_{i_2})-w(X_{i_1})-w(X_{i_2})\right)\middle| X_1,\ldots,X_m\right]
$$

$$
=\sum_{1\le i_1<i_2\le m}\left(h(X_{i_1},X_{i_2})-w(X_{i_1})-w(X_{i_2})\right)
$$

$$
+\mathbb{E}\left[\sum_{i=1}^m (h(X_i,X_{m+1})-w(X_i)-w(X_{m+1}))\middle| X_1,\ldots,X_m\right]
$$

$$
=\sum_{1\le i_1<i_2\le m}\left(h(X_{i_1},X_{i_2})-w(X_{i_1})-w(X_{i_2})\right)+\sum_{i=1}^m\left(\mathbb{E}\left[h(X_i,X_{m+1})|X_i\right]-w(X_i)\right)
$$

$$
=\sum_{1\le i_1<i_2\le m}\left(h(X_{i_1},X_{i_2})-w(X_{i_1})-w(X_{i_2})\right).
$$

Hence, Doob's inequalities give us, for every m, such that $1\le m\le n$:

$$
\mathbb{E}\left[\max_{1\le l\le m}\left|\sum_{1\le i_1<i_2\le l}\left(h(X_{i_1},X_{i_2})-w(X_{i_1})-w(X_{i_2})\right)\right|\right]^2 \overset{(5.27)}{\le} 4\binom{m}{2}(\sigma_h^2-2\sigma_w^2).
$$

Then, by [15, Theorem 1], applied with $\beta_i=\alpha_i=i$ and $r=2$, and using the fact that $\binom{m}{2}=\sum_{i=1}^m(i-1)$, we obtain:

$$
\mathbb{E}\|\mathbf{Y}_n-\tilde{\mathbf{Y}}_n\|^2=\frac{n^{-3}}{\sigma_w^2}\mathbb{E}\left[\sup_{t\in[0,1]}\left|t^{-1}\sum_{1\le i_1<i_2\le\lfloor nt\rfloor}\left(h(X_{i_1},X_{i_2})-w(X_{i_1})-w(X_{i_2})\right)\right|\right]^2
$$

$$
=\frac{n^{-1}}{\sigma_w^2}\mathbb{E}\left[\max_{1\le l\le n}l^{-1}\left|\sum_{1\le i_1<i_2\le l}\left(h(X_{i_1},X_{i_2})-w(X_{i_1})-w(X_{i_2})\right)\right|\right]^2
$$

$$
\le 16\left(\frac{\sigma_h^2}{\sigma_w^2}-2\right)\sum_{i=1}^n\frac{1}{i}n^{-1}\le 16\left(\frac{\sigma_h^2}{\sigma_w^2}-2\right)n^{-1}\log 3n. \tag{5.28}
$$

Also, by Doob's L^2 inequality:

$$\mathbb{E}\|\tilde{\mathbf{Y}}_n\|^2 = n^{-3}\mathbb{E}\left[\sup_{t\in[0,1]}\left|\frac{\lfloor nt\rfloor - 1}{t}\sum_{i=1}^{\lfloor nt\rfloor}\frac{w(X_i)}{\sigma_w}\right|\right]^2$$

$$= n^{-1}\mathbb{E}\left[\sup_{1\le l\le n}\left|\frac{l-1}{l}\sum_{i=1}^{l}\frac{w(X_i)}{\sigma_w}\right|\right]^2 \le 4. \tag{5.29}$$

Therefore:

$$|\mathbb{E}g(\mathbf{Y}_n) - \mathbb{E}g(\tilde{\mathbf{Y}}_n)|$$

$$\le \mathbb{E}\left[\sup_{c\in[0,1]}\|Dg\left((1-c)\tilde{\mathbf{Y}}_n + c\mathbf{Y}_n\right)\|\|\mathbf{Y}_n - \tilde{\mathbf{Y}}_n\|\right]$$

$$\le \|g\|_{M^2}\mathbb{E}\left[\sup_{c\in[0,1]}(1 + \|\tilde{\mathbf{Y}}_n + c(\mathbf{Y}_n - \hat{\mathbf{A}}_n)\|)\|\mathbf{Y}_n - \tilde{\mathbf{Y}}_n\|\right]$$

$$\le \|g\|_{M^2}\left(\mathbb{E}\|\mathbf{Y}_n - \tilde{\mathbf{Y}}_n\| + \mathbb{E}\|\mathbf{Y}_n - \tilde{\mathbf{Y}}_n\|^2 + \sqrt{\mathbb{E}\|\tilde{\mathbf{Y}}_n\|^2}\sqrt{\mathbb{E}\|\mathbf{Y}_n - \tilde{\mathbf{Y}}_n\|^2}\right)$$

$$\le \|g\|_{M^2}\left(12\left(\frac{\sigma_h^2}{\sigma_w^2} - 2\right)^{1/2}n^{-1/2}\sqrt{\log 3n} + 16\left(\frac{\sigma_h^2}{\sigma_w^2} - 2\right)n^{-1}\log 3n\right), \tag{5.30}$$

where the first inequality follows from the mean value theorem and the last one follows by (5.28) and (5.29).

We combine (5.22), (5.25), (5.26) and (5.30) to obtain the assertion. □

Remark 5.8. While, in the proof of Theorem 3.9, it is possible to obtain a bound on $|\mathbb{E}g(\tilde{\mathbf{Y}}_n) - \mathbb{E}g(\mathbf{Z})|$ for any $g \in M$, using methods analogous to those which let us prove Theorem 3.1, the situation becomes more complicated when it comes to approximating the remainder. This is because using Doob's L^3 inequality and [15, Corollary 1] for $\mathbb{E}\|\mathbf{Y}_n - \tilde{\mathbf{Y}}_n\|^3$ gives a bound which does not converge to 0 with n. Therefore, in (5.30) we cannot go beyond the second moment of $\|\mathbf{Y}_n - \tilde{\mathbf{Y}}_n\|$. Hence, for our technique of proof, it is necessary that we assume $g \in M^2$, as defined by (2.3).

Remark 5.9. The stronger assumption of $g \in M^1$ in Theorem 3.9 would simplify its proof. Namely, using the notation of the proof of Theorem 3.9, we could treat $\tilde{\mathbf{Y}}_n$ as a scaled sum of i.i.d. mean zero, variance 1 random variables $\frac{w(X_i)}{\sigma_w}$. Using (5.23) and applying Theorem 3.1 gives:

$$\left|\mathbb{E}g(\tilde{\mathbf{Y}}_n) - \mathbb{E}g(\mathbf{Z})\right| \le \frac{\|g\|_{M^1}}{2}n^{-1/2}\left(\frac{\mathbb{E}|w(X_1)|^3}{\sigma_w^3} + 8 + 10\sqrt{\log 2n}\right)$$

and (5.30) could be substituted with:

$$\left|\mathbb{E}g(\mathbf{Y}_n) - \mathbb{E}g(\tilde{\mathbf{Y}}_n)\right| \le \|g\|_{M^1}\mathbb{E}\|\mathbf{Y}_n - \tilde{\mathbf{Y}}_n\| \overset{(5.28)}{\le} \|g\|_M\left(\frac{\sigma_h^2}{\sigma_w^2} - 2\right)^{1/2}\frac{4\sqrt{\log 3n}}{n^{1/2}}.$$

Declaration of competing interest

The author declares that he has no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix. Proof of Proposition 2.3

As in the proof of [3, Proposition 3.1], we note that, by Skorokhod's representation theorem, $\mathbf{Z}_n$ and $\mathbf{Z}$ can be defined on the same probability space in such a way that $\|\mathbf{Z}_n - \mathbf{Z}\| \xrightarrow{n\to\infty} 0$ a.s. (as $\mathbf{Z}$ is continuous). The fact that $C([0,1], \mathbb{R}^p)$ equipped with norm $\|\cdot\|$ is separable, by the Stone–Weierstrass theorem, lets us use the argument of the proof of the Skorokhod representation theorem presented in [5, Chapter 5] and conclude that it is enough to show that $\mathbb{P}[\mathbf{Y}_n \in B] \to \mathbb{P}[\mathbf{Z} \in B]$ for all sets $B = \bigcap_{1 \le l \le L} B_l$, where $B_l = \{w \in D^p : \|w - s_l\| < \gamma_l\}$, $s_l \in C([0,1], \mathbb{R}^p)$ and γ_l is such that $\mathbb{P}[\mathbf{Z} \in \partial B_l] = 0$. Let us fix such a set B.

Let $\phi : \mathbb{R}^+ \to [0,1]$ be a non-increasing, three times continuously differentiable function satisfying, $\phi(x) = 1$ for $x \le 0$ and $\phi(x) = 0$ for $x \ge 1$ and fix some $0 < \epsilon, \eta_n \le 1, p_n \ge 4$. Define $g_{l,n} : D^p \to \mathbb{R}$ by:

$$g_{l,n}(w) = \phi\left(\frac{\left\| \sqrt{(\epsilon\gamma_l)^2 + \sum_{i=1}^{p}\left((w - s_l)^{(i)}\right)^2} \right\|_{p_n} - \gamma_l\sqrt{1 + \epsilon^2}}{\eta_n} \right), \tag{A.1}$$

where $\|w\|_{p_n} := \left(\int_0^1 |w(t)|^{p_n} dt \right)^{1/p_n}$ for any $w \in D^p$. We have the following result:

Lemma A.1. *For any finite L:*

$$\left\| \prod_{l=1}^{L} g_{l,n} \right\|_{M^0} \le \tilde{C} p_n^2 \eta_n^{-3}. \tag{A.2}$$

for a constant $\tilde{C}$ independent of p_n and η_n (which might depend on ϵ or γ_l's).

Proof. First, $\phi, \phi', \phi'', \phi'''$ are all everywhere continuous and constant outside of the compact interval $[0,1]$ and therefore bounded. Therefore also $\frac{|\phi''(x+h) - \phi''(x)|}{|h|}$ must be uniformly bounded.

Furthermore, let

$$f(w) = \frac{\left\| \sqrt{(\epsilon\gamma_l)^2 + \sum_{i=1}^{p}\left((w - s_l)^{(i)}\right)^2} \right\|_{p_n}}{\eta_n}, \tag{A.3}$$

and denote by $|\cdot|$ the Euclidean norm, and by $\langle \cdot \rangle$ the Euclidean inner product.

Step 1: Bounding the first derivative of f of (A.3)

We have that, for any $h \in D^p$,

$$
Df(w)[h] = \frac{1}{p_n \eta_n} \left(\int_0^1 ((\epsilon \gamma_l)^2 + |w - s_l|^2(t))^{p_n/2} dt \right)^{1/p_n - 1}
$$
$$
\cdot \frac{p_n}{2} \int_0^1 \left((\epsilon \gamma_l)^2 + |w - s_l|^2(t) \right)^{p_n/2 - 1} \cdot 2 \langle (w - s_l)(t), h(t) \rangle \, dt. \tag{A.4}
$$

Applying Hölder's inequality with coefficients $\frac{p_n}{p_n - 2k}$ and $\frac{p_n}{2k}$ and Cauchy–Schwarz inequality, we obtain that, for any $k = 1, 2, 3$, and $h_1, \ldots, h_k \in D^p$,

$$
\left| \int_0^1 \left((\epsilon \gamma_l)^2 + |w - s_l|^2(t) \right)^{p_n/2 - k} \langle (w - s_l)(t), h_1(t) \rangle \ldots \langle (w - s_l)(t), h_k(t) \rangle \, dt \right|
$$
$$
\leq \left(\int_0^1 \left((\epsilon \gamma_l)^2 + |w - s_l|^2(t) \right)^{p_n/2} dt \right)^{1 - 2k/p_n}
$$
$$
\cdot \left(\int_0^1 |w - s_l|^{p_n/2}(t) |h_1|^{p_n/(2k)}(t) \ldots |h_k|^{p_n/(2k)}(t) dt \right)^{2k/p_n}
$$
$$
\leq \left(\int_0^1 \left((\epsilon \gamma_l)^2 + |w - s_l|^2(t) \right)^{p_n/2} dt \right)^{1 - 2k/p_n} \cdot \left(\int_0^1 |w - s_l|^{p_n}(t) dt \right)^{k/p_n} \prod_{i=1}^k \|h_i\|_{p_n}. \tag{A.5}
$$

Applying (A.5) for $k = 1$, together with (A.4), we get

$$
|Df(w)[h]| \leq \frac{1}{\eta_n} \left(\frac{\int_0^1 |w - s_l|^{p_n}(t) dt}{\int_0^1 ((\epsilon \gamma_l)^2 + |w - s|^2(t))^{p_n/2} dt} \right)^{1/p_n} \|h\|_{p_n} \leq \frac{\|h\|_\infty}{\eta_n}
$$

and so

$$
\sup_{w \in D^p} \|Df(w)\| \leq \frac{1}{\eta_n}. \tag{A.6}
$$

Step 2: Bounding the second derivative of f of (A.3)

Note that, for any $h_1, h_2 \in D^p$,

$$
D^2 f(w)[h_1, h_2] = A + B \tag{A.7}
$$

for

$$
A = \frac{1}{\eta_n} \left[\int_0^1 \left((\epsilon \gamma_l)^2 + |w - s_l|^2(t) \right)^{p_n/2 - 1} \cdot \langle (w - s_l)(t), h_2(t) \rangle \, dt \right]
$$
$$
\cdot \frac{1 - p_n}{p_n} \left[\int_0^1 \left((\epsilon \gamma_l)^2 + |w - s_l|^2(t) \right)^{p_n/2} dt \right]^{1/p_n - 2}
$$
$$
\cdot \frac{p_n}{2} \int_0^1 \left((\epsilon \gamma_l)^2 + |w - s_l|^2(t) \right)^{p_n/2 - 1} \cdot 2 \langle (w - s_l)(t), h_1(t) \rangle \, dt
$$
$$
= \frac{1 - p_n}{\eta_n} \prod_{i=1}^2 \left\{ \left[\int_0^1 \left((\epsilon \gamma_l)^2 + |w - s_l|^2(t) \right)^{p_n/2 - 1} \cdot \langle (w - s_l)(t), h_i(t) \rangle \, dt \right] \right\}
$$
$$
\cdot \left[\int_0^1 \left((\epsilon \gamma_l)^2 + |w - s_l|^2(t) \right)^{p_n/2} dt \right]^{1/p_n - 2}
$$

$$B = \frac{1}{\eta_n} \left[\int_0^1 \left((\epsilon\gamma_l)^2 + |w - s_l|^2(t)\right)^{p_n/2} dt \right]^{1/p_n - 1}$$

$$\cdot \left[\int_0^1 \frac{p_n - 2}{2} \left((\epsilon\gamma_l)^2 + |w - s_l|^2(t)\right)^{p_n/2 - 2} \right.$$

$$\cdot 2 \langle (w - s_l)(t), h_1(t) \rangle \langle (w - s_l)(t), h_2(t) \rangle dt$$

$$\left. + \int_0^1 \left((\epsilon\gamma_l)^2 + |w - s_l|^2(t)\right)^{p_n/2 - 1} \langle h_1(t), h_2(t) \rangle dt \right]. \tag{A.8}$$

Notice that, by (A.5) with $k = 1$,

$$|A| \le \frac{p_n - 1}{\eta_n} \left(\frac{\left(\int_0^1 |w - s_l|^{p_n}(t) dt\right)^2}{\left(\int_0^1 ((\epsilon\gamma_l)^2 + |w - s_l|^2(t))^{p_n/2} dt\right)^3} \right)^{1/p_n} \|h_1\|_{p_n} \|h_2\|_{p_n}. \tag{A.9}$$

Furthermore, by Hölder's inequality with coefficients $\frac{p_n}{p_n - 2}$ and $\frac{p_n}{2}$ and by the Cauchy–Schwarz inequality,

$$\left| \int_0^1 \left((\epsilon\gamma_l)^2 + |w - s_l|^2(t)\right)^{p_n/2 - 1} \langle h_1(t), h_2(t) \rangle dt \right|$$

$$\le \left(\int_0^1 \left((\epsilon\gamma_l)^2 + |w - s_l|^2(t)\right)^{p_n/2} \right)^{1 - 2/p_n} \left(\int_0^1 \langle h_1(t), h_2(t) \rangle^{p_n/2} \right)^{2/p_n}$$

$$\le \left(\int_0^1 \left((\epsilon\gamma_l)^2 + |w - s_l|^2(t)\right)^{p_n/2} \right)^{1 - 2/p_n} \|h_1\|_{p_n} \|h_2\|_{p_n}. \tag{A.10}$$

By (A.5) and (A.10),

$$|B| \le \frac{p_n - 2}{\eta_n} \left(\frac{\left(\int_0^1 |w - s_l|^{p_n}(t) dt\right)^2}{\left(\int_0^1 ((\epsilon\gamma_l)^2 + |w - s_l|^2(t))^{p_n/2} dt\right)^3} \right)^{1/p_n} \|h_1\|_{p_n} \|h_2\|_{p_n}$$

$$+ \frac{1}{\eta_n} \left(\int_0^1 \left((\epsilon\gamma_l)^2 + |w - s_l|^2(t)\right)^{p_n/2} dt \right)^{-1/p_n} \|h_1\|_{p_n} \|h_2\|_{p_n}. \tag{A.11}$$

By (A.7), (A.9) and (A.11),

$$|D^2 f(w)[h_1, h_2]|$$

$$\le \left[\frac{2p_n - 3}{\eta_n} \left(\frac{\left(\int_0^1 |w - s_l|^{p_n}(t) dt\right)^2}{\left(\int_0^1 ((\epsilon\gamma_l)^2 + |w - s_l|^2(t))^{p_n/2} dt\right)^3} \right)^{1/p_n} \right.$$

$$\left. + \frac{1}{\eta_n} \left(\int_0^1 \left((\epsilon\gamma_l)^2 + |w - s_l|^2(t)\right)^{p_n/2} dt \right)^{-1/p_n} \right] \|h_1\|_{p_n} \|h_2\|_{p_n}$$

$$= \frac{1}{\eta_n} \left(\int_0^1 \left((\epsilon\gamma_l)^2 + |w - s_l|^2(t)\right)^{p_n/2} dt \right)^{-1/p_n}$$

$$\cdot \left[(2p_n - 3) \left(\frac{\left(\int_0^1 |w - s_l|^{p_n}(t) dt \right)^2}{\left(\int_0^1 ((\epsilon \gamma_l)^2 + |w - s_l|^2(t))^{p_n/2} dt \right)^2} \right)^{1/p_n} + 1 \right] \|h_1\|_{p_n} \|h_2\|_{p_n}$$

$$\leq \frac{2p_n - 2}{\eta_n} \left(\int_0^1 ((\epsilon \gamma_l)^2 + |w - s_l|^2(t))^{p_n/2} dt \right)^{-1/p_n} \|h_1\|_{p_n} \|h_2\|_{p_n}$$

$$\leq \frac{2p_n - 2}{\eta_n (\epsilon \gamma_l)} \|h_1\|_\infty \|h_2\|_\infty$$

and so

$$\sup_{w \in D^p} \|D^2 f(w)\| \leq 2 \frac{p_n - 1}{\eta_n (\epsilon \gamma_l)}. \tag{A.12}$$

Step 3: Bounding the third derivative of f of (A.3)

Finally, for any $h_1, h_2, h_3 \in D^p$,

$$D^3 f(w)[h_1, h_2, h_3] = C + D, \tag{A.13}$$

where C comes from differentiating A of (A.8) and is given by

$$C = E + F$$

for

$$E = \frac{1 - p_n}{\eta_n} \sum_{1 \leq i \neq j \leq 2} \left\{ \int_0^1 ((\epsilon \gamma_l)^2 + |w - s_l|^2(t))^{p_n/2-1} \langle (w - s_l)(t), h_i(t) \rangle \, dt \right.$$

$$\cdot \int_0^1 \left[\frac{p_n - 2}{2} ((\epsilon \gamma_l)^2 + |w - s_l|^2(t))^{p_n/2-2} \right.$$

$$\times \langle (w - s_l)(t), h_j(t) \rangle \cdot 2 \langle (w - s_l)(t), h_3(t) \rangle$$

$$\left. + ((\epsilon \gamma_l)^2 + |w - s_l|^2(t))^{p_n/2-1} \langle h_j(t), h_3(t) \rangle \right] dt$$

$$\left. \cdot \left[\int_0^1 ((\epsilon \gamma_l)^2 + |w - s_l|^2(t))^{p_n/2} dt \right]^{1/p_n-2} \right\}$$

$$F = \frac{(1 - p_n)(1 - 2p_n)}{p_n \eta_n} \left\{ \prod_{i=1}^3 \left[\int_0^1 ((\epsilon \gamma_l)^2 + |w - s_l|^2(t))^{p_n/2-1} \cdot \langle (w - s_l)(t), h_i(t) \rangle \, dt \right] \right\}$$

$$\cdot \left[\int_0^1 ((\epsilon \gamma_l)^2 + |w - s_l|^2(t))^{p_n/2} dt \right]^{1/p_n-3} \tag{A.14}$$

and D comes from differentiating B of (A.8) and is given by

$$D = G + H$$

for

$$G = \frac{1 - p_n}{\eta_n} \left[\int_0^1 ((\epsilon \gamma_l)^2 + |w - s_l|^2(t))^{p_n/2} dt \right]^{1/p_n-2}$$

$$\cdot \int_0^1 ((\epsilon \gamma_l)^2 + |w - s_l|^2(t))^{p_n/2-1} \langle (w - s_l)(t), h_3(t) \rangle \, dt$$

$$\cdot \left[\int_0^1 (p_n - 2) \left((\epsilon \gamma_l)^2 + |w - s_l|^2(t) \right)^{p_n/2 - 2} \right.$$

$$\cdot \langle (w - s_l)(t), h_1(t) \rangle \langle (w - s_l)(t), h_2(t) \rangle \, dt$$

$$\left. + \int_0^1 \left((\epsilon \gamma_l)^2 + |w - s_l|^2(t) \right)^{p_n/2 - 1} \langle h_1(t), h_2(t) \rangle \, dt \right]$$

$$H = \frac{p_n - 2}{\eta_n} \left[\int_0^1 \left((\epsilon \gamma_l)^2 + |w - s_l|^2(t) \right)^{p_n/2} dt \right]^{1/p_n - 1}$$

$$\cdot \left\{ \int_0^1 \left[(p_n - 2) \left((\epsilon \gamma_l)^2 + |w - s_l|^2(t) \right)^{p_n/2 - 2} \right. \right.$$

$$\left. \times \sum_{\substack{1 \le i,j,k \le 3 \\ i,j,k \text{ distinct}}} \langle (w - s_l)(t), h_i(t) \rangle \langle h_j(t), h_k(t) \rangle \right] dt$$

$$\left. + (p_n - 4) \int_0^1 \left[\left((\epsilon \gamma_l)^2 + |w - s_l|^2(t) \right)^{p_n/2 - 3} \prod_{i=1}^3 \langle (w - s_l)(t), h_i(t) \rangle \right] dt \right\}.$$

$$(A.15)$$

So

$$D^3 f(w)[h_1, h_2, h_3] = E + F + G + H \tag{A.16}$$

for E, F, G, H defined by (A.14) and (A.15). By (A.5) and (A.10),

$$|E| \le \frac{2(p_n - 1)\|h_1\|_{p_n} \|h_2\|_{p_n} \|h_3\|_{p_n}}{\eta_n}$$

$$\cdot \left(\frac{(p_n - 2) \left(\int_0^1 |w - s_l|^{p_n}(t) dt \right)^{3/p_n}}{\left(\int_0^1 \left((\epsilon \gamma_l)^2 + |w - s_l|^2(t) \right)^{p_n/2} dt \right)^{5/p_n}} \right.$$

$$\left. + \frac{\left(\int_0^1 |w - s_l|^{p_n}(t) dt \right)^{1/p_n}}{\left(\int_0^1 \left((\epsilon \gamma_l)^2 + |w - s_l|^2(t) \right)^{p_n/2} dt \right)^{3/p_n}} \right)$$

$$|F| \le \frac{(p_n - 1)(2p_n - 1)\|h_1\|_{p_n} \|h_2\|_{p_n} \|h_3\|_{p_n}}{p_n \eta_n} \cdot \frac{\left(\int_0^1 |w - s_l|^{p_n}(t) dt \right)^{3/p_n}}{\left(\int_0^1 \left((\epsilon \gamma_l)^2 + |w - s_l|^2(t) \right)^{p_n/2} dt \right)^{5/p_n}}$$

$$|G| \le \frac{(p_n - 1)\|h_1\|_{p_n} \|h_2\|_{p_n} \|h_3\|_{p_n}}{\eta_n}$$

$$
\cdot \left(\frac{(p_n - 2) \left(\int_0^1 |w - s_l|^{p_n}(t)dt \right)^{3/p_n}}{\left(\int_0^1 \left((\epsilon \gamma_l)^2 + |w - s_l|^2(t) \right)^{p_n/2} dt \right)^{5/p_n}} \right.
$$

$$
\left. + \frac{\left(\int_0^1 |w - s_l|^{p_n}(t)dt \right)^{1/p_n}}{\left(\int_0^1 \left((\epsilon \gamma_l)^2 + |w - s_l|^2(t) \right)^{p_n/2} dt \right)^{3/p_n}} \right)
$$

$$
|H| \le \frac{(p_n - 2)\|h_1\|_{p_n}\|h_2\|_{p_n}\|h_3\|_{p_n}}{\eta_n}
$$

$$
\cdot \left(\frac{(p_n - 4) \left(\int_0^1 |w - s_l|^{p_n}(t)dt \right)^{3/p_n}}{\left(\int_0^1 \left((\epsilon \gamma_l)^2 + |w - s_l|^2(t) \right)^{p_n/2} dt \right)^{5/p_n}} \right.
$$

$$
\left. + \frac{6 \left(\int_0^1 |w - s_l|^{p_n}(t)dt \right)^{1/p_n}}{\left(\int_0^1 \left((\epsilon \gamma_l)^2 + |w - s_l|^2(t) \right)^{p_n/2} dt \right)^{3/p_n}} \right), \tag{A.17}
$$

where the inequality for $|H|$ uses the following bound obtained by applying Hölder's inequality with coefficients $\frac{p_n}{p_n - 4}$ and $\frac{p_n}{4}$ and Cauchy–Schwarz inequality

$$
\left| \int_0^1 \left[\left((\epsilon \gamma_l)^2 + |w - s_l|^2(t) \right)^{p_n/2 - 2} \sum_{\substack{1 \le i,j,k \le 3 \\ i,j,k \text{ distinct}}} \langle (w - s_l)(t), h_i(t) \rangle \langle h_j(t), h_k(t) \rangle \right] dt \right|
$$

$$
\le \sum_{\substack{1 \le i,j,k \le 3 \\ i,j,k \text{ distinct}}} \left(\int_0^1 \left((\epsilon \gamma_l)^2 + |w - s_l|^2(t) \right)^{p_n/2} dt \right)^{1 - 4/p_n}
$$

$$
\times \left(\int_0^1 |w - s_l|^{p_n/4}(t) \prod_{i=1}^3 |h_i|^{p_n/4}(t)dt \right)^{4/p_n}
$$

$$
\le \sum_{\substack{1 \le i,j,k \le 3 \\ i,j,k \text{ distinct}}} \left(\int_0^1 \left((\epsilon \gamma_l)^2 + |w - s_l|^2(t) \right)^{p_n/2} dt \right)^{1 - 4/p_n}
$$

$$
\times \left(\int_0^1 |w - s_l|^{p_n}(t)dt \right)^{1/p_n} \prod_{i=1}^3 \|h_i\|_{p_n}.
$$

By (A.16) and (A.17),

$$
|D^3 f(w)[h_1, h_2, h_3]|
$$

$$
\le \frac{6 p_n^2 \left(\int_0^1 |w - s_l|^{p_n}(t)dt \right)^{3/p_n} \|h_1\|_{p_n}\|h_2\|_{p_n}\|h_3\|_{p_n}}{\eta_n \left(\int_0^1 \left((\epsilon \gamma_l)^2 + |w - s_l|^2(t) \right)^{p_n/2} dt \right)^{5/p_n}}
$$

$$+ \frac{9 p_n \left(\int_0^1 |w - s_l|^{p_n}(t) dt \right)^{1/p_n} \|h_1\|_{p_n} \|h_2\|_{p_n} \|h_3\|_{p_n}}{\eta_n \left(\int_0^1 \left((\epsilon \gamma_l)^2 + |w - s_l|^2(t) \right)^{p_n/2} dt \right)^{3/p_n}}$$

$$\leq \frac{15 p_n^2 \|h_1\|_{p_n} \|h_2\|_{p_n} \|h_3\|_{p_n}}{\eta_n} \left(\int_0^1 \left((\epsilon \gamma_l)^2 + |w - s_l|^2(t) \right)^{p_n/2} dt \right)^{-2/p_n}$$

$$\leq \frac{15 p_n^2}{(\epsilon \gamma_l)^2 \eta_n} \|h_1\|_\infty \|h_2\|_\infty \|h_3\|_\infty$$

and so

$$\| D^3 f(w) \| \leq \frac{15 p_n^2}{(\epsilon \gamma_l)^2 \eta_n}. \tag{A.18}$$

Step 4: Combining the bounds

The result now follows by combining (A.6), (A.12) and (A.18). Indeed, note that, by the chain rule,

$$
\begin{aligned}
D^3 & g_{l,n}(w)[h_1, h_2, h_3] \\
&= \phi''' \left(f(w) - \frac{\gamma_l \sqrt{1 + \epsilon^2}}{\eta_n} \right) \cdot \prod_{i=1}^3 Df(w)[h_i] \\
&\quad + \phi'' \left(f(w) - \frac{\gamma_l \sqrt{1 + \epsilon^2}}{\eta_n} \right) \cdot \sum_{\substack{1 \leq i,j,k \leq 3 \\ i,j,k \text{distinct}}} D^2 f(w)[h_i, h_j] Df(w)[h_k] \\
&\quad + \phi' \left(f(w) - \frac{\gamma_l \sqrt{1 + \epsilon^2}}{\eta_n} \right) D^3 f(w)[h_1, h_2, h_3].
\end{aligned}
$$

By (A.6), (A.12) and (A.18) and the fact that ϕ', ϕ'', ϕ''' are all bounded, we get that, for all $w \in D^p$,

$$\| D^3 g_{l,n}(w) \| \leq C_3 p_n^2 \eta_n^{-3},$$

for some constant C_3. Similar bounds may be obtained for the first and second derivative of $g_{l,n}$:

$$\| Dg_{l,n}(w) \| \leq C_1 \eta_n^{-1}, \quad \| D^2 g_{l,n}(w) \| \leq C_2 p_n \eta_n^{-1},$$

for constants C_1, C_2. Since ϕ is also bounded, the product rule yields the desired bound. $\square$

Now, we prove the following result:

Lemma A.2. *For the set B fixed at the beginning of this Appendix,*

$$\limsup_{n \to \infty} \mathbb{P}[\mathbf{Y}_n \in B] \leq \mathbb{P}[\mathbf{Z} \in B] \quad \text{and} \quad \liminf_{n \to \infty} \mathbb{P}[\mathbf{Y}_n \in B] \geq \mathbb{P}[\mathbf{Z} \in B].$$

Proof.

Step 1: Proving the first inequality

Note that

$$\mathbf{Y}_n \in B_l \implies \|\mathbf{Y}_n - s_l\| < \gamma_l \implies \sup_{t \in [0,1]} \sum_{i=1}^p \left((\mathbf{Y}_n(t) - s_l(t))^{(i)} \right)^2 < \gamma_l^2$$

$$\implies \sup_{t \in [0,1]} \left[\sum_{i=1}^{p} \left((\mathbf{Y}_n(t) - s_l(t))^{(i)} \right)^2 + (\epsilon \gamma_l)^2 \right] < \gamma_l^2 (1 + \epsilon^2)$$

$$\implies \left\| \sqrt{(\epsilon \gamma_l)^2 + \sum_{i=1}^{p} \left((\mathbf{Y}_n - s_l)^{(i)} \right)^2} \right\|_{p_n} \leq \gamma_l \sqrt{1 + \epsilon^2} \implies g_{l,n}(\mathbf{Y}_n) = 1. \tag{A.19}$$

Therefore, for all l,

$$\mathbb{1}_{[\mathbf{Y}_n \in B_l]} \leq g_{l,n}(\mathbf{Y}_n). \tag{A.20}$$

Also, note that, by Minkowski's inequality and the triangle inequality for the Euclidean norm:

$$\left\| \sqrt{(\epsilon \gamma_l)^2 + \sum_{i=1}^{p} \left((\mathbf{Z} - s_l)^{(i)} \right)^2} \right\|_{p_n} \leq \left\| \sqrt{(\epsilon \gamma_l)^2 + \sum_{i=1}^{p} \left((\mathbf{Z}_n - s_l)^{(i)} \right)^2} \right\|_{p_n} + \| \mathbf{Z}_n - \mathbf{Z} \|.$$

Therefore, if $\| \mathbf{Z} - s_l \| > \gamma_l$ then as $p_n \xrightarrow{n \to \infty} \infty$:

$$\liminf_{n \to \infty} \left\| \sqrt{(\epsilon \gamma_l)^2 + \sum_{i=1}^{p} \left((\mathbf{Z}_n - s_l)^{(i)} \right)^2} \right\|_{p_n}$$

$$\geq \liminf_{n \to \infty} \left\{ \left\| \sqrt{(\epsilon \gamma_l)^2 + \sum_{i=1}^{p} \left((\mathbf{Z} - s_l)^{(i)} \right)^2} \right\|_{p_n} - \| \mathbf{Z}_n - \mathbf{Z} \| \right\}$$

$$= \sup_{t \in [0,1]} \sqrt{(\epsilon \gamma_l)^2 + \sum_{i=1}^{p} \left((\mathbf{Z}(t) - s_l(t))^{(i)} \right)^2} > \gamma_l (1 + \epsilon^2)^{1/2}.$$

This, means that, if $p_n \xrightarrow{n \to \infty} \infty$, $\| \mathbf{Z} - s_l \| > \gamma_l$ and $\eta_n \xrightarrow{n \to \infty} 0$ then $g_{l,n}(\mathbf{Z}_n) = 0$ for sufficiently large n, i.e.

$$g_{l,n}(\mathbf{Z}_n) \leq \mathbb{1}_{[\| \mathbf{Z} - s_l \| \leq \gamma_l]}, \quad \text{as long as } p_n \xrightarrow{n \to \infty} \infty, \ \eta_n \xrightarrow{n \to \infty} 0 \text{ and } n \text{ is large.} \tag{A.21}$$

By those properties, taking $p_n \to \infty$ and $\eta_n \to 0$ such that $\kappa_n \eta_n^{-3} p_n^2 \to 0$, we obtain:

$$\limsup_{n \to \infty} \mathbb{P}[\mathbf{Y}_n \in B] \overset{\text{(A.20)}}{\leq} \limsup_{n \to \infty} \mathbb{E} \left[\prod_{l=1}^{L} g_{l,n}(\mathbf{Y}_n) \right]$$

$$\overset{\text{(2.4)}}{\leq} \limsup_{n \to \infty} \left\{ \mathbb{E} \left[\prod_{l=1}^{L} g_{l,n}(\mathbf{Z}_n) \right] + C \kappa_n \left\| \prod_{l=1}^{L} g_{l,n} \right\|_{M^0} \right\}$$

$$\overset{\text{Fatou, (A.2)}}{\leq} \mathbb{E} \left[\limsup_{n \to \infty} \prod_{l=1}^{L} g_{l,n}(\mathbf{Z}_n) \right] \overset{\text{(A.21)}}{\leq} \mathbb{P}[\mathbf{Z} \in B].$$

Step 2: Proving the second inequality

We define:

$$g_{l,n}^{*}(w) = \phi \left(\frac{\left\| \sqrt{(\epsilon \gamma_l)^2 + \sum_{i=1}^{p} \left((w - s_l)^{(i)} \right)^2} \right\|_{p_n} - \gamma_l \sqrt{\epsilon^2 + (1 - \theta)^2} (\delta \wedge \frac{r_n}{2})^{1/p_n} + \eta_n}{\eta_n} \right) \tag{A.22}$$

for $\theta > 0$ fixed and $\delta > 0$ such that:

$$\forall n \in \mathbb{N}: \quad \|\mathbf{Y}_n - s_l\| \geq \gamma_l \implies \text{leb}\{t : |\mathbf{Y}_n(t) - s_l(t)| \geq \gamma_l(1 - \theta)\} \geq \left(\delta \wedge \frac{r_n}{2}\right),$$

where leb denotes the Lebesgue measure. Such a δ exists for the following reason. The collection $(s_l, 1 \leq l \leq L)$ is uniformly equicontinuous and $\mathbf{Y}_n$ are constant on intervals of length at least r_n. The $\delta > 0$ we choose is such that:

$$|t_1 - t_2| \leq \delta \quad \implies \quad |s_l(t_1) - s_l(t_2)| \leq \frac{\theta \gamma_l}{2}.$$

If $\|\mathbf{Y}_n - s_l\| \geq \gamma_l$ then $|\mathbf{Y}_n(t_0) - s_l(t_0)| > \gamma_l\left(1 - \frac{\theta}{2}\right)$ for some t_0. Then, there exists an interval I_0 with t_0 being one of its endpoints and of length $\frac{r_n}{2} \wedge \delta$, such that $\mathbf{Y}_n$ is constant on I_0 and $|s_l(t) - s_l(t_0)| \leq \frac{\theta \gamma_l}{2}$ for all $t \in I_0$. Then, for $t \in I_0$ we obtain:

$$|\mathbf{Y}_n(t) - s_l(t)| \geq |\mathbf{Y}_n(t_0) - s_l(t_0)| - |\mathbf{Y}_n(t_0) - \mathbf{Y}_n(t)| - |s_l(t) - s_l(t_0)|$$
$$\geq \left(1 - \frac{\theta}{2}\right)\gamma_l - \frac{\theta \gamma_l}{2} = \gamma_l(1 - \theta).$$

It follows that:

$$\|\mathbf{Y}_n - s_l\| \geq \gamma_l \Longrightarrow \left\|\sqrt{\sum_{i=1}^{p}\left((\mathbf{Y}_n - s_l)^{(i)}\right)^2}\right\|_{p_n} \geq \gamma_l(1 - \theta)\left(\delta \wedge \frac{r_n}{2}\right)^{1/p_n} \Longrightarrow$$

$$\left\|\sqrt{(\epsilon \gamma_l)^2 + \sum_{i=1}^{p}\left((\mathbf{Y}_n - s_l)^{(i)}\right)^2}\right\|_{p_n} \geq \gamma_l\sqrt{\epsilon^2 + (1 - \theta)^2}\left(\delta \wedge \frac{r_n}{2}\right)^{1/p_n} \Longrightarrow g_{l,n}^*(\mathbf{Y}_n) = 0.$$

$$(A.23)$$

Therefore, for all l:

$$\mathbb{1}_{[\mathbf{Y}_n \in B_l]} \geq g_l^*(\mathbf{Y}_n). \tag{A.24}$$

Also, again, it can be shown that for any finite L and $\gamma := \min_{1 \leq l \leq L} \gamma_l$:

$$\left\|\prod_{l=1}^{L} g_{l,n}^*\right\|_{M^0} \leq C p_n^2 (\epsilon \gamma)^{-2} \eta_n^{-3} \quad \text{for some constant } C \text{ independent of } p_n, \epsilon, \gamma \text{ and } \eta_n. \tag{A.25}$$

Now suppose $\eta_n \to 0$, $p_n \to \infty$ and $r_n^{1/p_n} \to 1$. Also suppose that $\|\mathbf{Z} - s_l\| < \gamma_l(1 - \theta)$ so that there exists $\alpha > 0$ such that a.s. $\|\mathbf{Z}_n - s_l\| < \gamma_l(1 - \theta) - \alpha$ for n large enough. Then, for large n:

$$\left\|\sqrt{(\epsilon \gamma_l)^2 + \sum_{i=1}^{p}\left((\mathbf{Z}_n - s_l)^{(i)}\right)^2}\right\|_{p_n} \leq \sqrt{(\epsilon \gamma_l)^2 + \|\mathbf{Z}_n - s_l\|^2} \leq \gamma_l\sqrt{\epsilon^2 + (1 - \theta - \alpha \gamma_l^{-1})^2}$$

$$< \gamma_l\sqrt{\epsilon^2 + (1 - \theta)^2}\left(\delta \wedge \frac{r_n}{2}\right)^{1/p_n} - \eta_n$$

because $\left(\delta \wedge \frac{r_n}{2}\right)^{1/p_n} \xrightarrow{n \to \infty} 1$ and $\eta_n \xrightarrow{n \to \infty} 0$. So if $\eta_n \to 0$, $p_n \to \infty$ and $r_n^{1/p_n} \to 1$ then:

$$\|\mathbf{Z} - s_l\| < \gamma_l(1 - \theta) \Longrightarrow g_{l,n}^*(\mathbf{Z}_n) = 1$$

for n large enough, i.e.:

$$\mathbb{1}_{[\|\mathbf{Z}-s_l\|<\gamma_l(1-\theta)]} \leq g^*_{l,n}(\mathbf{Z}_n). \tag{A.26}$$

Let $\eta_n \to 0$ and $p_n \to \infty$ be such that $r_n^{1/p_n} \to 1$ and $\kappa_n p_n^2 \eta_n^{-3} \to 0$. This is possible by the assumption that $\kappa_n \log^2(1/r_n) \to 0$. Indeed, having $r_n^{1/p_n} \to 1$, all we require is that $\log\left(r_n^{1/p_n}\right)\eta_n^3 \to 0$ slower than $\kappa_n \log^2(1/r_n) \to 0$, because then:

$$\kappa_n p_n^2 \eta_n^{-3} = \frac{\kappa_n \left(\log(r_n)\right)^2}{\left(\frac{1}{p_n}\log(r_n)\right)^2 \eta_n^3} = \frac{\kappa_n \left(\log(1/r_n)\right)^2}{\left(\log\left(r_n^{1/p_n}\right)\right)^2 \eta_n^3} \to 0.$$

For instance, if $r_n \to 0$ and $\kappa_n \to 0$, we require p_n and η_n to be such that $\frac{\eta_n^3}{\kappa_n} \to \infty$ and $p_n^2 \to \infty$ faster than $(\log r_n)^2$ but slower than $\frac{\eta_n^3}{\kappa_n}$.

Then:

$$\liminf_{n\to\infty} \mathbb{P}\left[\mathbf{Y}_n \in B\right] \overset{(A.24)}{\geq} \liminf_{n\to\infty} \mathbb{E}\left[\prod_{l=1}^{L} g^*_{l,n}(\mathbf{Y}_n)\right]$$

$$\overset{(2.4)}{\geq} \liminf_{n\to\infty}\left\{\mathbb{E}\left[\prod_{l=1}^{L} g^*_{l,n}(\mathbf{Z}_n)\right] - C\kappa_n \left\|\prod_{l=1}^{L} g^*_{l,n}\right\|_{M^0}\right\}$$

$$\overset{\text{Fatou, }(A.25)}{\geq} \mathbb{E}\left[\liminf_{n\to\infty}\prod_{l=1}^{L} g^*_{l,n}(\mathbf{Z}_n)\right]$$

$$\overset{(A.26)}{\geq} \mathbb{P}\left[\bigcap_{1\leq l\leq L}\left(\|\mathbf{Z}-s_l\| < \gamma_l(1-\theta)\right)\right].$$

Since the choice of $\theta \in (0,1)$ was arbitrary, we conclude that: $\liminf_{n\to\infty} \mathbb{P}\left[\mathbf{Y}_n \in B\right] \geq \mathbb{P}(\mathbf{Z} \in B)$. $\square$

Lemma A.2 now implies that, for any set B described at the beginning of this Appendix, $\mathbb{P}[\mathbf{Y}_n \in B] \xrightarrow{n\to\infty} \mathbb{P}[\mathbf{Z} \in B]$, which finishes the proof of Proposition 2.3.

References

[1] A.D. Barbour, Stein's method for diffusion approximation, Probab. Theory Related Fields 84 (1990) 297–322.

[2] A.D. Barbour, L. Holst, S. Janson, Poisson Approximation, in: Oxford Studies in Probability, Clarendon Press, 1992.

[3] A.D. Barbour, S. Janson, A functional combinatorial central limit theorem, Electron. J. Probab. 14 (81) (2009) 2352–2370.

[4] E. Besançon, L. Decreusefond, P. Moyal, Stein's method for diffusive limit of Markov processes, 2018, arXiv:1805.01691.

[5] P. Billingsley, Convergence of Probability Measures, second ed., in: Wiley Series in Probability and Statistics, Wiley-Blackwell, 1999.

[6] L.H.Y. Chen, L. Goldstein, Q.-M. Shao, Normal Approximation By Stein's Method, in: Probability and Its Applications, Springer Verlag, 2011.

[7] L.H.Y. Chen, Q.-M. Shao, Normal approximation for nonlinear statistics using a concentration inequality approach, Bernoulli 13 (2) (2007) 581–599.

[8] T.C. Christofides, Maximal probability inequalities for multidimensionally indexed semimartingales and convergence theory of u-statistics (Ph.D. thesis), Johns Hopkins University, 1987.

[9] L. Coutin, L. Decreusefond, Stein's method for Brownian approximations, Commun. Stoch. Anal. 7 (3) (2013) 349–372.

[10] A. Dembo, S. Karlin, Poisson approximations for r-scan processes, Ann. Appl. Probab. 2 (2) (1992) 329–357.

[11] A. Dembo, Y. Rinott, Some examples of normal approximations by Stein's method, in: D. Aldous, R. Pemantle (Eds.), Random Discrete Structures, Springer New York, New York, NY, 1996, pp. 25–44.

[12] Ch. Döbler, G. Peccati, Quantitative de Jong theorems in any dimension, Electron. J. Probab. 22 (2017) 35 pp.

[13] M.D. Donsker, An invariance principle for certain probability limit theorems, Mem. Amer. Math. Soc. 6 (1951).

[14] P. Eichelsbacher, M. Löwe, Large deviations in partial sums of U-processes, Theory Probab. Appl. 43 (1) (1999) 26–41.

[15] I. Fazekas, On a general approach to the strong law of large numbers, J. Math. Sci. 200 (4) (2014) 411–423.

[16] M. Fischer, G. Nappo, On the moments of the modulus of continuity of ito processes, Stoch. Anal. Appl. 28 (1) (2010) 103–122.

[17] J. Glaz, J. Naus, S. Wallenstein, Scan Statistics, in: Springer Series in Statistics, Springer-Verlag New York, 2001.

[18] P. Hall, On the invariance principle for U-statistics, Stochastic Process. Appl. 9 (2) (1979) 163–174.

[19] W. Hoeffding, A class of statistics with asymptotically normal distribution, Ann. Math. Stat. 19 (3) (1948) 293–325.

[20] W. Hoeffding, The Strong Law of Large Numbers for U-Statistics, in: Institute of Statistics Mimeo Series 302, North Carolina State University, Dept. of Statistics, 1961.

[21] S. Janson, Gaussian Hilbert Spaces, in: Cambridge Tracts in Mathematics, Cambridge University Press, 1997.

[22] P. de Jong, A central limit theorem for generalized multilinear forms, J. Multivariate Anal. 34 (2) (1990) 275–289.

[23] S. Karlin, V. Brendel, Chance and statistical significance in protein and DNA sequence analysis, Science 257 (5066) (1992) 39–49.

[24] M.J. Kasprzak, Diffusion approximations via Stein's method and time changes, 2017, arXiv:1701.07633.

[25] M.J. Kasprzak, Functional approximations with Stein's method of exchangeable pairs, 2020, arXiv:1710.09263.

[26] M.J. Kasprzak, A.B. Duncan, S.J. Vollmer, Note on A. Barbour's paper on Stein's method for diffusion approximations, Electron. Commun. Probab. 22 (23) (2017) 1–8.

[27] S. Kotz, N.L. Johnson (Eds.), U-Statistics, in: Encyclopedia of Statistical Sciences, vol. 9, John Wiley and Sons, Inc., 1988, pp. 436–444.

[28] C. Ley, G. Reinert, Y. Swan, Stein's method for comparison of univariate distributions, Probab. Surv. 14 (2017) 1–52.

[29] J. Naus, Approximations for distributions of scan statistics, J. Amer. Statist. Assoc. 77 (377) (1982) 177–183.

[30] I. Nourdin, G. Peccati, Normal Approximations with Malliavin Calculus, in: Cambridge Tracts in Mathematics, Cambridge University Press, 2012.

[31] J. Pitman, M. Yor, A guide to Brownian motion and related stochastic processes, 2018, arXiv:1802.09679.

[32] G. Reinert, A. Röllin, Multivariate normal approximation with Stein's method of exchangeable pairs under a general linearity condition, Ann. Probab. 37 (6) (2009) 2150–2173.

[33] Y. Rinott, V. Rotar, On coupling constructions and rates in the CLT for dependent summands with applications to the antivoter model and weighted U-statistics, Ann. Appl. Probab. 7 (4) (1997) 1080–1105.

[34] N. Ross, Fundamentals of Stein's method, Probab. Surv. 8 (2011) 210–293.

[35] H. Rubin, R.A. Vitale, Asymptotic distribution of symmetric statistics, Ann. Statist. 8 (1) (1980) 165–170.

[36] R.J. Serfling, Approximation Theorems of Mathematical Statistics, in: Wiley Series in Probability and Statistics, John Wiley and Sons, Inc., 1980.

[37] H.-H. Shih, On Stein's method for infinite-dimensional Gaussian approximation in abstract Wiener spaces, J. Funct. Anal. 261 (5) (2011) 1236–1283.

[38] Ch. Stein, A bound for the error in the normal approximation to the distribution of a sum of dependent random variables, in: Proc. Sixth Berkeley Symp. on Math. Statist. and Prob., vol. 2, 1972, pp. 583–602.

[39] Y. Swan, A gateway to Stein's Method, 2016. https://sites.google.com/site/steinsmethod/home. (Accessed on 19 May 2016).

[40] A.R. Vitale, An Expansion for Symmetric Statistics and the Efron-Stein Inequality, in: Lecture Notes–Monograph Series, vol. 5, Institute of Mathematical Statistics, 1984, pp. 112–114.

[41] A. Wald, J. Wolfowitz, On a test whether two samples are from the same population, Ann. Math. Stat. 11 (2) (1940) 147–162.