

Stein’s method of exchangeable pairs in multivariate functional approximations

Christian Döbler¹ and Mikołaj J. Kasprzak²

¹ *University of Osnabrück
Institut für Mathematik
Albrechtstraße 28a
D-49076 Osnabrück
Germany*

e-mail: christian.doebler@uni-osnabrueck.de

² *University of Luxembourg
Department of Mathematics
Maison du Nombre
6 Avenue de la Fonte
L-4364 Esch-sur-Alzette
Luxembourg*

e-mail: mikolaj.kasprzak@uni.lu

Abstract: In this paper we develop a framework for multivariate functional approximation by a suitable Gaussian process via an exchangeable pairs coupling that satisfies a suitable approximate linear regression property, thereby building on work by Barbour (1990) and Kasprzak (2020). We demonstrate the applicability of our results by applying it to joint subgraph counts in an Erdős-Renyi random graph model on the one hand and to vectors of weighted, degenerate U -processes on the other hand. As a concrete instance of the latter class of examples, we provide a bound for the functional approximation of a vector of success runs of different lengths by a suitable Gaussian process which, even in the situation of just a single run, would be outside the scope of the existing theory.

MSC 2010 subject classifications: Primary 60B10, 60F17; secondary 60B12, 60J65, 60E05, 60E15.

Keywords and phrases: Stein’s method, functional convergence, exchangeable pairs, multivariate processes, U -statistics.

1. Introduction

In his seminal paper [Ste72], Charles Stein introduced a method for proving normal approximations and obtained a bound on the speed of convergence to the standard normal distribution. Later, Barbour [Bar88] and Götze [Göt91] developed the so-called generator approach to finding Stein’s equation, which made it possible to study approximations by many other probability laws. As a result, in [Bar90], the method was adapted to approximations by the (infinite-dimensional) Wiener measure.

Moreover, the exchangeable-pair approach, first developed by Stein in his monograph [Ste86] in the context of univariate normal approximations, has been at the heart of many results proved using Stein’s method. It was extended by [RR97a] and used in the context of non-normal approximations in [CDM05, Röl07, CFR11, Döb15, DP18]. The publication of [CM08, RR09, Mec09] brought a breakthrough in the understanding of the exchangeable-pair approach and made it available for applications to a wide array of multivariate normal approximation problems. In [Kas20a] the method was applied to the study of functional limit results and approximations by univariate Gaussian processes, using the setup of [Ste86, RR97a] and [Bar90].

In this paper we combine the functional approximation of [Bar90] and the multivariate exchangeable-pair method of [RR09, Mec09]. We obtain an abstract approximation theorem, which is applied in the context of weighted degenerate U -statistics, a particularly interesting example of which are homogeneous sums. The strength of the abstract approximation result is also presented in a random-graph-theoretic application.

1.1. Motivation

We are motivated by examples of multivariate quantities whose distance from the normal distribution can be established using Stein's method of exchangeable pairs, and whose functional equivalents have not been studied yet. Functional limit results play an important role in applied fields. Scaling limits of discrete processes can be studied using stochastic analysis and are often more robust to changes in the local details than the discrete processes themselves. That is why researchers often choose to describe discrete phenomena with continuous models. The error they make by doing this is measured by rates of convergence in functional limit results. The current paper contributes to solving the problem of bounding those rates.

The two main applications motivating the paper and considered therein are a continuous Gaussian-process approximation of a rescaled weighted U -statistic and the study of an Erdős-Renyi random graph process.

The first application deals with the approximation of so-called weighted U -processes, i.e. process analogues of the class of weighted U -statistics. This class of processes is very wide, containing the so-called homogeneous sum processes as well as symmetric, degenerate (complete or incomplete) U -processes. We derive a general result and successfully apply it to the case of homogeneous sum processes in Subsection 5.5. As a concrete example, in Subsection 5.6, we provide a bound for a Gaussian approximation of a process that is defined as a vector of success runs of different lengths. For functional limit theorems involving the class of symmetric, degenerate U -processes, we refer the reader to the recent paper [DKP19]. Moreover, we remark that, even in the univariate case of weighted U -statistics, the literature about limit theory for these random quantities is quite restricted. Indeed, apart from the abundance of references on limit theorems for homogeneous sums, the majority of articles focus on the limiting behavior of so-called reduced or incomplete U -statistics, i.e. weighted U -statistics whose weights only assume the values 0 and 1 (see e.g. [Blo76, Jan84, BK78]). Limit theorems for general weighted U -statistics can be found in references [RR97b, OR93, Maj94]. We stress, however, that the last two references focus on non-normal limiting distributions and that, in the degenerate case, [RR97b] only considers kernels of order 2.

The second example comes originally from [JN91] and was studied using exchangeable pairs in a finite-dimensional context in [RR10]. We look at a (dynamic) Erdős-Renyi random graph with $\lfloor nt \rfloor$ vertices, where t denotes the time, and study the distance from the asymptotic distribution of the joint law of the number of edges and the number of two-stars. Our approach can, however, be also extended to cover the number of triangles. Those statistics are often used when approximating the clustering coefficient of a network and applied in conditionally uniform graph tests.

1.2. Contribution of the paper

The main achievements of the paper are the following:

- (a) An abstract approximation theorem (Theorem 4.1), bounding the distance between a stochastic process \mathbf{Y}_n valued in \mathbb{R}^d , for a fixed positive integer d , and a Gaussian mixture process. The estimate is derived under the assumption that the process \mathbf{Y}_n satisfies the linear regression condition

$$Df(\mathbf{Y}_n)[\mathbf{Y}_n] = 2\mathbb{E} \left\{ Df(\mathbf{Y}_n) [(\mathbf{Y}_n - \mathbf{Y}'_n)\Lambda_n] \mid \mathbf{Y}_n \right\} + R_f,$$

for all functions f in a certain class of test functions, some $\Lambda_n \in \mathbb{R}^{d \times d}$ and some random variable $R_f = R_f(\mathbf{Y}_n)$. The class of test functions, with respect to which the bound is obtained, is so rich that the bound approaching zero fast enough implies weak convergence of the law of \mathbf{Y}_n in the Skorokhod and uniform topologies on the Skorokhod space. The exact conditions under which this happens are stated in Proposition 2.2.

- (b) A novel framework for continuous Gaussian process approximations of vectors of weighted, degenerate U -processes, presented in Section 5. Apart from proving a general result about those, we show how it may be applied in examples involving non-degenerate U -processes. In order to study such examples using our theory, one may decompose the given U -process into the vector of its degenerate Hoeffding components and prove a multivariate Gaussian limit theorem for this vector. Then, by applying a linear functional, one obtains a Gaussian limit for the original process. This strategy, in

a quantified fashion, is exemplified by the application to the r -runs process, discussed in Subsection 5.6. We stress that, even in the case of just one r -run process, the results about univariate functional approximations via exchangeable pairs from [Kas20a] would not be sufficient to obtain a Gaussian approximation. Thus, in this example, the multidimensionality of our approach proves to be absolutely vital. Moreover, both the kernels and the coefficients of the weighted U -processes we study in our general result may (and will in most cases) depend on the sample size n , hence yielding Gaussian limits even in degenerate situations. At the same time, our methods are flexible enough in order to yield bounds for the classical results on asymptotic Gaussianity, in non-degenerate situations, when the kernels are fixed.

- (c) A novel quantitative functional limit theorem for the edge counts and the number of two-stars in an Erdős-Renyi random graph $G(n, p)$ on n vertices with fixed edge probability p . Letting $I_{i,j}$, for $i, j = 1, \dots, n$ be the indicator that edge (i, j) is present in the graph, we consider the following statistics:

$$\mathbf{T}_n(t) = \frac{\lfloor nt \rfloor - 2}{2n^2} \sum_{i,j=1}^{\lfloor nt \rfloor} I_{i,j}, \quad \mathbf{V}_n(t) = \frac{1}{6n^2} \sum_{\substack{1 \leq i,j,k \leq \lfloor nt \rfloor \\ i,j,k \text{ distinct}}} I_{ij} I_{jk}, \quad t \in [0, 1],$$

corresponding to the number of edges and the number of two-stars, respectively. Theorem 6.2 provides a bound on the distance between the law of the process

$$t \mapsto (\mathbf{T}_n(t) - \mathbb{E}\mathbf{T}_n(t), \mathbf{V}_n(t) - \mathbb{E}\mathbf{V}_n(t)) \quad t \in [0, 1] \quad (1.1)$$

and the law of a piecewise constant Gaussian process. Theorem 6.4 estimates the distance between the law of (1.1) and that of a continuous Gaussian process. These results extend the result of [Kas20a] bounding the distance between the distribution of the edge counts and a univariate Gaussian process. As a corollary to our results, we immediately obtain weak convergence of the law of (1.1) in the Skorokhod and uniform topologies on the Skorokhod space to that of the continuous Gaussian process.

1.3. Stein's method in its generality

Stein's method in its generality is a powerful technique used to obtain bounds on quantities of the form $|\mathbb{E}_\nu h - \mathbb{E}_\mu h|$, where μ is the target (known) distribution, ν is an approximating measure and h is a real-valued test function chosen from a suitable class \mathcal{H} . The method is composed out of three main steps. First, one needs to find an operator \mathcal{A} acting on a class of real-valued functions, such that

$$(\forall f \in \text{Domain}(\mathcal{A}) \quad \mathbb{E}_\pi \mathcal{A}f = 0) \iff \pi = \mu.$$

Second, for a given function $h \in \mathcal{H}$, one solves the following Stein equation:

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

Finally, for $f = f_h$ solving the Stein equation, the following quantity:

$$|\mathbb{E}_\nu \mathcal{A}f_h| \quad (1.2)$$

needs to be bounded. This is achieved using various mathematical tools (Taylor's expansions, Malliavin calculus, as described in [NP12], coupling methods and others), applied in conjunction with smoothness properties of f_h . For an accessible account of the method we recommend the surveys [LRS17] and [Ros11] as well as the books [BHJ92] and [CGS11], which treat the cases of Poisson and normal approximation, respectively, in detail. A database of information and publications connected to Stein's method can also be found in [Swa16].

1.4. Stein's method of exchangeable pairs

The exchangeable-pair approach to Stein's method was first developed in [Ste86]. Therein, the author considered the setup in which, for a random variable W , one can construct another random variable W' such that (W, W') is an exchangeable pair and the following linear regression condition is satisfied

$$\mathbb{E}[W' - W|W] = -\lambda W \quad (1.3)$$

for some $\lambda > 0$. It follows from this assumption that

$$0 = \mathbb{E}[(f(W) + f(W'))(W - W')] = \mathbb{E}[(f(W') - f(W))(W - W')] + 2\lambda \mathbb{E}[Wf(W)]$$

and so

$$\mathbb{E}[Wf(W)] = \frac{1}{2\lambda} \mathbb{E}[(f(W) - f(W'))(W - W')].$$

Therefore, using Taylor's theorem, it can be proved that

$$|\mathbb{E}[f'(W)] - \mathbb{E}[Wf(W)]| \leq \frac{\|f'\|_\infty}{2\lambda} \sqrt{\text{Var}[\mathbb{E}[(W - W')^2|W]]} + \frac{\|f''\|_\infty}{2\lambda} \mathbb{E}|W - W'|^3,$$

which provides a bound on the quantity (1.2) for $\nu = \mathcal{L}(W)$ and \mathcal{A} being the canonical Stein operator corresponding to the standard normal law.

A multivariate version of the method was first described in [CM08] and then in [RR09]. In [RR09], for an exchangeable pair of d -dimensional vectors (W, W') , the following condition is used:

$$\mathbb{E}[W' - W|W] = -\Lambda W + R \quad (1.4)$$

for some invertible matrix Λ and a remainder term R . The approach of [RR09] was further reinterpreted and combined with the approach of [CM08] in [Mec09]. Extending this multivariate version of the exchangeable-pair method to multivariate functional approximations, with the linear regression condition taking form similar to (1.4), is the subject of the current paper.

1.5. Functional Stein's method

Approximations by laws of stochastic processes using Stein's method have been studied in [Bar90, BJ09, Shi11, CD13, CD14] and recently in [Kas17, Kas20b, Kas20a, BDM18, CD19, BC19]. These references can be divided into three groups.

The ones belonging to the first group, containing [Bar90, BJ09, Kas17, Kas20b, Kas20a], all use, adapt and extend the setup of [Bar90]. Therein, the author studied the rate of convergence in the celebrated functional central limit theorem, also called Donsker's theorem. 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. For each function g belonging to this class he provided a bound on the absolute difference between the expectation of g with respect to the law of a rescaled random walk and the expectation of g with respect to the Wiener measure. Crucially, he also proved that this class of functions g is so rich that his bounds imply weak convergence with respect to the Skorokhod topology of the considered rescaled random walk to Brownian Motion. This last property is vital for most applications of the limit theory for stochastic processes and may even be the main reason for the outstanding popularity of the Skorokhod topology. Indeed, by means of the continuous mapping theorem, limit theorems for many natural, non-linear functionals such as the supremum over time, immediately follow from a weak limit theorem in the Skorokhod topology.

On the other hand, the results of the second group of references, containing [CD13, CD14, BDM18, CD19], develop Stein's theory on a Hilbert space using a Besov-type topology. The bounds obtained therein, however, do not imply weak convergence in the Skorokhod topology. Therefore, the continuous mapping theorem does not apply in their setting. For instance, as opposed to the results of the first group of references, one cannot study convergence of the supremum of a process using the analysis of the second group of papers.

Finally, [Shi11] develops approximations by abstract Wiener measures on a real separable Banach space and [BC19] proves bounds on measure-determining distances from Gaussian random variables valued in Hilbert spaces. As for the second group, despite the elegant abstract theory used and developed in these references, the results do not imply convergence in the Skorohod topology on $D[0, 1]$.

In the current paper we shall follow the setup of the first group of references. We consider it more flexible than the one of the second group and more suited for applications to processes belonging to the widely-used (non-separable) Skorohod space than the ones of the third group.

In the context of these three groups of references and the present paper, we also mention the recent paper [DKP19] which, although not relying on functional approximation by Stein's method, provides functional limit theorems for the class of (degenerate and non-degenerate) symmetric U -processes with a kernel that may depend on the sample size n . Since it implicitly relies on a multivariate Gaussian limit theorem derived by Stein's method from [DP19], it is also naturally related to Stein's method.

Moreover, since one main class of applications in the present paper involves weighted U -processes, it is worthwhile to compare our results and their applicability to those of [DKP19]. Firstly, as mentioned above, the paper [DKP19] focuses on Gaussian limit theorems for symmetric U -processes, which constitute a narrower class than the weighted U -processes considered in the present work. Moreover, thanks to the finite-dimensional convergence results from [DP19], the conditions for convergence from [DKP19] are phrased in term of L^2 -norms of *contraction kernels* and, as such, can be considered as fourth moment conditions. Contrarily, as can be seen from the bounds and proofs of Section 5, the bounds and conditions in the present paper involve third moment quantities. This distinction is also clearly reflected in the respective applicability of the results proved in the present paper and those from [DKP19]. Indeed, whereas the symmetric U -processes considered in [DKP19] possess a global dependency structure, the results in Section 5 are most useful whenever the dependence of the weighted U -process is local in the sense that the involved array of weighting coefficients $(a_J)_J$ is sparse in some sense. The runs example in Subsection 5.6 provides an instructive showcase for this observation. Moreover, the methods used in the proofs of the main results necessitate that the quantities in the bounds involve the absolute values of both the kernels and the coefficients. Hence, no cancellation effect, typically occurring under fourth moment conditions, may be relied on in this case. We therefore consider our theorems as rather complementary to the ones in [DKP19].

1.6. Structure of the paper

Section 2 includes some introductory remarks about notation and the spaces of test functions with respect to which bounds on distances between probability laws in this paper will be derived. Section 3 gives a general form of the pre-limiting process to which all the processes of interest will be compared using Stein's method. It also presents the corresponding Stein equation, its solution and the smoothness properties of the solution. Section 4 contains the main abstract result of this paper providing a bound on the distance between a process valued in the Skorohod space $D([0, 1], \mathbb{R}^d)$ and the pre-limiting process described in the previous section. Section 5 discusses the application of the abstract theorem to weighted, degenerate U -statistics and presents a bound on their distance from a continuous Gaussian process. It furthermore explains how the bound simplifies in the context of homogeneous sums and applies it to the example of r -runs on the line. Section 6 discusses the example concerning an Erdős-Renyi random graph process and the bound on the distance between the number of its edges and two-stars and a continuous Gaussian process. Technical details of some of the proofs in this paper are postponed to Section 7.

2. Notation and spaces M and M^0

The following notation, similar to the one of [Bar90] and [Kas20b], is used throughout the paper. For a fixed positive integer d , let $D^d = D([0, 1], \mathbb{R}^d)$ be the Skorohod space of càdlàg \mathbb{R}^d -valued functions on $[0, 1]$. For $i = 1, \dots, d$, by e_i we denote the i th unit vector of the canonical basis of \mathbb{R}^d . The i th component of any $x \in \mathbb{R}^d$ will be denoted by $x^{(i)}$, so that $x = (x^{(1)}, \dots, x^{(d)})$. For a function w defined on $[0, 1]$ and taking values in a Euclidean space, we will also write

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

where $|\cdot|$ denotes the Euclidean norm. Moreover, the notation $\mathbb{E}^W[\cdot]$ will be used to represent $\mathbb{E}[\cdot|W]$.

By $D^k f$ we will always mean the k -th Fréchet derivative of f . The norm $\|\cdot\|$ of a k -linear form B on L will be taken to be

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

where $B[h_1, \dots, h_k]$ denotes B applied to arguments h_1, \dots, h_k .

Furthermore, we define

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

and let L be the Banach space of continuous functions $f : D^d \rightarrow \mathbb{R}$ such that $\|f\|_L < \infty$. As in [Bar90], we define $M \subset L$ as a subspace of L consisting of the twice Fréchet differentiable functions f , such that:

$$\|D^2 f(w + h) - D^2 f(w)\| \leq k_f \|h\|, \quad (2.1)$$

for some constant k_f , uniformly in $w, h \in D^d$. We have following lemma (whose proof we omit), which may be proved in an analogous way to that used to show (2.6) and (2.7) of [Bar90]:

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

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

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

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

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

and note that $M^0 \subset M$. Below, we present a d -dimensional version of [BJ09, Proposition 3.1] providing conditions, under which weak convergence of the approximating measure to the target one may be deduced from convergence of the corresponding expectations of functions $g \in M^0$. Its proof can be found in the appendix of [Kas20b].

Proposition 2.2. *Suppose that, for each $n \geq 1$, the random element \mathbf{Y}_n of D^d is piecewise constant with intervals of constancy of length at least r_n . Let $(\mathbf{Z}_n)_{n \geq 1}$ be random elements of D^p converging in distribution in D^d , with respect to the Skorokhod topology, to a random element $\mathbf{Z} \in C([0, 1], \mathbb{R}^d)$. If:*

$$|\mathbb{E}g(\mathbf{Y}_n) - \mathbb{E}g(\mathbf{Z}_n)| \leq C \mathcal{T}_n \|g\|_{M^0} \quad (2.2)$$

for each $g \in M^0$ and if $\mathcal{T}_n \log^2(1/r_n) \xrightarrow{n \rightarrow \infty} 0$, then the law of \mathbf{Y}_n converges weakly to that of \mathbf{Z} in D^d , in both the uniform and the Skorokhod topologies.

3. Setting up Stein's method for the pre-limiting approximation

We set up Stein's method in a fashion similar to [Bar90] and [Kas20b]. First, we define the process \mathbf{D}_n whose distribution will be treated as the target measure. We then construct a process $(\mathbf{W}_n(\cdot, u) : u \geq 0)$ for which the target measure is stationary. We subsequently calculate its infinitesimal generator \mathcal{A}_n and take it as our Stein operator. Next, we solve the Stein equation $\mathcal{A}_n f = g$, using the analysis of [KDV17], and prove several smoothness properties of the solution $f_n = \phi_n(g)$.

3.1. Target measure

Let

$$\mathbf{D}_n(t) = \sum_{i_1, \dots, i_m=1}^n \left(\tilde{Z}_{i_1, \dots, i_m}^{(1)} J_{i_1, \dots, i_m}^{(1)}(t), \dots, \tilde{Z}_{i_1, \dots, i_m}^{(d)} J_{i_1, \dots, i_m}^{(d)}(t) \right), \quad t \in [0, 1], \quad (3.1)$$

where $\tilde{Z}_{i_1, \dots, i_m}^{(k)}$'s for $k = 1, \dots, d$ are centred Gaussian and:

- A) the covariance matrix $\Sigma_n \in \mathbb{R}^{(n^m d) \times (n^m d)}$ of \tilde{Z} is positive definite, for $\tilde{Z} \in \mathbb{R}^{(n^m d)}$ built out of the $\tilde{Z}_{i_1, \dots, i_m}^{(k)}$'s in such a way that they appear in the lexicographic order with $\tilde{Z}_{i_1, \dots, i_m}^{(k)}$ appearing before $\tilde{Z}_{j_1, \dots, j_m}^{(k+1)}$'s for any $k = 1, \dots, d-1$ and $i_1, \dots, i_m, j_1, \dots, j_m = 1, \dots, n$;
- B) functions $J_{i_1, \dots, i_m}^{(k)} \in D([0, 1], \mathbb{R})$, for $i_1, \dots, i_m \in \{1, \dots, n\}$, $k \in \{1, \dots, p\}$, are independent of the $\tilde{Z}_{i_1, \dots, i_m}^{(k)}$'s; a natural example of those would be $J_{i_1, \dots, i_m}^{(k)} = \mathbb{1}_{A_{i_1, \dots, i_m}^{(k)}}$ for some measurable set $A_{i_1, \dots, i_m}^{(k)}$.

Remark 3.1. It is worth noting that processes \mathbf{D}_n of the form (3.1) are often approximations of interesting continuous Gaussian processes. An example is \mathbf{D}_n of (3.1), where all the $\tilde{Z}_{i_1, \dots, i_m}^{(k)}$'s are standard normal and independent, $m = 1$ and $J_i^{(k)} = \mathbb{1}_{[i/n, 1]}$ for all $k = 1, \dots, d$ and $i = 1, \dots, n$. By Donsker's theorem, it approximates the standard Brownian motion. By Proposition 2.2, under several assumptions, if a piecewise constant process \mathbf{Y}_n is close enough to process \mathbf{D}_n , then the law of \mathbf{Y}_n converges weakly to that of the continuous process that \mathbf{D}_n approximates.

Now consider an array of i.i.d. Ornstein-Uhlenbeck processes with stationary law $\mathcal{N}(0, 1)$, independent of the $J_{i_1, \dots, i_m}^{(k)}$'s, given by $\{(\mathcal{X}_{i_1, \dots, i_m}^{(k)}(u), u \geq 0) : i_1, \dots, i_m = 1, \dots, n, k = 1, \dots, d\}$. Let $\tilde{\mathcal{W}}(u) = (\Sigma_n)^{1/2} \mathcal{X}(u)$, where $\mathcal{X}(u) \in \mathbb{R}^{n^m d}$ is a vector composed out of the $\mathcal{X}_{i_1, \dots, i_m}^{(k)}(u)$'s in such a way that they are ordered exactly as $\tilde{Z}_{i_1, \dots, i_m}^{(k)}$'s are ordered in \tilde{Z} . Write $\mathcal{W}_{i_1, \dots, i_m}^{(k)}(u) = \left(\tilde{\mathcal{W}}(u)\right)_{I(k, i_1, \dots, i_m)}$ using the bijection $I : \{(k, i_1, \dots, i_m) : i_1, \dots, i_m = 1, \dots, n, k = 1, \dots, d\} \rightarrow \{1, \dots, dn^m\}$, given by:

$$I(k, i_1, \dots, i_m) = (k-1)n^m + (i_1-1)n^{m-1} + \dots + (i_{m-1}-1)n + i_m. \quad (3.2)$$

We will look at the process

$$\mathbf{W}_n(t, u) = \left(\mathbf{W}_n^{(1)}(t, u), \dots, \mathbf{W}_n^{(d)}(t, u)\right), \quad t \in [0, 1], u \geq 0,$$

where, for all $k = 1, \dots, d$:

$$\mathbf{W}_n^{(k)}(t, u) = \sum_{i_1, \dots, i_m=1}^n \mathcal{W}_{i_1, \dots, i_m}^{(k)}(u) J_{i_1, \dots, i_m}^{(k)}(t), \quad t \in [0, 1], u \geq 0.$$

It is easy to see that the stationary law of the process $(\mathbf{W}_n(\cdot, u))_{u \geq 0}$ is exactly the law of \mathbf{D}_n .

3.2. Stein equation

The following result follows immediately from [Kas20b, Propositions 4.1 and 4.4]:

Proposition 3.2. The infinitesimal generator of the process $(\mathbf{W}_n(\cdot, u))_{u \geq 0}$ acts on any $f \in M$ (for M defined in Section 2) in the following way:

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

Moreover, 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:

$$f_n = \phi_n(g) = - \int_0^\infty T_{n,u} g du, \quad (3.3)$$

where $(T_{n,u} f)(w) = \mathbb{E} [f(w e^{-u} + \sqrt{1 - e^{-2u}} \mathbf{D}_n(\cdot))]$. Furthermore, for $g \in M$:

$$\begin{aligned} A) \quad & \|D\phi_n(g)(w)\| \leq \|g\|_M \left(1 + \frac{2}{3}\|w\|^2 + \frac{4}{3}\mathbb{E}\|\mathbf{D}_n\|^2\right), \\ B) \quad & \|D^2\phi_n(g)(w)\| \leq \|g\|_M \left(\frac{1}{2} + \frac{\|w\|}{3} + \frac{\mathbb{E}\|\mathbf{D}_n\|}{3}\right), \\ C) \quad & \frac{\|D^2\phi_n(g)(w+h) - D^2\phi_n(g)(w)\|}{\|h\|} \\ & \leq \sup_{w, h \in D^p} \frac{\|D^2(g+c)(w+h) - D^2(g+c)(w)\|}{3\|h\|}, \end{aligned} \quad (3.4)$$

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

4. An abstract approximation theorem

The following result provides an expression for a bound on the distance between a process \mathbf{Y}_n and \mathbf{D}_n , defined by (3.1). It assumes that we can find some \mathbf{Y}'_n such that $(\mathbf{Y}_n, \mathbf{Y}'_n)$ is an exchangeable pair satisfying an appropriate condition. We explain in Remark 4.5 how our condition is similar to that of [RR09, (1.7)].

Theorem 4.1. *Assume that $(\mathbf{Y}_n, \mathbf{Y}'_n)$ is an exchangeable pair of $D([0, 1], \mathbb{R}^d)$ -valued random vectors such that:*

$$Df(\mathbf{Y}_n)[\mathbf{Y}_n] = 2\mathbb{E}^{\mathbf{Y}_n} Df(\mathbf{Y}_n) [(\mathbf{Y}_n - \mathbf{Y}'_n)\Lambda_n] + R_f, \quad (4.1)$$

where $\mathbb{E}^{\mathbf{Y}_n}[\cdot] := \mathbb{E}[\cdot | \mathbf{Y}_n]$, for all $f \in M$, some $\Lambda_n \in \mathbb{R}^{d \times d}$ and some random variable $R_f = R_f(\mathbf{Y}_n)$. Let \mathbf{D}_n be defined by (3.1). Then, for any $g \in M$:

$$|\mathbb{E}g(\mathbf{Y}_n) - \mathbb{E}g(\mathbf{D}_n)| \leq \epsilon_1 + \epsilon_2 + \epsilon_3,$$

where

$$\begin{aligned} \epsilon_1 &= \frac{\|g\|_M}{6} \mathbb{E} [\|(\mathbf{Y}_n - \mathbf{Y}'_n)\Lambda_n\| \|\mathbf{Y}_n - \mathbf{Y}'_n\|^2], \\ \epsilon_2 &= |\mathbb{E} D^2 f(\mathbf{Y}_n) [(\mathbf{Y}_n - \mathbf{Y}'_n)\Lambda_n, \mathbf{Y}_n - \mathbf{Y}'_n] - \mathbb{E} D^2 f(\mathbf{Y}_n) [\mathbf{D}_n, \mathbf{D}_n]|, \\ \epsilon_3 &= |\mathbb{E} R_f|, \end{aligned}$$

and $f = \phi_n(g)$, as defined by (3.3).

Remark 4.2 (Relevance of terms in the bound). *Term ϵ_1 measures how close \mathbf{Y}_n and \mathbf{Y}'_n are and how small (in a certain sense) Λ_n is. Term ϵ_2 quantifies the difference between the covariance structures of $\mathbf{Y}_n - \mathbf{Y}'_n$ and \mathbf{D}_n . This term may be estimated in several applications (see Theorems 5.1 and 6.2 below), yet this often requires some effort. Term ϵ_3 measures the error in the exchangeable-pair linear regression condition (4.1).*

Remark 4.3. *Condition (4.1) is always satisfied, for example with $\Lambda_n = 0$ and $R_f = Df(\mathbf{Y}_n)[\mathbf{Y}_n]$ for all $f \in M$. However, for the bound in Theorem 4.1 to be small, we require the expectation of R_f to be small in absolute value.*

Remark 4.4. *The term*

$$|\mathbb{E} D^2 f(\mathbf{Y}_n) [(\mathbf{Y}_n - \mathbf{Y}'_n)\Lambda_n, \mathbf{Y}_n - \mathbf{Y}'_n] - \mathbb{E} D^2 f(\mathbf{Y}_n) [\mathbf{D}_n, \mathbf{D}_n]|$$

in the bound obtained in Theorem 4.1 is an analogue of the second condition in [Mec09, Theorem 3]. The main result of that paper provides a bound on approximation by $\mathcal{N}(0, \Sigma)$ of a d -dimensional vector X . This is achieved by constructing an exchangeable pair (X, X') satisfying:

$$\mathbb{E}^X[X' - X] = \Lambda X + E \quad \text{and} \quad \mathbb{E}^X[(X' - X)(X' - X)^T] = 2\Lambda\Sigma + E'$$

for some invertible matrix Λ and some remainder terms E and E' . In the same spirit, Theorem 4.1 could be rewritten to assume (4.1) and:

$$\mathbb{E}^{\mathbf{Y}_n} D^2 f(\mathbf{Y}_n) [(\mathbf{Y}_n - \mathbf{Y}'_n)\Lambda_n, \mathbf{Y}_n - \mathbf{Y}'_n] = D^2 f(\mathbf{Y}_n) [\mathbf{D}_n, \mathbf{D}_n] + R_f^1.$$

The bound would then take the form:

$$|\mathbb{E}g(\mathbf{Y}_n) - \mathbb{E}g(\mathbf{D}_n)| \leq \frac{\|g\|_M}{6} \mathbb{E} [\|(\mathbf{Y}_n - \mathbf{Y}'_n)\Lambda_n\| \|\mathbf{Y}_n - \mathbf{Y}'_n\|^2] + |\mathbb{E} R_f| + |\mathbb{E} R_f^1|.$$

Remark 4.5. *The role of Λ_n in condition (4.1) is equivalent to that played by Λ^{-1} in [RR09] for Λ defined by (1.7) therein. In the functional setting, condition (4.1) is more appropriate than a straightforward adaptation of the setup of [RR09]. This is because, for general processes \mathbf{Y}_n , the properties of the*

Fréchet derivative do not allow us to treat evaluating the derivative in the direction of $\mathbf{Y}_n - \mathbf{Y}'_n$ as matrix multiplication. Indeed, multiplying both sides of the hypothetical condition:

$$-Df(\mathbf{Y}_n)[\Lambda \mathbf{Y}_n] = \mathbb{E}^{\mathbf{Y}_n} Df(\mathbf{Y}_n)[\mathbf{Y}_n - \mathbf{Y}'_n]$$

by Λ^{-1} does not yield:

$$-Df(\mathbf{Y}_n)[\mathbf{Y}_n] = \mathbb{E}^{\mathbf{Y}_n} Df(\mathbf{Y}_n)[\Lambda^{-1}(\mathbf{Y}_n - \mathbf{Y}'_n)].$$

Proof of Theorem 4.1. We will bound $|\mathbb{E}g(\mathbf{Y}_n) - \mathbb{E}g(\mathbf{D}_n)|$ by bounding $|\mathbb{E}\mathcal{A}_n f(\mathbf{Y}_n)|$, where f is the solution to the Stein equation:

$$\mathcal{A}_n f = g - \mathbb{E}g(\mathbf{D}_n),$$

for \mathcal{A}_n defined in Proposition 3.2. Note that, by exchangeability of $(\mathbf{Y}_n, \mathbf{Y}'_n)$ and (4.1):

$$\begin{aligned} 0 &= \mathbb{E}(Df(\mathbf{Y}'_n) + Df(\mathbf{Y}_n))[(\mathbf{Y}_n - \mathbf{Y}'_n)\Lambda_n] \\ &= \mathbb{E}(Df(\mathbf{Y}'_n) - Df(\mathbf{Y}_n))[(\mathbf{Y}_n - \mathbf{Y}'_n)\Lambda_n] + 2\mathbb{E}\{\mathbb{E}^{\mathbf{Y}_n} Df(\mathbf{Y}_n)[(\mathbf{Y}_n - \mathbf{Y}'_n)\Lambda_n]\} \\ &= \mathbb{E}(Df(\mathbf{Y}'_n) - Df(\mathbf{Y}_n))[(\mathbf{Y}_n - \mathbf{Y}'_n)\Lambda_n] + \mathbb{E}Df(\mathbf{Y}_n)[\mathbf{Y}_n] - \mathbb{E}R_f \end{aligned}$$

and so:

$$\mathbb{E}Df(\mathbf{Y}_n)[\mathbf{Y}_n] = \mathbb{E}(Df(\mathbf{Y}_n) - Df(\mathbf{Y}'_n))[(\mathbf{Y}_n - \mathbf{Y}'_n)\Lambda_n] + \mathbb{E}R_f.$$

Therefore:

$$\begin{aligned} |\mathbb{E}\mathcal{A}_n f(\mathbf{Y}_n)| &= |\mathbb{E}Df(\mathbf{Y}_n)[\mathbf{Y}_n] - \mathbb{E}D^2 f(\mathbf{Y}_n)[\mathbf{D}_n, \mathbf{D}_n]| \\ &= |\mathbb{E}(Df(\mathbf{Y}_n) - Df(\mathbf{Y}'_n))[(\mathbf{Y}_n - \mathbf{Y}'_n)\Lambda_n] - \mathbb{E}D^2 f(\mathbf{Y}_n)[\mathbf{D}_n, \mathbf{D}_n] + \mathbb{E}R_f| \\ &\leq |\mathbb{E}(Df(\mathbf{Y}_n) - Df(\mathbf{Y}'_n))[(\mathbf{Y}_n - \mathbf{Y}'_n)\Lambda_n] - \mathbb{E}D^2 f(\mathbf{Y}'_n)[(\mathbf{Y}_n - \mathbf{Y}'_n)\Lambda_n, \mathbf{Y}_n - \mathbf{Y}'_n]| \\ &\quad + |\mathbb{E}D^2 f(\mathbf{Y}_n)[(\mathbf{Y}_n - \mathbf{Y}'_n)\Lambda_n, \mathbf{Y}_n - \mathbf{Y}'_n] - \mathbb{E}D^2 f(\mathbf{Y}_n)[\mathbf{D}_n, \mathbf{D}_n]| + |\mathbb{E}R_f| \\ &\leq \frac{\|g\|_M}{6} \mathbb{E}[\|(\mathbf{Y}_n - \mathbf{Y}'_n)\Lambda_n\| \|\mathbf{Y}_n - \mathbf{Y}'_n\|^2] + |\mathbb{E}R_f| \\ &\quad + |\mathbb{E}D^2 f(\mathbf{Y}_n)[(\mathbf{Y}_n - \mathbf{Y}'_n)\Lambda_n, \mathbf{Y}_n - \mathbf{Y}'_n] - \mathbb{E}D^2 f(\mathbf{Y}_n)[\mathbf{D}_n, \mathbf{D}_n]|, \end{aligned}$$

where the last inequality follows by Taylor's theorem and Proposition 3.2. \square

5. Weighted, degenerate U -statistics

In this Section we will apply Theorem 4.1 in order to prove bounds for the approximation of a vector of weighted, degenerate U -processes by suitable Gaussian processes.

5.1. Introduction

The setup will be the following. We fix positive integers d, p_1, \dots, p_d and consider a sequence $(X_i)_{i \in \mathbb{N}}$ of i.i.d. random variables with distribution μ on some measurable space (E, \mathcal{E}) . Moreover, for $1 \leq i \leq d$, we let $\psi(i) \in L^2(\mu^{p_i})$ be a symmetric kernel such that $\mathbb{E}[\psi(i)^2(X_1, \dots, X_{p_i})] > 0$. We assume that $\psi(i)$ is **(completely) degenerate** with respect to μ , i.e. that

$$\mathbb{E}[\psi(i)(X_1, \dots, X_{p_i}) \mid X_1, \dots, X_{p_i-1}] = 0, \quad \text{a.s.}$$

We denote by $\mathcal{D}_p(n)$ the collection of p -subsets of the set $[n] := \{1, \dots, n\}$.

Furthermore, we fix an integer $n \geq \max(p_1, \dots, p_d)$ and let $\{a_J(i) : 1 \leq i \leq d, J \in \mathcal{D}_{p_i}(n)\}$, be a (given) set of real numbers (weights). We further let $\{\sigma_n(i) : 1 \leq i \leq d\}$ be a set of positive real numbers and, for $t \in [0, 1]$, define

$$\mathbf{Y}_n^{(i)}(t) := \frac{1}{\sigma_n(i)} \sum_{J \in \mathcal{D}_{p_i}([nt])} a_J(i) \psi(i)(X_j, j \in J).$$

In some applications it may be natural to take

$$\sigma_n(i)^2 = \sum_{J \in \mathcal{D}_{p_i}(n)} a_J(i)^2 \mathbb{E}[\psi(i)^2(X_1, \dots, X_{p_i})], \quad 1 \leq i \leq d,$$

i.e. equal to the variance of $\mathbf{Y}_n^{(i)}(1)$. This is, however, not necessary for our results.

With the above definitions, we let

$$\mathbf{Y}_n := (\mathbf{Y}_n^{(1)}, \dots, \mathbf{Y}_n^{(d)}),$$

which is, as one can easily observe, an element of $D([0, 1]; \mathbb{R}^d)$. We will write $X := (X_1, \dots, X_n)$ and construct an $X' := (X'_1, \dots, X'_n)$ such that the pair (X, X') is exchangeable. Specifically, we let X_0 be another random variable with distribution μ and let I be uniformly distributed on $[n]$ in such a way that $I, X_0, (X_j)_{j \in \mathbb{N}}$ are jointly independent. For $1 \leq j \leq n$, we let

$$X'_j := \begin{cases} X_j, & \text{if } j \neq I \\ X_0, & \text{if } j = I. \end{cases}$$

Then, for $t \in [0, 1]$ and $1 \leq i \leq d$, we define

$$(\mathbf{Y}_n^{(i)})'(t) := \frac{1}{\sigma_n(i)} \sum_{J \in \mathcal{D}_{p_i}(\lfloor nt \rfloor)} a_J(i) \psi(i)(X'_j, j \in J)$$

and

$$\mathbf{Y}'_n := ((\mathbf{Y}_n^{(1)})', \dots, (\mathbf{Y}_n^{(d)})').$$

The pair $(\mathbf{Y}_n, \mathbf{Y}'_n)$ is clearly exchangeable and, for $f \in M$, similarly as in the proof of [DP17, Lemma 2.3], one can use degeneracy to show that

$$Df(\mathbf{Y}_n)[\mathbf{Y}_n] = 2\mathbb{E}^{\mathbf{Y}_n} Df(\mathbf{Y}_n)[(\mathbf{Y}_n - \mathbf{Y}'_n)\Lambda_n],$$

where

$$\Lambda_n = \text{diag}\left(\frac{n}{2p_1}, \dots, \frac{n}{2p_d}\right). \quad (5.1)$$

Therefore condition (4.1) is satisfied for Λ_n of (5.1) and $R_f = 0$. In what follows we will assume that $1 \leq p_1 \leq p_2 \leq \dots \leq p_d$.

5.2. A pre-limiting process

We will construct a pre-limiting Gaussian process \mathbf{D}_n of the form (3.1) which has the same covariance structure as \mathbf{Y}_n . We take $\mathbf{D}_n = (\mathbf{D}_n^{(1)}, \dots, \mathbf{D}_n^{(d)})$ for

$$\mathbf{D}_n^{(i)}(t) = \frac{1}{\sigma_n(i)} \sum_{J \in \mathcal{D}_{p_i}(\lfloor nt \rfloor)} a_J(i) Z_J(i),$$

where, for $i = 1, \dots, d$ and $J \in \mathcal{D}_{p_i}(n)$, $Z_J(i)$ are jointly Gaussian random variables that are independent of X and satisfy

$$\mathbb{E}[Z_J(i)Z_K(l)] = \begin{cases} \mathbb{E}[\psi(i)(X_1, \dots, X_{p_i})\psi(l)(X_1, \dots, X_{p_l})], & \text{if } p_i = p_l \text{ and } K = J \\ 0, & \text{otherwise,} \end{cases}$$

for $i, l = 1, \dots, d$, $J \in \mathcal{D}_{p_i}(n)$ and $K \in \mathcal{D}_{p_l}(n)$.

5.3. Distance from the pre-limiting process

Having established the setup and defined the pre-limiting process above, we prove the following result:

Theorem 5.1. *Let \mathbf{Y}_n be defined as in Section 5 and \mathbf{D}_n be defined as in Section 5.2. Then, for any $g \in M$,*

$$\begin{aligned} \left| \mathbb{E}[g(\mathbf{Y}_n)] - \mathbb{E}[g(\mathbf{D}_n)] \right| &\leq \frac{2\sqrt{d}\|g\|_M}{3p_1} \sum_{i=1}^d \frac{\|\psi(i)\|_{L^3(\mu^{p_i})}^3}{\sigma_n(i)^3} \sum_{l=1}^n \left(\sum_{\substack{J \in \mathcal{D}_{p_i}(n): \\ l \in J}} |a_J(i)| \right)^3 \\ &+ \|g\|_M \sum_{i,j,k=1}^d \frac{\|\psi(i)\|_{L^3(\mu^{p_i})} \|\psi(j)\|_{L^3(\mu^{p_j})} \|\psi(k)\|_{L^3(\mu^{p_k})}}{\sigma_n(i)\sigma_n(j)\sigma_n(k)} \sum_{\substack{J \in \mathcal{D}_{p_i}(n), \\ K \in \mathcal{D}_{p_j}(n), \\ L \in \mathcal{D}_{p_k}(n): \\ J \cap K \neq \emptyset, \\ L \cap (J \cup K) \neq \emptyset}} |a_J(i)a_K(j)a_L(k)|. \end{aligned}$$

Proof.

Step 1. First note that, for ϵ_1 in Theorem 4.1,

$$\|(\mathbf{Y}_n - \mathbf{Y}'_n)\Lambda_n\| \|\mathbf{Y}_n - \mathbf{Y}'_n\|^2 \leq \frac{n}{2p_1} \|\mathbf{Y}_n - \mathbf{Y}'_n\|^3, \quad (5.2)$$

which follows directly from the definition of Λ_n in (5.1) and our assumption that $p_1 \leq \dots \leq p_d$. Now, note that

$$\begin{aligned} \|\mathbf{Y}_n - \mathbf{Y}'_n\|^3 &= \sup_{t \in [0,1]} \left[\left(\mathbf{Y}_n^{(1)}(t) - \left(\mathbf{Y}_n^{(1)} \right)'(t) \right)^2 + \dots + \left(\mathbf{Y}_n^{(d)}(t) - \left(\mathbf{Y}_n^{(d)} \right)'(t) \right)^2 \right]^{3/2} \\ &\leq \sqrt{d} \sup_{t \in [0,1]} \left[\left| \mathbf{Y}_n^{(1)}(t) - \left(\mathbf{Y}_n^{(1)} \right)'(t) \right|^3 + \dots + \left| \mathbf{Y}_n^{(d)}(t) - \left(\mathbf{Y}_n^{(d)} \right)'(t) \right|^3 \right] \\ &\leq \sqrt{d} \left[\left\| \mathbf{Y}_n^{(1)} - \left(\mathbf{Y}_n^{(1)} \right)' \right\|^3 + \dots + \left\| \mathbf{Y}_n^{(d)} - \left(\mathbf{Y}_n^{(d)} \right)' \right\|^3 \right]. \end{aligned} \quad (5.3)$$

Furthermore, for $i = 1, \dots, d$,

$$\begin{aligned} &\mathbb{E} \left\| \mathbf{Y}_n^{(i)} - \left(\mathbf{Y}_n^{(i)} \right)' \right\|^3 \\ &= \frac{1}{\sigma_n(i)^3} \mathbb{E} \left\{ \sup_{t \in [0,1]} \left| \sum_{\substack{J \in \mathcal{D}_{p_i}(\lfloor nt \rfloor): \\ I \in J}} a_J(i) (\psi(i)(X_j, j \in J) - \psi(i)(X_0, X_j, j \in J \setminus \{I\})) \mathbb{1}_{[\frac{\max(J)}{n}, 1]}(t) \right|^3 \right\} \\ &\leq \frac{1}{\sigma_n(i)^3} \mathbb{E} \left(\sum_{\substack{J \in \mathcal{D}_{p_i}(n): \\ I \in J}} |a_J(i)| |\psi(i)(X_j, j \in J) - \psi(i)(X_0, X_j, j \in J \setminus \{I\})| \right)^3 \\ &\leq \frac{1}{n\sigma_n(i)^3} \sum_{l=1}^n \sum_{\substack{J, K, L \in \mathcal{D}_{p_i}(n): \\ l \in J \cap K \cap L}} |a_J(i)a_K(i)a_L(i)| \mathbb{E} \left[|\psi(i)(X_j, j \in J) - \psi(i)(X_0, X_j, j \in J \setminus \{l\})| \right. \\ &\quad \left. |\psi(i)(X_j, j \in K) - \psi(i)(X_0, X_j, j \in K \setminus \{l\})| |\psi(i)(X_j, j \in L) - \psi(i)(X_0, X_j, j \in L \setminus \{l\})| \right] \end{aligned}$$

$$\leq \frac{\mathbb{E}|\psi(i)(X_1, \dots, X_{p_i}) - \psi(i)(X_2, \dots, X_{p_{i+1}})|^3}{n\sigma_n(i)^3} \sum_{l=1}^n \sum_{\substack{J, K, L \in \mathcal{D}_{p_i}(n): \\ l \in J \cap K \cap L}} |a_J(i)a_K(i)a_L(i)| \quad (5.4)$$

$$\leq \frac{8\mathbb{E}|\psi(i)(X_1, \dots, X_{p_i})|^3}{n\sigma_n(i)^3} \sum_{l=1}^n \left(\sum_{\substack{J \in \mathcal{D}_{p_i}(n): \\ l \in J}} |a_J(i)| \right)^3. \quad (5.5)$$

Combining (5.2) -(5.5) we obtain

$$\epsilon_1 \leq \frac{\sqrt{d}\|g\|_M}{12p_1} \sum_{i=1}^d \frac{\mathbb{E}|\psi(i)(X_1, \dots, X_{p_i}) - \psi(i)(X_2, \dots, X_{p_{i+1}})|^3}{\sigma_n(i)^3} \sum_{l=1}^n \sum_{\substack{J, K, L \in \mathcal{D}_{p_i}(n): \\ l \in J \cap K \cap L}} |a_J(i)a_K(i)a_L(i)| \quad (5.6)$$

$$\leq \frac{2\sqrt{d}\|g\|_M}{3p_1} \sum_{i=1}^d \frac{\mathbb{E}|\psi(i)(X_1, \dots, X_{p_i})|^3}{\sigma_n(i)^3} \sum_{l=1}^n \left(\sum_{\substack{J \in \mathcal{D}_{p_i}(n): \\ l \in J}} |a_J(i)| \right)^3. \quad (5.7)$$

Step 2. We will now bound ϵ_2 of Theorem 4.1. Denoting by e_i the i th element of the canonical basis of \mathbb{R}^d , for $i = 1, \dots, d$, for any $f \in M$, we have

$$\begin{aligned} & D^2 f(\mathbf{Y}_n) [(\mathbf{Y}_n - \mathbf{Y}'_n) \Lambda_n, \mathbf{Y}_n - \mathbf{Y}'_n] \\ &= D^2 f(\mathbf{Y}_n) \left[\sum_{i=1}^d \frac{n}{2p_i} \left(\mathbf{Y}_n^{(i)} - (\mathbf{Y}_n^{(i)})' \right) e_i, \sum_{i=1}^d \left(\mathbf{Y}_n^{(i)} - (\mathbf{Y}_n^{(i)})' \right) e_i \right] \\ &= \sum_{i,j=1}^d \frac{n}{2p_i} D^2 f(\mathbf{Y}_n) \left[\left(\mathbf{Y}_n^{(i)} - (\mathbf{Y}_n^{(i)})' \right) e_i, \left(\mathbf{Y}_n^{(j)} - (\mathbf{Y}_n^{(j)})' \right) e_j \right] \end{aligned} \quad (5.8)$$

We now let $f = \phi_n(g)$, as defined by (3.3), and fix some $i, j \in \{1, \dots, d\}$. We have that

$$\begin{aligned} & \left| \frac{n}{2p_i} \mathbb{E} D^2 f(\mathbf{Y}_n) \left[\left(\mathbf{Y}_n^{(i)} - (\mathbf{Y}_n^{(i)})' \right) e_i, \left(\mathbf{Y}_n^{(j)} - (\mathbf{Y}_n^{(j)})' \right) e_j \right] - \mathbb{E} D^2 f(\mathbf{Y}_n) [\mathbf{D}_n^{(i)} e_i, \mathbf{D}_n^{(j)} e_j] \right| \\ &= \frac{1}{\sigma_n(i)\sigma_n(j)} \left| \frac{n}{2p_i} \sum_{\substack{J \in \mathcal{D}_{p_i}(n), \\ K \in \mathcal{D}_{p_j}(n)}} a_J(i)a_K(j) \mathbb{E} \left[(\psi(i)(X_u, u \in J) - \psi(i)(X'_u, u \in J)) \right. \right. \\ & \quad \cdot (\psi(j)(X_u, u \in K) - \psi(j)(X'_u, u \in K)) D^2 f(\mathbf{Y}_n) [\mathbb{1}_{[\frac{\max(J)}{n}, 1]} e_i, \mathbb{1}_{[\frac{\max(K)}{n}, 1]} e_j] \Big] \\ & \quad \left. - \sum_{\substack{J \in \mathcal{D}_{p_i}(n), \\ K \in \mathcal{D}_{p_j}(n)}} a_J(i)a_K(j) \mathbb{1}_{\{J=K\}} \mathbb{E} [\psi(i)(X_1, \dots, X_{p_i}) \psi(j)(X_1, \dots, X_{p_j})] \mathbb{E} [D^2 f(\mathbf{Y}_n) [\mathbb{1}_{[\frac{\max(J)}{n}, 1]} e_i, \mathbb{1}_{[\frac{\max(K)}{n}, 1]} e_j]] \right| \\ &= \frac{1}{2p_i \sigma_n(i) \sigma_n(j)} \left| \sum_{l=1}^n \sum_{\substack{J \in \mathcal{D}_{p_i}(n), \\ K \in \mathcal{D}_{p_j}(n), \\ l \in J \cap K}} a_J(i)a_K(j) \mathbb{E} \left[(\psi(i)(X_u, u \in J) - \psi(i)(X_0, X_u, u \in J \setminus \{l\})) \right. \right. \\ & \quad \cdot (\psi(j)(X_u, u \in K) - \psi(j)(X_0, X_u, u \in K \setminus \{l\})) D^2 f(\mathbf{Y}_n) [\mathbb{1}_{[\frac{\max(J)}{n}, 1]} e_i, \mathbb{1}_{[\frac{\max(K)}{n}, 1]} e_j] \Big] \\ & \quad \left. - 2 \sum_{l=1}^n \sum_{\substack{J \in \mathcal{D}_{p_i}(n), \\ K \in \mathcal{D}_{p_j}(n), \\ l \in J \cap K}} a_J(i)a_K(j) \mathbb{1}_{\{J=K\}} \mathbb{E} [\psi(i)(X_1, \dots, X_{p_i}) \psi(j)(X_1, \dots, X_{p_j})] \mathbb{E} [D^2 f(\mathbf{Y}_n) [\mathbb{1}_{[\frac{\max(J)}{n}, 1]} e_i, \mathbb{1}_{[\frac{\max(K)}{n}, 1]} e_j]] \right| \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{2p_i\sigma_n(i)\sigma_n(j)} \left| \sum_{l=1}^n \sum_{\substack{J \in \mathcal{D}_{p_i}(n), \\ K \in \mathcal{D}_{p_j}(n), \\ l \in J \cap K}} a_J(i)a_K(j) \mathbb{E} \left[\left((\psi(i)(X_u, u \in J) - \psi(i)(X_0, X_u, u \in J \setminus \{l\})) \right. \right. \right. \\
&\quad \cdot (\psi(j)(X_u, u \in K) - \psi(j)(X_0, X_u, u \in K \setminus \{l\})) - 2\mathbb{1}_{\{J=K\}} \mathbb{E}[\psi(i)(X_1, \dots, X_{p_i})\psi(j)(X_1, \dots, X_{p_j})] \Big) \\
&\quad \cdot D^2 f(\mathbf{Y}_n) [\mathbb{1}_{[\frac{\max(J)}{n}, 1]} e_i, \mathbb{1}_{[\frac{\max(K)}{n}, 1]} e_j] \Big] \Big| \quad (5.9)
\end{aligned}$$

Now, we define

$$\mathbf{Y}_n^{J,K} := \left((\mathbf{Y}_n^{J,K})^{(1)}, \dots, (\mathbf{Y}_n^{J,K})^{(d)} \right)$$

via

$$(\mathbf{Y}_n^{J,K})^{(i)}(t) := \frac{1}{\sigma_n(i)} \sum_{\substack{L \in \mathcal{D}_{p_i}(\lfloor nt \rfloor): \\ L \cap (J \cup K) = \emptyset}} a_J(i) \psi(i)(X_j, j \in L), \quad 1 \leq i \leq d, t \in [0, 1].$$

Then, using independence, we obtain that

$$\begin{aligned}
&\left| \frac{n}{2p_i} \mathbb{E} D^2 f(\mathbf{Y}_n) \left[\left(\mathbf{Y}_n^{(i)} - (\mathbf{Y}_n^{(i)})' \right) e_i, \left(\mathbf{Y}_n^{(j)} - (\mathbf{Y}_n^{(j)})' \right) e_j \right] - \mathbb{E} D^2 f(\mathbf{Y}_n) [\mathbf{D}_n^{(i)} e_i, \mathbf{D}_n^{(j)} e_j] \right| \\
&= \frac{1}{2p_i\sigma_n(i)\sigma_n(j)} \left| \sum_{l=1}^n \sum_{\substack{J \in \mathcal{D}_{p_i}(n), \\ K \in \mathcal{D}_{p_j}(n), \\ l \in J \cap K}} a_J(i)a_K(j) \mathbb{E} \left[\left((\psi(i)(X_u, u \in J) - \psi(i)(X_0, X_u, u \in J \setminus \{l\})) \right. \right. \right. \\
&\quad \cdot (\psi(j)(X_u, u \in K) - \psi(j)(X_0, X_u, u \in K \setminus \{l\})) - 2\mathbb{1}_{\{J=K\}} \mathbb{E}[\psi(i)(X_1, \dots, X_{p_i})\psi(j)(X_1, \dots, X_{p_j})] \Big) \\
&\quad \cdot \left(D^2 f(\mathbf{Y}_n) - D^2 f(\mathbf{Y}_n^{J,K}) \right) [\mathbb{1}_{[\frac{\max(J)}{n}, 1]} e_i, \mathbb{1}_{[\frac{\max(K)}{n}, 1]} e_j] \Big] \Big| \\
&\stackrel{(3.4)^C}{\leq} \frac{\|g\|_M}{6p_i\sigma_n(i)\sigma_n(j)} \sum_{l=1}^n \sum_{\substack{J \in \mathcal{D}_{p_i}(n), \\ K \in \mathcal{D}_{p_j}(n), \\ l \in J \cap K}} |a_J(i)a_K(j)| \mathbb{E} \left[\left| (\psi(i)(X_u, u \in J) - \psi(i)(X_0, X_u, u \in J \setminus \{l\})) \right. \right. \\
&\quad \cdot (\psi(j)(X_u, u \in K) - \psi(j)(X_0, X_u, u \in K \setminus \{l\})) - 2\mathbb{1}_{\{J=K\}} \mathbb{E}[\psi(i)(X_1, \dots, X_{p_i})\psi(j)(X_1, \dots, X_{p_j})] \Big| \|\mathbf{Y}_n - \mathbf{Y}_n^{J,K}\| \Big]. \quad (5.10)
\end{aligned}$$

Now, we observe that

$$\begin{aligned}
\|\mathbf{Y}_n - \mathbf{Y}_n^{J,K}\| &\leq \sum_{k=1}^d \frac{1}{\sigma_n(k)} \|\mathbf{Y}_n^{(k)} - (\mathbf{Y}_n^{J,K})^{(k)}\| \\
&\leq \sum_{k=1}^d \frac{1}{\sigma_n(k)} \sum_{\substack{L \in \mathcal{D}_{p_k}(n): \\ L \cap (J \cup K) \neq \emptyset}} |a_J(k)| |\psi(k)(X_u, u \in L)|.
\end{aligned}$$

Hence,

$$\begin{aligned}
&\left| \frac{n}{2p_i} \mathbb{E} D^2 f(\mathbf{Y}_n) \left[\left(\mathbf{Y}_n^{(i)} - (\mathbf{Y}_n^{(i)})' \right) e_i, \left(\mathbf{Y}_n^{(j)} - (\mathbf{Y}_n^{(j)})' \right) e_j \right] - \mathbb{E} D^2 f(\mathbf{Y}_n) [\mathbf{D}_n^{(i)} e_i, \mathbf{D}_n^{(j)} e_j] \right| \\
&\leq \sum_{k=1}^d \frac{\|g\|_M}{6p_i\sigma_n(i)\sigma_n(j)\sigma_n(k)} \sum_{l=1}^n \sum_{\substack{J \in \mathcal{D}_{p_i}(n), \\ K \in \mathcal{D}_{p_j}(n), \\ l \in J \cap K}} \sum_{\substack{L \in \mathcal{D}_{p_k}(n): \\ L \cap (J \cup K) \neq \emptyset}} |a_J(i)a_K(j)a_L(k)| \mathbb{E} \left[\left| (\psi(i)(X_u, u \in J) - \psi(i)(X_0, X_u, u \in J \setminus \{l\})) \right. \right.
\end{aligned}$$

$$\begin{aligned}
& \cdot \left(\psi(j)(X_u, u \in K) - \psi(j)(X_0, X_u, u \in K \setminus \{l\}) \right) - 2\mathbb{1}_{\{J=K\}} \mathbb{E}[\psi(i)(X_1, \dots, X_{p_i}) \psi(j)(X_1, \dots, X_{p_j})] \left| \psi(k)(X_u, u \in L) \right| \\
& \leq \sum_{k=1}^d \frac{\|g\|_M}{p_i \sigma_n(i) \sigma_n(j) \sigma_n(k)} \sum_{l=1}^n \sum_{\substack{J \in \mathcal{D}_{p_i}(n), \\ K \in \mathcal{D}_{p_j}(n), \\ l \in J \cap K}} \sum_{L \in \mathcal{D}_{p_k}(n): \\ L \cap (J \cup K) \neq \emptyset} |a_J(i) a_K(j) a_L(k)| \|\psi(i)\|_{L^3(\mu^{p_i})} \|\psi(j)\|_{L^3(\mu^{p_j})} \|\psi(k)\|_{L^3(\mu^{p_k})} \\
& \leq \sum_{k=1}^d \frac{\|g\|_M \|\psi(i)\|_{L^3(\mu^{p_i})} \|\psi(j)\|_{L^3(\mu^{p_j})} \|\psi(k)\|_{L^3(\mu^{p_k})}}{\sigma_n(i) \sigma_n(j) \sigma_n(k)} \sum_{\substack{J \in \mathcal{D}_{p_i}(n), \\ K \in \mathcal{D}_{p_j}(n), \\ L \in \mathcal{D}_{p_k}(n): \\ J \cap K \neq \emptyset, \\ L \cap (J \cup K) \neq \emptyset}} |a_J(i) a_K(j) a_L(k)| \quad (5.11)
\end{aligned}$$

Hence, we obtain that

$$\begin{aligned}
\epsilon_2 & \leq \sum_{i,j=1}^d \left| \frac{n}{2p_i} \mathbb{E} D^2 f(\mathbf{Y}_n) \left[\left(\mathbf{Y}_n^{(i)} - \left(\mathbf{Y}_n^{(i)} \right)' \right) e_i, \left(\mathbf{Y}_n^{(j)} - \left(\mathbf{Y}_n^{(j)} \right)' \right) e_j \right] - \mathbb{E} D^2 f(\mathbf{Y}_n) \left[\mathbf{D}_n^{(i)} e_i, \mathbf{D}_n^{(j)} e_j \right] \right| \\
& \leq \sum_{i,j,k=1}^d \frac{\|g\|_M \|\psi(i)\|_{L^3(\mu^{p_i})} \|\psi(j)\|_{L^3(\mu^{p_j})} \|\psi(k)\|_{L^3(\mu^{p_k})}}{\sigma_n(i) \sigma_n(j) \sigma_n(k)} \sum_{\substack{J \in \mathcal{D}_{p_i}(n), \\ K \in \mathcal{D}_{p_j}(n), \\ L \in \mathcal{D}_{p_k}(n): \\ J \cap K \neq \emptyset, \\ L \cap (J \cup K) \neq \emptyset}} |a_J(i) a_K(j) a_L(k)|. \quad (5.12)
\end{aligned}$$

□

5.4. Distance from a continuous process

We now prove a following theorem, which bounds the distance between the law of \mathbf{Y}_n and that of a continuous Gaussian process. Let us introduce some notation first.

Theorem 5.2. *Let $\Sigma_n^{(m)} \in \mathbb{R}^{d \times d}$ be given by*

$$\left(\Sigma_n^{(m)} \right)_{i,l} = \begin{cases} \frac{n}{\sigma_n(i) \sigma_n(l)} \sum_{\substack{J \in \mathcal{D}_{p_i}(m): \\ m = \max(J)}} a_J(i) a_J(l) \mathbb{E} [\psi(i)(X_1, \dots, X_{p_i}) \psi(l)(X_1, \dots, X_{p_l})], & \text{if } p_i = p_l \\ 0, & \text{otherwise,} \end{cases}$$

for $i, l = 1, \dots, d$. For $i = 1, \dots, d$, let

$$\delta_n^{(i)} = \frac{1}{(\sigma_n(i))^2} \sup_{m \in [n]} \sum_{\substack{J \in \mathcal{D}_{p_i}(m): \\ m = \max(J)}} a_J(i)^2 \mathbb{E} [\psi(i)^2(X_1, \dots, X_{p_i})]$$

and

$$T_n^{(i)} = \frac{1}{(\sigma_n(i))^2} \sum_{J \in \mathcal{D}_{p_i}(n)} a_J(i)^2 \mathbb{E} [\psi(i)^2(X_1, \dots, X_{p_i})].$$

Furthermore, let

$$\varphi_n(s) = \sum_{m=p_1}^n \left(\Sigma_n^{(m)} \right)^{1/2} \mathbb{1}_{\left(\frac{m-1}{n}, \frac{m}{n} \right]}(s), \quad s \in [0, 1]$$

and suppose that $\varphi : [0, 1] \rightarrow \mathbb{R}^{d \times d}$ is such that, for all $i, j = 1, \dots, d$,

$$\lim_{n \rightarrow \infty} \int_0^1 \left| (\varphi_n(s) - \varphi(s))_{i,j} \right|^2 ds = 0.$$

Let $\|\cdot\|_F$ denote the Frobenius norm. Suppose that \mathbf{W} is a d -dimensional standard Brownian motion. Let

$$\mathbf{Z}(t) = \int_0^t \varphi(s) d\mathbf{W}(s)$$

and \mathbf{Y}_n be defined as in Section 5.1. Then, for any $g \in M$,

$$|\mathbb{E}g(\mathbf{Y}_n) - \mathbb{E}g(\mathbf{Z})| \leq \|g\|_M(\gamma_1 + \gamma_2 + \gamma_3 + \gamma_4 + \gamma_5),$$

and, for any $g \in M^0$,

$$|\mathbb{E}g(\mathbf{Y}_n) - \mathbb{E}g(\mathbf{Z})| \leq \|g\|_{M^0}(\gamma_1 + \gamma_2 + \gamma_3),$$

where

$$\begin{aligned} \gamma_1 &= \frac{2\sqrt{d}}{3p_1} \sum_{i=1}^d \frac{\|\psi(i)\|_{L^3(\mu^{p_i})}^3}{\sigma_n(i)^3} \sum_{l=1}^n \left(\sum_{\substack{J \in \mathcal{D}_{p_i}(n): \\ l \in J}} |a_J(i)| \right)^3; \\ \gamma_2 &= \sum_{i,j,k=1}^d \frac{\|\psi(i)\|_{L^3(\mu^{p_i})} \|\psi(j)\|_{L^3(\mu^{p_j})} \|\psi(k)\|_{L^3(\mu^{p_k})}}{\sigma_n(i)\sigma_n(j)\sigma_n(k)} \sum_{\substack{J \in \mathcal{D}_{p_i}(n), \\ K \in \mathcal{D}_{p_j}(n), \\ L \in \mathcal{D}_{p_k}(n): \\ J \cap K \neq \emptyset, \\ L \cap (J \cup K) \neq \emptyset}} |a_J(i)a_K(j)a_L(k)|; \\ \gamma_3 &= 2\sqrt{\int_0^1 \|\varphi_n(s) - \varphi(s)\|_F^2 ds} + 12\sqrt{\sum_{i=1}^d \delta_n^{(i)} \log \left(\frac{2T_n^{(i)}}{\delta_n^{(i)}} \right)}; \\ \gamma_4 &= \sqrt{d} \sum_{i=1}^d \left[8447 \left(\delta_n^{(i)} \log \left(\frac{2T_n^{(i)}}{\delta_n^{(i)}} \right) \right)^{3/2} + 44 \left(\sum_{j=1}^d \int_0^1 [(\varphi_n(s) - \varphi(s))_{i,j}]^2 ds \right)^{3/2} \right]; \\ \gamma_5 &= \sqrt{d} \left(\int_0^1 \|\varphi(s)\|_F^2 ds \right) \sum_{i=1}^d \left[50\sqrt{\delta_n^{(i)} \log \left(\frac{2T_n^{(i)}}{\delta_n^{(i)}} \right)} + 19\sqrt{\sum_{j=1}^d \int_0^1 [(\varphi_n(s) - \varphi(s))_{i,j}]^2 ds} \right]. \end{aligned}$$

Proof. Let us write $\mathbf{W} = (\mathbf{W}^{(1)}, \dots, \mathbf{W}^{(d)})$, where $\mathbf{W}^{(1)}, \dots, \mathbf{W}^{(d)}$ are i.i.d. standard Brownian motions in \mathbb{R} .

Step 1. Consider process \mathbf{D}_n defined in Section 5.2. Note that, for $i = 1, \dots, d$,

$$\begin{aligned} \mathbf{D}_n^{(i)}(t) &= \frac{1}{\sigma_n^{(i)}} \sum_{J \in \mathcal{D}_{p_i}(\lfloor nt \rfloor)} a_J(i) Z_J(i) \\ &= \frac{1}{\sigma_n^{(i)}} \sum_{m=p_i}^{\lfloor nt \rfloor} \sum_{\substack{J \in \mathcal{D}_{p_i}(\lfloor m \rfloor): \\ m = \max(J)}} a_J(i) Z_J(i) \\ &= \frac{1}{\sigma_n^{(i)}} \sum_{m=p_i}^{\lfloor nt \rfloor} \tilde{Z}_m(i), \end{aligned}$$

where $\{\tilde{Z}_m(i) : m \in [n], i \in [d]\}$ is a jointly Gaussian collection of centred random variables with the following covariance structure:

$$\mathbb{E} [\tilde{Z}_{m_1}(i) \tilde{Z}_{m_2}(l)] = \begin{cases} \sum_{\substack{J \in \mathcal{D}_{p_i}(\lfloor m_1 \rfloor): \\ m_1 = \max(J)}} a_J^{(i)} a_J^{(l)} \mathbb{E} [\psi(i)(X_1, \dots, X_{p_i}) \psi(l)(X_1, \dots, X_{p_l})], & \text{if } p_i = p_l \text{ and } m_1 = m_2 \\ 0, & \text{otherwise.} \end{cases}$$

Using this observation, note that \mathbf{D}_n has the same distribution as $\tilde{\mathbf{Z}}_n$ given by

$$\tilde{\mathbf{Z}}_n(t) := \frac{1}{\sqrt{n}} \sum_{m=p_1}^n \int_0^{\lfloor nt \rfloor} \left(\Sigma_n^{(m)} \right)^{1/2} \mathbb{1}_{(m-1, m]}(s) d\mathbf{W}(s), \quad t \in [0, 1],$$

whose distribution, by a simple change of variables, is equal to that of

$$\mathbf{Z}_n(t) := \sum_{m=p_1}^n \int_0^{\lfloor nt \rfloor / n} \left(\Sigma_n^{(m)} \right)^{1/2} \mathbb{1}_{(\frac{m-1}{n}, \frac{m}{n}]}(s) d\mathbf{W}(s) = \int_0^{\lfloor nt \rfloor / n} \varphi_n(s) d\mathbf{W}(s), \quad t \in [0, 1].$$

Step 2. By Doob's L^2 inequality and Itô's isometry, we note that

$$\begin{aligned} \mathbb{E} \sup_{t \in [0, 1]} \left| \int_0^t (\varphi_n(s) - \varphi(s)) d\mathbf{W}(s) \right|^2 &= \mathbb{E} \left[\sup_{t \in [0, 1]} \sum_{i=1}^d \left(\sum_{j=1}^d \int_0^t (\varphi_n(s) - \varphi(s))_{i,j} d\mathbf{W}^{(j)}(s) \right)^2 \right] \\ &\leq 4 \sum_{i=1}^d \mathbb{E} \left[\left(\sum_{j=1}^d \int_0^1 (\varphi_n(s) - \varphi(s))_{i,j} d\mathbf{W}^{(j)}(s) \right)^2 \right] \\ &= 4 \sum_{i,j=1}^d \mathbb{E} \left[\left(\int_0^1 (\varphi_n(s) - \varphi(s))_{i,j} d\mathbf{W}^{(j)}(s) \right)^2 \right] \\ &= 4 \int_0^1 \|\varphi_n(s) - \varphi(s)\|_F^2 ds \end{aligned} \quad (5.13)$$

Similarly, by Doob's L^3 inequality, the formula for Gaussian moments and Itô's isometry,

$$\begin{aligned} \mathbb{E} \sup_{t \in [0, 1]} \left| \int_0^t (\varphi_n(s) - \varphi(s)) d\mathbf{W}(s) \right|^3 &= \mathbb{E} \left[\sup_{t \in [0, 1]} \left(\sum_{i=1}^d \left(\sum_{j=1}^d \int_0^t (\varphi_n(s) - \varphi(s))_{i,j} d\mathbf{W}^{(j)}(s) \right)^2 \right)^{3/2} \right] \\ &\leq \frac{27\sqrt{d}}{8} \sum_{i=1}^d \mathbb{E} \left[\left| \sum_{j=1}^d \int_0^1 (\varphi_n(s) - \varphi(s))_{i,j} d\mathbf{W}^{(j)}(s) \right|^3 \right] \\ &= \frac{27\sqrt{d}}{2\sqrt{2\pi}} \sum_{i=1}^d \left(\mathbb{E} \left[\left(\sum_{j=1}^d \int_0^1 (\varphi_n(s) - \varphi(s))_{i,j} d\mathbf{W}^{(j)}(s) \right)^2 \right] \right)^{3/2} \\ &= \frac{27\sqrt{d}}{2\sqrt{2\pi}} \sum_{i=1}^d \left(\sum_{j=1}^d \int_0^1 [(\varphi_n(s) - \varphi(s))_{i,j}]^2 ds \right)^{3/2}. \end{aligned} \quad (5.14)$$

Step 3. We now apply an argument similar to that of [FN10, Theorem 1]. Note that

$$\mathbf{M}_n(t) = \int_0^{t \wedge 1} \varphi_n(s) d\mathbf{W}(s) + (\mathbf{W}(t) - \mathbf{W}(1)) \mathbb{1}_{[t > 1]}$$

is a martingale vanishing at zero. In particular, so are the coordinate processes

$$\mathbf{M}_n^{(i)}(t) = \int_0^{t \wedge 1} \sum_{j=1}^d (\varphi_n)_{i,j} d\mathbf{W}^{(j)}(s) + (\mathbf{W}^{(i)}(t) - \mathbf{W}^{(i)}(1)) \mathbb{1}_{[t > 1]}$$

Note that, by the Dambis-Dubins-Schwarz theorem, for each $i = 1, \dots, d$, there exists a Wiener process $\tilde{\mathbf{W}}^{(i)}$, such that

$$\mathbf{M}_n^{(i)}(t) = \tilde{\mathbf{W}}^{(i)} \left(\left\langle \mathbf{M}_n^{(i)} \right\rangle_t \right), \quad t \geq 0,$$

where $\langle \mathbf{M}_n^{(i)} \rangle_t$ is the quadratic variation of $\mathbf{M}_n^{(i)}$, i.e.

$$\langle \mathbf{M}_n^{(i)} \rangle_t = \sum_{j=1}^d \int_0^{t \wedge 1} ((\varphi_n)_{i,j})^2 ds + (t-1) \vee 0.$$

Note that

$$\langle \mathbf{M}_n^{(i)} \rangle_1 = \sum_{m=p_1}^n \int_0^1 (\Sigma_n^{(m)})_{i,i} \mathbb{1}_{(\frac{m-1}{n}, \frac{m}{n}]}(s) ds = \frac{1}{n} \sum_{m=p_1}^n (\Sigma_n^{(m)})_{i,i} = T_n^{(i)}$$

and

$$\begin{aligned} \sup_{t \in [0,1]} \left(\langle \mathbf{M}_n^{(i)} \rangle_t - \langle \mathbf{M}_n^{(i)} \rangle_{\lfloor nt \rfloor / n} \right) &= \sup_{t \in [0,1]} \sum_{j=1}^d \int_{\lfloor nt \rfloor / n}^t ((\varphi_n)_{i,j}(s))^2 ds \\ &= \sup_{t \in [0,1]} \sum_{j=1}^d \int_{\lfloor nt \rfloor / n}^t \left(\left((\Sigma_n^{((\lfloor nt \rfloor + 1) \wedge n)})^{1/2} \right)_{i,j}^2 ds \right. \\ &= \sup_{t \in [0,1]} \left(t - \frac{\lfloor nt \rfloor}{n} \right) (\Sigma_n^{((\lfloor nt \rfloor + 1) \wedge n)})_{i,i} \\ &\leq \frac{1}{(\sigma_n^{(i)})^2} \sup_{m \in [n]} \sum_{\substack{J \in \mathcal{D}_{p_i}(m): \\ m = \max(J)}} a_J(i)^2 \mathbb{E} [\psi(i)^2 (X_1, \dots, X_{p_i})] \\ &= \delta_n^{(i)}. \end{aligned}$$

Therefore, using [FN10, Lemma 3], we have that

$$\begin{aligned} &\mathbb{E} \sup_{t \in [0,1]} \left| \left(\int_{\lfloor nt \rfloor / n}^t \varphi_n(s) d\mathbf{W}(s) \right)_i \right|^2 \\ &\leq \mathbb{E} \sup \left\{ \left| \tilde{\mathbf{W}}^{(i)}(u) - \tilde{\mathbf{W}}^{(i)}(v) \right|^2 : u, v \in [0, \langle \mathbf{M}_n^{(i)} \rangle_1], |u - v| \leq \sup_{t \in [0,1]} \left(\langle \mathbf{M}_n^{(i)} \rangle_t - \langle \mathbf{M}_n^{(i)} \rangle_{\lfloor nt \rfloor / n} \right) \right\} \\ &\leq \mathbb{E} \sup \left\{ \left| \tilde{\mathbf{W}}^{(i)}(u) - \tilde{\mathbf{W}}^{(i)}(v) \right|^2 : u, v \in [0, T_n^{(i)}], |u - v| \leq \delta_n^{(i)} \right\} \\ &\leq \frac{5 \cdot 6^2}{2 \log 2} \left(\delta_n^{(i)} \log \frac{2T_n^{(i)}}{\delta_n^{(i)}} \right) \end{aligned}$$

and

$$\begin{aligned} &\mathbb{E} \sup_{t \in [0,1]} \left| \left(\int_{\lfloor nt \rfloor / n}^t \varphi_n(s) d\mathbf{W}(s) \right)_i \right|^3 \\ &\leq \mathbb{E} \sup \left\{ \left| \tilde{\mathbf{W}}^{(i)}(u) - \tilde{\mathbf{W}}^{(i)}(v) \right|^3 : u, v \in [0, \langle \mathbf{M}_n^{(i)} \rangle_1], |u - v| \leq \sup_{t \in [0,1]} \left(\langle \mathbf{M}_n^{(i)} \rangle_t - \langle \mathbf{M}_n^{(i)} \rangle_{\lfloor nt \rfloor / n} \right) \right\} \\ &\leq \mathbb{E} \sup \left\{ \left| \tilde{\mathbf{W}}^{(i)}(u) - \tilde{\mathbf{W}}^{(i)}(v) \right|^3 : u, v \in [0, T_n^{(i)}], |u - v| \leq \delta_n^{(i)} \right\} \\ &\leq \frac{5 \cdot 6^3}{\sqrt{\pi} (\log 2)^{3/2}} \left(\delta_n^{(i)} \log \left(\frac{2T_n^{(i)}}{\delta_n^{(i)}} \right) \right)^{3/2}. \end{aligned}$$

Finally, it follows that

$$\mathbb{E} \sup_{t \in [0,1]} \left| \int_{\lfloor nt \rfloor/n}^t \varphi_n(s) d\mathbf{W}(s) \right| \leq \frac{6\sqrt{5}}{\sqrt{2\log 2}} \sqrt{\sum_{i=1}^d \delta_n^{(i)} \log \left(\frac{2T_n^{(i)}}{\delta_n^{(i)}} \right)}; \quad (5.15)$$

$$\mathbb{E} \sup_{t \in [0,1]} \left| \int_{\lfloor nt \rfloor/n}^t \varphi_n(s) d\mathbf{W}(s) \right|^3 \leq \sqrt{d} \sum_{i=1}^d \mathbb{E} \sup_{t \in [0,1]} \left| \left(\int_{\lfloor nt \rfloor/n}^t \varphi_n(s) d\mathbf{W}(s) \right)_i \right|^3 \leq \frac{5 \cdot 6^3 \sqrt{d}}{\sqrt{\pi} (\log 2)^{3/2}} \sum_{i=1}^d \left(\delta_n^{(i)} \log \left(\frac{2T_n^{(i)}}{\delta_n^{(i)}} \right) \right)^{3/2}. \quad (5.16)$$

Step 3. Using the calculations above, we note that

$$\begin{aligned} \mathbb{E} \|\mathbf{Z}_n - \mathbf{Z}\| &\stackrel{(5.13), (5.15)}{\leq} 2 \sqrt{\int_0^1 \|\varphi_n(s) - \varphi(s)\|_F^2 ds} + \frac{6\sqrt{5}}{\sqrt{2\log 2}} \sqrt{\sum_{i=1}^d \delta_n^{(i)} \log \left(\frac{2T_n^{(i)}}{\delta_n^{(i)}} \right)}; \\ \mathbb{E} \|\mathbf{Z}_n - \mathbf{Z}\|^3 &\stackrel{(5.14), (5.16)}{\leq} \frac{20 \cdot 6^3 \sqrt{d}}{\sqrt{\pi} (\log 2)^{3/2}} \left(\delta_n^{(i)} \log \left(\frac{2T_n^{(i)}}{\delta_n^{(i)}} \right) \right)^{3/2} + \frac{54\sqrt{d}}{\sqrt{2\pi}} \sum_{i=1}^d \left(\sum_{j=1}^d \int_0^1 [(\varphi_n(s) - \varphi(s))_{i,j}]^2 ds \right)^{3/2}. \end{aligned}$$

We furthermore note that, using Doob's L^3 inequality, the formula for Gaussian moments and Itô's isometry,

$$\begin{aligned} \mathbb{E} \|\mathbf{Z}\|^3 &= \mathbb{E} \left[\sup_{t \in [0,1]} \left(\sum_{i=1}^d \left(\sum_{j=1}^d \int_0^t (\varphi(s))_{i,j} d\mathbf{W}^{(j)}(s) \right)^2 \right)^{3/2} \right] \\ &\leq \frac{27\sqrt{d}}{8} \sum_{i=1}^d \mathbb{E} \left[\left| \sum_{j=1}^d \int_0^1 (\varphi(s))_{i,j} d\mathbf{W}^{(j)}(s) \right|^3 \right] \\ &= \frac{27\sqrt{d}}{2\sqrt{2\pi}} \sum_{i=1}^d \left(\mathbb{E} \left[\left(\sum_{j=1}^d \int_0^1 (\varphi(s))_{i,j} d\mathbf{W}^{(j)}(s) \right)^2 \right] \right)^{3/2} \\ &= \frac{27\sqrt{d}}{2\sqrt{2\pi}} \sum_{i=1}^d \left(\sum_{j=1}^d \int_0^1 |(\varphi(s))_{i,j}|^2 ds \right)^{3/2}. \end{aligned}$$

Therefore, using the mean value theorem

$$\begin{aligned} |\mathbb{E}g(\mathbf{D}_n) - \mathbb{E}g(\mathbf{Z})| &\leq \mathbb{E} \left[\sup_{c \in [0,1]} \|Dg(\mathbf{Z} + c(\mathbf{Z}_n - \mathbf{Z}))\| \|\mathbf{Z} - \mathbf{Z}_n\| \right] \\ &\leq \|g\|_M \mathbb{E} \left[\sup_{c \in [0,1]} (1 + \|\mathbf{Z} + c(\mathbf{Z}_n - \mathbf{Z})\|^2) \|\mathbf{Z} - \mathbf{Z}_n\| \right] \\ &\stackrel{\text{Hölder}}{\leq} \|g\|_M \left\{ \mathbb{E} \|\mathbf{Z} - \mathbf{Z}_n\| + 2\mathbb{E} \|\mathbf{Z} - \mathbf{Z}_n\|^3 + 2(\mathbb{E} \|\mathbf{Z}\|^3)^{2/3} (\mathbb{E} \|\mathbf{Z} - \mathbf{Z}_n\|^3)^{1/3} \right\} \\ &\leq \|g\|_M (\gamma_3 + \gamma_4 + \gamma_5) \end{aligned}$$

and

$$|\mathbb{E}g(\mathbf{D}_n) - \mathbb{E}g(\mathbf{Z})| \leq \|g\|_{M^0} \mathbb{E} \|\mathbf{Z}_n - \mathbf{Z}\| \leq \|g\|_{M^0} \gamma_3.$$

The result now follows by Theorem 5.1 and the triangle inequality. \square

Remark 5.3. The approximation results in this Section are merely stated for vectors of degenerate weighted U -processes. In many applications, however, the given weighted U -process might involve non-degenerate kernels. If

$$\mathbf{U}_n(t) = \sum_{J \in \mathcal{D}_p(\lfloor nt \rfloor)} a_J \psi(X_j, j \in J)$$

is such a non-degenerate, weighted U -process, then it can be written in its Hoeffding decomposition as a sum of degenerate, weighted U -processes as follows:

$$\begin{aligned} \mathbf{U}_n(t) &= \int_{E^p} \psi d\mu^p \sum_{J \in \mathcal{D}_p(\lfloor nt \rfloor)} a_J + \sum_{q=1}^p \sum_{K \in \mathcal{D}_q(\lfloor nt \rfloor)} \left(\sum_{\substack{J \in \mathcal{D}_p(\lfloor nt \rfloor): \\ K \subseteq J}} a_J \right) \psi_q(X_i, i \in K) \\ &=: \int_{E^p} \psi d\mu^p \sum_{J \in \mathcal{D}_p(\lfloor nt \rfloor)} a_J + \sum_{q=1}^p \mathbf{U}_n^{(q)}(t), \end{aligned}$$

where the kernels ψ_q , $1 \leq q \leq p$, are degenerate kernels which are expressible in terms of ψ . Hence, the results of this Section for the vector $(\mathbf{U}_n^{(1)}, \dots, \mathbf{U}_n^{(p)})$ together with the application of a linear functional immediately yield bounds on the approximation of \mathbf{U}_n by a suitable Gaussian process. For simplicity we do not state the resulting bounds explicitly but leave their derivation to the interested reader. In the very particular example of d -runs on the line, however, we will work out this procedure in full detail.

5.5. Homogeneous sum processes

In this subsection we consider an important subclass of weighted degenerate U -statistics, namely the processes given as so-called *homogeneous sums*. In this case, the random variables $X_i, i \in \mathbb{N}$, are real-valued such that $\mathbb{E}|X_1|^3 < \infty$, $\mathbb{E}[X_1] = 0$ and $\mathbb{E}[X_1^2] = 1$. Moreover, for each $1 \leq i \leq d$, the kernel $\psi(i)$ is given by

$$\psi(i)(x_1, \dots, x_{p_i}) = \prod_{j=1}^{p_i} x_j.$$

In particular, $\psi(i)$ does not depend on n . Hence, for $1 \leq i \leq d$ and $t \in [0, 1]$ we have that

$$\mathbf{Y}_n^{(i)}(t) = \frac{1}{\sigma_n(i)} \sum_{J \in \mathcal{D}_{p_i}(\lfloor nt \rfloor)} a_J(i) \prod_{j \in J} X_j,$$

where the $\sigma_n(i)$ are positive reals and the random variables $Z_J(i)$ making up the processes $\mathbf{D}_n^{(i)}$, defined in Subsection 5.2, are standard normally distributed. In this situation we have the following results, which are direct consequences of Theorems 5.1 and 5.2, respectively.

Corollary 5.4. *With the above definitions and notation we have that*

$$\begin{aligned} \left| \mathbb{E}[g(\mathbf{Y}_n)] - \mathbb{E}[g(\mathbf{D}_n)] \right| &\leq \frac{2\sqrt{d}\|g\|_M}{3p_1} \sum_{i=1}^d \frac{(\mathbb{E}|X_1|^3)^{p_i}}{\sigma_n(i)^3} \sum_{l=1}^n \left(\sum_{\substack{J \in \mathcal{D}_{p_i}(n): \\ l \in J}} |a_J(i)| \right)^3 \\ &+ \|g\|_M \sum_{i,j,k=1}^d \frac{(\mathbb{E}|X_1|^3)^{(p_i+p_j+p_k)/3}}{\sigma_n(i)\sigma_n(j)\sigma_n(k)} \sum_{\substack{J \in \mathcal{D}_{p_i}(n), \\ K \in \mathcal{D}_{p_j}(n), \\ L \in \mathcal{D}_{p_k}(n): \\ J \cap K \neq \emptyset, \\ L \cap (J \cup K) \neq \emptyset}} |a_J(i)a_K(j)a_L(k)|. \end{aligned}$$

Corollary 5.5. Let $\Sigma_n^{(m)} \in \mathbb{R}^{d \times d}$ be given by

$$\left(\Sigma_n^{(m)}\right)_{i,l} = \begin{cases} \frac{n}{\sigma_n(i)\sigma_n(l)} \sum_{\substack{J \in \mathcal{D}_{p_i}(m): \\ m=\max(J)}} a_J(i)a_J(l), & \text{if } p_i = p_l \\ 0, & \text{otherwise,} \end{cases}$$

for $i, l = 1, \dots, d$.

For $i = 1, \dots, d$, let

$$\delta_n^{(i)} = \frac{1}{(\sigma_n(i))^2} \sup_{m \in [n]} \sum_{\substack{J \in \mathcal{D}_{p_i}(m): \\ m=\max(J)}} a_J(i)^2$$

and

$$T_n^{(i)} = \frac{1}{(\sigma_n(i))^2} \sum_{J \in \mathcal{D}_{p_i}(n)} a_J(i)^2.$$

Furthermore, let

$$\varphi_n(s) = \sum_{m=p_1}^n \left(\Sigma_n^{(m)}\right)^{1/2} \mathbb{1}_{\left(\frac{m-1}{n}, \frac{m}{n}\right]}(s), \quad s \in [0, 1]$$

and suppose that $\varphi : [0, 1] \rightarrow \mathbb{R}^{d \times d}$ is such that, for any $i, j = 1, \dots, d$,

$$\lim_{n \rightarrow \infty} \int_0^1 \left| (\varphi_n(s) - \varphi(s))_{i,j} \right|^2 ds = 0,$$

Let \mathbf{Y}_n be defined as in Section 5.1 and $\|\cdot\|_F$ denote the Frobenius norm. Suppose that \mathbf{W} is a d -dimensional standard Brownian motion and

$$\mathbf{Z}(t) = \int_0^t \varphi(s) d\mathbf{W}(s).$$

Then, for any $g \in M$,

$$|\mathbb{E}g(\mathbf{Y}_n) - \mathbb{E}g(\mathbf{Z})| \leq \|g\|_M(\gamma_1 + \gamma_2 + \gamma_3 + \gamma_4 + \gamma_5)$$

and for any $g \in M^0$,

$$|\mathbb{E}g(\mathbf{Y}_n) - \mathbb{E}g(\mathbf{Z})| \leq \|g\|_{M^0}(\gamma_1 + \gamma_2 + \gamma_3),$$

where

$$\begin{aligned} \gamma_1 &= \frac{2\sqrt{d}}{3p_1} \sum_{i=1}^d \frac{(\mathbb{E}|X_1|^3)^{p_i}}{\sigma_n(i)^3} \sum_{l=1}^n \left(\sum_{\substack{J \in \mathcal{D}_{p_i}(n): \\ l \in J}} |a_J(i)| \right)^3; \\ \gamma_2 &= \sum_{i,j,k=1}^d \frac{(\mathbb{E}|X_1|^3)^{(p_i+p_j+p_k)/3}}{\sigma_n(i)\sigma_n(j)\sigma_n(k)} \sum_{\substack{J \in \mathcal{D}_{p_i}(n), \\ K \in \mathcal{D}_{p_j}(n), \\ L \in \mathcal{D}_{p_k}(n): \\ J \cap K \neq \emptyset, \\ L \cap (J \cup K) \neq \emptyset}} |a_J(i)a_K(j)a_L(k)|; \\ \gamma_3 &= 2\sqrt{\int_0^1 \|\varphi_n(s) - \varphi(s)\|_F^2 ds} + 12\sqrt{\sum_{i=1}^d \delta_n^{(i)} \log \left(\frac{2T_n^{(i)}}{\delta_n^{(i)}} \right)}; \\ \gamma_4 &= \sqrt{d} \sum_{i=1}^d \left[8447 \left(\delta_n^{(i)} \log \left(\frac{2T_n^{(i)}}{\delta_n^{(i)}} \right) \right)^{3/2} + 44 \left(\sum_{j=1}^d \int_0^1 [(\varphi_n(s) - \varphi(s))_{i,j}]^2 ds \right)^{3/2} \right]; \\ \gamma_5 &= \sqrt{d} \left(\int_0^1 \|\varphi(s)\|_F^2 ds \right) \sum_{i=1}^d \left[50\sqrt{\delta_n^{(i)} \log \left(\frac{2T_n^{(i)}}{\delta_n^{(i)}} \right)} + 19\sqrt{\sum_{j=1}^d \int_0^1 [(\varphi_n(s) - \varphi(s))_{i,j}]^2 ds} \right]. \end{aligned}$$

5.6. Example: runs on the line

Let ξ_1, \dots, ξ_n be i.i.d. random variables, such that $\mathbb{P}[\xi_1 = 1] = p = 1 - \mathbb{P}[\xi_1 = 0]$, for $p \in (0, 1)$. For any $1 \leq r < n$ let $\sigma_n(r) = \sqrt{np^r(1-p)}$ and V_r be the rescaled centred number of r -runs given by

$$\mathbf{V}_n^{(r)}(t) := \frac{1}{\sigma_n(r)} \sum_{m=1}^{\lfloor nt \rfloor} (\xi_m \cdot \xi_{m+1} \cdots \xi_{m+r-1} - p^r), \quad t \in [0, 1],$$

where we adopt the torus convention, i.e. that $\xi_{n+1} = \xi_1, \xi_{n+2} = \xi_2$ and so on.

A similar setup was considered in [RR09], where the authors studied the rate of the (finite-dimensional) weak convergence of the law of $\mathbf{V}_n^{(r)}(1)$ to the normal distribution. The authors of [RR09] note that the standard exchangeable-pair construction of [RR97b] does not lead to a bound going to zero as $n \rightarrow \infty$. In order to solve this problem, they apply their *embedding method* and study the joint convergence of $(\mathbf{V}_n^{(1)}(1), \dots, \mathbf{V}_n^{(r)}(1))$ to a multivariate normal law, using a slightly unusual construction of the exchangeable pair. Our propositions in this subsection provide bounds on the rate of the *functional* convergence of $(\mathbf{V}_n^{(r_1)}, \dots, \mathbf{V}_n^{(r_d)})$ to a Gaussian process for *any* collection $\{r_1, \dots, r_d\}$. They implicitly use the standard exchangeable-pair construction of Subsection 5.1. Our bounds are of the same order as the bound on the rate of the (finite-dimensional) convergence provided in [RR09].

We start with the following result on the pre-limiting approximation:

Proposition 5.6. *Adopt the notation from above. Let $d \geq 1$ and $\frac{n}{2} > r_1 \geq r_2 \geq \dots \geq r_d \geq 1$. Let*

$$\mathbf{V}_n = (\mathbf{V}_n^{(r_1)}, \dots, \mathbf{V}_n^{(r_d)}).$$

Let $\{Z_J : J \in \mathcal{D}_j(n), j = 1, \dots, r_1\}$ be a collection of i.i.d. standard normal random variables. For $i = 1, \dots, d$, let furthermore

$$\mathbf{D}_n^{(r_i)}(t) = \frac{1}{\sigma_n(r_i)} \sum_{m=1}^{\lfloor nt \rfloor} \sum_{j=1}^{r_i} \sum_{0 \leq i_1 < \dots < i_j \leq r_i - 1} p^{r_i - j} Z_{m+i_1, \dots, m+i_j}, \quad t \in [0, 1].$$

and

$$\mathbf{D}_n = (\mathbf{D}_n^{(r_1)}, \dots, \mathbf{D}_n^{(r_d)}).$$

Then, for any $g \in M^0$,

$$|\mathbb{E}g(\mathbf{V}_n) - \mathbb{E}g(\mathbf{D}_n)| \leq \|g\|_{M^0} (\gamma_1 + \gamma_2) n^{-1/2},$$

where

$$\begin{aligned} \gamma_1 &= \frac{2\sqrt{dr_1} \left(\sum_{i=1}^d r_i \right)^{3/2}}{3r_d} \sum_{i=1}^d \sum_{j=1}^{r_i} \frac{(1+p^3-2p^4)^j p^{3r_i/2-3j} \binom{r_i-1}{j-1}^3}{(1-p)^{3/2}}; \\ \gamma_2 &= 2\sqrt{dr_1} \left(\sum_{i=1}^d r_i \right) \sum_{u,v,w=1}^d \sum_{j_1=1}^{r_u} \sum_{j_2=1}^{r_v} \sum_{j_3=1}^{r_w} \frac{(1+p^3-2p^4)^{(j_1+j_2+j_3)/3} p^{(r_u+r_v+r_w)/2-j_1-j_2-j_3}}{(1-p)^{3/2}} r_w (r_u \vee r_v)^2 \\ &\quad \cdot \binom{r_u-1}{j_1-1} \binom{r_v-1}{j_2-1} \binom{r_w-1}{j_3-1}. \end{aligned}$$

Proof.

Step 1. For $i = 1, 2, \dots$, let $X_i = \xi_i - p$. It is easy to prove, by induction on r , that

$$\mathbf{V}_n^{(r)}(t) = \frac{1}{\sigma_n(r)} \sum_{m=1}^{\lfloor nt \rfloor} \sum_{j=1}^r \sum_{0 \leq i_1 < \dots < i_j \leq r-1} p^{r-j} X_{m+i_1} \cdots X_{m+i_j}, \quad t \in [0, 1] \quad (5.17)$$

Indeed, for any $m = 1, \dots, n$,

$$\xi_m - p = X_m$$

and, assuming that

$$\xi_m \xi_{m+1} \dots \xi_{m+r-1} - p^r = \sum_{j=1}^r \sum_{0 \leq i_1 < \dots < i_j \leq r-1} p^{r-j} X_{m+i_1} \dots X_{m+i_j} \quad (5.18)$$

we have,

$$\begin{aligned} & \xi_m \xi_{m+1} \dots \xi_{m+r} - p^{r+1} \\ &= (\xi_m \xi_{m+1} \dots \xi_{m+r-1} - p^r) (\xi_{m+r} - p) + p (\xi_m \xi_{m+1} \dots \xi_{m+r-1} - p^r) + p^r (\xi_{m+r} - p) \\ &\stackrel{(5.18)}{=} \sum_{j=1}^r \sum_{0 \leq i_1 < \dots < i_j \leq r-1} p^{r-j} X_{m+i_1} \dots X_{m+i_j} X_{m+r} + \sum_{j=1}^r \sum_{0 \leq i_1 < \dots < i_j \leq r-1} p^{r+1-j} X_{m+i_1} \dots X_{m+i_j} + p^r X_{m+r} \\ &= \sum_{j=2}^{r+1} \sum_{0 \leq i_1 < \dots < i_j = r} p^{r+1-j} X_{m+i_1} \dots X_{m+i_j} + \sum_{j=1}^r \sum_{0 \leq i_1 < \dots < i_j \leq r-1} p^{r+1-j} X_{m+i_1} \dots X_{m+i_j} + p^r X_{m+r} \\ &= \sum_{j=1}^{r+1} \sum_{0 \leq i_1 < \dots < i_j \leq r} p^{r+1-j} X_{m+i_1} \dots X_{m+i_j}, \end{aligned}$$

as required.

Step 2. Now, for any $r = 1, 2, \dots, r_1$ and $j = 1, \dots, r$, note that

$$\begin{aligned} & \frac{p^{r-j}}{\sigma_n(r)} \sum_{m=1}^{\lfloor nt \rfloor} \sum_{0 \leq i_1 < \dots < i_j \leq r-1} X_{m+i_1} \dots X_{m+i_j} \\ &= \frac{p^{r-j}}{\sigma_n(r)} \sum_{m=1}^{\lfloor nt \rfloor} \sum_{m \leq i_1 < \dots < i_j \leq m+r-1} X_{i_1} \dots X_{i_j} \\ &= \frac{p^{r-j}}{\sigma_n(r)} \sum_{1 \leq i_1 < \dots < i_j \leq \lfloor nt \rfloor + r - 1} ((r - i_j + i_1) \vee 0) X_{i_1} \dots X_{i_j} \\ &= \frac{p^{r-j}}{\sigma_n(r)} \sum_{J \in \mathcal{D}_j((\lfloor nt \rfloor + r - 1) \wedge n)} a_J(r) X_{i_1} \dots X_{i_j}, \end{aligned}$$

for

$$\begin{aligned} a_J(r) &= p^{r-j} \max(r - \max(J) + \min(J), 0) \\ &\quad + p^{r-j} \max(r + \min(J \cap (n/2, n]) - \max(J \cap [1, n/2)) - n, 0) \mathbb{1}_{\{J \cap [1, n/2) \neq \emptyset \neq J \cap (n/2, n]\}}. \end{aligned}$$

Furthermore, let

$$\mathbf{U}_n^{(r,j)}(t) = \frac{1}{\sigma_n(r)} \sum_{J \in \mathcal{D}_j(\lfloor nt \rfloor)} a_J(r) \prod_{i \in J} X_i, \quad t \in [0, 1].$$

and define function $f : (D([0, 1], \mathbb{R}))^{r_1 + \dots + r_d} \rightarrow D([0, 1], \mathbb{R}^d)$, given by

$$\begin{aligned} & f(x_{1,1}, \dots, x_{1,r_1}, x_{2,2}, \dots, x_{2,r_2}, \dots, x_{d,1}, \dots, x_{d,r_d}) \\ &= \left(\left(\sum_{j=1}^{r_1} x_{1,j} \left(\left(t + \frac{r_1 - 1}{n} \right) \wedge 1 \right), \dots, \sum_{j=1}^{r_d} x_{d,j} \left(\left(t + \frac{r_d - 1}{n} \right) \wedge 1 \right) \right), t \in [0, 1] \right). \end{aligned}$$

Hence, note that, by (5.17),

$$g(\mathbf{V}_n) = g \circ f \left(\mathbf{U}_n^{(r_1,1)}, \dots, \mathbf{U}_n^{(r_1,r_1)}, \dots, \mathbf{U}_n^{(r_d,1)}, \dots, \mathbf{U}_n^{(r_d,r_d)} \right).$$

It is proved in Lemma 7.1 in Section 7.1 of the Appendix that

$$\|g \circ f\|_{M^0} \leq \|g\|_{M^0} \sqrt{dr_1} \sum_{i=1}^d r_i. \quad (5.19)$$

Step 3. Now, note that, for $r, r_u, r_v, r_w = 1, 2, \dots, r_1$,

$$\begin{aligned} 1) \quad & \sum_{l=1}^n \left(\sum_{\substack{J \in \mathcal{D}_j(n): \\ l \in J}} |a_J(r)| \right)^3 \leq p^{3r-3j} \sum_{l=1}^n \left(\sum_{m=l-r+1}^l \binom{r-1}{j-1} \right)^3 = p^{3r-3j} r^3 \binom{r-1}{j-1}^3 n \\ 2) \quad & \sum_{\substack{J \in \mathcal{D}_{j_1}(n), \\ K \in \mathcal{D}_{j_2}(n), \\ L \in \mathcal{D}_{j_3}(n): \\ J \cap K \neq \emptyset, \\ L \cap (J \cup K) \neq \emptyset}} |a_J(r_u) a_K(r_v) a_L(r_w)| \\ & \leq \frac{p^{r_u+r_v+r_w-j_1-j_2-j_3}}{r_u \wedge r_v} \sum_{l=1}^n \sum_{m_1=l-r_u+1}^l \sum_{m_2=l-r_v+1}^l \sum_{k=l-r_u \vee r_v+1}^{l+r_u \vee r_v-1} \sum_{m_3=k-r_w+1}^k \binom{r_u-1}{j_1-1} \binom{r_v-1}{j_2-1} \binom{r_w-1}{j_3-1} \\ & \leq 2p^{r_u+r_v+r_w-j_1-j_2-j_3} r_w (r_u \vee r_v)^2 \binom{r_u-1}{j_1-1} \binom{r_v-1}{j_2-1} \binom{r_w-1}{j_3-1} n. \end{aligned} \quad (5.20)$$

and so, using (5.20) and (5.19), for any $g \in M^0$,

$$\begin{aligned} A) \quad & \|g \circ f\|_{M^0} \frac{2\sqrt{\sum_{i=1}^d r_i}}{3r_d} \sum_{i=1}^d \sum_{j=1}^{r_i} \frac{(\mathbb{E}|X_1|^3)^j}{\sigma_n(r_i)^3} \sum_{l=1}^n \left(\sum_{\substack{J \in \mathcal{D}_j(n): \\ l \in J}} |a_J(r_i)| \right)^3 \\ & \leq \|g\|_{M^0} \frac{2\sqrt{dr_1} \left(\sum_{i=1}^d r_i \right)^{3/2}}{3r_d} \sum_{i=1}^d \sum_{j=1}^{r_i} \frac{(1+p^3-2p^4)^j p^{3r_i/2-3j}}{(1-p)^{3/2}} \binom{r_i-1}{j-1}^3 n^{-1/2}; \\ & \leq \|g\|_{M^0} \frac{2\sqrt{dr_1} \left(\sum_{i=1}^d r_i \right)^{3/2}}{3r_d} \sum_{i=1}^d \sum_{j=1}^{r_i} \frac{(1+p^3-2p^4)^j p^{3r_i/2-3j}}{(1-p)^{3/2}} \binom{r_i-1}{j-1}^3 n^{-1/2}; \\ B) \quad & \|g \circ f\|_{M^0} \sum_{u,v,w=1}^d \sum_{j_1=1}^{r_u} \sum_{j_2=1}^{r_v} \sum_{j_3=1}^{r_w} \frac{(\mathbb{E}|X_1|^3)^{(j_1+j_2+j_3)/3}}{\sigma_n(r_u) \sigma_n(r_v) \sigma_n(r_w)} \sum_{\substack{J \in \mathcal{D}_{j_1}(n), \\ K \in \mathcal{D}_{j_2}(n), \\ L \in \mathcal{D}_{j_3}(n): \\ J \cap K \neq \emptyset, \\ L \cap (J \cup K) \neq \emptyset}} |a_J(r_u) a_K(r_v) a_L(r_w)| \\ & \leq 2\sqrt{dr_1} \left(\sum_{i=1}^d r_i \right) \sum_{u,v,w=1}^d \sum_{j_1=1}^{r_u} \sum_{j_2=1}^{r_v} \sum_{j_3=1}^{r_w} \frac{(1+p^3-2p^4)^{(j_1+j_2+j_3)/3} p^{(r_u+r_v+r_w)/2-j_1-j_2-j_3}}{(1-p)^{3/2}} r_w (r_u \vee r_v)^2 \\ & \quad \cdot \binom{r_u-1}{j_1-1} \binom{r_v-1}{j_2-1} \binom{r_w-1}{j_3-1} n^{-1/2}. \end{aligned}$$

The result now follows by Corollary 5.4. \square

Next, we deal with the continuous process approximation as given in Corollary 5.5. For this, we need to either compute or estimate the quantities $\delta_n^{(i)}$, $T_n^{(i)}$ and $\Sigma_n^{(m)}$. After rearranging the entries of the random vector according to their order as homogeneous sums, we can write $\Sigma_n^{(m)}$ as a block diagonal matrix. More precisely, for $1 \leq q \leq r_1$ letting

$$N(q) := \max\{1 \leq j \leq d : r_j \geq q\}, \quad (5.21)$$

we can write $\Sigma_n^{(m)}$ as a block diagonal matrix with blocks $\Sigma_n^{(m)}(1), \dots, \Sigma_n^{(m)}(r_1)$, where, for fixed $q = 1, \dots, r_1$, $\Sigma_n^{(m)}(q)$ is an $N(q) \times N(q)$ matrix, namely the covariance matrix of the random vector

$$\left(\sqrt{n} \mathbf{U}_n^{(r_1, q)}(m/n) - \sqrt{n} \mathbf{U}_n^{(r_1, q)}((m-1)/n), \dots, \sqrt{n} \mathbf{U}_n^{(r_{N(q)}, q)}(m/n) - \sqrt{n} \mathbf{U}_n^{(r_{N(q)}, q)}((m-1)/n) \right)^T.$$

A simple computation shows that, for $q > 1$ and $r_i \wedge r_l \leq m \leq n+1 - r_i \wedge r_l$,

$$\begin{aligned} \Sigma_n^{(m)}(q)(i, l) &= \frac{n}{\sigma_n(r_i) \sigma_n(r_l)} \sum_{\substack{J \in \mathcal{D}_q(n): \\ \max(J)=m}} a_J(r_i) a_J(r_l) \\ &= \frac{p^{\frac{r_i+r_l}{2}-q}}{1-p} \sum_{k=q-1}^{r_i \wedge r_l - 1} \binom{k-1}{q-2} (r_i - k)(r_l - k). \end{aligned}$$

Otherwise, for $q > 1$ and $m \geq n+2 - r_i \wedge r_l$,

$$\Sigma_n^{(m)}(q)(i, l) = \frac{p^{\frac{r_i+r_l}{2}-q}}{1-p} \left[\sum_{k=q-1}^{r_i \wedge r_l - 1} \binom{k-1}{q-2} (r_i - k)(r_l - k) + \sum_{u=n+2-r_i \wedge r_l}^m \sum_{k=(q-1) \vee (n-u-1)}^{r_i \wedge r_l - 1} \binom{k-1}{q-2} (r_i - k)(r_l - k) \right],$$

for $q > 1$ and $m \leq r_i \wedge r_l - 1$,

$$\Sigma_n^{(m)}(q)(i, l) = \frac{p^{\frac{r_i+r_l}{2}-q}}{1-p} \sum_{k=q-1}^{m-1} \binom{k-1}{q-2} (r_i - k)(r_l - k),$$

and, for all $1 \leq m \leq n$

$$\Sigma_n^{(m)}(1)(i, l) = \frac{p^{\frac{r_i+r_l}{2}-1}}{1-p} r_i r_l.$$

Hence, we let Σ be a block diagonal matrix with blocks $\Sigma(1) \in \mathbb{R}^{N(1) \times N(1)}, \dots, \Sigma(r_1) \in \mathbb{R}^{N(r_1) \times N(r_1)}$, where

$$\Sigma(1)(i, l) = \frac{p^{\frac{r_i+r_l}{2}-1}}{1-p} r_i r_l \quad (5.22)$$

and for any $q = 2, \dots, r_1$ and $i, l = 1, \dots, N(q)$,

$$\Sigma(q)(i, l) = \frac{p^{\frac{r_i+r_l}{2}-q}}{1-p} \sum_{k=q-1}^{r_i \wedge r_l - 1} \binom{k-1}{q-2} (r_i - k)(r_l - k). \quad (5.23)$$

Note that, for $\varphi(s) \equiv \Sigma^{1/2}$ and $\varphi_n(s) = \sum_{m=1}^n \left(\Sigma_n^{(m)} \right)^{1/2} \mathbb{1}_{\left[\frac{m-1}{n}, \frac{m}{n} \right]}(s)$, $s \in [0, 1]$,

$$\begin{aligned} & \int_0^1 \|\varphi_n(s) - \varphi(s)\|_F^2 ds \\ & \leq \frac{2(r_1-1)}{n} \left[\sum_{m=1}^{r_1-1} \left\| \left(\Sigma_n^{(m)} \right)^{1/2} - \Sigma^{1/2} \right\|_F^2 + \sum_{m=n+2-r_1}^n \left\| \left(\Sigma_n^{(m)} \right)^{1/2} - \Sigma^{1/2} \right\|_F^2 \right] \\ & \leq \frac{4(r_1-1)}{n} \sum_{k=1}^d \sum_{i=1}^{r_k} \left[\sum_{m=1}^{r_1-1} \left(\left| \left(\Sigma_n^{(m)} \right)_{i,i} \right| + |\Sigma_{i,i}| \right) + \sum_{m=n+2-r_1}^n \left(\left| \left(\Sigma_n^{(m)} \right)_{i,i} \right| + |\Sigma_{i,i}| \right) \right] \\ & \leq \frac{24(r_1)^3}{n} \sum_{q=1}^{r_1} \sum_{i=1}^{N(q)} \sum_{k=q-1}^{r_i-1} \left(\binom{k-1}{q-2} \mathbb{1}_{[q>1]} + \mathbb{1}_{[q=1]} \right) \frac{p^{r_i-q}}{1-p} (r_i - k)^2. \end{aligned}$$

Moreover, with obvious notation,

$$T_n^{(i)}(q) = \frac{1}{(\sigma_n(r_i))^2} \sum_{J \in \mathcal{D}_q(n)} a_J(r_i)^2 = \frac{1}{n} \sum_{m=1}^n \Sigma_n^{(m)}(q)(i, i) = \begin{cases} \frac{p^{r_i-1}}{1-p} r_i^2, & \text{if } q = 1, \\ \frac{p^{r_i-q}}{1-p} \sum_{k=q-1}^{r_i-1} \binom{k-1}{q-2} (r_i - k)^2, & \text{if } q > 1. \end{cases}$$

Furthermore, for $q > 1$,

$$\begin{aligned}\delta_n^{(i)}(q) &= \frac{1}{(\sigma_n(r_i))^2} \sup_{m \in [n]} \sum_{\substack{J \in \mathcal{D}_q(m): \\ m = \max(J)}} a_J(i)^2 \\ &= \frac{p^{r_i-q}}{n(1-p)} \sum_{k=q-1}^{r_i-1} \binom{k-1}{q-2} (r_i-k)^2 + \frac{p^{r_i-q}}{n(1-p)} \sum_{u=n+2-r_i}^n \sum_{k=(q-1) \vee (n-u-1)}^{r_i-1} \binom{k-1}{q-2} (r_i-k)^2\end{aligned}$$

and

$$\delta_n^{(i)}(1) = \frac{p^{r_i-1}}{n(1-p)} r_i^2$$

and so, for all $q = 1, \dots, r_1$,

$$\frac{1}{n} T_n^{(i)}(q) \leq \delta_n^{(i)}(q) \leq \frac{r_i}{n} T_n^{(i)}(q).$$

Thus, taking (5.19) into account, we note that

$$\begin{aligned}\|g \circ f\|_{M^0} &\left[12 \sqrt{\sum_{q=1}^{r_1} \sum_{i=1}^{N(q)} \delta_n^{(i)}(q) \log \left(\frac{2T_n^{(i)}(q)}{\delta_n^{(i)}(q)} \right)} + 2 \sqrt{\int_0^1 \|\varphi_n(s) - \varphi(s)\|_F^2 ds} \right] \\ &\leq \|g\|_{M^0} \sqrt{dr_1} \sum_{j=1}^d r_j \left[12 \frac{\sqrt{\log n}}{\sqrt{n}} \left(\sum_{q=2}^{r_1} \sum_{i=1}^{N(q)} \sum_{k=q-1}^{r_i-1} \binom{k-1}{q-2} \frac{p^{r_i-q} r_i}{(1-p)} (r_i-k)^2 + \sum_{i=1}^d \frac{p^{r_i-1}}{1-p} r_i^3 \right)^{1/2} \right. \\ &\quad \left. + \frac{4\sqrt{6}(r_1)^{3/2}}{\sqrt{n}} \left(\sum_{q=2}^{r_1} \sum_{i=1}^{N(q)} \sum_{k=q-1}^{r_i-1} \binom{k-1}{q-2} \frac{p^{r_i-q}}{1-p} (r_i-k)^2 + \sum_{i=1}^d \frac{p^{r_i-1}}{1-p} r_i^2 \right)^{1/2} \right] \\ &\leq \|g\|_{M^0} 4\sqrt{dr_1^2} \left(\sum_{j=1}^d r_j \right) \left(\sum_{q=2}^{r_1} \sum_{i=1}^{N(q)} \sum_{k=q-1}^{r_i-1} \binom{k-1}{q-2} \frac{p^{r_i-q}}{1-p} (r_i-k)^2 + \sum_{i=1}^d \frac{p^{r_i-1}}{1-p} r_i^2 \right)^{1/2} \cdot \frac{3\sqrt{\log n} + \sqrt{6}}{\sqrt{n}}.\end{aligned}$$

Hence, using Corollary 5.5 and Proposition 5.6 (and noting that reordering the arguments of function f does not change the bound on $\|g \circ f\|_{M^0}$ obtained in Lemma 7.1), we obtain the following result:

Proposition 5.7. *Adopt the notation form above. In particular, let N be as in (5.21), \mathbf{V}_n be defined as in Proposition 5.6 and Σ be the block diagonal matrix with blocks $\Sigma(1) \in \mathbb{R}^{N(1) \times N(1)}, \dots, \Sigma(r_1) \in \mathbb{R}^{N(r_1) \times N(r_1)}$ defined by (5.22) and (5.23). Let $\mathbf{Z}' = \Sigma^{1/2} \mathbf{W}$, where \mathbf{W} is a $(\sum_{i=1}^d r_i)$ -dimensional standard Brownian motion and write $\mathbf{Z}' = ((\mathbf{Z}')^{(1)}, (\mathbf{Z}')^{(2)}, \dots)$. Set $N(0) = 0$. For $i = 1, \dots, d$ and $t \in [0, 1]$, define*

$$\mathbf{Z}^{(i)}(t) = \left((\mathbf{Z}')^{(i)} + (\mathbf{Z}')^{(N(1)+i)} + (\mathbf{Z}')^{(N(1)+N(2)+i)} + \dots + (\mathbf{Z}')^{(N(1)+N(2)+\dots+N(r_{i-1})+i)} \right) \left(\left(t + \frac{r_i-1}{n} \right) \wedge 1 \right)$$

and let

$$\mathbf{Z} = (\mathbf{Z}^{(1)}, \dots, \mathbf{Z}^{(d)}).$$

Then, for any $g \in M^0$, we have

$$|\mathbb{E}g(\mathbf{V}_n) - \mathbb{E}g(\mathbf{Z})| \leq n^{-1/2} \|g\|_{M^0} \left(\gamma_1 + \gamma_2 + \gamma_3 \sqrt{\log n} \right),$$

where γ_1 and γ_2 are as in Proposition 5.6 and

$$\gamma_3 = 22\sqrt{dr_1^2} \left(\sum_{j=1}^d r_j \right) \left(\sum_{q=2}^{r_1} \sum_{i=1}^{N(q)} \sum_{k=q-1}^{r_i-1} \binom{k-1}{q-2} \frac{p^{r_i-q}}{1-p} (r_i-k)^2 + \sum_{i=1}^d \frac{p^{r_i-1}}{1-p} r_i^2 \right)^{1/2}.$$

Remark 5.8. Assuming that d, r_1, \dots, r_d are all fixed and do not depend on n , the bound in Proposition 5.7 is of order $\sqrt{\frac{\log n}{n}}$. Therefore, by Proposition 2.2, weak convergence of the law of \mathbf{V}_n to that of \mathbf{Z} , in both the Skorokhod and the uniform topologies on the Skorokhod space, follows immediately from Proposition 5.7 as a corollary.

Remark 5.9. It is possible to obtain bounds similar to those in Propositions 5.6 and 5.7 for the larger class of test functions M . It would, however, require some more involved computations, which would make the discussion of this example rather long.

6. Edge and two-star counts in Erdős-Renyi random graphs

In this section we study an Erdős-Renyi random graph with a fixed edge probability p and $\lfloor nt \rfloor$ edges for $t \in [0, 1]$. We analyse the asymptotic behaviour of the joint law of its (rescaled) number of edges and its (rescaled) number of two-stars (i.e. subgraphs which are trees with one internal node and 2 leaves). Hence, we extend the result of [Kas20a], where the univariate process convergence of the rescaled number of edges is studied. We also extend the analysis of [RR10], whose authors provide a bound on the distance between the (three-dimensional) joint law of the (rescaled) number of edges, two-stars and triangles in a $G(n, p)$ graph and a Gaussian vector. In Theorem 6.2, we establish a bound on the distance between our process and a pre-limiting Gaussian processes with paths in $D([0, 1], \mathbb{R}^2)$. Then, in Theorem 6.4, a bound on the quality of a continuous Gaussian process approximation is provided.

It is worth noting that the analysis of a three-dimensional process representing the number of edges, triangles and two-stars in a $G(\lfloor nt \rfloor, p)$ graph does not pose any additional challenges except that it makes the algebraic computations more involved. The only reason we do not do it here is that it would make this section rather lengthy.

6.1. Introduction

Consider an Erdős-Renyi random graph $G(\lfloor nt \rfloor, p)$ on $\lfloor nt \rfloor$ vertices, for $t \in [0, 1]$, with a fixed edge probability p . Let $I_{i,j} = I_{j,i}$'s be i.i.d. Bernoulli(p) random variables indicating that edge (i, j) is present in this graph. We consider the following process, representing the re-scaled total number of edges

$$\mathbf{T}_n(t) = \frac{\lfloor nt \rfloor - 2}{2n^2} \sum_{1 \leq i \neq j \leq \lfloor nt \rfloor} I_{i,j} = \frac{\lfloor nt \rfloor - 2}{n^2} \sum_{1 \leq i < j \leq \lfloor nt \rfloor} I_{i,j}, \quad (6.1)$$

and a re-scaled statistic related to the number of two-stars

$$\mathbf{V}_n(t) = \frac{1}{6n^2} \sum_{\substack{1 \leq i, j, k \leq \lfloor nt \rfloor \\ i, j, k \text{ distinct}}} I_{ij} I_{jk} = \frac{1}{n^2} \sum_{1 \leq i < j < k \leq \lfloor nt \rfloor} (I_{i,j} I_{j,k} + I_{i,j} I_{i,k} + I_{j,k} I_{i,k}). \quad (6.2)$$

Furthermore, let $\mathbf{Y}_n(t) = (\mathbf{T}_n(t) - \mathbb{E}\mathbf{T}_n(t), \mathbf{V}_n(t) - \mathbb{E}\mathbf{V}_n(t))$ for $t \in [0, 1]$.

Remark 6.1. Note that, for all $t \in [0, 1]$, $\mathbb{E}\mathbf{T}_n(t) = \frac{\lfloor nt \rfloor - 2}{n^2} \binom{\lfloor nt \rfloor}{2} p$ and $\mathbb{E}\mathbf{V}_n(t) = \frac{3}{n^2} \binom{\lfloor nt \rfloor}{3} p^2$ and, by an argument similar to that of [RR10, Section 5], the covariance matrix of $(\mathbf{T}_n(t) - \mathbb{E}\mathbf{T}_n(t), \mathbf{V}_n(t) - \mathbb{E}\mathbf{V}_n(t))$ is given by

$$3 \frac{\binom{\lfloor nt \rfloor}{3}}{n^4} p(1-p) \begin{pmatrix} (\lfloor nt \rfloor - 2) & 2p(\lfloor nt \rfloor - 2) \\ 2p(\lfloor nt \rfloor - 2) & 4p^2(\lfloor nt \rfloor - 2) + p(1-p) \end{pmatrix}.$$

The scaling therefore ensures that the covariances are of the same order in n .

6.2. Exchangeable pair setup

In order to construct a suitable exchangeable pair, following [RR10], we pick (I, J) according to $\mathbb{P}[I = i, J = j] = \frac{1}{\binom{n}{2}}$ for $1 \leq i < j \leq n$. If $I = i, J = j$, we replace $I_{i,j} = I_{j,i}$ by an independent copy $I'_{i,j} = I'_{j,i}$

and set:

$$\begin{aligned}\mathbf{T}'_n(t) &= \mathbf{T}_n(t) - \frac{\lfloor nt \rfloor - 2}{n^2} (I_{I,J} - I'_{I,J}) \mathbb{1}_{[I/n,1] \cap [J/n,1]}(t) \\ \mathbf{V}'_n(t) &= \mathbf{V}_n(t) - \frac{1}{n^2} \sum_{k:k \neq I,J} (I_{I,J} - I'_{I,J}) (I_{J,k} + I_{I,k}) \mathbb{1}_{[I/n,1] \cap [J/n,1] \cap [k/n,1]}(t).\end{aligned}$$

We, similarly, let $\mathbf{Y}'_n(t) = (\mathbf{T}'_n(t) - \mathbb{E}\mathbf{T}_n(t), \mathbf{V}'_n(t) - \mathbb{E}\mathbf{V}_n(t))$ and note that, for $\mathbf{Y}_n = (\mathbf{Y}_n(t), t \in [0, 1])$ and $\mathbf{Y}'_n = (\mathbf{Y}'_n(t), t \in [0, 1])$, $(\mathbf{Y}_n, \mathbf{Y}'_n)$ forms an exchangeable pair. Note that, for any $m = 1, 2$, any $f \in M$, as defined in Section 2, and e_1, e_2 denoting the canonical basis vectors $(1, 0)$ and $(0, 1)$, respectively, we have

$$\begin{aligned}& \mathbb{E}^{\mathbf{Y}_n} \{Df(\mathbf{Y}_n) [(\mathbf{T}'_n - \mathbf{T}_n) e_m]\} \\ &= \mathbb{E}^{\mathbf{Y}_n} \left\{ Df(\mathbf{Y}_n) \left[\frac{\lfloor n \cdot \rfloor - 2}{n^2} (I'_{I,J} - I_{I,J}) \mathbb{1}_{[I/n,1] \cap [J/n,1]} e_m \right] \right\} \\ &= \frac{2}{n^3(n-1)} \sum_{i < j} \mathbb{E}^{\mathbf{Y}_n} \{Df(\mathbf{Y}_n) [(\lfloor n \cdot \rfloor - 2) (I'_{i,j} - I_{i,j}) \mathbb{1}_{[i/n,1] \cap [j/n,1]} e_m] \mid I = i, J = j\} \\ &= -\frac{1}{\binom{n}{2}} Df(\mathbf{Y}_n) [\mathbf{T}_n e_m] + \frac{2}{n^3(n-1)} p \sum_{i < j} Df(\mathbf{Y}_n) [(\lfloor n \cdot \rfloor - 2) \mathbb{1}_{[i/n,1] \cap [j/n,1]} e_m] \\ &= -\frac{1}{\binom{n}{2}} Df(\mathbf{Y}_n) [(\mathbf{T}_n(\cdot) - \mathbb{E}\mathbf{T}_n(\cdot)) e_m].\end{aligned}$$

Also:

$$\begin{aligned}& \mathbb{E}^{\mathbf{Y}_n} Df(\mathbf{Y}_n) [(\mathbf{V}_n - \mathbf{V}'_n) e_m] \\ &= \frac{1}{n^2 \binom{n}{2}} \sum_{i < j} \mathbb{E}^{\mathbf{Y}_n} \left\{ \sum_{k:k \neq i,j} Df(\mathbf{Y}_n) \left[(I_{i,j} - I'_{i,j}) (I_{j,k} + I_{i,k}) \mathbb{1}_{[i/n,1] \cap [j/n,1] \cap [k/n,1]} e_m \right] \mid I = i, J = j \right\} \\ &= \frac{2}{\binom{n}{2}} Df(\mathbf{Y}_n) [\mathbf{V}_n e_m] - \frac{p}{n^2 \binom{n}{2}} \sum_{i < j} \sum_{k:k \neq i,j} \mathbb{E}^{\mathbf{Y}_n} Df(\mathbf{Y}_n) [(I_{j,k} + I_{i,k}) \mathbb{1}_{[i/n,1] \cap [j/n,1] \cap [k/n,1]} e_m] \\ &= \frac{2}{\binom{n}{2}} Df(\mathbf{Y}_n) [\mathbf{V}_n e_m] - \frac{p}{n^2 \binom{n}{2}} \sum_{\substack{1 \leq i,j,k \leq n \\ i,j,k \text{ distinct}}} \mathbb{E}^{\mathbf{Y}_n} Df(\mathbf{Y}_n) [I_{i,j} \mathbb{1}_{[i/n,1] \cap [j/n,1] \cap [k/n,1]} e_m] \\ &= \frac{2}{\binom{n}{2}} Df(\mathbf{Y}_n) [(\mathbf{V}_n - \mathbb{E}\mathbf{V}_n(\cdot)) e_m] - \frac{p}{n^2 \binom{n}{2}} \sum_{\substack{1 \leq i,j,k \leq n \\ i,j,k \text{ distinct}}} \mathbb{E}^{\mathbf{Y}_n} Df(\mathbf{Y}_n) [(I_{i,j} - p) \mathbb{1}_{[i/n,1] \cap [j/n,1] \cap [k/n,1]} e_m] \\ &= \frac{2}{\binom{n}{2}} Df(\mathbf{Y}_n) [(\mathbf{V}_n - \mathbb{E}\mathbf{V}_n(\cdot)) e_m] - \frac{2p}{\binom{n}{2}} Df(\mathbf{Y}_n) \left[\frac{1}{\lfloor n \cdot \rfloor - 2} (\mathbf{T}_n - \mathbb{E}\mathbf{T}_n(\cdot)) e_m \left(\sum_{k=1}^n \mathbb{1}_{[k/n,1]} - 2 \right) \right] \\ &= \frac{2}{\binom{n}{2}} Df(\mathbf{Y}_n) [(\mathbf{V}_n - \mathbb{E}\mathbf{V}_n(\cdot)) e_m] - \frac{2p}{\binom{n}{2}} Df(\mathbf{Y}_n) [(\mathbf{T}_n - \mathbb{E}\mathbf{T}_n(\cdot)) e_m].\end{aligned}$$

Therefore, for any $m = 1, 2$:

$$\begin{aligned}\text{A) } Df(\mathbf{Y}_n) [(\mathbf{T}_n - \mathbb{E}\mathbf{T}_n) e_m] &= \frac{n(n-1)}{2} \mathbb{E}^{\mathbf{Y}_n} \{Df(\mathbf{Y}_n) [(\mathbf{T}_n - \mathbf{T}'_n) e_m]\} \\ \text{B) } Df(\mathbf{Y}_n) [(\mathbf{V}_n - \mathbb{E}\mathbf{V}_n) e_m] &= \frac{n(n-1)}{4} \mathbb{E}^{\mathbf{Y}_n} \{Df(\mathbf{Y}_n) [(\mathbf{V}_n - \mathbf{V}'_n) e_m] + p Df(\mathbf{Y}_n) [(\mathbf{T}_n - \mathbb{E}\mathbf{T}_n) e_m]\} \\ &= \frac{n(n-1)}{4} \mathbb{E}^{\mathbf{Y}_n} \{Df(\mathbf{Y}_n) [(2p(\mathbf{T}_n - \mathbf{T}'_n) + \mathbf{V}_n - \mathbf{V}'_n) e_m]\}\end{aligned}$$

and so

$$Df(\mathbf{Y}_n) [\mathbf{Y}_n] = 2 \mathbb{E}^{\mathbf{Y}_n} Df(\mathbf{Y}_n) [(\mathbf{Y}_n - \mathbf{Y}'_n) \Lambda_n],$$

where

$$\Lambda_n = \frac{n(n-1)}{8} \begin{pmatrix} 2 & 2p \\ 0 & 1 \end{pmatrix}. \quad (6.3)$$

Therefore, condition (4.1) is satisfied with Λ_n of (6.3) and $R_f = 0$.

6.3. A pre-limiting process

Suppose that the collection $\{Z_{i,j}^{(1)} : i, j \in [n], i < j\} \cup \{Z_{i,j,k}^{(2)} : i, j, k \in [n], i < j < k\}$ is jointly centred Gaussian with the following covariance structure:

$$\begin{aligned} \mathbb{E}Z_{ij}^{(1)}Z_{kl}^{(1)} &= \begin{cases} \frac{p(1-p)}{n^4}, & \text{if } (i, j) = (k, l), \\ 0, & \text{otherwise,} \end{cases} \\ \mathbb{E}Z_{i,j,k}^{(2)}Z_{l,m}^{(1)} &= \begin{cases} \frac{2p^2(1-p)}{n^4}, & \text{if } \{l, m\} \subset \{i, j, k\}, \\ 0, & \text{otherwise,} \end{cases} \\ \mathbb{E}Z_{i,j,k}^{(2)}Z_{r,s,t}^{(2)} &= \begin{cases} \frac{3p^2(1+2p-3p^2)}{n^4}, & \text{if } (i, j, k) = (r, s, t), \\ \frac{4p^3(1-p)}{n^4}, & \text{if } |\{i, j, k\} \cap \{r, s, t\}| = 2, \\ 0, & \text{otherwise.} \end{cases} \end{aligned}$$

Let $\mathbf{D}_n = (\mathbf{D}_n^{(1)}, \mathbf{D}_n^{(2)})$ be defined in the following way:

$$\begin{aligned} \mathbf{D}_n^{(1)}(t) &= (\lfloor nt \rfloor - 2) \sum_{1 \leq i < j \leq \lfloor nt \rfloor} Z_{i,j}^{(1)}, \quad t \in [0, 1] \\ \mathbf{D}_n^{(2)}(t) &= \sum_{1 \leq i < j < k \leq \lfloor nt \rfloor} Z_{i,j,k}^{(2)}, \quad t \in [0, 1]. \end{aligned}$$

Note that the covariance structure of the collection $\{Z_{i,j}^{(1)} : i, j \in [n], i < j\} \cup \{Z_{i,j,k}^{(2)} : i, j, k \in [n], i < j < k\}$ is the same as the covariance structure of the summands in the formulas (6.1) and (6.2).

6.4. Distance from the pre-limiting process

We provide an estimate of the distance between \mathbf{Y}_n and the pre-limiting piecewise constant Gaussian process.

Theorem 6.2. *Let \mathbf{Y}_n be defined as in Section 6.1 and \mathbf{D}_n be defined as in Section 6.3. Then, for any $g \in M$,*

$$|\mathbb{E}g(\mathbf{Y}_n) - \mathbb{E}g(\mathbf{D}_n)| \leq 23\|g\|_M n^{-1}.$$

Remark 6.3. *Our bound in Theorem 6.2 is of the same order as the analogous bound obtained in [RR10] on the distance between the (finite-dimensional) distributions of $\mathbf{Y}_n(1)$ and $\mathbf{D}_n(1)$.*

The proof is based on Theorem 4.1. In **Step 1** we estimate term ϵ_1 , which involves bounding $\|\Lambda_n\|_2$ of (6.3) and the third moment of $\|\mathbf{Y}_n - \mathbf{Y}'_n\|$. In **Step 2** we treat ϵ_2 , using involved computations, which are, to a large extent, postponed to the appendix. Term ϵ_3 is equal to zero as R_f of Section 6.2 is equal to zero.

Proof of Theorem 6.2. We adopt the notation of sections 6.1, 6.2, 6.3 and apply Theorem 4.1.

Step 1. First note that, for ϵ_1 in Theorem 4.1,

$$|(\mathbf{Y}_n - \mathbf{Y}'_n)\Lambda_n| \leq \|\Lambda_n\|_2 \|\mathbf{Y}_n - \mathbf{Y}'_n\|,$$

where $\|\cdot\|$ denotes the Euclidean norm in \mathbb{R}^2 and $\|\cdot\|_2$ is the induced operator 2-norm. Furthermore,

$$\|\Lambda_n\|_2 \leq \|\Lambda_n\|_F = \frac{n(n-1)}{8} \sqrt{2^2 + (2p)^2 + 0^2 + 1^2} \leq \frac{3n(n-1)}{8},$$

for $\|\cdot\|_F$ denoting the Frobenius norm (which, for $\Theta \in \mathbb{R}^{d_1 \times d_2}$ is defined by $\|\Theta\|_F = \sqrt{\sum_{i=1}^{d_1} \sum_{j=1}^{d_2} |\Theta_{i,j}|^2}$). Therefore:

$$\begin{aligned}
& \mathbb{E} \left[\|(\mathbf{Y}_n - \mathbf{Y}'_n) \Lambda_n\| \|\mathbf{Y}_n - \mathbf{Y}'_n\|^2 \right] \\
& \leq \frac{3n(n-1)}{8} \mathbb{E} \|\mathbf{Y}_n - \mathbf{Y}'_n\|^3 \\
& \leq \frac{3n(n-1)}{8} \mathbb{E} \left[\frac{(n-2)^2}{n^4} (I_{I,J} - I'_{I,J})^2 + \frac{1}{n^4} \left(\sum_{k:k \neq I,J} (I_{I,J} - I'_{I,J}) (I_{J,k} + I_{I,k}) \right)^2 \right]^{3/2} \\
& \leq \frac{3n(n-1)}{8} \left[\frac{(n-2)^2}{n^4} + \frac{(2(n-2))^2}{n^4} \right]^{3/2} \\
& \leq \frac{5}{n},
\end{aligned}$$

where the third inequality follows because $|I_{I,J} - I'_{I,J}| \leq 1$ and $|I_{J,k} + I_{I,k}| \leq 2$ for all k and

$$\epsilon_1 \leq \frac{5\|g\|_M}{6n}. \quad (6.4)$$

Step 2. In order to deal with ϵ_2 in Theorem 4.1, we need to bound

$$\begin{aligned}
& |\mathbb{E} D^2 f(\mathbf{Y}_n) [(\mathbf{Y}_n - \mathbf{Y}'_n) \Lambda_n, \mathbf{Y}_n - \mathbf{Y}'_n] - \mathbb{E} D^2 f(\mathbf{Y}_n) [\mathbf{D}_n, \mathbf{D}_n]| \\
& = \left| \frac{n(n-1)}{8} \mathbb{E} D^2 f(\mathbf{Y}_n) [(2(\mathbf{T}_n - \mathbf{T}'_n), 2p(\mathbf{T}_n - \mathbf{T}'_n) + (\mathbf{V}_n - \mathbf{V}'_n)), (\mathbf{T}_n - \mathbf{T}'_n, \mathbf{V}_n - \mathbf{V}'_n)] \right. \\
& \quad \left. - \mathbb{E} D^2 f(\mathbf{Y}_n) [\mathbf{D}_n, \mathbf{D}_n] \right| \\
& \leq S_1 + S_2 + S_3 + S_4,
\end{aligned} \quad (6.5)$$

where:

$$\begin{aligned}
S_1 &= \left| \frac{n(n-1)}{8} \mathbb{E} D^2 f(\mathbf{Y}_n) [(2(\mathbf{T}_n - \mathbf{T}'_n), 0), (\mathbf{T}_n - \mathbf{T}'_n, 0)] - \mathbb{E} D^2 f(\mathbf{Y}_n) [(\mathbf{D}_n^{(1)}, 0), (\mathbf{D}_n^{(1)}, 0)] \right| \\
S_2 &= \left| \frac{n(n-1)}{8} \mathbb{E} D^2 f(\mathbf{Y}_n) [(0, 2p(\mathbf{T}_n - \mathbf{T}'_n) + \mathbf{V}_n - \mathbf{V}'_n), (\mathbf{T}_n - \mathbf{T}'_n, 0)] \right. \\
& \quad \left. - \mathbb{E} D^2 f(\mathbf{Y}_n) [(0, \mathbf{D}_n^{(2)}), (\mathbf{D}_n^{(1)}, 0)] \right| \\
S_3 &= \left| \frac{n(n-1)}{8} \mathbb{E} D^2 f(\mathbf{Y}_n) [(2(\mathbf{T}_n - \mathbf{T}'_n), 0), (0, \mathbf{V}_n - \mathbf{V}'_n)] - \mathbb{E} D^2 f(\mathbf{Y}_n) [(\mathbf{D}_n^{(1)}, 0), (0, \mathbf{D}_n^{(2)})] \right| \\
S_4 &= \left| \frac{n(n-1)}{8} \mathbb{E} D^2 f(\mathbf{Y}_n) [(0, 2p(\mathbf{T}_n - \mathbf{T}'_n) + \mathbf{V}_n - \mathbf{V}'_n), (0, \mathbf{V}_n - \mathbf{V}'_n)] \right. \\
& \quad \left. - \mathbb{E} D^2 f(\mathbf{Y}_n) [(0, \mathbf{D}_n^{(2)}), (0, \mathbf{D}_n^{(2)})] \right|.
\end{aligned} \quad (6.6)$$

In Lemma 7.2, in the appendix, we obtain the following estimates:

$$S_1 \leq \frac{\sqrt{5}\|g\|_M}{12n}, \quad S_2 \leq \frac{\sqrt{178}\|g\|_M}{6n}, \quad S_3 \leq \frac{\sqrt{178}\|g\|_M}{6n}, \quad S_4 \leq \frac{(\sqrt{612} + \sqrt{178})\|g\|_M}{3n}. \quad (6.7)$$

Note that, therefore, by (6.5) and (6.7),

$$\epsilon_2 = |\mathbb{E} D^2 f(\mathbf{Y}_n) [(\mathbf{Y}_n - \mathbf{Y}'_n) \Lambda_n, \mathbf{Y}_n - \mathbf{Y}'_n] - \mathbb{E} D^2 f(\mathbf{Y}_n) [\mathbf{D}_n, \mathbf{D}_n]| \leq 18\|g\|_M n^{-1}. \quad (6.8)$$

Using Theorem 4.1 together with (6.8) and (6.4) gives the desired result. \square

6.5. Distance from the continuous process

We now study the approximation of \mathbf{Y}_n by a continuous Gaussian process with covariance equal to the limit of the covariance of \mathbf{D}_n . We obtain a bound on the quality of this approximation. This is achieved by applying Theorem 6.2 and by bounding the distance between \mathbf{D}_n and the continuous process via the Brownian modulus of continuity.

Theorem 6.4. *Let \mathbf{Y}_n be defined as in Subsection 6.1 and let $\mathbf{Z} = (\mathbf{Z}^{(1)}, \mathbf{Z}^{(2)})$ be defined by:*

$$\begin{cases} \mathbf{Z}^{(1)}(t) = \frac{\sqrt{p(1-p)}}{\sqrt{2+8p^2}} t \mathbf{B}_1(t^2) + \frac{p\sqrt{2p(1-p)}}{\sqrt{1+4p^2}} t \mathbf{B}_2(t^2), \\ \mathbf{Z}^{(2)}(t) = \frac{p\sqrt{2p(1-p)}}{\sqrt{1+4p^2}} t \mathbf{B}_1(t^2) + \frac{2p^2\sqrt{2p(1-p)}}{\sqrt{1+4p^2}} t \mathbf{B}_2(t^2) \end{cases},$$

where $\mathbf{B}_1, \mathbf{B}_2$ are independent standard Brownian Motions. Then, for any $g \in M$:

$$|\mathbb{E}g(\mathbf{Y}_n) - \mathbb{E}g(\mathbf{Z})| \leq \|g\|_M \left(16422n^{-1/2} \sqrt{\log n} + 138n^{-1/2} \right).$$

Remark 6.5.

Theorem 6.4, together with Proposition 2.2, implies that \mathbf{Y}_n converges to \mathbf{Z} in distribution with respect to the Skorokhod and uniform topologies.

Remark 6.6. Theorem 6.4 can be adapted to situations in which $p = p_n$ varies with n . More precisely, as indicated by the necessary and sufficient conditions for approximate normality of the marginal distributions given in [Ruc88], Theorem 6.4 can be modified to yield a quantitative functional CLT in the case that $n^3 p_n^2 \rightarrow \infty$ and $n^2(1 - p_n) \rightarrow \infty$.

In **Step 1** of the proof of Theorem 6.4, we use i.i.d standard Brownian Motions to construct a process \mathbf{Z}_n having the same distribution as \mathbf{D}_n . In **Step 2** we couple \mathbf{Z}_n and \mathbf{Z} and use the Brownian modulus of continuity to bound moments of the supremum distance between them. In **Step 3** we combine those bounds with the mean value theorem to obtain the desired final estimate.

Proof of Theorem 6.4.

Step 1. Let \mathbf{B}_3 be another standard Brownian Motion, mutually independent with \mathbf{B}_1 and \mathbf{B}_2 . Let $\mathbf{Z}_n = (\mathbf{Z}_n^{(1)}, \mathbf{Z}_n^{(2)})$ be defined by:

$$\begin{aligned} \text{A) } \mathbf{Z}_n^{(1)}(t) &= \frac{(\lfloor nt \rfloor - 2)\sqrt{p(1-p)}}{n^2\sqrt{2+8p^2}} \mathbf{B}_1(\lfloor nt \rfloor(\lfloor nt \rfloor - 1)) + \frac{(\lfloor nt \rfloor - 2)p\sqrt{2p(1-p)}}{n^2\sqrt{1+4p^2}} \mathbf{B}_2(\lfloor nt \rfloor(\lfloor nt \rfloor - 1)); \\ \text{B) } \mathbf{Z}_n^{(2)}(t) &= \frac{(\lfloor nt \rfloor - 2)p\sqrt{2p(1-p)}}{n^2\sqrt{1+4p^2}} \mathbf{B}_1(\lfloor nt \rfloor(\lfloor nt \rfloor - 1)) \\ &\quad + \frac{(\lfloor nt \rfloor - 2)2p^2\sqrt{2p(1-p)}}{n^2\sqrt{1+4p^2}} \mathbf{B}_2(\lfloor nt \rfloor(\lfloor nt \rfloor - 1)) \\ &\quad + \frac{p(1-p)}{n^2\sqrt{2}} \mathbf{B}_3(\lfloor nt \rfloor(\lfloor nt \rfloor - 1)(\lfloor nt \rfloor - 2)). \end{aligned}$$

Now, note that $(\mathbf{D}_n^{(1)}, \mathbf{D}_n^{(2)}) \stackrel{\mathcal{D}}{=} (\mathbf{Z}_n^{(1)}, \mathbf{Z}_n^{(2)})$. To see this, observe that for all $u, t \in [0, 1]$,

$$\begin{aligned} \text{A) } \mathbb{E} \mathbf{D}_n^{(1)}(t) \mathbf{D}_n^{(1)}(u) &= (\lfloor nt \rfloor - 2)(\lfloor nu \rfloor - 2) \lfloor n(t \wedge u) \rfloor (\lfloor n(t \wedge u) \rfloor - 1) \frac{p(1-p)}{2n^4} \\ &= \mathbb{E} \mathbf{Z}_n^{(1)}(t) \mathbf{Z}_n^{(1)}(u); \\ \text{B) } \mathbb{E} \mathbf{D}_n^{(2)}(t) \mathbf{D}_n^{(2)}(u) &= \binom{\lfloor n(t \wedge u) \rfloor}{3} \frac{3p^2(1+2p-3p^2)}{n^4} + \binom{\lfloor n(t \wedge u) \rfloor}{2} [(\lfloor nt \rfloor - 2)(\lfloor nu \rfloor - 2) - (\lfloor n(t \wedge u) \rfloor - 2)] \frac{4p^3(1-p)}{n^4} \end{aligned}$$

$$\begin{aligned}
&= \lfloor n(t \wedge u) \rfloor (\lfloor n(t \wedge u) \rfloor - 1) \frac{4p^3(1-p)(\lfloor nt \rfloor - 2)(\lfloor nu \rfloor - 2) + (\lfloor n(t \wedge u) \rfloor - 2)p^2(1-p)^2}{2n^4} \\
&= \mathbb{E} \mathbf{Z}_n^{(2)}(t) \mathbf{Z}_n^{(2)}(u); \\
\text{C) } & \mathbb{E} \mathbf{D}_n^{(1)}(t) \mathbf{D}_n^{(2)}(u) \\
&= (\lfloor nt \rfloor - 2)(\lfloor nu \rfloor - 2) \lfloor n(t \wedge u) \rfloor (\lfloor n(t \wedge u) \rfloor - 1) \frac{p^2(1-p)}{n^4} \\
&= \mathbb{E} \mathbf{Z}_n^{(1)}(t) \mathbf{Z}_n^{(2)}(u).
\end{aligned} \tag{6.9}$$

Step 2. We now let \mathbf{Z} be constructed as in Theorem 6.4, using the same Brownian Motions $\mathbf{B}_1, \mathbf{B}_2$, as the ones used in the construction of \mathbf{Z}_n . In Lemma 7.3, proved in the appendix, we obtain the following bounds:

$$\begin{aligned}
\mathbb{E} \|\mathbf{Z}_n - \mathbf{Z}\| &\leq \frac{8}{\sqrt{n}} + \frac{39\sqrt{\log n}}{\sqrt{n}} \\
\mathbb{E} \|\mathbf{Z}_n - \mathbf{Z}\|^3 &\leq \frac{49}{n^{3/2}} + \frac{8167(\log n)^{3/2}}{n^{3/2}} \\
\mathbb{E} \|\mathbf{Z}\|^2 &\leq \frac{4}{3}.
\end{aligned} \tag{6.10}$$

Step 3. We note that, by (6.10):

$$\begin{aligned}
|\mathbb{E}g(\mathbf{Z}) - \mathbb{E}g(\mathbf{D}_n)| &\stackrel{\text{MVT}}{\leq} \mathbb{E} \left[\sup_{c \in [0,1]} \|Dg(\mathbf{Z} + c(\mathbf{Z}_n - \mathbf{Z}))\| \|\mathbf{Z} - \mathbf{Z}_n\| \right] \\
&\leq \|g\|_M \mathbb{E} \left[\sup_{c \in [0,1]} (1 + \|\mathbf{Z} + c(\mathbf{Z}_n - \mathbf{Z})\|^2) \|\mathbf{Z} - \mathbf{Z}_n\| \right] \\
&\leq \|g\|_M \mathbb{E} [\|\mathbf{Z} - \mathbf{Z}_n\| + \|\mathbf{Z}\| \|\mathbf{Z} - \mathbf{Z}_n\| + \|\mathbf{Z} - \mathbf{Z}_n\|^2] \\
&\leq \|g\|_M [\mathbb{E} \|\mathbf{Z} - \mathbf{Z}_n\| + 2\mathbb{E} \|\mathbf{Z} - \mathbf{Z}_n\|^3 + 2(\mathbb{E} \|\mathbf{Z}\|^3)^{2/3} (\mathbb{E} \|\mathbf{Z} - \mathbf{Z}_n\|^3)^{1/3}] \\
&\leq \|g\|_M \left(\frac{115}{\sqrt{n}} + \frac{16422\sqrt{\log n}}{\sqrt{n}} \right),
\end{aligned}$$

which, together with Theorem 6.2 gives the desired estimate. \square

Remark 6.7. The representation of \mathbf{Z} in terms of two independent Brownian Motions comes from a careful analysis of the limiting covariance of \mathbf{D}_n , which may be derived using (6.9).

7. Appendix - technical details of the proofs of Proposition 5.6 and Theorems 6.2 and 6.4

7.1. Technical details of the proof of Proposition 5.6

Lemma 7.1. Let $n, d \in \mathbb{N}$ and $r_1 \geq r_2 \geq \dots \geq r_d \geq 1$. Define function $f : (D([0,1], \mathbb{R}))^{r_1 + \dots + r_d} \rightarrow D([0,1], \mathbb{R}^d)$, given by

$$\begin{aligned}
&f(x_{1,1}, \dots, x_{1,r_1}, x_{2,2}, \dots, x_{2,r_2}, \dots, x_{d,1}, \dots, x_{d,r_d}) \\
&= \left(\left(\sum_{j=1}^{r_1} x_{1,j} \left(\left(t + \frac{r_1-1}{n} \right) \wedge 1 \right), \dots, \sum_{j=1}^{r_d} x_{d,j} \left(\left(t + \frac{r_d-1}{n} \right) \wedge 1 \right) \right), t \in [0,1] \right).
\end{aligned}$$

Then, for any $g \in M^0$,

$$\|g \circ f\|_{M^0} \leq \|g\|_{M^0} \sqrt{dr_1} \sum_{i=1}^d r_i.$$

Proof. Note that function f is twice Fréchet differentiable with

$$(A) \quad Df(w) [(x_{1,1}, \dots, x_{1,r_1}, x_{2,1}, \dots, x_{2,r_2}, \dots, x_{d,1}, \dots, x_{d,r_d})] \\ = \left(\left(\sum_{j=1}^{r_1} x_{1,j} \left(\left(t + \frac{r_1-1}{n} \right) \wedge 1 \right) \right), \dots, \sum_{j=1}^{r_d} x_{d,j} \left(\left(t + \frac{r_d-1}{n} \right) \wedge 1 \right) \right), t \in [0, 1] \\ (B) \quad D^2 f(w)[x^{(1)}, x^{(2)}] = 0$$

for all $w, x^{(1)}, x^{(2)}, (x_{1,1}, \dots, x_{1,r_1}, x_{2,1}, \dots, x_{2,r_2}, \dots, x_{d,1}, \dots, x_{d,r_d}) \in (D([0, 1], \mathbb{R}))^{r_1 + \dots + r_d}$. Furthermore, for any $w \in (D([0, 1], \mathbb{R}))^{r_1 + \dots + r_d}$,

$$a) \quad \|f(w)\| \leq \sqrt{\sup_{t \in [0,1]} \left(\sum_{j=1}^{r_1} w_{1,j} \left(\left(t + \frac{r_1-1}{n} \right) \wedge 1 \right) \right)^2 + \dots + \sup_{t \in [0,1]} \left(\sum_{j=1}^{r_d} w_{d,j} \left(\left(t + \frac{r_d-1}{n} \right) \wedge 1 \right) \right)^2} \\ \leq \sqrt{\sum_{i=1}^d \sup_{t \in [0,1]} \left| \sum_{j=1}^{r_i} w_{i,j}(t) \right|^2} \\ b) \quad \|Df(w)\| \leq \sqrt{\sum_{i=1}^d r_i}$$

Therefore, for any $w, h \in (D([0, 1], \mathbb{R}))^{r_1 + \dots + r_d}$,

$$A) \quad |g \circ f(w)| \leq \|g\|_{M^0}; \\ B) \quad \|D(g \circ f)(w)\| = \|Dg(f(w))[Df(w)[\cdot]]\| \leq \|g\|_{M^0} \|Df(w)\| \leq \|g\|_{M^0} \sqrt{\sum_{i=1}^d r_i}; \\ C) \quad \|D^2(g \circ f)(w)\| = \|D^2g(f(w))[Df(w), Df(w)]\| \leq \|g\|_{M^0} \|Df(w)\|^2 \leq \|g\|_{M^0} \sum_{i=1}^d r_i; \\ D) \quad \|D^2(g \circ f)(w+h) - D^2(g \circ f)(w)\| \\ = \|D^2g(f(w+h))[Df(w+h), Df(w+h)] - D^2g(f(w))[Df(w), Df(w)]\| \\ \leq \|D^2g(f(w+h))[Df(w+h), Df(w+h)] - D^2g(f(w))[Df(w+h), Df(w+h)]\| \\ + \|D^2g(f(w))[Df(w+h), Df(w+h)] - D^2g(f(w))[Df(w), Df(w)]\| \\ \leq \|g\|_{M^0} \|f(w+h) - f(w)\| \|Df(w+h)\|^2 \\ \leq \|g\|_{M^0} \left(\sum_{i=1}^d \sup_{t \in [0,1]} \left| \sum_{j=1}^{r_i} h_{i,j}(t) \right|^2 \right)^{1/2} \sum_{i=1}^d r_i, \quad (7.1)$$

where $D)$ follows from the fact that $Df(w) = Df(w+h)$. Moreover,

$$\frac{\left(\sum_{i=1}^d \sup_{t \in [0,1]} \left| \sum_{j=1}^{r_i} h_{i,j}(t) \right|^2 \right)^{1/2}}{\sup_{t \in [0,1]} \left(\sum_{i=1}^d \sum_{j=1}^{r_i} h_{i,j}^2(t) \right)^{1/2}} \leq \frac{\sup_{t \in [0,1]} \left(dr_1 \sum_{i=1}^d \sum_{j=1}^{r_i} h_{i,j}^2(t) \right)^{1/2}}{\sup_{t \in [0,1]} \left(\sum_{i=1}^d \sum_{j=1}^{r_i} h_{i,j}^2(t) \right)^{1/2}} = \sqrt{dr_1}. \quad (7.2)$$

Therefore, using (7.1) and (7.2),

$$\|g \circ f\|_{M^0} \leq \|g\|_{M^0} \sqrt{dr_1} \sum_{i=1}^d r_i.$$

□

7.2. Technical details of the proof of Theorem 6.2

Lemma 7.2. For $S_i, i = 1, 2, 3, 4$ of (6.6), we have the following estimates:

$$S_1 \leq \frac{\sqrt{5}\|g\|_M}{12n}, \quad S_2 \leq \frac{\sqrt{178}\|g\|_M}{6n}, \quad S_3 \leq \frac{\sqrt{178}\|g\|_M}{6n}, \quad S_4 \leq \frac{(\sqrt{612} + \sqrt{178})\|g\|_M}{3n}.$$

Proof. For S_1 , for fixed $i, j \in \{1, \dots, n\}$, let \mathbf{Y}_n^{ij} be equal to \mathbf{Y}_n except for the fact that I_{ij} is replaced by an independent copy, i.e. for all $t \in [0, 1]$ let:

$$\begin{aligned} \mathbf{T}_n^{ij}(t) &= \mathbf{T}_n(t) - \frac{\lfloor nt \rfloor - 2}{n^2} (I_{ij} - I'_{ij}) \mathbb{1}_{[i/n, 1] \cap [j/n, 1]}(t) \\ \mathbf{V}_n^{ij}(t) &= \mathbf{V}_n(t) - \frac{1}{n^2} \sum_{k: k \neq i, j} (I_{ij} - I'_{ij}) (I_{jk} + I_{ik}) \mathbb{1}_{[i/n, 1] \cap [j/n, 1] \cap [k/n, 1]}(t) \end{aligned}$$

and let $\mathbf{Y}_n^{ij}(t) = (\mathbf{T}_n^{ij}(t) - \mathbb{E}\mathbf{T}_n(t), \mathbf{V}_n^{ij}(t) - \mathbb{E}\mathbf{V}_n(t))$.

By noting that the mean zero $Z_{i,k}^{(1)}$ and $Z_{i',j}^{(1)}$ are independent for $i \neq i'$, we obtain:

$$\begin{aligned} S_1 &= \left| \frac{n(n-1)}{8} \mathbb{E} D^2 f(\mathbf{Y}_n) [(\mathbf{T}_n - \mathbf{T}'_n)(2, 0), (\mathbf{T}_n - \mathbf{T}'_n)(1, 0)] \right. \\ &\quad \left. - \mathbb{E} D^2 f(\mathbf{Y}_n) \left[\sum_{1 \leq i < j \leq n} Z_{i,j}^{(1)} (\lfloor n \cdot \rfloor - 2)(1, 0) \mathbb{1}_{[i/n, 1] \cap [j/n, 1]}, \sum_{1 \leq i < j \leq n} Z_{i,j}^{(1)} (\lfloor n \cdot \rfloor - 2)(1, 0) \mathbb{1}_{[i/n, 1] \cap [j/n, 1]} \right] \right| \\ &= \left| \frac{1}{2n^4} \sum_{1 \leq i < j \leq n} \mathbb{E} \{ (I_{i,j} - 2pI_{i,j} + p) \right. \\ &\quad \cdot D^2 f(\mathbf{Y}_n) [(\lfloor n \cdot \rfloor - 2)(1, 0) \mathbb{1}_{[i/n, 1] \cap [j/n, 1]}, (\lfloor n \cdot \rfloor - 2)(1, 0) \mathbb{1}_{[i/n, 1] \cap [j/n, 1]}] \} \\ &\quad \left. - \sum_{1 \leq i < j \leq n} \left\{ \mathbb{E} \left(Z_{i,j}^{(1)} \right)^2 \right. \right. \\ &\quad \left. \cdot \mathbb{E} D^2 f(\mathbf{Y}_n) [(\lfloor n \cdot \rfloor - 2)(1, 0) \mathbb{1}_{[i/n, 1] \cap [j/n, 1]}, (\lfloor n \cdot \rfloor - 2)(1, 0) \mathbb{1}_{[i/n, 1] \cap [j/n, 1]}] \right\} \right| \\ &= \left| \sum_{1 \leq i < j \leq n} \mathbb{E} \left\{ \left(\frac{1}{2n^4} (I_{i,j} - 2pI_{i,j} + p) - \mathbb{E} \left(Z_{i,j}^{(1)} \right)^2 \right) \right. \right. \\ &\quad \left. \cdot D^2 f(\mathbf{Y}_n) [(\lfloor n \cdot \rfloor - 2)(1, 0) \mathbb{1}_{[i/n, 1] \cap [j/n, 1]}, (\lfloor n \cdot \rfloor - 2)(1, 0) \mathbb{1}_{[i/n, 1] \cap [j/n, 1]}] \right\} \right| \\ &= \left| \sum_{1 \leq i < j \leq n} \mathbb{E} \left\{ \frac{1}{2n^4} (I_{i,j} - 2pI_{i,j} + p) \right. \right. \\ &\quad \left. \cdot (D^2 f(\mathbf{Y}_n) - D^2 f(\mathbf{Y}_n^{ij})) [(\lfloor n \cdot \rfloor - 2)(1, 0) \mathbb{1}_{[i/n, 1] \cap [j/n, 1]}, (\lfloor n \cdot \rfloor - 2)(1, 0) \mathbb{1}_{[i/n, 1] \cap [j/n, 1]}] \right\} \right| \\ &\stackrel{(3.4)}{\leq} \frac{\|g\|_M}{6n^2} \sum_{1 \leq i < j \leq n} \mathbb{E} |(I_{i,j} - 2pI_{i,j} + p)| \|\mathbf{Y}_n - \mathbf{Y}_n^{ij}\|, \end{aligned} \tag{7.3}$$

where (7.3) follows from Proposition 3.2. Now,

$$\|\mathbf{Y}_n - \mathbf{Y}_n^{ij}\| \leq \frac{1}{n^2} \sqrt{(\lfloor n \cdot \rfloor - 2)^2 (I_{ij} - I'_{ij})^2 + \left(\sum_{k: k \neq i, j} |I_{ij} - I'_{ij}| (I_{jk} + I_{ik}) \right)^2}$$

and so, by (7.3),

$$\begin{aligned}
S_1 &\leq \frac{\|g\|_M}{6n^4} \sum_{1 \leq i < j \leq n} \mathbb{E} \left\{ |I_{i,j} - 2pI_{i,j} + p| \cdot \sqrt{(n-2)^2(I_{ij} - I'_{ij})^2 + \left(\sum_{k \neq i,j} |I_{ij} - I'_{ij}|(I_{jk} + I_{ik}) \right)^2} \right\} \\
&\leq \frac{\|g\|_M}{6n^3} \sum_{1 \leq i < j \leq n} \mathbb{E} \left\{ |I_{i,j} - 2pI_{i,j} + p| \cdot \sqrt{(I_{ij} - I'_{ij})^2 + (|I_{ij} - I'_{ij}|(I_{jk} + I_{ik}))^2} \right\} \\
&\leq \frac{\sqrt{5}\|g\|_M}{12n}, \tag{7.4}
\end{aligned}$$

where the last inequality holds because $|I_{ij} - 2pI_{ij} + p| \leq 1$, $|I_{ij} - I'_{ij}| \leq 1$ and $I_{jk} + I_{ik} \leq 2$ for all $k \in \{1, \dots, n\}$.

For S_2 , let \mathbf{Y}_n^{ijk} equal to \mathbf{Y}_n except that I_{ij}, I_{jk}, I_{ik} are replaced by $I'_{ij}, I'_{jk}, I'_{ik}$, i.e. for all $t \in [0, 1]$ let

$$\begin{aligned}
\mathbf{T}_n^{ijk}(t) &= \mathbf{T}_n(t) - \frac{\lfloor nt \rfloor - 2}{n^2} [(I_{ij} - I'_{ij}) \mathbb{1}_{[i/n, 1] \cap [j/n, 1]}(t) \\
&\quad + (I_{jk} - I'_{jk}) \mathbb{1}_{[j/n, 1] \cap [k/n, 1]}(t) + (I_{ik} - I'_{ik}) \mathbb{1}_{[i/n, 1] \cap [k/n, 1]}(t)] \\
\mathbf{V}_n^{ijk}(t) &= \mathbf{V}_n(t) - \frac{1}{n^2} \sum_{l: l \neq i, j, k} [(I_{ij} - I'_{ij})(I_{jl} + I_{il}) \mathbb{1}_{[i/n, 1] \cap [j/n, 1] \cap [l/n, 1]}(t) \\
&\quad + (I_{jk} - I'_{jk})(I_{jl} + I_{kl}) \mathbb{1}_{[k/n, 1] \cap [j/n, 1] \cap [l/n, 1]}(t) + (I_{ik} - I'_{ik})(I_{jl} + I_{il}) \mathbb{1}_{[i/n, 1] \cap [k/n, 1] \cap [l/n, 1]}(t)] \\
&\quad - \frac{1}{n^2} [(I_{ij}I_{jk} - I'_{ij}I'_{jk}) + (I_{ij}I_{ik} - I'_{ij}I'_{ik}) + (I_{ik}I_{jk} - I'_{ik}I'_{jk})] \mathbb{1}_{[i/n, 1] \cap [j/n, 1] \cap [k/n, 1]}(t). \tag{7.5}
\end{aligned}$$

Let $\mathbf{Y}_n^{ijk}(t) = (\mathbf{T}_n^{ijk}(t) - \mathbb{E}\mathbf{T}_n(t), \mathbf{V}_n^{ijk}(t) - \mathbb{E}\mathbf{V}_n(t))$ for all $t \in [0, 1]$. Then

$$\begin{aligned}
S_2 &= \left| \frac{n(n-1)}{8} \mathbb{E} \{ D^2 f(\mathbf{Y}_n) [(\mathbf{T}_n - \mathbf{T}'_n)(0, 2p) + (\mathbf{V}_n - \mathbf{V}'_n)(0, 1), (\mathbf{T}_n - \mathbf{T}'_n)(1, 0)] \} \right. \\
&\quad \left. - \mathbb{E} D^2 f(\mathbf{Y}_n) \left[\sum_{1 \leq i < j < k \leq n} Z_{i,j,k}^{(2)}(0, 1) \mathbb{1}_{[i/n, 1] \cap [j/n, 1] \cap [k/n, 1]}, \sum_{1 \leq i < j \leq n} Z_{i,j}^{(1)}(\lfloor n \cdot \rfloor - 2)(1, 0) \mathbb{1}_{[i/n, 1] \cap [j/n, 1]} \right] \right| \\
&= \left| \frac{1}{4n^4} \sum_{1 \leq i < j \leq n} \sum_{\substack{1 \leq k \leq n \\ k \notin \{i, j\}}} \mathbb{E} \{ [2p(I_{ij} - 2pI_{ij} + p) + (I_{ij} - 2pI_{ij} + p)(I_{jk} + I_{ik}) - 8p^2(1-p)] \right. \\
&\quad \left. \cdot (D^2 f(\mathbf{Y}_n) - D^2 f(\mathbf{Y}_n^{ijk})) [(0, 1) \mathbb{1}_{[i/n, 1] \cap [j/n, 1] \cap [k/n, 1]}, (\lfloor n \cdot \rfloor - 2)(1, 0) \mathbb{1}_{[i/n, 1] \cap [j/n, 1]}] \} \right| \\
&\stackrel{(3.4)}{\leq} \frac{\|g\|_M}{12n^3} \sum_{1 \leq i < j \leq n} \sum_{\substack{1 \leq k \leq n \\ k \notin \{i, j\}}} \mathbb{E} \{ |2p(I_{ij} - 2pI_{ij} + p) + (I_{ij} - 2pI_{ij} + p)(I_{jk} + I_{ik})| \|\mathbf{Y}_n - \mathbf{Y}_n^{ijk}\| \} \\
&\leq \frac{\|g\|_M}{3n^3} \sum_{1 \leq i < j \leq n} \sum_{\substack{1 \leq k \leq n \\ k \notin \{i, j\}}} \mathbb{E} \|\mathbf{Y}_n - \mathbf{Y}_n^{ijk}\|. \tag{7.6}
\end{aligned}$$

Now, by (7.5), we note that:

$$\begin{aligned}
&\|\mathbf{Y}_n - \mathbf{Y}_n^{ijk}\| \\
&\leq \frac{1}{n^2} \left\{ (n-2)^2(|I_{ij} - I'_{ij}| + |I_{jk} - I'_{jk}| + |I_{ik} - I'_{ik}|)^2 \right.
\end{aligned}$$

$$\begin{aligned}
& + \left[\sum_{l:l \neq i,j,k} (|I_{ij} - I'_{ij}|(I_{jl} + I_{il}) + |I_{jk} - I'_{jk}|(I_{jl} + I_{kl}) + |I_{ik} - I'_{ik}|(I_{jl} + I_{il}) + |I_{ik} - I'_{ik}|(I_{jl} + I_{il})) \right. \\
& \quad \left. + |I_{ij}I_{jk} - I'_{ij}I'_{jk}| + |I_{ij}I_{ik} - I'_{ij}I'_{ik}| + |I_{ij}I_{jk} - I'_{ij}I'_{jk}| \right]^2 \Bigg\}^{1/2} \\
& \leq \frac{1}{n^2} \sqrt{9(n-2)^2 + (8(n-3) + 3)^2} \\
& = \frac{\sqrt{73n^2 - 372n + 477}}{n^2}, \tag{7.7}
\end{aligned}$$

where the second inequality follows from the fact that for all $a, b, c \in \{1, \dots, n\}$, $|I_{ab} - I'_{ab}| \leq 1$, $(I_{ab} + I_{bc}) \leq 2$ and $|I_{ab}I_{bc} - I'_{ab}I'_{bc}| \leq 1$. Therefore, by (7.6):

$$S_2 \leq \frac{\|g\|_M n(n-1)(n-2)\sqrt{73n^2 - 372n + 477}}{6n^5} \leq \frac{\sqrt{178}\|g\|_M}{6n}. \tag{7.8}$$

Similarly, for S_3 ,

$$\begin{aligned}
S_3 & = \left| \frac{n(n-1)}{8} \mathbb{E} \{ D^2 f(\mathbf{Y}_n) [(\mathbf{T}_n - \mathbf{T}'_n)(2, 0), (\mathbf{V}_n - \mathbf{V}'_n)(0, 1)] \} \right. \\
& \quad \left. - \mathbb{E} D^2 f(\mathbf{Y}_n) \left[\sum_{1 \leq i < j < k \leq n} Z_{i,j,k}^{(2)}(0, 1) \mathbb{1}_{[i/n, 1] \cap [j/n, 1] \cap [k/n, 1]}, \sum_{1 \leq i < j \leq n} Z_{i,j}^{(1)}(\lfloor n \cdot \rfloor - 2)(1, 0) \mathbb{1}_{[i/n, 1] \cap [j/n, 1]} \right] \right| \\
& = \left| \frac{1}{4n^4} \sum_{1 \leq i < j \leq n} \sum_{\substack{1 \leq k \leq n \\ k \notin \{i, j\}}} \mathbb{E} \{ [2(I_{ij} - 2pI_{ij} + p)(I_{jk} + I_{ik}) - 8p^2(1-p)] \right. \\
& \quad \left. \cdot (D^2 f(\mathbf{Y}_n) - D^2 f(\mathbf{Y}_n^{ijk})) [(0, 1) \mathbb{1}_{[i/n, 1] \cap [j/n, 1] \cap [k/n, 1]}, (\lfloor n \cdot \rfloor - 2)(1, 0) \mathbb{1}_{[i/n, 1] \cap [j/n, 1]}] \} \right| \\
& \stackrel{(3.4)}{\leq} \frac{\|g\|_M}{12n^3} \sum_{1 \leq i < j \leq n} \sum_{\substack{1 \leq k \leq n \\ k \notin \{i, j\}}} \mathbb{E} \{ |2(I_{ij} - 2pI_{ij} + p)(I_{jk} + I_{ik})| \|\mathbf{Y}_n - \mathbf{Y}_n^{ijk}\| \} \\
& \leq \frac{\|g\|_M}{3n^3} \sum_{1 \leq i < j \leq n} \sum_{\substack{1 \leq k \leq n \\ k \notin \{i, j\}}} \mathbb{E} \|\mathbf{Y}_n - \mathbf{Y}_n^{ijk}\| \\
& \stackrel{(7.7)}{\leq} \frac{\sqrt{178}\|g\|_M}{6n} \tag{7.9}
\end{aligned}$$

Now, for S_4 , let \mathbf{Y}_n^{ijkl} be equal to \mathbf{Y}_n except that $I_{ij}, I_{ik}, I_{il}, I_{jk}, I_{jl}, I_{kl}$ are replaced with independent copies $I'_{ij}, I'_{ik}, I'_{il}, I'_{jk}, I'_{jl}, I'_{kl}$, i.e. for all $t \in [0, 1]$ let

$$\begin{aligned}
\mathbf{T}_n^{ijkl}(t) & = \mathbf{T}_n(t) - \frac{\lfloor nt \rfloor - 2}{n^2} [(I_{ij} - I'_{ij}) \mathbb{1}_{[i/n, 1] \cap [j/n, 1]}(t) + (I_{ik} - I'_{ik}) \mathbb{1}_{[i/n, 1] \cap [k/n, 1]}(t) \\
& \quad + (I_{il} - I'_{il}) \mathbb{1}_{[i/n, 1] \cap [l/n, 1]}(t) + (I_{jk} - I'_{jk}) \mathbb{1}_{[j/n, 1] \cap [k/n, 1]}(t) \\
& \quad + (I_{jl} - I'_{jl}) \mathbb{1}_{[j/n, 1] \cap [l/n, 1]}(t) + (I_{kl} - I'_{kl}) \mathbb{1}_{[k/n, 1] \cap [l/n, 1]}(t)] \\
\mathbf{V}_n^{ijkl}(t) & = \mathbf{V}_n(t) - \frac{1}{n^2} \sum_{m:m \neq i,j,k,l} [(I_{ij} - I'_{ij})(I_{im} + I_{jm}) \mathbb{1}_{[i/n, 1] \cap [j/n, 1] \cap [m/n, 1]}(t) \\
& \quad + (I_{ik} - I'_{ik})(I_{im} + I_{km}) \mathbb{1}_{[i/n, 1] \cap [k/n, 1] \cap [m/n, 1]}(t)
\end{aligned}$$

$$\begin{aligned}
& + (I_{il} - I'_{il}) (I_{im} + I_{lm}) \mathbb{1}_{[i/n,1] \cap [l/n,1] \cap [m/n,1]}(t) \\
& + (I_{jk} - I'_{jk}) (I_{jm} + I_{km}) \mathbb{1}_{[j/n,1] \cap [k/n,1] \cap [m/n,1]}(t) \\
& + (I_{jl} - I'_{jl}) (I_{jm} + I_{lm}) \mathbb{1}_{[j/n,1] \cap [l/n,1] \cap [m/n,1]}(t) \\
& + (I_{kl} - I'_{kl}) (I_{km} + I_{lm}) \mathbb{1}_{[k/n,1] \cap [l/n,1] \cap [m/n,1]}(t) \\
& - \frac{1}{n^2} [(I_{ij}I_{jk} - I'_{ij}I'_{jk}) + (I_{ij}I_{ik} - I'_{ij}I'_{ik}) + (I_{ik}I_{jk} - I'_{ij}I'_{jk})] \mathbb{1}_{[i/n,1] \cap [j/n,1] \cap [k/n,1]}(t) \\
& - \frac{1}{n^2} [(I_{ij}I_{jl} - I'_{ij}I'_{jl}) + (I_{ij}I_{il} - I'_{ij}I'_{il}) + (I_{il}I_{jl} - I'_{ij}I'_{jl})] \mathbb{1}_{[i/n,1] \cap [j/n,1] \cap [l/n,1]}(t) \\
& - \frac{1}{n^2} [(I_{ik}I_{kl} - I'_{ik}I'_{kl}) + (I_{ik}I_{il} - I'_{ik}I'_{il}) + (I_{il}I_{kl} - I'_{ik}I'_{kl})] \mathbb{1}_{[i/n,1] \cap [k/n,1] \cap [l/n,1]}(t) \\
& - \frac{1}{n^2} [(I_{jk}I_{jl} - I'_{jk}I'_{jl}) + (I_{jl}I_{kl} - I'_{jk}I'_{kl}) + (I_{kl}I_{jk} - I'_{kl}I'_{jk})] \mathbb{1}_{[j/n,1] \cap [k/n,1] \cap [l/n,1]}(t) \quad (7.10)
\end{aligned}$$

and for all $t \in [0, 1]$ let $\mathbf{Y}_n^{ijkl}(t) = (\mathbf{T}_n^{ijkl}(t) - \mathbb{E}\mathbf{T}_n, \mathbf{V}_n^{ijkl}(t) - \mathbb{E}\mathbf{V}_n(t))$. Note that:

$$\begin{aligned}
S_4 &= \left| \frac{n(n-1)}{8} \mathbb{E} \{ D^2 f(\mathbf{Y}_n) [(\mathbf{T}_n - \mathbf{T}'_n)(0, 2p) + (\mathbf{V}_n - \mathbf{V}'_n)(0, 1), (\mathbf{V}_n - \mathbf{V}'_n)(0, 1)] \} \right. \\
&\quad \left. - \mathbb{E} D^2 f(\mathbf{Y}_n) \left[\sum_{1 \leq i < j < k \leq n} Z_{i,j,k}^{(2)}(0, 1) \mathbb{1}_{[i/n,1] \cap [j/n,1] \cap [k/n,1]}, \sum_{1 \leq i < j < k \leq n} Z_{i,j,k}^{(2)}(0, 1) \mathbb{1}_{[i/n,1] \cap [j/n,1] \cap [k/n,1]} \right] \right| \\
&\leq \left| \frac{1}{4n^4} \sum_{1 \leq i < j \leq n} \sum_{\substack{1 \leq k \neq l \leq n \\ \{k,l\} \cap \{i,j\} = \emptyset}} \mathbb{E} \{ [2p(I_{ij} - 2pI_{ij} + p)(I_{jk} + I_{ik}) + (I_{ij} - 2pI_{ij} + p)(I_{ik}I_{il} + I_{ik}I_{jl} + I_{jk}I_{il} + I_{jk}I_{jl}) \right. \right. \\
&\quad \left. \left. - 16p^3(1-p)] \cdot (D^2 f(\mathbf{Y}_n) - D^2 f(\mathbf{Y}_n^{ijkl})) [(0, 1) \mathbb{1}_{[i/n,1] \cap [j/n,1] \cap [l/n,1]}, (0, 1) \mathbb{1}_{[i/n,1] \cap [j/n,1] \cap [k/n,1]}] \right\} \right| \\
&\quad + \left| \frac{1}{4n^4} \sum_{1 \leq i < j \leq n} \sum_{\substack{1 \leq k \leq n \\ k \notin \{i,j\}}} \mathbb{E} \{ [2p(I_{ij} - 2pI_{ij} + p)(I_{jk} + I_{ik}) + (I_{ij} - 2pI_{ij} + p)(I_{ik} + 2I_{ik}I_{jk} + I_{jk}) \right. \right. \\
&\quad \left. \left. - 4p^2(1 + 2p - 3p^2)] \cdot (D^2 f(\mathbf{Y}_n) - D^2 f(\mathbf{Y}_n^{ijk})) [(0, 1) \mathbb{1}_{[i/n,1] \cap [j/n,1] \cap [k/n,1]}, (0, 1) \mathbb{1}_{[i/n,1] \cap [j/n,1] \cap [k/n,1]}] \right\} \right| \\
&\leq \frac{\|g\|_M}{12n^4} \sum_{1 \leq i < j \leq n} \sum_{\substack{1 \leq k \neq l \leq n \\ \{k,l\} \cap \{i,j\} = \emptyset}} \mathbb{E} \{ |(I_{ij} - 2pI_{ij} + p)(2pI_{jk} + 2pI_{ik} + I_{ik}I_{il} + I_{ik}I_{jl} + I_{jk}I_{il} + I_{jk}I_{jl})| \cdot \|\mathbf{Y}_n - \mathbf{Y}_n^{ijkl}\| \} \\
&\quad + \frac{\|g\|_M}{12n^4} \sum_{1 \leq i < j \leq n} \sum_{\substack{1 \leq k \leq n \\ k \notin \{i,j\}}} \mathbb{E} \{ |(I_{ij} - 2pI_{ij} + p)(2pI_{jk} + 2pI_{ik} + I_{ik} + 2I_{ik}I_{jk} + I_{jk})| \cdot \|\mathbf{Y}_n - \mathbf{Y}_n^{ijk}\| \} \\
&\leq \frac{2\|g\|_M}{3n^4} \sum_{1 \leq i < j \leq n} \sum_{\substack{1 \leq k \neq l \leq n \\ \{k,l\} \cap \{i,j\} = \emptyset}} \mathbb{E} \|\mathbf{Y}_n - \mathbf{Y}_n^{ijkl}\| + \frac{2\|g\|_M}{3n^4} \sum_{1 \leq i < j \leq n} \sum_{\substack{1 \leq k \leq n \\ k \notin \{i,j\}}} \mathbb{E} \|\mathbf{Y}_n - \mathbf{Y}_n^{ijk}\|. \quad (7.11)
\end{aligned}$$

Now, by (7.10), note that:

$$\|\mathbf{Y}_n - \mathbf{Y}_n^{ijkl}\|$$

$$\begin{aligned}
&\leq \frac{1}{n^2} \left\{ (n-2)^2 (|I_{ij} - I'_{ij}| + |I_{ik} - I'_{ik}| + |I_{il} - I'_{il}| + |I_{jk} - I'_{jk}| + |I_{jl} - I'_{jl}| + |I_{kl} - I'_{kl}|)^2 \right. \\
&\quad + \left[\sum_{m:m \neq i,j,k,l} [|I_{ij} - I'_{ij}| (I_{im} + I_{jm}) + |I_{ik} - I'_{ik}| (I_{im} + I_{km}) + |I_{il} - I'_{il}| (I_{im} + I_{lm}) \right. \\
&\quad \quad + |I_{jk} - I'_{jk}| (I_{jm} + I_{km}) + |I_{jl} - I'_{jl}| (I_{jm} + I_{lm}) + |I_{kl} - I'_{kl}| (I_{km} + I_{lm})] \\
&\quad + |I_{ij}I_{jk} - I'_{ij}I'_{jk}| + |I_{ij}I_{ik} - I'_{ij}I'_{ik}| + |I_{ik}I_{jk} - I'_{ij}I'_{jk}| + |I_{ij}I_{jl} - I'_{ij}I'_{jl}| \\
&\quad + |I_{ij}I_{il} - I'_{ij}I'_{il}| + |I_{il}I_{jl} - I'_{ij}I'_{jl}| + |I_{ik}I_{kl} - I'_{ik}I'_{kl}| + |I_{ik}I_{il} - I'_{ik}I'_{il}| \\
&\quad \left. \left. + |I_{il}I_{kl} - I'_{ik}I'_{kl}| + |I_{jk}I_{jl} - I'_{jk}I'_{jl}| + |I_{jl}I_{kl} - I'_{jl}I'_{kl}| + |I_{kl}I_{jk} - I'_{kl}I'_{jk}| \right] \right\}^{1/2} \\
&\leq \frac{\sqrt{36(n-2)^2 + (12(n-4) + 12)^2}}{n^2} \\
&= \frac{\sqrt{180n^2 - 1008n + 1440}}{n^2}.
\end{aligned}$$

Therefore, by (7.11) and (7.7),

$$S_4 \leq \frac{\|g\|_M \cdot \sqrt{180n^2 - 1008n + 1440} + \sqrt{73n^2 - 372n + 477}}{3n^2} \leq \frac{(\sqrt{612} + \sqrt{178}) \|g\|_M}{3n}. \quad (7.12)$$

The result now follows by (7.6), (7.8), (7.9), (7.12). \square

7.3. Technical details of the proof of Theorem 6.4

Lemma 7.3. *Using the notation of Step 2 of the proof of Theorem 6.4,*

$$\begin{aligned}
\mathbb{E} \|\mathbf{Z}_n - \mathbf{Z}\| &\leq \frac{8}{\sqrt{n}} + \frac{39\sqrt{\log n}}{\sqrt{n}} \\
\mathbb{E} \|\mathbf{Z}_n - \mathbf{Z}\|^3 &\leq \frac{49}{n^{3/2}} + \frac{8167(\log n)^{3/2}}{n^{3/2}} \\
\mathbb{E} \|\mathbf{Z}\|^2 &\leq \frac{4}{3}.
\end{aligned}$$

Proof. Note the following:

1. By Doob's L^2 and L^3 inequalities,

$$\begin{aligned}
\text{A) } \mathbb{E} \left[\sup_{t \in [0,1]} \left| \mathbf{B}_3 \left(\frac{\lfloor nt \rfloor (\lfloor nt \rfloor - 1) (\lfloor nt \rfloor - 2)}{n^3} \right) \right| \right] &\leq 2 \sqrt{\mathbb{E} \left[\left| \mathbf{B}_3 \left(\frac{n(n-1)(n-2)}{n^3} \right) \right|^2 \right]} \leq 2; \\
\text{B) } \mathbb{E} \left[\sup_{t \in [0,1]} \left| \mathbf{B}_3 \left(\frac{\lfloor nt \rfloor (\lfloor nt \rfloor - 1) (\lfloor nt \rfloor - 2)}{n^3} \right) \right|^3 \right] &\leq \frac{27}{8} \mathbb{E} \left[\left| \mathbf{B}_3 \left(\frac{n(n-1)(n-2)}{n^3} \right) \right|^3 \right] \leq \frac{27}{8}.
\end{aligned} \quad (7.13)$$

2. By Doob's L^2 and L^3 inequality,

$$\mathbb{E} \left[\sup_{t \in [0,1]} |\mathbf{B}_1(t^2)| \right] \leq 2, \quad \mathbb{E} \left[\sup_{t \in [0,1]} |\mathbf{B}_1(t^2)|^3 \right] \leq \frac{27}{8} \quad \text{and} \quad \left| \frac{\lfloor nt \rfloor - 2}{n} - t \right| \leq \frac{3}{n} \quad \text{for all } t \in [0,1]. \quad (7.14)$$

3. Using [FN10, Lemma 3] and the fact that

$$\left| \frac{\lfloor nt \rfloor (\lfloor nt \rfloor - 1)}{n^2} - t^2 \right| \leq \left| \frac{(nt - \lfloor nt \rfloor)(nt + \lfloor nt \rfloor)}{n^2} \right| + \frac{1}{n^2} \leq \frac{3}{n},$$

we obtain

$$\begin{aligned} \mathbb{E} \left[\sup_{t \in [0,1]} \left| \mathbf{B}_1 \left(\frac{\lfloor nt \rfloor (\lfloor nt \rfloor - 1)}{n^2} \right) - \mathbf{B}_1(t^2) \right| \right] &\leq \frac{30\sqrt{3 \log(\frac{2n}{3})}}{n^{1/2} \sqrt{\pi \log(2)}}; \\ \mathbb{E} \left[\sup_{t \in [0,1]} \left| \mathbf{B}_1 \left(\frac{\lfloor nt \rfloor (\lfloor nt \rfloor - 1)}{n^2} \right) - \mathbf{B}_1(t^2) \right|^3 \right] &\leq \frac{1080 (3 \log(\frac{2n}{3}))^{3/2}}{n^{3/2} (\pi \log(2))^{3/2}} \end{aligned} \quad (7.15)$$

Now, we can bound $\mathbb{E} \|\mathbf{Z}_n - \mathbf{Z}\|$ in the following way:

$$\begin{aligned} &\mathbb{E} \|\mathbf{Z}_n - \mathbf{Z}\| \\ &\leq \frac{\sqrt{p(1-p)}}{\sqrt{2+8p^2}} \mathbb{E} \left[\sup_{t \in [0,1]} \left| \frac{\lfloor nt \rfloor - 2}{n} \mathbf{B}_1 \left(\frac{\lfloor nt \rfloor (\lfloor nt \rfloor - 1)}{n^2} \right) - t \mathbf{B}_1(t^2) \right| \right] \\ &\quad + \frac{p\sqrt{2p(1-p)}}{\sqrt{1+4p^2}} \mathbb{E} \left[\sup_{t \in [0,1]} \left| \frac{\lfloor nt \rfloor - 2}{n} \mathbf{B}_2 \left(\frac{\lfloor nt \rfloor (\lfloor nt \rfloor - 1)}{n^2} \right) - t \mathbf{B}_2(t^2) \right| \right] \\ &\quad + \frac{p\sqrt{2p(1-p)}}{\sqrt{1+4p^2}} \mathbb{E} \left[\sup_{t \in [0,1]} \left| \frac{\lfloor nt \rfloor - 2}{n} \mathbf{B}_1 \left(\frac{\lfloor nt \rfloor (\lfloor nt \rfloor - 1)}{n^2} \right) - t \mathbf{B}_1(t^2) \right| \right] \\ &\quad + \frac{2p^2\sqrt{2p(1-p)}}{\sqrt{1+4p^2}} \mathbb{E} \left[\sup_{t \in [0,1]} \left| \frac{\lfloor nt \rfloor - 2}{n} \mathbf{B}_2 \left(\frac{\lfloor nt \rfloor (\lfloor nt \rfloor - 1)}{n^2} \right) - t \mathbf{B}_2(t^2) \right| \right] \\ &\quad + \frac{p(1-p)}{n^{1/2}} \mathbb{E} \left[\sup_{t \in [0,1]} \left| \mathbf{B}_3 \left(\frac{\lfloor nt \rfloor (\lfloor nt \rfloor - 1)(\lfloor nt \rfloor - 2)}{n^3} \right) \right| \right] \\ &\stackrel{(7.13)}{\leq} \frac{(1+4p+4p^2)\sqrt{p(1-p)}}{\sqrt{2+8p^2}} \mathbb{E} \left[\sup_{t \in [0,1]} \left| \frac{\lfloor nt \rfloor - 2}{n} \mathbf{B}_1 \left(\frac{\lfloor nt \rfloor (\lfloor nt \rfloor - 1)}{n^2} \right) - t \mathbf{B}_1(t^2) \right| \right] + \frac{2p(1-p)}{n^{1/2}} \\ &\leq \frac{(1+4p+4p^2)\sqrt{p(1-p)}}{\sqrt{2+8p^2}} \left(\mathbb{E} \left[\sup_{t \in [0,1]} \left| \left(\frac{\lfloor nt \rfloor - 2}{n} - t \right) \mathbf{B}_1(t^2) \right| \right] \right. \\ &\quad \left. + \mathbb{E} \left[\sup_{t \in [0,1]} \left| \mathbf{B}_1 \left(\frac{\lfloor nt \rfloor (\lfloor nt \rfloor - 1)}{n^2} \right) - \mathbf{B}_1(t^2) \right| \right] \right) + \frac{2p(1-p)}{n^{1/2}} \\ &\stackrel{(7.14), (7.15)}{\leq} \frac{(1+4p+4p^2)\sqrt{p(1-p)}}{\sqrt{2+8p^2}} \left(\frac{6}{n} + \frac{30\sqrt{3 \log n}}{n^{1/2} \sqrt{\pi \log(2)}} \right) + \frac{2p(1-p)}{n^{1/2}} \\ &\leq \frac{8}{\sqrt{n}} + \frac{39\sqrt{\log n}}{\sqrt{n}}. \end{aligned}$$

Similarly,

$$\begin{aligned} &\mathbb{E} \|\mathbf{Z}_n - \mathbf{Z}\|^3 \\ &\leq 4\sqrt{3} \left(\frac{p(1-p)}{2+8p^2} \right)^{3/2} \mathbb{E} \left[\sup_{t \in [0,1]} \left| \frac{\lfloor nt \rfloor - 2}{n} \mathbf{B}_1 \left(\frac{\lfloor nt \rfloor (\lfloor nt \rfloor - 1)}{n^2} \right) - t \mathbf{B}_1(t^2) \right|^3 \right] \\ &\quad + 4\sqrt{3} \left(\frac{2p^3(1-p)}{1+4p^2} \right)^{3/2} \mathbb{E} \left[\sup_{t \in [0,1]} \left| \frac{\lfloor nt \rfloor - 2}{n} \mathbf{B}_2 \left(\frac{\lfloor nt \rfloor (\lfloor nt \rfloor - 1)}{n^2} \right) - t \mathbf{B}_2(t^2) \right|^3 \right] \\ &\quad + 9\sqrt{3} \left(\frac{2p^3(1-p)}{1+4p^2} \right)^{3/2} \mathbb{E} \left[\sup_{t \in [0,1]} \left| \frac{\lfloor nt \rfloor - 2}{n} \mathbf{B}_1 \left(\frac{\lfloor nt \rfloor (\lfloor nt \rfloor - 1)}{n^2} \right) - t \mathbf{B}_1(t^2) \right|^3 \right] \end{aligned}$$

$$\begin{aligned}
& + 9\sqrt{3} \left(\frac{8p^5(1-p)}{1+4p^2} \right)^{3/2} \mathbb{E} \left[\sup_{t \in [0,1]} \left| \frac{\lfloor nt \rfloor - 2}{n} \mathbf{B}_2 \left(\frac{\lfloor nt \rfloor (\lfloor nt \rfloor - 1)}{n^2} \right) - t \mathbf{B}_2(t^2) \right|^3 \right] \\
& + 9\sqrt{3} \frac{p^3(1-p)^3}{n^{3/2}} \mathbb{E} \left[\sup_{t \in [0,1]} \left| \mathbf{B}_3 \left(\frac{\lfloor nt \rfloor (\lfloor nt \rfloor - 1)(\lfloor nt \rfloor - 2)}{n^2} \right) \right|^3 \right] \\
& \stackrel{(7.13)}{\leq} \frac{\sqrt{6}p^{3/2}(1-p)^{3/2}(1+26p^3+126p^6)}{(1+4p^2)^{3/2}} \mathbb{E} \left[\sup_{t \in [0,1]} \left| \frac{\lfloor nt \rfloor - 2}{n} \mathbf{B}_1 \left(\frac{\lfloor nt \rfloor (\lfloor nt \rfloor - 1)}{n^2} \right) - t \mathbf{B}_1(t^2) \right|^3 \right] \\
& \quad + \frac{243\sqrt{3}}{512n^{3/2}} \\
& \leq \frac{4\sqrt{6}p^{3/2}(1-p)^{3/2}(1+26p^3+126p^6)}{(1+4p^2)^{3/2}} \left(\mathbb{E} \left[\sup_{t \in [0,1]} \left| \left(\frac{\lfloor nt \rfloor - 2}{n} - t \right) \mathbf{B}_1(t^2) \right|^3 \right] \right. \\
& \quad \left. + \mathbb{E} \left[\sup_{t \in [0,1]} \left| \mathbf{B}_1 \left(\frac{\lfloor nt \rfloor (\lfloor nt \rfloor - 1)}{n^2} \right) - \mathbf{B}_1(t^2) \right|^3 \right] \right) + \frac{243\sqrt{3}}{512n^{3/2}} \\
& \stackrel{(7.14), (7.15)}{\leq} \frac{4\sqrt{6}p^{3/2}(1-p)^{3/2}(1+26p^3+126p^6)}{(1+4p^2)^{3/2}} \left(\frac{81}{8n^3} + \frac{1080(3 \log n)^{3/2}}{n^{3/2}(\pi \log(2))^{3/2}} \right) + \frac{243\sqrt{3}}{512n^{3/2}} \\
& \leq \frac{49}{n^{3/2}} + \frac{8167(\log n)^{3/2}}{n^{3/2}}.
\end{aligned}$$

Furthermore,

$$\begin{aligned}
\mathbb{E} \|\mathbf{Z}\|^3 & \leq \sqrt{2} \mathbb{E} \left[\sup_{t \in [0,1]} \left(\frac{\sqrt{p(1-p)}}{\sqrt{2+8p^2}} t \mathbf{B}_1(t^2) + \frac{p\sqrt{2p(1-p)}}{\sqrt{1+4p^2}} t \mathbf{B}_2(t^2) \right)^3 \right] \\
& \quad + \sqrt{2} \mathbb{E} \left[\sup_{t \in [0,1]} \left(\frac{p\sqrt{2p(1-p)}}{\sqrt{1+4p^2}} t \mathbf{B}_1(t^2) + \frac{2p^2\sqrt{2p(1-p)}}{\sqrt{1+4p^2}} t \mathbf{B}_2(t^2) \right)^3 \right] \\
& \leq \frac{2p^{3/2}(1-p)^{3/2}(1+2^{7/2}p^3+2^{11/2}p^6)}{(1+4p^2)^{3/2}} \mathbb{E} \left[\sup_{t \in [0,1]} |\mathbf{B}_1(t^2)|^3 \right] \\
& \leq \frac{27p^{3/2}(1-p)^{3/2}(1+2^{7/2}p^3+2^{11/2}p^6)}{4(1+4p^2)^{3/2}} \leq \frac{4}{3}.
\end{aligned}$$

This finishes the proof. \square

Acknowledgements

The authors would like to thank Gesine Reinert, Giovanni Peccati and Alison Etheridge for helpful discussions and comments on the early versions of this work.

Mikołaj Kasprzak was supported by the EPSRC doctoral studentship at Oxford University and the **FNR grant FoRGES (R-AGR- 3376-10)** at Luxembourg University.

References

- [Bar88] A.D. Barbour. Stein's method and Poisson process convergence. *Journal of Applied Probability*, 25:175–184, 1988.
- [Bar90] A.D. Barbour. Stein's Method for Diffusion Approximation. *Probability Theory and Related Fields*, 84:297–322, 1990.
- [BC19] S. Bourguin and S. Campese. Approximation of Hilbert-valued Gaussian measures on Dirichlet structures. arXiv:1905.05127, 2019.
- [BDM18] E. Besançon, L. Decreusefond, and P. Moyal. Stein's method for diffusive limit of Markov processes. arXiv:1805.01691, 2018.

- [BHJ92] A.D. Barbour, Lars Holst, and Svante Janson. *Poisson Approximation*. Oxford Studies in Probability. Clarendon Press, 1992.
- [BJ09] A.D. Barbour and S. Janson. A functional combinatorial central limit theorem. *Electronic Journal of Probability*, 14(81):2352–2370, 2009.
- [BK78] B. M. Brown and D. G. Kildea. Reduced U -statistics and the Hodges-Lehmann estimator. *Ann. Statist.*, 6(4):828–835, 1978.
- [Blo76] G. Blom. Some properties of incomplete U -statistics. *Biometrika*, 63(3):573–580, 1976.
- [CD13] L. Coutin and L. Decreusefond. Stein's method for Brownian Approximations. *Communications on Stochastic Analysis*, 7(3):349–372, 2013.
- [CD14] L. Coutin and L. Decreusefond. Higher order expansions via Stein's method. *Communications on Stochastic Analysis*, 8(2):155–168, 2014.
- [CD19] L. Coutin and L. Decreusefond. Stein's method for rough paths. *Potential Analysis*, 2019. in press.
- [CDM05] S. Chatterjee, P. Diaconis, and E. Meckes. Exchangeable pairs and poisson approximation. *Probab. Surveys*, 2:64–106, 2005.
- [CFR11] S. Chatterjee, J. Fulman, and A. Röllin. Exponential approximation by stein's method and spectral graph theory. *ALEA Lat. Am. J. Probab. Math. Stat*, 2011.
- [CGS11] L.H.Y Chen, L. Goldstein, and Q.-M. Shao. *Normal Approximation by Stein's Method*. Probability and Its Applications. Springer Verlag, 2011.
- [CM08] S. Chatterjee and E. Meckes. Multivariate normal approximation using exchangeable pairs. *ALEA Lat. Am. J. Probab. Math. Stat.*, 4:257–283, 2008.
- [DKP19] C. Döbler, M. Kasprzak, and G. Peccati. Functional convergence of U -processes with size-dependent kernels. [arXiv:1912.02705](https://arxiv.org/abs/1912.02705), 2019.
- [Döb15] C. Döbler. Stein's method of exchangeable pairs for the beta distribution and generalizations. *Electron. J. Probab.*, 20:34 pp., 2015.
- [DP17] C. Döbler and G. Peccati. Quantitative de Jong theorems in any dimension. *Electron. J. Probab.*, 22:no. 2, 1–35, 2017.
- [DP18] C. Döbler and G. Peccati. The gamma Stein equation and noncentral de Jong theorems. *Bernoulli*, 24(4B):3384–3421, 2018.
- [DP19] C. Döbler and G. Peccati. Quantitative CLTs for symmetric U -statistics using contractions. *Electron. J. Probab.*, 24:Paper No. 5, 43, 2019.
- [FN10] M. Fischer and G. Nappo. On the Moments of the Modulus of Continuity of Ito Processes. *Stochastic Analysis and Applications*, 28(1):103–122, 2010.
- [Göt91] F. Götze. On the rate of convergence in the multivariate CLT. *The Annals of Probability*, 19(2):724–739, 1991.
- [Jan84] S. Janson. The asymptotic distributions of incomplete U -statistics. *Z. Wahrsch. Verw. Gebiete*, 66(4):495–505, 1984.
- [JN91] S. Janson and K. Nowicki. The asymptotic distributions of generalized U -statistics with applications to random graphs. *Probability Theory and Related Fields*, 90(3):341–375, Sep 1991.
- [Kas17] M.J. Kasprzak. Diffusion approximations via Stein's method and time changes. [arXiv:1701.07633](https://arxiv.org/abs/1701.07633), 2017.
- [Kas20a] M.J. Kasprzak. Functional approximations via Stein's method of exchangeable pairs. *Ann. Inst. H. Poincaré Probab. Statist.*, 2020. in press, [arxiv:1710.09263](https://arxiv.org/abs/1710.09263).
- [Kas20b] M.J. Kasprzak. Stein's method for multivariate Brownian approximations of sums under dependence. *Stochastic Processes and their Applications*, 2020. in press, [arXiv:1708.02521](https://arxiv.org/abs/1708.02521).
- [KDV17] M.J. Kasprzak, A. B. Duncan, and S.J. Vollmer. Note on A. Barbour's paper on Stein's method for diffusion approximations. *Electron. Commun. Probab.*, 22(23):1–8, 2017.
- [LRS17] C. Ley, G. Reinert, and Y. Swan. Stein's method for comparison of univariate distributions. *Probability Surveys*, 14:1–52, 2017.
- [Maj94] P. Major. Asymptotic distributions for weighted U -statistics. *Ann. Probab.*, 22(3):1514–1535, 1994.
- [Mec09] E. Meckes. *On Stein's method for multivariate normal approximation*, volume Volume 5 of *Collections*, pages 153–178. Institute of Mathematical Statistics, Beachwood, Ohio, USA, 2009.
- [NP12] I. Nourdin and G. Peccati. *Normal Approximations with Malliavin Calculus*. Cambridge tracts in Mathematics. Cambridge University Press, 2012.

- [OR93] K. A. O'Neil and R. A. Redner. Asymptotic distributions of weighted U -statistics of degree 2. *Ann. Probab.*, 21(2):1159–1169, 1993.
- [Röl07] A. Röllin. Translated poisson approximation using exchangeable pair couplings. *The Annals of Applied Probability*, 17(5/6):1596–1614, 2007.
- [Ros11] N. Ross. Fundamentals of Stein's Method. *Probability Surveys*, 8:210–293, 2011.
- [RR97a] Y. Rinott and 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):1080–1105, 11 1997.
- [RR97b] Y. Rinott and 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):1080–1105, 1997.
- [RR09] G. Reinert and A. Röllin. Multivariate normal approximation with Stein's method of exchangeable pairs under a general linearity condition. *The Annals of Probability*, 37(6):2150–2173, 2009.
- [RR10] G. Reinert and A. Röllin. Random subgraph counts and u -statistics: Multivariate normal approximation via exchangeable pairs and embedding. *Journal of Applied Probability*, 47(2):378–393, 2010.
- [Ruc88] A. Ruciński. When are small subgraphs of a random graph normally distributed? *Probab. Theory Related Fields*, 78(1):1–10, 1988.
- [Shi11] H.-H. Shih. On Stein's method for infinite-dimensional Gaussian approximation in abstract Wiener spaces. *Journal of Functional Analysis*, 261(5):1236 – 1283, 2011.
- [Ste72] Ch. Stein. A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. *Proc. Sixth Berkeley Symp. on Math. Statist. and Prob.*, 2:583–602, 1972.
- [Ste86] Ch. Stein. *Approximate Computation of Expectations*. Institute of Mathematical Statistics Lecture Notes, Monograph Series, 7. Hayward, Calif. Institute of Mathematical Statistics, 1986.
- [Swa16] Y. Swan. A gateway to Stein's Method. <https://sites.google.com/site/steinsmethod/home>, 2016. Accessed on 19/05/2016.