

Functional approximations via Stein’s method of exchangeable pairs

Mikołaj J. Kasprzak,

*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: We combine the method of exchangeable pairs with Stein’s method for functional approximation. As a result, we give a general linearity condition under which an abstract Gaussian approximation theorem for stochastic processes holds. We apply this approach to estimate the distance of a sum of random variables, chosen from an array according to a random permutation, from a Gaussian mixture process. This result lets us prove a functional combinatorial central limit theorem. We also consider a graph-valued process and bound the speed of convergence of the distribution of its rescaled edge counts to a continuous Gaussian process.

Résumé: Nous combinons la méthode des paires échangeables avec la méthode d’approximation fonctionnelle de Stein. De cette façon, nous obtenons une condition générale de linéarité sous laquelle un résultat abstrait d’approximation Gaussienne est valide. Nous appliquons cette approche à l’estimation de la distance entre une somme de variables aléatoires, choisies dans un tableau par le biais d’une permutation aléatoire, et un mélange de processus Gaussiens. À partir de ce résultat, nous prouvons un théorème central limite fonctionnel combinatoire. Nous considérons également un graphe aléatoire et fournissons des bornes pour la vitesse de convergence de la loi de son nombre d’arêtes (après un changement d’échelle) vers un processus Gaussien continu.

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

Keywords and phrases: Stein’s method, functional convergence, exchangeable pairs, stochastic processes.

1. Introduction

In [32] Stein observed that a random variable Z has the standard normal law if and only if

$$\mathbb{E}Zf(Z) = \mathbb{E}f'(Z) \tag{1.1}$$

for all smooth functions f . Therefore, if, for a random variable W with mean zero and variance 1,

$$|\mathbb{E}f'(W) - \mathbb{E}Wf(W)| \tag{1.2}$$

is close to zero for a large class of functions f , then the law of W should be approximately Gaussian. In [33], Stein combined this observation with his *exchangeable-pair* approach. Therein, for a centred and scaled random variable W , its copy W' is constructed in such a way that (W, W') forms an exchangeable pair and the linear regression condition:

$$\mathbb{E}[W' - W|W] = -\lambda W \tag{1.3}$$

is satisfied for some $\lambda > 0$. This, in many cases, simplifies the process of obtaining bounds on the distance of W from the normal distribution.

This approach was extended in [29] to examples in which an approximate linear regression condition holds:

$$\mathbb{E}[W' - W|W] = -\lambda W + R$$

for some remainder R . A multivariate version of the method was first described in [9] and then in [27]. In [27], for an exchangeable pair of d -dimensional vectors (W, W') the following condition is used:

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

for some invertible matrix Λ and a remainder term R . The approach of [27] was further reinterpreted and combined with the approach of [9] in [24].

On the other hand, in the seminal paper [1], Barbour addressed the problem of providing bounds on the rate of convergence in functional limit results (or invariance principles as they are often called in the literature). He observed that Stein's logic of [32] may also be used in the setup of the Functional Central Limit Theorem. He found a condition, similar to (1.1), characterising the distribution of a standard real Wiener process. Combined with Taylor's theorem, it allowed Barbour to obtain a bound on the rate of convergence in the celebrated Donsker's invariance principle.

This paper is an attempt to combine the method of exchangeable pairs with functional approximations. We provide a novel approach to bounding distances of stochastic processes from Gaussian processes. Our approach is influenced by the setup of [29] and [1].

1.1. Motivation

We are motivated by a number of (finite-dimensional) examples studied in Stein's method literature using exchangeable pairs, which could be extended to the functional setting. Functional limit results play an important role in applied fields. Researchers often choose to model discrete phenomena with continuous processes arising as scaling limits of discrete ones. The reason is that those scaling limits may be studied using stochastic analysis and are more robust to changes in local details. Questions about the rate of convergence in functional limit results are equivalent to ones about the error those researchers make. Obtaining bounds on a certain distance between the scaled discrete and the limiting continuous processes provides a way of quantifying this error.

We consider two main examples. The first one is a combinatorial functional central limit theorem. The second one considers a process representing edge counts in a graph-valued process created by unveiling subsequent vertices of a Bernoulli random graph as time progresses.

The former is a functional version of the result proved qualitatively in [19] and quantitatively in [10] and an extension of the main result of [2]. It considers an array $\{X_{i,j} : i, j = 1, \dots, n\}$ of i.i.d. random variables, which are then used to create a stochastic process:

$$t \mapsto \frac{1}{s_n} \sum_{i=1}^{\lfloor nt \rfloor} X_{i\pi(i)}, \quad (1.5)$$

where s_n^2 is the variance of $\sum_{i=1}^n X_{i\pi(i)}$ and π is a uniform random permutation on $\{1, \dots, n\}$. The motivation for studying this and similar topics comes from permutation tests in non-parametric statistics. Similar setups, yet with a deterministic array of numbers, and in a finite-dimensional context, have also been considered by other authors (see [35] for one of the first works on this topic and [5], [18], [25] for quantitative results).

The second example, which considers Bernoulli random graphs, goes back to [20]. It was first studied using exchangeable pairs in a finite-dimensional context in [28], where a random vector whose components represent statistics corresponding to the number of edges, two-stars and triangles is studied. The authors bound its distance from a normal distribution. We consider a functional analogue of this result, concentrating, for simplicity, only on the number of edges. Our approach can, however, be also extended to encompass the number of two-stars and triangles using a multivariate functional exchangeable-pair methodology which we leave for future work. All of those statistics are often of interest in applications, for example, when approximating the clustering coefficient of a network or in conditional 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), providing a bound on the distance between a stochastic process \mathbf{Y}_n valued in \mathbb{R} and a Gaussian mixture process. The theorem assumes that the process \mathbf{Y}_n satisfies the linear regression condition

$$\mathbb{E} \{Df(\mathbf{Y}_n) [\mathbf{Y}'_n - \mathbf{Y}_n] | \mathbf{Y}_n\} = -\lambda_n Df(\mathbf{Y}_n) [\mathbf{Y}_n] + R_f,$$

for all functions f in a certain class of test functions, some $\lambda_n > 0$ and some random variable $R_f = R_f(\mathbf{Y}_n)$. Crucially, the bound in Theorem 4.1 is derived with respect to a class of test functions so rich that the bound approaching zero fast enough, under certain assumptions, implies the law of \mathbf{Y}_n converging weakly in the Skorokhod and uniform topologies on the Skorokhod space. The exact conditions under which this convergence occurs are stated in Proposition 2.2. Theorem 4.1 is used in the derivation of the remaining results of this paper.

- (b) A novel functional combinatorial central limit theorem. In Theorem 5.1, we establish a bound on the distance between process (1.5) and a Gaussian mixture, piecewise constant process. Furthermore, a qualitative result showing convergence in distribution of process (1.5) to a continuous Gaussian limiting process is provided in Theorem 5.5. Thus, we extend [2], where similar results were proved under the assumption that all the $X_{i,j}$'s for $i, j = 1, \dots, n$ are deterministic. Our bound is also an extension of [10], where a bound on the rate of weak convergence of the law of $\frac{1}{s_n} \sum_{i=1}^n X_{i\pi(i)}$ to the standard normal distribution is obtained.
- (c) A novel functional limit theorem for a statistic corresponding to edge counts in a Bernoulli random graph, together with a bound on the rate of convergence. We consider a Bernoulli random graph $G(n, p)$ on n vertices with edge probabilities p . Letting $I_{i,j}$, for $i, j = 1, \dots, n$ be the indicator that edge (i, j) is present in the graph, we study a scaled statistic representing the number of edges:

$$\mathbf{T}_n(t) = \frac{\lfloor nt \rfloor - 2}{2n^2} \sum_{i,j=1}^{\lfloor nt \rfloor} I_{i,j}, \quad t \in [0, 1].$$

Theorem 6.3 provides a bound on the distance between the law of the process

$$t \mapsto \mathbf{T}_n(t) - \mathbb{E}\mathbf{T}_n(t), \quad t \in [0, 1] \tag{1.6}$$

and the law of a piecewise constant Gaussian process. Theorem 6.4 bounds a distance between the law of (1.6) and the distribution of a continuous Gaussian process. Weak convergence of the law of (1.6) in the Skorokhod and uniform topologies on the Skorokhod space to that of the continuous Gaussian process follows from the bound as a corollary.

1.3. Stein's method of exchangeable pairs

The idea behind the exchangeable-pair approach of [33] was the following. In order to obtain a bound on a distance between the distribution of a centred and scaled random variable W and the standard normal law, one can bound (1.2) for functions f coming from a suitable class. Supposing that we can construct a W' such that (W, W') is an exchangeable pair and (1.3) is satisfied, we can write

$$\begin{aligned} 0 &= \mathbb{E}[(f(W) + f(W'))(W - W')] \\ &= \mathbb{E}[(f(W') - f(W))(W - W')] + 2\mathbb{E}[f(W)\mathbb{E}[W - W'|W]] \\ &= \mathbb{E}[(f(W') - f(W))(W - W')] + 2\lambda\mathbb{E}[Wf(W)]. \end{aligned}$$

It follows that

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

Therefore, using Taylor's theorem,

$$\begin{aligned} &|\mathbb{E}[f'(W)] - \mathbb{E}[Wf(W)]| \\ &= \left| \mathbb{E}[f'(W)] + \frac{1}{2\lambda} \mathbb{E}[(f(W') - f(W))(W - W')] \right| \\ &\leq \left| \mathbb{E}[f'(W)] - \frac{1}{2\lambda} \mathbb{E}[f'(W)(W - W')^2] \right| + \frac{\|f''\|_\infty}{2\lambda} \mathbb{E}|W - W'|^3 \\ &\leq \|f'\|_\infty \mathbb{E} \left| \mathbb{E} \left[\frac{1}{2\lambda} (W - W')^2 | W \right] - 1 \right| + \frac{\|f''\|_\infty}{2\lambda} \mathbb{E}|W - W'|^3 \\ &\leq \frac{\|f'\|_\infty}{2\lambda} \sqrt{\text{Var}[\mathbb{E}[(W - W')^2 | W]]} + \frac{\|f''\|_\infty}{2\lambda} \mathbb{E}|W - W'|^3, \end{aligned}$$

which provides the desired bound.

Before the publication of [9, 27, 24] the method was restricted to one-dimensional approximations. It was, however, also used in the context of non-normal approximations (e.g [7, 8, 30]). More recently several authors have extended and applied the method. Döbler extended it to Beta distribution in [14] and Chen and Fang used it for the combinatorial CLT [10].

1.4. Stein's method in its generality

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

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

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

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

is solved. Finally, using properties of the solution and various mathematical tools (among which the most popular are Taylor's expansions in the continuous case, Malliavin calculus, as described in [26], and coupling methods), a bound is sought for the quantity $|\mathbb{E}_{\nu_n} \mathcal{A}f_h|$.

1.5. Functional Stein's method

Approximations by laws of stochastic processes have not been covered in the Stein's method literature very widely, with the notable exceptions of [1, 2, 31, 11, 12] and recently [21, 22, 3, 13, 6]. In [31], Stein's method is developed for approximations by abstract Wiener measures on a real separable Banach space. References [21, 3] establish a method for bounding the speed of weak convergence of continuous-time Markov chains satisfying certain assumptions to diffusion processes. Reference [22], on the other hand, treats multi-dimensional processes represented by scaled sums of random variables with different dependence structures and establishes bounds on their distances from continuous Gaussian processes. The recent reference [6] introduces a Dirichlet structure and the corresponding Gamma calculus in an infinite-dimensional context and combines it with the methodology of [31] to derive bounds on distances from Gaussian random variables valued in Hilbert spaces.

The bounds in references [1, 2, 21, 22] are obtained with respect to convergence-determining classes of test functions and weak convergence results in the Skorokhod topology follow from the bounds as corollaries. These references all use and adapt the setup of [1]. Therein, the author studies Donsker's theorem saying that for a sequence of i.i.d. real random variables $(X_n)_{n=1}^\infty$ with mean zero and unit variance, the random process

$$\mathbf{Y}_n(t) = n^{-1/2} \sum_{i=1}^{\lfloor nt \rfloor} X_i, \quad t \in [0, 1] \tag{1.7}$$

converges in distribution to the standard Brownian motion with respect to the Skorokhod topology. Through a careful and technical application of the general Stein's method philosophy explained above, and a subsequent repetitive use of Taylor's theorem, Barbour [1] proved a powerful estimate on a distance between the law of \mathbf{Y}_n in (1.7) and the Wiener measure. Specifically, he considered test functions g acting on the Skorokhod space $D([0, 1], \mathbb{R})$ of càdlàg real-valued maps on $[0, 1]$, such that g takes values in the reals, does not grow faster than a cubic, is twice Fréchet differentiable and its second derivative is Lipschitz. Denoting by \mathbf{Z} Brownian motion on $[0, 1]$ and adopting the notation of (1.7), his result says that

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

where C_g is a constant, independent of n , yet depending on the (carefully defined) *smoothness properties* of g . Among the applications and extensions considered by Barbour are an analysis of the empirical distribution function of i.i.d. random variables and the Wald-Wolfowitz theorem often used to construct tests in non-parametric statistics [34].

The results of [11, 12, 3, 13] consider a totally different setup. Therein, the authors develop a Stein theory on a Hilbert space using a Besov-type topology. Their bounds do not imply weak convergence in the Skorokhod topology and, for most of the natural examples one can consider, the continuous mapping theorem does not apply in the setting they work with. For instance, as opposed to the results of [1], those of [11] do not allow one to draw conclusions about convergence of the supremum of a process. For these reasons, the setup we use to obtain the bounds in the current paper is analogous to that of the former class of references [1, 2, 21, 22]. We consider it more flexible than the one of the class of references [11, 12, 3, 13]. It is also more suited for applications to processes belonging to the widely-used Skorokhod space than the one of [31], which only caters for separable Banach spaces.

1.6. Structure of the paper

The paper is organised as follows. In Section 2 we introduce the spaces of test functions which will be used in the main results. We also quote a proposition showing that, under certain assumptions, they determine convergence in distribution under the uniform topology. In Section 3 we set up the Stein equation for approximation by a pre-limiting process and provide properties of the solutions. In Section 4 we provide an exchangeable-pair condition and prove an abstract exchangeable-pair-type approximation theorem. Section 5 is devoted to the functional combinatorial central limit theorem example and Section 6 discusses the graph-valued process example.

2. Spaces M , M^1 , M^2 , M^0

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

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

We also let $D = D([0, 1], \mathbb{R})$ be the Skorokhod space of all càdlàg functions on $[0, 1]$ taking values in \mathbb{R} . We will often write $\mathbb{E}^W[\cdot]$ instead of $\mathbb{E}[\cdot|W]$.

Let us define:

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

and let L be the Banach space of continuous functions $f : D \rightarrow \mathbb{R}$ such that $\|f\|_L < \infty$. Following [1], we now let $M \subset L$ consist 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 $w, h \in D$. By $D^k f$ we mean the k -th Fréchet derivative of f and the norm of a k -linear form B on L is defined to be $\|B\| = \sup_{\{h: \|h\|=1\}} |B[h, \dots, h]|$. Note the following lemma, which can be proved in an analogous way to that used to show (2.6) and (2.7) of [1]. We omit the proof here.

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

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

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

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

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

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

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

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

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

We note that $M^0 \subset M^1 \subset M^2 \subset M$. The next proposition is [2, Proposition 3.1] and shows conditions, under which convergence of the sequence of expectations of a functional g under the approximating measures to the expectation of g under the target measure for all $g \in M^0$ implies weak convergence of the measures of interest.

Proposition 2.2 (Proposition 3.1 of [2]). *Suppose that, for each $n \geq 1$, the random element \mathbf{Y}_n of 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 converging in distribution in D , with respect to the Skorokhod topology, to a random element $\mathbf{Z} \in C([0, 1], \mathbb{R})$. If:*

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

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

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

The steps of the construction presented in this section will be similar to those used to set up Stein's method in [1] and [22]. After defining the process \mathbf{D}_n whose distribution will be the target measure in Stein's method, we will construct a process $(\mathbf{W}_n(\cdot, u) : u \geq 0)$ for which the target measure is stationary. We will then 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 [23] and prove some properties of the solution $f_n = \phi_n(g)$, with the most important one being that its second Fréchet derivative is Lipschitz.

3.1. Target measure

Let

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

where $\tilde{Z}_{i_1, \dots, i_m}$'s are centred Gaussian and:

- A) the covariance matrix $\Sigma_n \in \mathbb{R}^{(n^m) \times (n^m)}$ of the vector \tilde{Z} is positive definite, where $\tilde{Z} \in \mathbb{R}^{(n^m)}$ is formed out of the $\tilde{Z}_{i_1, \dots, i_m}$'s in such a way that they appear in the lexicographic order.
- B) The collection $\{J_{i_1, \dots, i_m} \in D([0, 1], \mathbb{R}) : i_1, \dots, i_m \in \{1, \dots, n\}\}$ is independent of the collection $\{\tilde{Z}_{i_1, \dots, i_m} : i_1, \dots, i_m \in \{1, \dots, n\}\}$. A typical example would be $J_{i_1, \dots, i_m} = \mathbb{1}_{A_{i_1, \dots, i_m}}$ for some measurable set A_{i_1, \dots, i_m} .

Remark 3.1. *It is worth noting that processes \mathbf{D}_n taking the form (3.1) often approximate interesting continuous Gaussian processes very well. An example is a Gaussian scaled random walk, i.e. \mathbf{D}_n of (3.1), where all the $\tilde{Z}_{i_1, \dots, i_m}$'s are standard normal and independent, $m = 1$ and $J_i = \mathbb{1}_{[i/n, 1]}$ for all $i = 1, \dots, n$. It approximates Brownian Motion. By Proposition 2.2, under several assumptions, proving by Stein's method that a piece-wise constant process \mathbf{Y}_n is close enough to process \mathbf{D}_n proves \mathbf{Y}_n 's convergence in law to the continuous process that \mathbf{D}_n approximates.*

Now let $\{(\mathcal{X}_{i_1, \dots, i_m}(u), u \geq 0) : i_1, \dots, i_m = 1, \dots, n\}$ be 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} 's. Consider $\tilde{\mathcal{Z}}(u) = (\Sigma_n)^{1/2} \mathcal{X}(u)$, where $\mathcal{X}(u) \in \mathbb{R}^m$ is formed out of the $\mathcal{X}_{i_1, \dots, i_m}(u)$'s in such a way that they appear in the same order as the $\tilde{Z}_{i_1, \dots, i_m}$'s appear in \tilde{Z} . Write $\mathcal{U}_{i_1, \dots, i_m}(u) = \left(\tilde{\mathcal{Z}}(u)\right)_{I(i_1, \dots, i_m)}$ using the bijection $I : \{(i_1, \dots, i_m) : i_1, \dots, i_m = 1, \dots, n\} \rightarrow \{1, \dots, n^m\}$, given by:

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

Consider a process:

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

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

By [22, Propositions 4.1 and 4.4], the following result is immediate:

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

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

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

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

for any constant function $c : D \rightarrow \mathbb{R}$ and for all $w, h \in D$. Moreover, for all $g \in M^1$, as defined in (2.2),

$$\begin{aligned} A) \quad & \|D\phi_n(g)(w)\| \leq \|g\|_{M^1}, \\ B) \quad & \|D^2\phi_n(g)(w)\| \leq \frac{1}{2}\|g\|_{M^1} \end{aligned}$$

and for all $g \in M^2$, as defined in (2.3),

$$\|D\phi_n(g)(w)\| \leq \|g\|_{M^2}.$$

4. An abstract approximation theorem

We now present a theorem which provides an expression for a bound on the distance between some process \mathbf{Y}_n and \mathbf{D}_n , defined by (3.1), provided that we can find some \mathbf{Y}'_n such that $(\mathbf{Y}_n, \mathbf{Y}'_n)$ is an exchangeable pair satisfying an appropriate condition.

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

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

where $\mathbb{E}^{\mathbf{Y}_n}[\cdot] := \mathbb{E}[\cdot | \mathbf{Y}_n]$, for all $f \in M$, some $\lambda_n > 0$ 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

$$\begin{aligned} \epsilon_1 &= \frac{\|g\|_M}{12\lambda_n} \mathbb{E} [\|\mathbf{Y}_n - \mathbf{Y}'_n\|^3], \\ \epsilon_2 &= \left| \frac{1}{2\lambda_n} \mathbb{E} D^2 f(\mathbf{Y}_n) [\mathbf{Y}_n - \mathbf{Y}'_n, \mathbf{Y}_n - \mathbf{Y}'_n] - \mathbb{E} D^2 f(\mathbf{Y}_n) [\mathbf{D}_n, \mathbf{D}_n] \right|, \\ \epsilon_3 &= \frac{1}{\lambda_n} |\mathbb{E} R_f|, \end{aligned}$$

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 large λ_n is. Term ϵ_2 corresponds to the comparison of the covariance structure of $\mathbf{Y}_n - \mathbf{Y}'_n$ and \mathbf{D}_n . Estimating this term usually requires some effort yet is possible in several applications (see Theorem 5.1 and 6.3 below). Term ϵ_3 measures the error in the exchangeable-pair linear regression condition (4.1).

Remark 4.3. The term

$$\left| \frac{1}{2\lambda_n} \mathbb{E} D^2 f(\mathbf{Y}_n) [\mathbf{Y}_n - \mathbf{Y}'_n, \mathbf{Y}_n - \mathbf{Y}'_n] - \mathbb{E} D^2 f(\mathbf{Y}_n) [\mathbf{D}_n, \mathbf{D}_n] \right|$$

in the bound obtained in Theorem 4.1 is an analogue of the second condition in [24, Theorem 3]. Therein, a bound on approximation by $\mathcal{N}(0, \Sigma)$ of a d -dimensional vector X is obtained 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:

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

The bound would then take the form:

$$|\mathbb{E}g(\mathbf{Y}_n) - \mathbb{E}g(\mathbf{D}_n)| \leq \frac{\|g\|_M}{12\lambda_n} \mathbb{E} [\|\mathbf{Y}_n - \mathbf{Y}'_n\|^3] + \frac{1}{\lambda_n} |\mathbb{E} R_f| + \frac{1}{2\lambda_n} |\mathbb{E} R_f^1|.$$

Remark 4.4. While Theorem 4.1 is formulated for univariate stochastic processes belonging to the Skorokhod space $D([0, 1], \mathbb{R})$, its proof below is valid for processes belonging to $D([0, 1], \mathbb{R}^p)$, for any $p \geq 1$.

We only present the univariate case here since in the modern approach to multivariate Stein's method of exchangeable pairs, presented in [27, 28, 24], the exchangeable-pair condition for multivariate approximations is written not in terms of a real λ_n , as in (4.1), but in terms of a matrix-valued coefficient Λ , as in (1.4).

While it is possible to rewrite condition (4.1) appropriately, in terms of a matrix Λ_n , prove a corresponding bound and use it in interesting applications, we think of it as a separate problem and leave it for future research.

Proof of Theorem 4.1. Our aim is to 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):

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

and so:

$$\mathbb{E} Df(\mathbf{Y}_n)[\mathbf{Y}_n] = \frac{1}{2\lambda_n} \{\mathbb{E}(Df(\mathbf{Y}_n) - Df(\mathbf{Y}'_n))[\mathbf{Y}_n - \mathbf{Y}'_n] + 2\mathbb{E} R_f\}.$$

Therefore:

$$\begin{aligned} |\mathbb{E} \mathcal{A}_n f(\mathbf{Y}_n)| &= |\mathbb{E} Df(\mathbf{Y}_n)[\mathbf{Y}_n] - \mathbb{E} D^2 f(\mathbf{Y}_n)[\mathbf{D}_n, \mathbf{D}_n]| \\ &= \left| \frac{1}{2\lambda_n} \mathbb{E}(Df(\mathbf{Y}_n) - Df(\mathbf{Y}'_n))[\mathbf{Y}_n - \mathbf{Y}'_n] - \mathbb{E} D^2 f(\mathbf{Y}_n)[\mathbf{D}_n, \mathbf{D}_n] + \frac{1}{\lambda_n} \mathbb{E} R_f \right| \\ &\leq \left| \frac{1}{2\lambda_n} \mathbb{E}(Df(\mathbf{Y}_n) - Df(\mathbf{Y}'_n))[\mathbf{Y}_n - \mathbf{Y}'_n] - \frac{1}{2\lambda_n} \mathbb{E} D^2 f(\mathbf{Y}'_n)[\mathbf{Y}_n - \mathbf{Y}'_n, \mathbf{Y}_n - \mathbf{Y}'_n] \right| \\ &\quad + \left| \frac{1}{2\lambda_n} \mathbb{E} D^2 f(\mathbf{Y}_n)[\mathbf{Y}_n - \mathbf{Y}'_n, \mathbf{Y}_n - \mathbf{Y}'_n] - \mathbb{E} D^2 f(\mathbf{Y}_n)[\mathbf{D}_n, \mathbf{D}_n] \right| + \frac{1}{\lambda_n} |\mathbb{E} R_f| \\ &\leq \frac{\|g\|_M}{12\lambda_n} \mathbb{E} [\|\mathbf{Y}_n - \mathbf{Y}'_n\|^3] + \frac{1}{\lambda_n} |\mathbb{E} R_f| \\ &\quad + \left| \frac{1}{2\lambda_n} \mathbb{E} D^2 f(\mathbf{Y}_n)[\mathbf{Y}_n - \mathbf{Y}'_n, \mathbf{Y}_n - \mathbf{Y}'_n] - \mathbb{E} D^2 f(\mathbf{Y}_n)[\mathbf{D}_n, \mathbf{D}_n] \right|, \end{aligned}$$

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

5. A Functional Combinatorial Central Limit Theorem

In this section we consider a functional version of the result proved in [19]. Our object of interest is a stochastic process represented by a scaled sum of independent random variables chosen from an $n \times n$ array. Only one random variable is picked from each row and for row i , the corresponding random variable is picked from column $\pi(i)$, where π is a random permutation on $[n] = \{1, \dots, n\}$. Theorem 5.1 establishes a bound on the distance between this process and a pre-limiting process and Theorem 5.5 shows convergence of this process, under certain assumptions, to a continuous Gaussian process.

Our analysis in this section is similar to that of [2], where the summands in the scaled sums are chosen from a deterministic array. The authors therein also establish bounds on the approximation by a pre-limit Gaussian process and show convergence to a continuous Gaussian process. Furthermore, they establish a bound on the distance from the continuous Gaussian process for a restricted class of test functions. For random arrays the situation is more involved.

Our setup is analogous to the one considered in [10], where a bound on the rate of convergence in the one-dimensional combinatorial central limit theorem is obtained using Stein's method of exchangeable pairs.

5.1. Introduction

Let $\mathbb{X} = \{X_{ij} : i, j \in [n]\}$, where $[n] = \{1, 2, \dots, n\}$, be an $n \times n$ array of independent \mathbb{R} -valued random variables, where $n \geq 2$, $\mathbb{E} X_{ij} = c_{ij}$, $\text{Var} X_{ij} = \sigma_{ij}^2 \geq 0$ and $\mathbb{E} |X_{ij}|^3 < \infty$. Suppose that $c_i = c_j = 0$ where $c_i = \sum_{j=1}^n \frac{c_{ij}}{n} = \mathbb{E} X_{i\pi(i)}$, $c_j = \sum_{i=1}^n \frac{c_{ij}}{n}$. Let π be a uniform random permutation of $[n]$, independent of \mathbb{X} and for

$$s_n^2 = \frac{1}{n} \sum_{i,j=1}^n \sigma_{ij}^2 + \frac{1}{n-1} \sum_{i,j=1}^n c_{ij}^2. \quad (5.1)$$

let

$$\mathbf{Y}_n(t) = \frac{1}{s_n} \sum_{i=1}^{\lfloor nt \rfloor} X_{i\pi(i)} = \frac{1}{s_n} \sum_{i=1}^n X_{i\pi(i)} \mathbb{1}_{[i/n, 1]}(t), \quad t \in [0, 1].$$

We note that $s_n^2 = \text{Var} \left[\sum_{i=1}^n X_{i\pi(i)} \right]$, by the first part of [10, Theorem 1.1]. The process \mathbf{Y}_n is similar to the process Y considered in [2] and defined by (1.4) therein with the most important difference being that we allow the $X_{i,j}$'s to be random, whereas the authors in [2] assumed them to be deterministic. Bounds on the distance between one-dimensional distributions of \mathbf{Y}_n and a normal distribution have been obtained via Stein's method in [10, Theorem 1.1].

5.2. Exchangeable pair setup

Select uniformly at random two different indices $I, J \in [n]$ and let:

$$\mathbf{Y}'_n = \mathbf{Y}_n - \frac{1}{s_n} X_{I\pi(I)} \mathbb{1}_{[I/n,1]} - \frac{1}{s_n} X_{J\pi(J)} \mathbb{1}_{[J/n,1]} + \frac{1}{s_n} X_{I\pi(J)} \mathbb{1}_{[I/n,1]} + \frac{1}{s_n} X_{J\pi(I)} \mathbb{1}_{[J/n,1]}.$$

Note that $(\mathbf{Y}_n, \mathbf{Y}'_n)$ is an exchangeable pair and that for all $f \in M$:

$$\begin{aligned} & \mathbb{E}^{\mathbf{Y}_n} \{Df(\mathbf{Y}_n) [\mathbf{Y}_n - \mathbf{Y}'_n]\} \\ &= \frac{1}{s_n} \mathbb{E}^{\mathbf{Y}_n} \left\{ Df(\mathbf{Y}_n) \left[X_{I\pi(I)} \mathbb{1}_{[I/n,1]} + X_{J\pi(J)} \mathbb{1}_{[J/n,1]} - X_{I\pi(J)} \mathbb{1}_{[I/n,1]} - X_{J\pi(I)} \mathbb{1}_{[J/n,1]} \right] \right\} \\ &= \frac{1}{n(n-1)s_n} \sum_{i,j=1}^n \mathbb{E}^{\mathbf{Y}_n} \left\{ Df(\mathbf{Y}_n) \left[X_{i\pi(i)} \mathbb{1}_{[i/n,1]} + X_{j\pi(j)} \mathbb{1}_{[j/n,1]} - X_{i\pi(j)} \mathbb{1}_{[i/n,1]} - X_{j\pi(i)} \mathbb{1}_{[j/n,1]} \right] \right\} \\ &= \frac{2}{n-1} Df(\mathbf{Y}_n) [\mathbf{Y}_n] - \frac{2}{n(n-1)s_n} \sum_{i,j=1}^n \mathbb{E}^{\mathbf{Y}_n} Df(\mathbf{Y}_n) [X_{i,\pi(j)} \mathbb{1}_{[i/n,1]}]. \end{aligned}$$

Therefore:

$$\mathbb{E}^{\mathbf{Y}_n} \{Df(\mathbf{Y}_n) [\mathbf{Y}'_n - \mathbf{Y}_n]\} = -\frac{2}{n-1} Df(\mathbf{Y}_n) [\mathbf{Y}_n] + \frac{2}{n(n-1)s_n} \sum_{i,j=1}^n Df(\mathbf{Y}_n) \left[\mathbb{E}^{\mathbf{Y}_n} [X_{i,\pi(j)}] \mathbb{1}_{[i/n,1]} \right]. \quad (5.2)$$

So condition (4.1) is satisfied with

$$\lambda_n = \frac{2}{n-1} \quad \text{and} \quad R_f = \frac{2}{n(n-1)s_n} \sum_{i,j=1}^n Df(\mathbf{Y}_n) \left[\mathbb{E}^{\mathbf{Y}_n} [X_{i,\pi(j)}] \mathbb{1}_{[i/n,1]} \right].$$

5.3. Pre-limiting process

Now let $\hat{Z}_i = \frac{1}{\sqrt{n-1}} \sum_{l=1}^n X_{il}'' \left(Z_{il} - \frac{1}{n} \sum_{j=1}^n Z_{jl} \right)$, for $\mathbb{X}'' = \{X_{ij}'' : i, j \in [n]\}$ being an independent copy of \mathbb{X} and Z_{il} 's i.i.d. standard normal, independent of $\mathbb{X} \cup \mathbb{X}'$. Then, let

$$\mathbf{D}_n(t) = \frac{1}{s_n} \sum_{i=1}^{\lfloor nt \rfloor} \hat{Z}_i, \quad t \in [0, 1]. \quad (5.3)$$

We will compare the distribution of \mathbf{Y}_n with the distribution of \mathbf{D}_n . \mathbf{D}_n is a conceptually easy process with the same covariance structure as \mathbf{Y}_n . It is constructed in a way similar to the process in [2, (3.13)]. Note

that \hat{Z}_i has mean zero for all i and

$$\begin{aligned}
\mathbb{E}\hat{Z}_i^2 &= \frac{1}{n-1} \sum_{l=1}^n \mathbb{E}[X_{il}^2] \mathbb{E} \left[\left(Z_{il} - \frac{1}{n} \sum_{j=1}^n Z_{jl} \right)^2 \right] \\
&\quad + \frac{1}{n-1} \sum_{1 \leq l \neq k \leq n} \mathbb{E}[X_{il}X_{ik}] \mathbb{E} \left[\left(Z_{il} - \frac{1}{n} \sum_{j=1}^n Z_{jl} \right) \left(Z_{ik} - \frac{1}{n} \sum_{j=1}^n Z_{jk} \right) \right] \\
&= \frac{1}{n-1} \sum_{l=1}^n \mathbb{E}[X_{il}^2] \left(1 - \frac{2}{n} + \frac{1}{n} \right) \\
&= \frac{1}{n} \sum_{l=1}^n \mathbb{E}X_{il}^2 \\
&= \frac{1}{2n^2} \left(2(n-1) \sum_{l=1}^n \mathbb{E}X_{il}^2 + 2 \sum_{r=1}^n \mathbb{E}X_{ir}^2 \right) \\
&= \frac{1}{2n^2} \left(\sum_{1 \leq k \neq l \leq n} \mathbb{E}[(X_{ik} - X_{il})^2] + 2 \sum_{1 \leq k \neq l \leq n} \mathbb{E}X_{ik}\mathbb{E}X_{il} + 2 \sum_{r=1}^n \mathbb{E}X_{ir}^2 \right) \\
&= \frac{1}{2n^2} \left(\sum_{1 \leq k \neq l \leq n} \mathbb{E}[(X_{ik} - X_{il})^2] + 2 \sum_{r=1}^n \sigma_{ir}^2 \right) \tag{5.4}
\end{aligned}$$

as $c_i = 0$, and, for $i \neq j$,

$$\begin{aligned}
\mathbb{E}\hat{Z}_i\hat{Z}_j &= \frac{1}{n-1} \sum_{k,l=1}^n \mathbb{E}(X_{ik}X_{jl}) \mathbb{E} \left[\left(Z_{ik} - \frac{1}{n} \sum_{r=1}^n Z_{rk} \right) \left(Z_{jl} - \frac{1}{n} \sum_{r=1}^n Z_{rl} \right) \right] \\
&= - \frac{1}{n(n-1)} \sum_{k=1}^n c_{ik}c_{jk} \\
&= \frac{1}{2n^2(n-1)} \left(2 \sum_{k=1}^n (-\mathbb{E}X_{ik}) \mathbb{E}X_{jk} - 2(n-1) \sum_{k=1}^n \mathbb{E}X_{ik}\mathbb{E}X_{jk} \right) \\
&= \frac{1}{2n^2(n-1)} \left(2 \sum_{1 \leq k \neq l \leq n} \mathbb{E}X_{il}\mathbb{E}X_{jk} - 2 \sum_{1 \leq k \neq l \leq n} \mathbb{E}X_{ik}\mathbb{E}X_{jk} \right) \\
&= \frac{1}{2n^2(n-1)} \sum_{1 \leq k \neq l \leq n} \mathbb{E}(X_{ik} - X_{il})(X_{jl} - X_{jk}). \tag{5.5}
\end{aligned}$$

5.4. Pre-limiting approximation

We have the following theorem, comparing the distribution of \mathbf{Y}_n and \mathbf{D}_n :

Theorem 5.1. For \mathbf{Y}_n defined in Subsection 5.1, \mathbf{D}_n defined in Subsection 5.3 and any $g \in M^1$, as defined in (2.2),

$$\begin{aligned}
&|\mathbb{E}g(\mathbf{Y}_n) - \mathbb{E}g(\mathbf{D}_n)| \\
&\leq \frac{\|g\|_{M^1}}{2n^2(n-1)^2s_n^3} \sum_{1 \leq i,j,k,l,u \leq n} \left\{ \mathbb{E}|X_{ik}|^3 + 5\mathbb{E}|X_{ik}|^2\mathbb{E}|X_{il}| + 9\mathbb{E}|X_{ik}|^2\mathbb{E}|X_{jl}| + 5\mathbb{E}|X_{ik}|^2\mathbb{E}|X_{jk}| \right. \\
&\quad + 14\mathbb{E}|X_{ik}|\mathbb{E}|X_{il}|\mathbb{E}|X_{jl}| + 2\mathbb{E}|X_{ik}|\mathbb{E}|X_{il}|\mathbb{E}|X_{iu}| + 6\mathbb{E}|X_{ik}|\mathbb{E}|X_{jl}|\mathbb{E}|X_{iu}| + 2\mathbb{E}[|X_{uk}||X_{ik}|^2] \\
&\quad \left. + 4\mathbb{E}|X_{uk}X_{ik}|\mathbb{E}|X_{il}| + 4\mathbb{E}|X_{uk}X_{ik}|\mathbb{E}|X_{jl}| + 8\mathbb{E}|X_{uk}|\mathbb{E}|X_{ik}|\mathbb{E}|X_{jl}| + 2\mathbb{E}|X_{uk}X_{ik}X_{jk}| \right\}
\end{aligned}$$

$$\begin{aligned}
& + 2\mathbb{E}|X_{uk}|\mathbb{E}|X_{ik}|\mathbb{E}|X_{jk}| + \frac{2}{n}\mathbb{E}(2|X_{ik}| + 6|X_{jl}|) \cdot \left(\sum_{r=1}^n \mathbb{E}|X_{ir}|^2 + \left| \sum_{r=1}^n c_{ir}c_{jr} \right| \right) \Bigg\}. \\
& + \frac{2\|g\|_{M^1}}{\sqrt{n}} + \frac{\|g\|_{M^1}}{n^2 s_n^2} \sum_{i,j=1}^n \sigma_{ij}^2.
\end{aligned}$$

Remark 5.2 (Relevance of terms in the bound). *The first long sum in the bound corresponds to ϵ_1 and (to a large extent) ϵ_2 of Theorem 4.1. It represents the usual Berry-Esseen third moment estimate arising as a result of applying Taylor's theorem. Term $\frac{2\|g\|_{M^1}}{\sqrt{n}}$ also comes from the estimation of ϵ_2 . The last term corresponds to ϵ_3 .*

Remark 5.3. *Assuming that $s_n = O(\sqrt{n})$, we obtain that the bound in Theorem 5.1 is of order $\frac{1}{\sqrt{n}}$.*

Remark 5.4. *If we assume that $\mathbb{E}|X_{ik}|^3 \leq \beta_3$ for all $i, k = 1, \dots, n$ then the bound simplifies in the following way*

$$|\mathbb{E}g(\mathbf{Y}_n) - \mathbb{E}g(\mathbf{D}_n)| \leq \|g\|_{M^1} \left(\frac{40\beta_3 n^2}{(n-1)s_n^3} + \frac{8\beta_3^{1/3}}{n(n-1)s_n^3} \left| \sum_{i,j,r=1}^n c_{ir}c_{jr} \right| + \frac{2}{\sqrt{n}} + \frac{1}{n^2 s_n^2} \sum_{i,j=1}^n \sigma_{ij}^2 \right).$$

We will use Theorem 4.1 to prove Theorem 5.1. In the proof, in **Step 1**, we justify why Theorem 4.1 may indeed be used in this case. In other words, we check that \mathbf{D}_n of (5.3) satisfies the conditions \mathbf{D}_n of Theorem 4.1 is supposed to satisfy and that the exchangeable-pair condition for \mathbf{Y}_n holds. In **Step 2** we bound terms ϵ_1 and ϵ_3 coming from Theorem 4.1. This is relatively straightforward due to the \mathbf{Y}_n and \mathbf{Y}'_n of Subsection 5.2 being constructed in such a way that they are close to each other and R_f of the same subsection being small. Then, in **Step 3**, we treat the remaining term using a strategy analogous to that of the proof of [2, Theorem 2.1]. The strategy is based on Taylor's expansions and considering copies of \mathbf{Y}_n which are independent of some of the summands in \mathbf{Y}_n . Finally, we combine the estimates obtained in the previous steps to obtain the assertion.

Proof of theorem 5.1. We adopt the notation of Subsections 5.1, 5.2 and 5.3. Furthermore, we fix a function $g \in M^1$, as defined in (2.2) and let $f = \phi_n(g)$, a solution to the Stein equation for \mathbf{D}_n , as defined in (3.3).

Step 1. We note that \mathbf{D}_n can be expressed in the following way:

$$\mathbf{D}_n = \sum_{i,l=1}^n \left(Z_{il} - \frac{1}{n} \sum_{j=1}^n Z_{jl} \right) J_{i,l}, \quad \text{where} \quad J_{i,l}(t) = \frac{X''_{il}}{s_n \sqrt{n-1}} \mathbb{1}_{[i/n, 1]}(t),$$

which, together with (5.2), lets us apply Theorem 4.1.

Step 2. For the first term in Theorem 4.1, for any $g \in M^1$:

$$\epsilon_1 = \frac{\|g\|_{M^1}}{12\lambda_n} \mathbb{E}\|\mathbf{Y}_n - \mathbf{Y}'_n\|^3 = \frac{(n-1)\|g\|_{M^1}}{24} \mathbb{E}\|\mathbf{Y}_n - \mathbf{Y}'_n\|^3.$$

We note that:

$$\begin{aligned}
\mathbb{E}\|\mathbf{Y}_n - \mathbf{Y}'_n\|^3 & \leq \frac{16}{s_n^3} (\mathbb{E}|X_{I\pi(I)}|^3 + \mathbb{E}|X_{J\pi(J)}|^3 + \mathbb{E}|X_{I\pi(J)}|^3 + \mathbb{E}|X_{J\pi(I)}|^3) \\
& = \frac{16}{n(n-1)s_n^3} \sum_{i \neq j} (\mathbb{E}|X_{i\pi(i)}|^3 + \mathbb{E}|X_{j\pi(j)}|^3 + \mathbb{E}|X_{i\pi(j)}|^3 + \mathbb{E}|X_{j\pi(i)}|^3) \\
& = \frac{32}{n(n-1)s_n^3} \sum_{i \neq j} (\mathbb{E}|X_{i\pi(i)}|^3 + \mathbb{E}|X_{i\pi(j)}|^3) \\
& = \frac{64}{n^2 s_n^3} \sum_{i,j=1}^n \mathbb{E}|X_{ij}|^3.
\end{aligned}$$

Hence,

$$\epsilon_1 \leq \frac{8\|g\|_{M^1}}{3ns_n^3} \sum_{i,j=1}^n \mathbb{E}|X_{ij}|^3. \quad (5.6)$$

Furthermore, by Proposition 3.2:

$$\begin{aligned} \epsilon_3 &= \left| \frac{1}{ns_n} \sum_{i,j=1}^n \mathbb{E} Df(\mathbf{Y}_n) [X_{i\pi(j)} \mathbb{1}_{[i/n,1]}] \right| = \left| \frac{1}{ns_n} \sum_{i,j=1}^n \mathbb{E} Df(\mathbf{Y}_n) [X_{ij} \mathbb{1}_{[i/n,1]}] \right| \\ &\leq \|g\|_{M^1} \frac{1}{ns_n} \mathbb{E} \left\| \sum_{i,j=1}^n X_{ij} \mathbb{1}_{[i/n,1]} \right\| \\ &\leq \frac{2\|g\|_{M^1}}{ns_n} \sqrt{\mathbb{E} \left| \sum_{i,j=1}^n X_{ij} \right|^2} \\ &\leq \frac{2\|g\|_{M^1}}{ns_n} \sqrt{\sum_{i,j=1}^n \sigma_{ij}^2} \\ &\leq \frac{2\|g\|_{M^1}}{\sqrt{n}}, \end{aligned} \quad (5.7)$$

where we have used Doob's L^2 inequality in the second inequality and (5.1) in the last one.

Step 3. We will follow a strategy similar to that of [10]. Define a new permutation π_{ijkl} coupled with π , such that:

$$\mathcal{L}(\pi_{ijkl}) = \mathcal{L}(\pi | \pi(i) = k, \pi(j) = l),$$

where $\mathcal{L}(\cdot)$ denotes the law. As noted in [10, (3.14)], we can construct it in the following way. For τ_{ij} denoting the transposition of i, j :

$$\pi_{ijkl} = \begin{cases} \pi, & \text{if } l = \pi(j), k = \pi(i) \\ \pi \cdot \tau_{\pi^{-1}(k), i}, & \text{if } l = \pi(j), k \neq \pi(i) \\ \pi \cdot \tau_{\pi^{-1}(l), j}, & \text{if } l \neq \pi(j), k = \pi(i) \\ \pi \cdot \tau_{\pi^{-1}(l), i} \cdot \tau_{\pi^{-1}(k), j} \cdot \tau_{ij}, & \text{if } l \neq \pi(j), k \neq \pi(i). \end{cases}$$

We also let

$$\mathbf{Y}_{n,ijkl} = \frac{1}{s_n} \sum_{i'=1}^n X_{i'\pi_{ijkl}(i')} \mathbb{1}_{[i'/n,1]}.$$

Then $\mathcal{L}(\mathbf{Y}_{n,ijkl}) = \mathcal{L}(\mathbf{Y}_n | \pi(i) = k, \pi(j) = l)$ (recalling that $\mathcal{L}(\cdot)$ denotes the law). Also, for each choice of $i \neq j, k \neq l$ let $\mathbb{X}^{ijkl} := \{X_{i'j'}^{ijkl} : i', j' \in [n]\}$ be the same as $\mathbb{X} := \{X_{ij} : i, j \in [n]\}$ except that $\{X_{ik}, X_{il}, X_{jk}, X_{jl}\}$ has been replaced by an independent copy $\{X'_{ik}, X'_{il}, X'_{jk}, X'_{jl}\}$. Then let

$$\mathbf{Y}_n^{ijkl} = \frac{1}{s_n} \sum_{i'=1}^n X_{i'\pi(i')}^{ijkl} \mathbb{1}_{[i'/n,1]}$$

and note that \mathbf{Y}_n^{ijkl} is independent of $\{X_{ik}, X_{il}, X_{jk}, X_{jl}\}$ and $\mathcal{L}(\mathbf{Y}_n^{ijkl}) = \mathcal{L}(\mathbf{Y}_n)$ (where \mathcal{L} denotes the law).

Now, by Lemma 7.1, proved in the appendix, for ϵ_2 of Theorem 4.1,

$$\begin{aligned} \epsilon_2 &= \left| \frac{1}{2\lambda_n} \mathbb{E} D^2 f(\mathbf{Y}_n) [(\mathbf{Y}_n - \mathbf{Y}'_n), \mathbf{Y}_n - \mathbf{Y}'_n] - \mathbb{E} D^2 f(\mathbf{Y}_n) [\mathbf{D}_n, \mathbf{D}_n] \right| \\ &\leq A + B \end{aligned} \quad (5.8)$$

where

$$A = \left| \frac{1}{n(n-1)s_n^2} \sum_{\substack{1 \leq i,j,k,l \leq n \\ i \neq j, k \neq l}} \mathbb{E} \left\{ \left[\frac{(X_{ik} - X_{il})^2}{2n} - \frac{\hat{Z}_i^2}{n-1} \right] (D^2 f(\mathbf{Y}_{n,ijkl}) - D^2 f(\mathbf{Y}_n^{ijkl})) [\mathbb{1}_{[i/n,1]} \mathbb{1}_{[i/n,1]}] \right\} \right. \\ \left. + \frac{1}{n(n-1)s_n^2} \sum_{\substack{1 \leq i,j,k,l \leq n \\ i \neq j, k \neq l}} \mathbb{E} \left\{ \left[\frac{(X_{ik} - X_{il})(X_{jl} - X_{jk})}{2n} - \hat{Z}_i \hat{Z}_j \right] \right. \right. \\ \left. \left. \cdot (D^2 f(\mathbf{Y}_{n,ijkl}) - D^2 f(\mathbf{Y}_n^{ijkl})) [\mathbb{1}_{[i/n,1]}, \mathbb{1}_{[j/n,1]}] \right\} \right|, \quad (5.9)$$

$$B = \left| \frac{1}{n(n-1)s_n^2} \sum_{\substack{1 \leq i,j,k,l \leq n \\ i \neq j, k \neq l}} \mathbb{E} \left\{ \left[\frac{(X_{ik} - X_{il})^2}{2n} - \frac{\hat{Z}_i^2}{n-1} \right] D^2 f(\mathbf{Y}_n^{ijkl}) [\mathbb{1}_{[i/n,1]} \mathbb{1}_{[i/n,1]}] \right\} \right. \\ \left. + \frac{1}{n(n-1)s_n^2} \sum_{\substack{1 \leq i,j,k,l \leq n \\ i \neq j, k \neq l}} \mathbb{E} \left\{ \left[\frac{(X_{ik} - X_{il})(X_{jl} - X_{jk})}{2n} - \hat{Z}_i \hat{Z}_j \right] \right. \right. \\ \left. \left. \cdot D^2 f(\mathbf{Y}_n^{ijkl}) [\mathbb{1}_{[i/n,1]}, \mathbb{1}_{[j/n,1]}] \right\} \right|.$$

Recalling that \mathbf{Y}_n^{ijkl} is independent of $\{X_{ik}, X_{il}, X_{jk}, X_{jl}\}$ and $\mathcal{L}(\mathbf{Y}_n^{ijkl}) = \mathcal{L}(\mathbf{Y}_n)$,

$$B = \left| \frac{1}{n(n-1)s_n^2} \sum_{\substack{1 \leq i,j,k,l \leq n \\ i \neq j, k \neq l}} \mathbb{E} \left[\frac{(X_{ik} - X_{il})^2}{2n} - \frac{\hat{Z}_i^2}{n-1} \right] \mathbb{E} \{ D^2 f(\mathbf{Y}_n) [\mathbb{1}_{[i/n,1]} \mathbb{1}_{[i/n,1]}] \} \right. \\ \left. + \frac{1}{n(n-1)s_n^2} \sum_{\substack{1 \leq i,j,k,l \leq n \\ i \neq j, k \neq l}} \mathbb{E} \left[\frac{(X_{ik} - X_{il})(X_{jl} - X_{jk})}{2n} - \hat{Z}_i \hat{Z}_j \right] \mathbb{E} \left\{ D^2 f(\mathbf{Y}_n) [\mathbb{1}_{[i/n,1]}, \mathbb{1}_{[j/n,1]}] \right\} \right| \\ \leq \frac{\|g\|_{M^1}}{n^2(n-1)s_n^2} \sum_{1 \leq i \neq j \leq n} \sum_{r=1}^n \sigma_{ir}^2 \\ = \frac{\|g\|_{M^1}}{n^2 s_n^2} \sum_{i,j=1}^n \sigma_{ij}^2, \quad (5.10)$$

where the inequality follows by (5.4), (5.5) and Proposition 3.2. Furthermore, by Lemma 7.2, proved in the appendix,

$$A \leq \frac{\|g\|_{M^1}}{2n^2(n-1)^2 s_n^3} \sum_{1 \leq i,j,k,l,u \leq n} \left\{ \mathbb{E}|X_{ik}|^3 + 5\mathbb{E}|X_{ik}|^2 \mathbb{E}|X_{il}| + 9\mathbb{E}|X_{ik}|^2 \mathbb{E}|X_{jl}| + 5\mathbb{E}|X_{ik}|^2 \mathbb{E}|X_{jk}| \right. \\ + 14\mathbb{E}|X_{ik}| \mathbb{E}|X_{il}| \mathbb{E}|X_{jl}| + 2\mathbb{E}|X_{ik}| \mathbb{E}|X_{il}| \mathbb{E}|X_{iu}| + 6\mathbb{E}|X_{ik}| \mathbb{E}|X_{jl}| \mathbb{E}|X_{iu}| + 2\mathbb{E}[|X_{uk}| |X_{ik}|^2] \\ + 4\mathbb{E}|X_{uk} X_{ik}| \mathbb{E}|X_{il}| + 4\mathbb{E}|X_{uk} X_{ik}| \mathbb{E}|X_{jl}| + 8\mathbb{E}|X_{uk}| \mathbb{E}|X_{ik}| \mathbb{E}|X_{jl}| + 2\mathbb{E}|X_{uk} X_{ik} X_{jk}| \\ \left. + 2\mathbb{E}|X_{uk}| \mathbb{E}|X_{ik}| \mathbb{E}|X_{jk}| + \frac{2}{n} \mathbb{E}(2|X_{ik}| + 6|X_{jl}|) \cdot \left(\sum_{r=1}^n \mathbb{E}|X_{ir}|^2 + \left| \sum_{r=1}^n c_{ir} c_{jr} \right| \right) \right\}. \quad (5.11)$$

We now use (5.6),(5.7),(5.8),(5.10),(5.11) to obtain the assertion. \square

5.5. Convergence to a continuous Gaussian process

Theorem 5.5. *Let \mathbb{X} and \mathbf{Y}_n be as defined in Subsection 5.1 and suppose that for all $u, t \in [0, 1]$,*

$$\frac{1}{s_n^2(n-1)} \sum_{i=1}^{\lfloor nt \rfloor} \sum_{j=1}^{\lfloor nu \rfloor} \sum_{k=1}^n \mathbb{E} X_{ik} X_{jk} \left(\delta_{i,j} - \frac{1}{n} \right) \xrightarrow{n \rightarrow \infty} \sigma(u, t) \quad (5.12)$$

and

$$\frac{1}{s_n^2} \sum_{i=1}^{\lfloor nt \rfloor} \sum_{j=1}^{\lfloor nu \rfloor} \sum_{l=1}^n \mathbb{E} X_{il} X_{jl} \xrightarrow{n \rightarrow \infty} \sigma^{(2)}(u, t) \quad (5.13)$$

pointwise for some functions $\sigma, \sigma^{(2)} : [0, 1]^2 \rightarrow \mathbb{R}_+$. Suppose furthermore that

$$\sup_{n \in \mathbb{N}} \frac{1}{n^2 s_n^4} \sum_{l=1}^n \sum_{i=1}^n \text{Var} [X_{il}^2] < \infty. \quad (5.14)$$

and

$$\frac{1}{s_n^2(n-1)} \sum_{i=1}^{\lfloor nt \rfloor} \left(\sum_{l=1}^n X_{il}'' Z_{il} \right)^2 \xrightarrow{P} c(t) \quad (5.15)$$

pointwise for some function $c : [0, 1] \rightarrow \mathbb{R}_+$ and

$$\lim_{n \rightarrow \infty} \frac{1}{s_n \sqrt{n-1}} \mathbb{E} \left[\sup_{i=1, \dots, n} |X_{il}'' Z_{il}| \right] = 0. \quad (5.16)$$

Then $(\mathbf{Y}_n(t), t \in [0, 1])$ converges weakly in the uniform topology to a continuous Gaussian process $(\mathbf{Z}(t), t \in [0, 1])$ with the covariance function σ .

Remark 5.6. Assumption (5.13) could also say that

$$\frac{1}{s_n^2} \sum_{i=1}^{\lfloor nt \rfloor} \sum_{j=1}^{\lfloor nu \rfloor} \sum_{l=1}^n \mathbb{E} X_{il} X_{jl}$$

simply converges pointwise rather than giving the limit a name. However, we will use $\sigma^{(2)}$ in the proof so it is convenient to use it in the formulation of the Theorem as well.

Remark 5.7. Assumption (5.15) is necessary for the limiting process in Theorem 5.5 to be continuous. It essentially corresponds to the assumption that the quadratic variation of the following process

$$\mathbf{D}_n^{(1)}(t) = \frac{1}{s_n \sqrt{n-1}} \sum_{i=1}^{\lfloor nt \rfloor} \sum_{l=1}^n X_{il}'' Z_{il}$$

converges to the function c pointwise in probability, which then implies the weak convergence of the process $\mathbf{D}_n^{(1)}$ to a continuous process. While it is relatively easy to show that $\mathbf{D}_n^{(2)} = \mathbf{D}_n - \mathbf{D}_n^{(1)}$ converges to a continuous limit, we had to explicitly add this assumption to ensure that \mathbf{D}_n does as well.

The proof of Theorem 5.5 will be similar to the proof of [2, Theorem 3.3]. The pre-limiting approximand \mathbf{D}_n , defined in Subsection 5.3, will be expressed as a sum of two parts. In **Steps 1** and **2** we prove that each of those parts is C-tight (i.e. they are tight and for each of them any convergent subsequence converges to a process with continuous sample paths). In **Step 3** we show that the assumptions of Theorem 5.5 trivially imply the convergence of the covariance function of \mathbf{D}_n , which together with C-tightness implies the convergence of \mathbf{D}_n to a continuous process. Theorem 6.3 will then be combined with Proposition 2.2 to show convergence of \mathbf{Y}_n to the same limiting process. Finally, the combinatorial central limit theorem for random arrays, proved in [19] and analysed in [10], will imply that \mathbf{Z} is Gaussian.

Proof of Theorem 5.5. We will use the notation of Subsections 5.1 and 5.3.

Step 1. Note that $\mathbf{D}_n = \mathbf{D}_n^{(1)} + \mathbf{D}_n^{(2)}$, where:

$$\mathbf{D}_n^{(1)}(t) = \frac{1}{s_n \sqrt{n-1}} \sum_{i=1}^{\lfloor nt \rfloor} \sum_{l=1}^n X''_{il} Z_{il}, \quad \mathbf{D}_n^{(2)}(t) = \frac{1}{s_n \sqrt{n-1}} \sum_{i=1}^{\lfloor nt \rfloor} \sum_{l=1}^n X''_{il} \bar{Z}_l$$

for $\bar{Z}_l = \frac{1}{n} \sum_{j=1}^n Z_{jl}$.

Now, note that, by (5.15):

$$\left\langle \mathbf{D}_n^{(1)} \right\rangle_t \xrightarrow{P} c(t)$$

pointwise, where $\langle \cdot \rangle$ denotes quadratic variation. Therefore, by [16, Chapter 7, Theorem 1.4] and using (5.16), we obtain that $\mathbf{D}_n^{(1)}$ converges weakly in the Skorokhod topology on $D[0, 1]$ to a continuous Gaussian process with independent increments.

We now note that the Skorokhod space equipped with the metric (topologically equivalent to the Skorokhod metric) with respect to which it is complete. It is also universally measurable by the discussion at the beginning of [15, Chapter 11.5]. Since it is also separable and $\mathbf{D}_n^{(1)} \Rightarrow \mathbf{Z}_1$, for some continuous process \mathbf{Z}_1 , in the Skorokhod topology, [15, Theorem 11.5.3] implies that $(\mathbf{D}_n^{(1)})_{n \geq 1}$ is C-tight.

Step 2. Also, note that for $u > t$ s.t. $\lfloor nu \rfloor \geq \lfloor nt \rfloor + 1$,

$$\begin{aligned} \mathbb{E} \left[\left| \mathbf{D}_n^{(2)}(u) - \mathbf{D}_n^{(2)}(t) \right|^2 \middle| X''_{il}, i, l \in [n] \right] &= \frac{1}{n(n-1)s_n^2} \sum_{l=1}^n \left(\sum_{i=\lfloor nt \rfloor + 1}^{\lfloor nu \rfloor} X_{il} \right)^2 \\ &\leq \frac{\lfloor nu \rfloor - \lfloor nt \rfloor}{n(n-1)s_n^2} \sum_{l=1}^n \sum_{i=\lfloor nt \rfloor + 1}^{\lfloor nu \rfloor} X_{il}^2 \end{aligned}$$

and

$$\mathbb{E} \left[\left| \mathbf{D}_n^{(2)}(u) - \mathbf{D}_n^{(2)}(t) \right|^2 \middle| X''_{il}, i, l \in [n] \right] = 0, \quad \text{for } u > t \text{ s.t. } \lfloor nu \rfloor = \lfloor nt \rfloor.$$

Since $(\mathbf{D}_n^{(2)} \mid X''_{il}, i, l \in [n])$ is Gaussian for u , such that $\lfloor nu \rfloor \geq \lfloor nt \rfloor + 1$,

$$\begin{aligned} \mathbb{E} \left| \mathbf{D}_n^{(2)}(u) - \mathbf{D}_n^{(2)}(t) \right|^4 &= 3 \mathbb{E} \left\{ \left(\mathbb{E} \left[\left| \mathbf{D}_n^{(2)}(u) - \mathbf{D}_n^{(2)}(t) \right|^2 \middle| X''_{il}, i, l \in [n] \right] \right)^2 \right\} \\ &\leq 3 \left(\frac{\lfloor nu \rfloor - \lfloor nt \rfloor}{n(n-1)s_n^2} \right)^2 \mathbb{E} \left(\sum_{l=1}^n \sum_{i=\lfloor nt \rfloor + 1}^{\lfloor nu \rfloor} X_{il}^2 \right)^2 \\ &= 3 \left(\frac{\lfloor nu \rfloor - \lfloor nt \rfloor}{n(n-1)s_n^2} \right)^2 \left[\left(\sum_{l=1}^n \sum_{i=\lfloor nt \rfloor + 1}^{\lfloor nu \rfloor} \mathbb{E} X_{il}^2 \right)^2 + \sum_{l=1}^n \sum_{i=\lfloor nt \rfloor + 1}^{\lfloor nu \rfloor} \left(\mathbb{E} X_{il}^4 - (\mathbb{E} X_{il}^2)^2 \right) \right] \\ &\leq C \left(\frac{\lfloor nu \rfloor - \lfloor nt \rfloor}{(n-1)} \right)^2 \end{aligned} \tag{5.17}$$

for some constant C , by (5.14). Now, note that:

$$\text{Cov} \left(\mathbf{D}_n^{(2)}(t), \mathbf{D}_n^{(2)}(u) \right) = \frac{1}{s_n^2} \sum_{i=1}^{\lfloor nt \rfloor} \sum_{j=1}^{\lfloor nu \rfloor} \sum_{l=1}^n \mathbb{E} X_{il} X_{jl} \xrightarrow{n \rightarrow \infty} \sigma^{(2)}(t, u),$$

by (5.13). Consider a mean zero Gaussian process \mathbf{Z}_2 with covariance function $\mathbb{E} \mathbf{Z}_2(t) \mathbf{Z}_2(u) = \sigma^{(2)}(t, u)$. The finite dimensional distributions of $\mathbf{D}_n^{(2)}$ converge to those of \mathbf{Z}_2 . We can now construct $\mathbf{D}_n^{(2)}$ and \mathbf{Z}_2 on the

same probability space and use Skorokhod's representation theorem, Fatou's lemma and (5.17) to conclude that:

$$\mathbb{E} \left(|\mathbf{Z}_2(u) - \mathbf{Z}_2(t)|^4 \right) \leq \lim_{n \rightarrow \infty} \mathbb{E} \left(\left| \mathbf{D}_n^{(2)}(u) - \mathbf{D}_n^{(2)}(t) \right|^4 \right) \leq C(u-t)^2.$$

By [4, Theorem 12.4], we can assume that $\mathbf{Z}_2 \in C[0, 1]$. Now, note that for $0 \leq t \leq v \leq u \leq 1$:

$$\begin{aligned} \mathbb{E} \left| \mathbf{D}_n^{(2)}(v) - \mathbf{D}_n^{(2)}(t) \right|^2 \left| \mathbf{D}_n^{(2)}(v) - \mathbf{D}_n^{(2)}(u) \right|^2 &\leq \sqrt{\mathbb{E} \left| \mathbf{D}_n^{(2)}(v) - \mathbf{D}_n^{(2)}(t) \right|^4 \mathbb{E} \left| \mathbf{D}_n^{(2)}(v) - \mathbf{D}_n^{(2)}(u) \right|^4} \\ &\stackrel{(5.17)}{\leq} C \frac{(\lfloor nv \rfloor - \lfloor nt \rfloor)(\lfloor nu \rfloor - \lfloor nv \rfloor)}{(n-1)^2} \\ &\leq \bar{C}(u-t)^2; \end{aligned}$$

for some constant \bar{C} . Therefore, by [4, Theorem 15.6], $\mathbf{D}_n^{(2)} \Rightarrow \mathbf{Z}_2$ in the Skorokhod and uniform topologies and so, by [15, Theorem 11.5.3], $\mathbf{D}_n^{(2)}$ is C-tight.

Step 3. Since both $\mathbf{D}_n^{(1)}$ and $\mathbf{D}_n^{(2)}$ are C-tight, so is their difference \mathbf{D}_n . Now:

$$\begin{aligned} \text{Cov}(\mathbf{D}_n(t), \mathbf{D}_n(u)) &= \frac{1}{s_n^2(n-1)} \sum_{i=1}^{\lfloor nt \rfloor} \sum_{j=1}^{\lfloor nu \rfloor} \sum_{k,l=1}^n \mathbb{E} \{ X_{ik} X_{jl} (Z_{ik} - \bar{Z}_k) (Z_{jl} - \bar{Z}_l) \} \\ &= \frac{1}{s_n^2(n-1)} \sum_{i=1}^{\lfloor nt \rfloor} \sum_{j=1}^{\lfloor nu \rfloor} \sum_{k=1}^n \mathbb{E} \{ X_{ik} X_{jk} (Z_{ik} - \bar{Z}_k) (Z_{jk} - \bar{Z}_k) \} \\ &= \frac{1}{s_n^2(n-1)} \sum_{i=1}^{\lfloor nt \rfloor} \sum_{j=1}^{\lfloor nu \rfloor} \sum_{k=1}^n \mathbb{E} X_{ik} X_{jk} \left(\delta_{i,j} - \frac{1}{n} \right) \xrightarrow{n \rightarrow \infty} \sigma(u, t), \end{aligned}$$

by (5.12) and we obtain that \mathbf{D}_n converges to a random element $\mathbf{Z} \in C[0, 1]$ with covariance function σ in distribution with respect to the uniform and Skorokhod topologies.

Proposition 2.2 and Theorem 5.1 therefore imply that $(\mathbf{Y}_n(t), t \in [0, 1])$ converges weakly to $(\mathbf{Z}(t), t \in [0, 1])$ in the uniform topology. Using, for example, [10, Theorem 1.1], we conclude that \mathbf{Z} is a Gaussian process. \square

6. Edge and two-star counts in Bernoulli random graphs

In this section we consider a process representing a properly rescaled number of edges in a Bernoulli random graph with a fixed edge probability and $\lfloor nt \rfloor$ edges for $t \in [0, 1]$. A similar setup has been considered in [28], where the authors established a bound on the distance between a three-dimensional vector consisting of a rescaled number of edges, a rescaled number of two-stars and a rescaled number of triangles in a $G(n, p)$ graph and a three-dimensional Gaussian vector. We first compare our process to a pre-limiting Gaussian processes with paths in $D([0, 1], \mathbb{R})$ and bound the distance between the two in Theorem 6.3. Then, in Theorem 6.4, we bound the distance of our process from a continuous Gaussian process.

6.1. Introduction

Let us consider a Bernoulli random graph $G(n, p)$ on n vertices with edge probabilities p .

Let $I_{i,j} = I_{j,i}$ be the Bernoulli(p)-indicator that edge (i, j) is present in this graph. These indicators, for $(i, j) \in \{1, \dots, n\}^2$ are independent. We will look at a process representing at each $t \in [0, 1]$ the re-scaled total number of edges in the graph formed out of the given Bernoulli random graph by considering only its first $\lfloor nt \rfloor$ vertices and the edges between them:

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

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

Remark 6.1. Note that, for all $t \in [0, 1]$, $\mathbb{E}\mathbf{T}_n(t) = \frac{\lfloor nt \rfloor - 2}{n^2} \binom{\lfloor nt \rfloor}{2} p$. Furthermore, note that, by an argument similar to that of [28, Section 5], the variance of $\mathbf{T}_n(t) - \mathbb{E}\mathbf{T}_n(t)$ is given by

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

Remark 6.2. While one might often be interested in the asymptotic properties of the number of two-stars or the number of triangles in the above mentioned Bernoulli random graph, it is difficult to prove useful exchangeable-pair conditions for those. Specifically, one encounters several problems which were overcome in [28] (in the finite-dimensional setting) with the so-called embedding method (introduced in [28, Section 3]). Indeed, the problem becomes easier when the distribution of the number of edges, the number of two-stars and the number of triangles is considered jointly rather than considering the one-dimensional marginal laws separately. We therefore leave the problem of establishing bounds on the functional approximations of those statistics for further research, related to the multivariate method of exchangeable pairs (cf. Remark 4.4).

6.2. Exchangeable pair setup

We now construct an exchangeable pair, as in [28], by picking (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 put:

$$\mathbf{T}'_n(t) = \mathbf{T}_n(t) - \frac{\lfloor nt \rfloor - 2}{n^2} (I_{I,J} - I'_{I,J}) \mathbb{1}_{[I/n,1] \cap [J/n,1]}(t).$$

We also let $\mathbf{Y}'_n(t) = \mathbf{T}'_n(t) - \mathbb{E}\mathbf{T}_n(t)$ 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. We note that for any $f \in M$, as defined in Section 2,

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

and so

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

where

$$\lambda_n = \frac{2}{n(n-1)}. \quad (6.1)$$

Therefore, condition (4.1) is satisfied with λ_n of (6.1) and $R_f = 0$.

6.3. A pre-limiting process

Let $Z_{ii} = 0$ for all $i \in [n] = \{1, \dots, n\}$. Furthermore, assume that the collection $\{Z_{ij} : i, j \in [n], i \neq j\}$ is jointly centred Gaussian with the following covariance structure:

$$\mathbb{E}Z_{ij}Z_{kl} = \begin{cases} \frac{p(1-p)}{2n^4}, & i = k, j = l, i \neq j \\ 0, & \text{otherwise.} \end{cases}$$

Let

$$\mathbf{D}_n = (\lfloor nt \rfloor - 2) \sum_{i,j=1}^{\lfloor nt \rfloor} Z_{ij}, \quad t \in [0, 1].$$

6.4. Distance from the pre-limiting process

We first give a theorem providing a bound on the distance between \mathbf{Y}_n and the pre-limiting piecewise constant Gaussian process.

Theorem 6.3. *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 $g \in M$,*

$$|\mathbb{E}g(\mathbf{Y}_n) - \mathbb{E}g(\mathbf{D}_n)| \leq \frac{1}{8} \|g\|_M n^{-1}.$$

In **Step 1** of the proof, which is based on Theorem 4.1, we estimate term ϵ_1 thereof. It involves bounding the third moment of $\|\mathbf{Y}_n - \mathbf{Y}'_n\|$ for \mathbf{Y}'_n constructed in Section 6.2. In **Step 2** we treat ϵ_2 , which requires some more involved calculations. Term ϵ_3 is equal to zero as R_f of Section 6.2 is equal to zero.

Proof of Theorem 6.3. We adopt the notation of sections 6.1, 6.2, 6.3. We will apply Theorem 4.1. Let $g \in M$ and let $f = \phi(g)$, as defined in (3.3).

Step 1. First note that

$$\mathbb{E} [\|\mathbf{Y}_n - \mathbf{Y}'_n\|^3] \leq \mathbb{E} \left[\frac{(n-2)^3}{n^6} |I_{I,J} - I'_{I,J}|^3 \right] \leq \frac{(n-2)^3}{n^6}$$

where the second inequality follows because $|I_{I,J} - I'_{I,J}| \leq 1$. Therefore,

$$\epsilon_1 \leq \frac{\|g\|_M}{12\lambda_n} \mathbb{E} [\|\mathbf{Y}_n - \mathbf{Y}'_n\|^3] \leq \frac{\|g\|_M}{24n}. \quad (6.2)$$

Step 2. For ϵ_2 in Theorem 4.1, we wish to bound

$$\begin{aligned} \epsilon_2 &= \left| \frac{1}{2\lambda_n} \mathbb{E} D^2 f(\mathbf{Y}_n) [\mathbf{Y}_n - \mathbf{Y}'_n, \mathbf{Y}_n - \mathbf{Y}'_n] - \mathbb{E} D^2 f(\mathbf{Y}_n) [\mathbf{D}_n, \mathbf{D}_n] \right| \\ &= \left| \frac{n(n-1)}{4} \mathbb{E} D^2 f(\mathbf{Y}_n) [\mathbf{T}_n - \mathbf{T}'_n, \mathbf{T}_n - \mathbf{T}'_n] - \mathbb{E} D^2 f(\mathbf{Y}_n) [\mathbf{D}_n, \mathbf{D}_n] \right|. \end{aligned}$$

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:

$$\mathbf{T}_n^{ij}(t) = \mathbf{T}_n(t) - \frac{\lfloor nt \rfloor - 2}{n^2} (I_{ij} - I'_{ij}) \mathbb{1}_{[i/n, 1] \cap [j/n, 1]}(t)$$

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

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

$$\begin{aligned} \epsilon_2 &= \left| \frac{n(n-1)}{4} \mathbb{E} D^2 f(\mathbf{Y}_n) [\mathbf{T}_n - \mathbf{T}'_n, \mathbf{T}_n - \mathbf{T}'_n] \right. \\ &\quad \left. - \sum_{j,k=1}^n \mathbb{E} D^2 f(\mathbf{Y}_n) \left[\sum_{i=1}^n Z_{ik} (\lfloor n \cdot \rfloor - 2) \mathbb{1}_{[i/n, 1] \cap [k/n, 1]}, \sum_{i=1}^n Z_{ij} (\lfloor n \cdot \rfloor - 2) \mathbb{1}_{[i/n, 1] \cap [j/n, 1]} \right] \right| \\ &= \left| \frac{1}{4n^4} \sum_{1 \leq i \neq j \leq n} \mathbb{E} \left\{ (I_{ij} - 2pI_{ij} + p) D^2 f(\mathbf{Y}_n) [(\lfloor n \cdot \rfloor - 2) \mathbb{1}_{[i/n, 1] \cap [j/n, 1]}, (\lfloor n \cdot \rfloor - 2) \mathbb{1}_{[i/n, 1] \cap [j/n, 1]}] \right\} \right. \\ &\quad \left. - \sum_{i,j=1}^n \left\{ \mathbb{E} (Z_{ij})^2 \mathbb{E} D^2 f(\mathbf{Y}_n) [(\lfloor n \cdot \rfloor - 2) \mathbb{1}_{[i/n, 1] \cap [j/n, 1]}, (\lfloor n \cdot \rfloor - 2) \mathbb{1}_{[i/n, 1] \cap [j/n, 1]}] \right\} \right| \\ &= \left| \sum_{1 \leq i \neq j \leq n} \mathbb{E} \left\{ \left(\frac{1}{4n^4} (I_{ij} - 2pI_{ij} + p) - \mathbb{E} (Z_{ij})^2 \right) \right\} \right| \end{aligned} \quad (6.3)$$

$$\begin{aligned}
& \cdot D^2 f(\mathbf{Y}_n) \left[(\lfloor n \cdot \rfloor - 2) \mathbb{1}_{[i/n, 1] \cap [j/n, 1]}, (\lfloor n \cdot \rfloor - 2) \mathbb{1}_{[i/n, 1] \cap [j/n, 1]} \right] \Bigg\} \Bigg| \\
& = \left| \sum_{1 \leq i \neq j \leq n} \mathbb{E} \left\{ \frac{1}{4n^4} (I_{ij} - 2pI_{ij} + p) \right. \right. \\
& \quad \cdot (D^2 f(\mathbf{Y}_n) - D^2 f(\mathbf{Y}_n^{ij})) \left. \left. [(\lfloor n \cdot \rfloor - 2) \mathbb{1}_{[i/n, 1] \cap [j/n, 1]}, (\lfloor n \cdot \rfloor - 2) \mathbb{1}_{[i/n, 1] \cap [j/n, 1]}] \right\} \right| \\
& \leq \frac{\|g\|_M}{12n^2} \sum_{1 \leq i \neq j \leq n} \mathbb{E} [(I_{ij} - 2pI_{ij} + p) \|\mathbf{Y}_n - \mathbf{Y}_n^{ij}\|], \tag{6.4}
\end{aligned}$$

where (6.4) follows from Proposition 3.2. Now,

$$\|\mathbf{Y}_n - \mathbf{Y}_n^{ij}\| \leq \frac{1}{n} |I_{ij} - I'_{ij}|$$

and so, by (6.4),

$$\epsilon_2 \leq \frac{\|g\|_M}{12n^3} \sum_{1 \leq i \neq j \leq n} \mathbb{E} [|I_{i,j} - 2pI_{i,j} + p| |I_{ij} - I'_{ij}|] \leq \frac{\|g\|_M}{12n}, \tag{6.5}$$

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

Using Theorem 4.1 together with (6.5) and (6.2) gives the desired result. \square

6.5. Distance from the continuous process

We now establish a bound on the rate of convergence of \mathbf{Y}_n to a continuous Gaussian process whose covariance is the limit of the covariance of \mathbf{D}_n . We do this by bounding the distance between \mathbf{D}_n and the continuous process via the Brownian modulus of continuity and using Theorem 6.3.

Theorem 6.4. *Let \mathbf{Y}_n be defined as in Subsection 6.1 and let \mathbf{Z} be defined by:*

$$\mathbf{Z}(t) = \left(\frac{p(1-p)}{2} \right)^{1/2} t \mathbf{B}(t^2), \quad t \in [0, 1],$$

where \mathbf{B} is a standard Brownian Motion. Then, for any $g \in M^2$:

$$|\mathbb{E}g(\mathbf{Y}_n) - \mathbb{E}g(\mathbf{Z})| \leq \|g\|_{M^2} \left(9n^{-1} + 90n^{-1/2} \sqrt{\log n} \right).$$

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.*

In **Step 1** of the proof of Theorem 6.4, we provide a coupling between \mathbf{D}_n and a standard Brownian motion. Using this Brownian motion, we 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 bound the first two moments of the supremum distance between them, using the Brownian modulus of continuity. In **Step 3** we use those bounds together with the Mean Value Theorem to obtain Theorem 6.4.

Proof of Theorem 6.4.

Step 1. Let \mathbf{B} be a standard Brownian motion and let \mathbf{Z}_n be defined by:

$$\mathbf{Z}_n(t) = \frac{(\lfloor nt \rfloor - 2) \sqrt{p(1-p)}}{n\sqrt{2}} \mathbf{B} \left(\frac{\lfloor nt \rfloor (\lfloor nt \rfloor - 1)}{n^2} \right).$$

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

$$\begin{aligned} \mathbb{E} \mathbf{D}_n(t) \mathbf{D}_n(u) &= (\lfloor nt \rfloor - 2)(\lfloor nu \rfloor - 2) \lfloor n(t \wedge u) \rfloor (\lfloor n(t \wedge u) \rfloor - 1) \frac{p(1-p)}{2n^4} \\ &= \mathbb{E} \mathbf{Z}_n(t) \mathbf{Z}_n(u). \end{aligned} \quad (6.6)$$

Step 2. We let \mathbf{Z} and \mathbf{Z}_n be coupled in such a way that \mathbf{Z} is constructed as in Theorem 6.4, using the same Brownian Motion \mathbf{B} , as the one used in the construction of \mathbf{Z}_n . In Lemma 7.3, proved in the appendix, we derive bounds for moments of the supremum distance between \mathbf{Z} and \mathbf{Z}_n :

$$\begin{aligned} \mathbb{E} \|\mathbf{Z}_n - \mathbf{Z}\| &\leq \frac{3}{\sqrt{2n}} + \frac{36\sqrt{\log n}}{\sqrt{n}} \\ \mathbb{E} \|\mathbf{Z}_n - \mathbf{Z}\|^2 &\leq \frac{5}{n^2} + \frac{49 \log n}{n} \\ \mathbb{E} \|\mathbf{Z}\|^2 &\leq \frac{1}{2}. \end{aligned} \quad (6.7)$$

Step 3. We note that $\|Dg(w)\| \leq \|g\|_{M^2}(1 + \|w\|)$ and therefore, by (6.7):

$$\begin{aligned} |\mathbb{E}g(\mathbf{Z}) - \mathbb{E}g(\mathbf{D}_n)| &\stackrel{\text{MVT}}{\leq} \mathbb{E} \left[\sup_{c \in [0,1]} \|Dg(\mathbf{Z} + c(\mathbf{Z}_n - \mathbf{Z}))\| \|\mathbf{Z} - \mathbf{Z}_n\| \right] \\ &\leq \|g\|_{M^2} \mathbb{E} \left[\sup_{c \in [0,1]} (1 + \|\mathbf{Z} + c(\mathbf{Z}_n - \mathbf{Z})\|) \|\mathbf{Z} - \mathbf{Z}_n\| \right] \\ &\leq \|g\|_{M^2} \mathbb{E} [\|\mathbf{Z} - \mathbf{Z}_n\| + \|\mathbf{Z}\| \|\mathbf{Z} - \mathbf{Z}_n\| + \|\mathbf{Z} - \mathbf{Z}_n\|^2] \\ &\leq \|g\|_{M^2} \left[\mathbb{E} \|\mathbf{Z} - \mathbf{Z}_n\| + \sqrt{\mathbb{E} \|\mathbf{Z}\|^2 \mathbb{E} \|\mathbf{Z} - \mathbf{Z}_n\|^2} + \mathbb{E} \|\mathbf{Z} - \mathbf{Z}_n\|^2 \right] \\ &\leq \|g\|_{M^2} \left(\left(\frac{3 + \sqrt{5}}{\sqrt{2}} + 5 \right) n^{-1} + 90n^{-1/2} \sqrt{\log n} \right), \end{aligned}$$

which, together with Theorem 6.3 gives the desired result. \square

Remark 6.6. The representation of \mathbf{Z} in terms of a Brownian motion comes from a careful analysis of the limiting covariance of \mathbf{D}_n . Indeed, (6.6) provides an explicit derivation of the covariance, which converges to the covariance of \mathbf{Z} .

Acknowledgements

The author would like to thank Gesine Reinert for many helpful discussions and comments on the early versions of this work. The author is also grateful to Andrew Barbour, Christian Döbler, Ivan Nourdin, Giovanni Peccati and Yvik Swan for many interesting discussions. This research has been supported by an EPSRC PhD Studentship at the University of Oxford and the FNR grant FoRGES (R-AGR- 3376-10) at the University of Luxembourg.

References

- [1] A.D. Barbour. Stein's Method for Diffusion Approximation. *Probability Theory and Related Fields*, 84:297–322, 1990.
- [2] A.D. Barbour and S. Janson. A functional combinatorial central limit theorem. *Electronic Journal of Probability*, 14(81):2352–2370, 2009.
- [3] E. Besançon, L. Decreusefond, and P. Moyal. Stein's method for diffusive limit of Markov processes. arXiv:1805.01691, 2018.
- [4] P. Billingsley. *Convergence of Probability Measures*. Wiley, New York, 1968.

- [5] E. Bolthausen. An estimate of the remainder in a combinatorial central limit theorem. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 66(3):379–386, 1984.
- [6] S. Bourguin and S. Campese. Approximation of Hilbert-valued Gaussian measures on Dirichlet structures. arXiv:1905.05127, 2019.
- [7] S. Chatterjee, P. Diaconis, and E. Meckes. Exchangeable pairs and Poisson approximation. *Probability Surveys*, 2:64–106, 2005.
- [8] 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.
- [9] S. Chatterjee and E. Meckes. Multivariate normal approximation using exchangeable pairs. *ALEA Lat. Am. J. Probab. Math. Stat.*, 4:257–283, 2008.
- [10] Louis H.Y. Chen and Xiao Fang. On the error bound in a combinatorial central limit theorem. *Bernoulli*, 21(1):335–359, 02 2015.
- [11] L. Coutin and L. Decreusefond. Stein's method for Brownian Approximations. *Communications on Stochastic Analysis*, 7(3):349–372, 2013.
- [12] L. Coutin and L. Decreusefond. Higher order expansions via Stein's method. *Communications on Stochastic Analysis*, 8(2):155–168, 2014.
- [13] L. Coutin and L. Decreusefond. Stein's method for rough paths. *Potential Analysis*, 2019. in press.
- [14] Ch. Döbler. Stein's method of exchangeable pairs for the Beta distribution and generalizations. *Electronic Journal of Probability*, 20:34 pp., 2015.
- [15] R. M. Dudley. *Real Analysis and Probability*:. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2 edition, 2002.
- [16] S.N. Ethier and T.G. Kurtz. *Markov processes: characterization and convergence*. Wiley, New York, 1986.
- [17] 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.
- [18] L. Goldstein. Berry-Esseen bounds for combinatorial central limit theorems and pattern occurrences, using zero and size biasing. *Journal of Applied Probability*, 42(3):661–683, 2005.
- [19] S.-T. Ho and L.H.Y. Chen. An L_p Bound for the remainder in a combinatorial central limit theorem. *The Annals of Probability*, 6(2):231–249, 04 1978.
- [20] 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.
- [21] M.J. Kasprzak. Diffusion approximations via Stein's method and time changes. arXiv:1701.07633, 2017.
- [22] M.J. Kasprzak. Stein's method for multivariate Brownian approximations of sums under dependence. *Stochastic Processes and their Applications*, 2020. in press.
- [23] M.J. Kasprzak, A.B. Duncan, and S.J. Vollmer. Note on A. Barbour's paper on Stein's method for diffusion approximations. *Electronic Communications in Probability*, 22(23):1–8, 2017.
- [24] E. Meckes. *On Stein's method for multivariate normal approximation*, volume 5 of *Collections*, pages 153–178. Institute of Mathematical Statistics, Beachwood, Ohio, USA, 2009.
- [25] K. Neammanee and N. Rerkruthairat. An improvement of a uniform bound on a combinatorial central limit theorem. *Communications in Statistics. Theory and Methods*, 41(9):1590–1602, 2012.
- [26] I. Nourdin and G. Peccati. *Normal Approximations with Malliavin Calculus*. Cambridge tracts in Mathematics. Cambridge University Press, 2012.
- [27] 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.
- [28] 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.
- [29] 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. *The Annals of Applied Probability*, 7(4):1080–1105, 1997.
- [30] A. Röllin. Translated poisson approximation using exchangeable pair couplings. *The Annals of Applied Probability*, 17(5/6):1596–1614, 2007.
- [31] 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.
- [32] 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.

- [33] Ch. Stein. *Approximate Computation of Expectations*. Institute of Mathematical Statistics Lecture Notes, Monograph Series, 7. Hayward, Calif. Institute of Mathematical Statistics, 1986.
- [34] A. Wald and J. Wolfowitz. On a test whether two samples are from the same population. *Annals of Mathematical Statistics*, 11(2):147–162, 06 1940.
- [35] A. Wald and J. Wolfowitz. Statistical tests based on permutations of the observations. *Annals of Mathematical Statistics*, 15:358–372, 1944.

7. Appendix - technical details of the proofs of Theorems 5.1, and 6.4

7.1. Technical details of the proof of Theorem 5.1

Lemma 7.1. *In the setup of Theorem 6.3 and for ϵ_2 defined by Theorem 4.1,*

$$\begin{aligned} \epsilon_2 &= \left| \frac{1}{2\lambda_n} \mathbb{E} D^2 f(\mathbf{Y}_n) [\mathbf{Y}_n - \mathbf{Y}'_n, \mathbf{Y}_n - \mathbf{Y}'_n] - \mathbb{E} D^2 f(\mathbf{Y}_n) [\mathbf{D}_n, \mathbf{D}_n] \right| \\ &\leq A + B, \end{aligned}$$

for

$$\begin{aligned} A &= \left| \frac{1}{n(n-1)s_n^2} \sum_{\substack{1 \leq i,j,k,l \leq n \\ i \neq j, k \neq l}} \mathbb{E} \left\{ \left[\frac{(X_{ik} - X_{il})^2}{2n} - \frac{\hat{Z}_i^2}{n-1} \right] (D^2 f(\mathbf{Y}_{n,ijkl}) - D^2 f(\mathbf{Y}_n^{ijkl})) [\mathbb{1}_{[i/n,1]}, \mathbb{1}_{[i/n,1]}] \right\} \right. \\ &\quad \left. + \frac{1}{n(n-1)s_n^2} \sum_{\substack{1 \leq i,j,k,l \leq n \\ i \neq j, k \neq l}} \mathbb{E} \left\{ \left[\frac{(X_{ik} - X_{il})(X_{jl} - X_{jk})}{2n} - \hat{Z}_i \hat{Z}_j \right] \right. \right. \\ &\quad \left. \left. \cdot (D^2 f(\mathbf{Y}_{n,ijkl}) - D^2 f(\mathbf{Y}_n^{ijkl})) [\mathbb{1}_{[i/n,1]}, \mathbb{1}_{[j/n,1]}] \right\} \right|, \\ B &= \left| \frac{1}{n(n-1)s_n^2} \sum_{\substack{1 \leq i,j,k,l \leq n \\ i \neq j, k \neq l}} \mathbb{E} \left\{ \left[\frac{(X_{ik} - X_{il})^2}{2n} - \frac{\hat{Z}_i^2}{n-1} \right] D^2 f(\mathbf{Y}_n^{ijkl}) [\mathbb{1}_{[i/n,1]}, \mathbb{1}_{[i/n,1]}] \right\} \right. \\ &\quad \left. + \frac{1}{n(n-1)s_n^2} \sum_{\substack{1 \leq i,j,k,l \leq n \\ i \neq j, k \neq l}} \mathbb{E} \left\{ \left[\frac{(X_{ik} - X_{il})(X_{jl} - X_{jk})}{2n} - \hat{Z}_i \hat{Z}_j \right] D^2 f(\mathbf{Y}_n^{ijkl}) [\mathbb{1}_{[i/n,1]}, \mathbb{1}_{[j/n,1]}] \right\} \right|. \end{aligned}$$

Proof. Note that

$$\begin{aligned} \epsilon_2 &= \left| \frac{1}{2\lambda_n} \mathbb{E} D^2 f(\mathbf{Y}_n) [(\mathbf{Y}_n - \mathbf{Y}'_n), \mathbf{Y}_n - \mathbf{Y}'_n] - \mathbb{E} D^2 f(\mathbf{Y}_n) [\mathbf{D}_n, \mathbf{D}_n] \right| \\ &= \left| \frac{n-1}{4} \mathbb{E} D^2 f(\mathbf{Y}_n) [\mathbf{Y}_n - \mathbf{Y}'_n, \mathbf{Y}_n - \mathbf{Y}'_n] - \mathbb{E} D^2 f(\mathbf{Y}_n) [\mathbf{D}_n, \mathbf{D}_n] \right| \end{aligned} \tag{7.1}$$

and

$$\begin{aligned}
& \frac{n-1}{4} \mathbb{E} D^2 f(\mathbf{Y}_n) [\mathbf{Y}_n - \mathbf{Y}'_n, \mathbf{Y}_n - \mathbf{Y}'_n] - \mathbb{E} D^2 f(\mathbf{Y}_n) [\mathbf{D}_n, \mathbf{D}_n] \\
&= \frac{1}{2ns_n^2} \sum_{i,j=1}^n \mathbb{E} \{ (X_{i\pi(i)} - X_{i\pi(j)})^2 D^2 f(\mathbf{Y}_n) [\mathbb{1}_{[i/n,1]} \mathbb{1}_{[i/n,1]}] \} \\
& \quad + \frac{1}{2ns_n^2} \sum_{i,j=1}^n \mathbb{E} \{ (X_{i\pi(i)} - X_{i\pi(j)})(X_{j\pi(j)} - X_{j\pi(i)}) D^2 f(\mathbf{Y}_n) [\mathbb{1}_{[i/n,1]}, \mathbb{1}_{[j/n,1]}] \} - \mathbb{E} D^2 f(\mathbf{Y}_n) [\mathbf{D}_n, \mathbf{D}_n] \\
&= \frac{1}{2n^2(n-1)s_n^2} \sum_{\substack{1 \leq i,j,k,l \leq n \\ i \neq j, k \neq l}} \mathbb{E} \{ (X_{ik} - X_{il})^2 \cdot D^2 f(\mathbf{Y}_n) [\mathbb{1}_{[i/n,1]} \mathbb{1}_{[i/n,1]}] \mid \pi(i) = k, \pi(j) = l \} \\
& \quad + \frac{1}{n(n-1)s_n^2} \sum_{\substack{1 \leq i,j,k,l \leq n \\ i \neq j, k \neq l}} \mathbb{E} \left\{ \left[\frac{(X_{ik} - X_{il})(X_{jl} - X_{jk})}{2n} \right] D^2 f(\mathbf{Y}_n) [\mathbb{1}_{[i/n,1]}, \mathbb{1}_{[j/n,1]}] \mid \pi(i) = k, \pi(j) = l \right\} \\
& \quad - \frac{1}{s_n^2} \sum_{1 \leq i \neq j \leq n} \mathbb{E} [\hat{Z}_i \hat{Z}_j] \mathbb{E} D^2 f(\mathbf{Y}_n) [\mathbb{1}_{[i/n,1]}, \mathbb{1}_{[j/n,1]}] - \frac{1}{(n-1)s_n^2} \sum_{1 \leq i \neq j \leq n} \mathbb{E} [\hat{Z}_i^2] \mathbb{E} D^2 f(\mathbf{Y}_n) [\mathbb{1}_{[i/n,1]}, \mathbb{1}_{[i/n,1]}] \\
&= \frac{1}{2n^2(n-1)s_n^2} \sum_{\substack{1 \leq i,j,k,l \leq n \\ i \neq j, k \neq l}} \mathbb{E} \{ (X_{ik} - X_{il})^2 D^2 f(\mathbf{Y}_{n,ijkl}) [\mathbb{1}_{[i/n,1]} \mathbb{1}_{[i/n,1]}] \} \\
& \quad + \frac{1}{n(n-1)s_n^2} \sum_{\substack{1 \leq i,j,k,l \leq n \\ i \neq j, k \neq l}} \mathbb{E} \left\{ \frac{(X_{ik} - X_{il})(X_{jl} - X_{jk})}{2n} D^2 f(\mathbf{Y}_{n,ijkl}) [\mathbb{1}_{[i/n,1]}, \mathbb{1}_{[j/n,1]}] \right\} \\
& \quad - \frac{1}{n(n-1)s_n^2} \sum_{\substack{1 \leq i,j,k,l \leq n \\ i \neq j, k \neq l}} \mathbb{E} [\hat{Z}_i \hat{Z}_j] \mathbb{E} D^2 f(\mathbf{Y}_{n,ijkl}) [\mathbb{1}_{[i/n,1]}, \mathbb{1}_{[j/n,1]}] \\
& \quad - \frac{1}{n(n-1)s_n^2} \sum_{\substack{1 \leq i,j,k,l \leq n \\ i \neq j, k \neq l}} \mathbb{E} [\hat{Z}_i^2] \mathbb{E} D^2 f(\mathbf{Y}_{n,ijkl}) [\mathbb{1}_{[i/n,1]}, \mathbb{1}_{[i/n,1]}]. \tag{7.2}
\end{aligned}$$

Now, the lemma follows by taking the absolute value in (7.2) and combining it with (7.1). \square

Lemma 7.2. For A of (5.9),

$$\begin{aligned}
A \leq & \frac{\|g\|_{M^1}}{n^3(n-1)s_n^3} \sum_{1 \leq i,j,k,l,u \leq n} \left\{ \mathbb{E} |X_{ik}|^3 + 5\mathbb{E} |X_{ik}| \mathbb{E} |X_{il}|^2 + 7\mathbb{E} |X_{ik}|^2 \mathbb{E} |X_{jl}| \right. \\
& + 5\mathbb{E} |X_{ik}|^2 \mathbb{E} |X_{jk}| + 16\mathbb{E} |X_{ik}| \mathbb{E} |X_{il}| \mathbb{E} |X_{jl}| + 2\mathbb{E} |X_{iu}| \mathbb{E} |X_{ik}| \mathbb{E} |X_{il}| \\
& + 4\mathbb{E} |X_{iu}| \mathbb{E} |X_{il}| \mathbb{E} |X_{jk}| + 6\mathbb{E} |X_{uk}| \mathbb{E} |X_{ik}| \mathbb{E} |X_{jl}| + 2\mathbb{E} |X_{uk}| \mathbb{E} |X_{ik}| \mathbb{E} |X_{jk}| \\
& \left. + \frac{1}{n} (2\mathbb{E} |X_{ik}| + 2\mathbb{E} |X_{jl}| + 2\mathbb{E} |X_{uk}| + 2\mathbb{E} |X_{ul}|) \cdot \sum_{r=1}^n (\mathbb{E} |X_{ir}|^2 + |c_{ir} c_{jr}|) \right\}.
\end{aligned}$$

Proof. Let us adopt the notation of the proof of Theorem 5.1. Define index sets $\mathcal{I} = \{i, j, \pi^{-1}(k), \pi^{-1}(l)\}$ and $\mathcal{J} = \{k, l, \pi(i), \pi(j)\}$. Then, letting $\mathbf{S} = \frac{1}{s_n} \sum_{i' \notin \mathcal{I}} X_{i'\pi(i')} \mathbb{1}_{[i'/n,1]}$, we can write:

$$\mathbf{Y}_{n,ijkl} = \mathbf{S} + \frac{1}{s_n} \sum_{i' \in \mathcal{I}} X_{i'\pi_{ijkl}(i')} \mathbb{1}_{[i'/n,1]}, \quad \mathbf{Y}_n^{ijkl} = \mathbf{S} + \frac{1}{s_n} \sum_{i' \in \mathcal{I}} X_{i'\pi(i')}^{ijkl} \mathbb{1}_{[i'/n,1]}.$$

Since \mathbf{S} depends only on the components of \mathbb{X} outside the square $\mathcal{I} \times \mathcal{J}$ and $\{\pi(i) : i \notin \mathcal{I}\}$, \mathbf{S} is independent of:

$$\left\{ X_{il}, X_{jk}, X_{ik}, X_{jl}, \sum_{i' \in \mathcal{I}} X_{i'\pi_{ijkl}(i')}, \sum_{i' \in \mathcal{I}} X_{i'\pi(i')}^{ijkl} \right\},$$

given $\pi^{-1}(k), \pi^{-1}(l), \pi(i), \pi(j)$.

Note that, by Proposition 3.2,

$$\begin{aligned}
A &\leq \frac{\|g\|_{M^1}}{n(n-1)s_n^2} \sum_{\substack{1 \leq i,j,k,l \leq n \\ i \neq j, k \neq l}} \mathbb{E} \left\{ \left\| \mathbf{Y}_{n,ijkl} - \mathbf{Y}_n^{ijkl} \right\| \left(\left| \frac{(X_{ik} - X_{il})^2}{2n} - \frac{\mathbb{E}\hat{Z}_i^2}{n-1} \right| \right. \right. \\
&\quad \left. \left. + \left| \frac{(X_{ik} - X_{il})(X_{jl} - X_{jk})}{2n} - \mathbb{E}(\hat{Z}_i \hat{Z}_j) \right| \right) \right\} \\
&\leq \frac{\|g\|_{M^1}}{n(n-1)s_n^3} \sum_{\substack{1 \leq i,j,k,l \leq n \\ i \neq j, k \neq l}} \mathbb{E} \left\{ \sum_{i' \in \mathcal{I}} \left| X_{i', \pi_{ijkl}(i')} - X_{i' \pi(i')}^{ijkl} \right| \right. \\
&\quad \cdot \left(\left| \frac{(X_{ik} - X_{il})^2}{2n} - \frac{\mathbb{E}\hat{Z}_i^2}{n-1} \right| + \left| \frac{(X_{ik} - X_{il})(X_{jl} - X_{jk})}{2n} - \mathbb{E}(\hat{Z}_i \hat{Z}_j) \right| \right) \Big\} \\
&\leq \frac{\|g\|_{M^1}}{n(n-1)s_n^3} \sum_{\substack{1 \leq i,j,k,l \leq n \\ i \neq j, k \neq l}} \mathbb{E} \left\{ \left(|X_{ik} - X_{i\pi(i)}^{ijkl}| + |X_{jl} - X_{j\pi(j)}^{ijkl}| + |X_{\pi^{-1}(k),k} - X_{\pi^{-1}(k),k}^{ijkl}| + |X_{\pi^{-1}(l),l} - X_{\pi^{-1}(l),l}^{ijkl}| \right) \right. \\
&\quad \cdot \left(\left| \frac{(X_{ik} - X_{il})^2}{2n} - \frac{\mathbb{E}\hat{Z}_i^2}{n-1} \right| + \left| \frac{(X_{ik} - X_{il})(X_{jl} - X_{jk})}{2n} - \mathbb{E}(\hat{Z}_i \hat{Z}_j) \right| \right) \Big\} \\
&\leq \frac{\|g\|_{M^1}}{2n(n-1)^2 s_n^3} \sum_{\substack{1 \leq i,j,k,l \leq n \\ i \neq j, k \neq l}} \mathbb{E} \left\{ \left(|X_{ik}| + |X_{i,\pi(i)}^{ijkl}| + |X_{j,l}| + |X_{j,\pi(j)}^{ijkl}| + |X_{\pi^{-1}(k),k}| + |X_{\pi^{-1}(k),k}^{ijkl}| \right. \right. \\
&\quad + |X_{\pi^{-1}(l),l}| + |X_{\pi^{-1}(l),l}^{ijkl}| \Big) \left(|X_{ik}|^2 + |X_{il}|^2 + 2|X_{ik}X_{il}| + 2\mathbb{E}|\hat{Z}_i|^2 + |X_{ik}X_{jl}| + |X_{ik}X_{jk}| \right. \\
&\quad \left. \left. + |X_{il}X_{jl}| + |X_{il}X_{jk}| + 2(n-1) \left| \mathbb{E}(\hat{Z}_i \hat{Z}_j) \right| \right) \right\} \\
&\leq \frac{\|g\|_{M^1}}{2n^2(n-1)^2 s_n^3} \sum_{\substack{1 \leq i,j,k,l,u \leq n \\ i \neq j, k \neq l}} \mathbb{E} \left\{ \left(|X_{ik}| + \mathbb{E}|X_{iu}| + |X_{j,l}| + \mathbb{E}|X_{ju}| + |X_{uk}| + \mathbb{E}|X_{uk}| + |X_{ul}| + \mathbb{E}|X_{ul}| \right) \right. \\
&\quad \cdot \left(|X_{ik}|^2 + |X_{il}|^2 + 2|X_{ik}X_{il}| + \frac{2}{n} \sum_{r=1}^n \mathbb{E}|X_{ir}|^2 + |X_{ik}X_{jl}| + |X_{ik}X_{jk}| \right. \\
&\quad \left. \left. + |X_{il}X_{jl}| + |X_{il}X_{jk}| + \frac{2}{n} \left| \sum_{r=1}^n c_{ir}c_{jr} \right| \right) \right\} \\
&\leq \frac{\|g\|_{M^1}}{2n^2(n-1)^2 s_n^3} \sum_{1 \leq i,j,k,l,u \leq n} \left\{ \mathbb{E}|X_{ik}|^3 + 5\mathbb{E}|X_{ik}|^2\mathbb{E}|X_{il}| + 9\mathbb{E}|X_{ik}|^2\mathbb{E}|X_{jl}| + 5\mathbb{E}|X_{ik}|^2\mathbb{E}|X_{jk}| \right. \\
&\quad + 14\mathbb{E}|X_{ik}|\mathbb{E}|X_{il}|\mathbb{E}|X_{jl}| + 2\mathbb{E}|X_{ik}|\mathbb{E}|X_{il}|\mathbb{E}|X_{iu}| + 6\mathbb{E}|X_{ik}|\mathbb{E}|X_{jl}|\mathbb{E}|X_{iu}| + 2\mathbb{E}[|X_{uk}||X_{ik}|^2] \\
&\quad + 4\mathbb{E}|X_{uk}X_{ik}|\mathbb{E}|X_{il}| + 4\mathbb{E}|X_{uk}X_{ik}|\mathbb{E}|X_{jl}| + 8\mathbb{E}|X_{uk}|\mathbb{E}|X_{ik}|\mathbb{E}|X_{jl}| + 2\mathbb{E}|X_{uk}X_{ik}X_{jk}| \\
&\quad \left. + 2\mathbb{E}|X_{uk}|\mathbb{E}|X_{ik}|\mathbb{E}|X_{jk}| + \frac{2}{n} \mathbb{E}(2|X_{ik}| + 6|X_{jl}|) \cdot \left(\sum_{r=1}^n \mathbb{E}|X_{ir}|^2 + \left| \sum_{r=1}^n c_{ir}c_{jr} \right| \right) \right\}.
\end{aligned}$$

which finishes the proof. \square

7.2. 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,*

$$\mathbb{E} \|\mathbf{Z}_n - \mathbf{Z}\| \leq \frac{3}{\sqrt{2n}} + \frac{36\sqrt{\log n}}{\sqrt{n}}$$

$$\begin{aligned}\mathbb{E} \|\mathbf{Z}_n - \mathbf{Z}\|^2 &\leq \frac{5}{n^2} + \frac{49 \log n}{n} \\ \mathbb{E} \|\mathbf{Z}\|^2 &\leq \frac{1}{2}.\end{aligned}$$

Proof. Note the following

1. By Doob's L^2 inequality,

$$\mathbb{E} \left[\sup_{t \in [0,1]} |\mathbf{B}(t^2)| \right] \leq 2 \quad \text{and} \quad \left| \frac{\lfloor nt \rfloor - 2}{n} - t \right| \leq \frac{3}{n}, \quad \text{for all } t \in [0, 1]. \quad (7.3)$$

2. Using [17, Lemma 3] and the fact that

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

we obtain

$$\mathbb{E} \left[\sup_{t \in [0,1]} \left| \mathbf{B} \left(\frac{\lfloor nt \rfloor (\lfloor nt \rfloor - 1)}{n^2} \right) - \mathbf{B}(t^2) \right| \right] \leq \frac{30\sqrt{3 \log(\frac{2n}{3})}}{n^{1/2} \sqrt{\pi \log(2)}}. \quad (7.4)$$

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

$$\begin{aligned}\mathbb{E} \|\mathbf{Z}_n - \mathbf{Z}\| &\leq \frac{\sqrt{p(1-p)}}{\sqrt{2}} \mathbb{E} \left[\sup_{t \in [0,1]} \left| \frac{\lfloor nt \rfloor - 2}{n} \mathbf{B} \left(\frac{\lfloor nt \rfloor (\lfloor nt \rfloor - 1)}{n^2} \right) - t \mathbf{B}(t^2) \right| \right] \\ &\leq \frac{\sqrt{p(1-p)}}{\sqrt{2}} \left\{ \mathbb{E} \left[\sup_{t \in [0,1]} \left| \left(\frac{\lfloor nt \rfloor - 2}{n} - t \right) \mathbf{B}(t^2) \right| \right] \right. \\ &\quad \left. + \mathbb{E} \left[\sup_{t \in [0,1]} \frac{\lfloor nt \rfloor - 2}{n} \left| \mathbf{B} \left(\frac{\lfloor nt \rfloor (\lfloor nt \rfloor - 1)}{n^2} \right) - \mathbf{B}(t^2) \right| \right] \right\} \\ &\stackrel{(7.3), (7.4)}{\leq} \frac{\sqrt{p(1-p)}}{\sqrt{2}} \left(\frac{6}{n} + \frac{30\sqrt{3 \log n}}{n^{1/2} \sqrt{\pi \log(2)}} \right) \\ &\leq \frac{3}{\sqrt{2}n} + \frac{36\sqrt{\log n}}{\sqrt{n}}.\end{aligned}$$

Similarly, using Doob's L^2 inequality and [17, Lemma 3],

$$\begin{aligned}\mathbb{E} \|\mathbf{Z}_n - \mathbf{Z}\|^2 &\leq \frac{p(1-p)}{2} \mathbb{E} \left[\sup_{t \in [0,1]} \left| \frac{\lfloor nt \rfloor - 2}{n} \mathbf{B} \left(\frac{\lfloor nt \rfloor (\lfloor nt \rfloor - 1)}{n^2} \right) - t \mathbf{B}(t^2) \right|^2 \right] \\ &\leq \frac{p(1-p)}{2} \left(\mathbb{E} \left[\sup_{t \in [0,1]} \left| \left(\frac{\lfloor nt \rfloor - 2}{n} - t \right) \mathbf{B}(t^2) \right|^2 \right] + \mathbb{E} \left[\sup_{t \in [0,1]} \left| \mathbf{B} \left(\frac{\lfloor nt \rfloor (\lfloor nt \rfloor - 1)}{n^2} \right) - \mathbf{B}(t^2) \right|^2 \right] \right) \\ &\leq \frac{18p(1-p)}{n^2} + \frac{135p(1-p) \log(\frac{2}{3}n)}{\log(2) n} \\ &\leq \frac{5}{n^2} + \frac{49 \log n}{n}.\end{aligned}$$

Furthermore, by Doob's L^2 inequality,

$$\mathbb{E} \|\mathbf{Z}\|^2 \leq \frac{p(1-p)}{2} \mathbb{E} \left[\sup_{t \in [0,1]} t^2 |\mathbf{B}_1(t^2)|^2 \right] \leq 2p(1-p) \leq \frac{1}{2}.$$

This finishes the proof. \square