Electronic Journal of Probability

Electron. J. Probab. **24** (2019), no. 27, 1–35.
ISSN: 1083-6489 https://doi.org/10.1214/19-EJP284

Multivariate approximation in total variation using local dependence

A. D. Barbour* A. Xia†

Abstract

We establish two theorems for assessing the accuracy in total variation of multivariate discrete normal approximation to the distribution of an integer valued random vector W. The first is for sums of random vectors whose dependence structure is local. The second applies to random vectors W resulting from integrating the $\mathbb{Z}^d$-valued marks of a marked point process with respect to its ground process. The error bounds are of magnitude comparable to those given in [Rinott & Rotar (1996)], but now with respect to the stronger total variation distance. Instead of requiring the summands to be bounded, we make third moment assumptions. We demonstrate the use of the theorems in four applications: monochrome edges in vertex coloured graphs, induced triangles and 2-stars in random geometric graphs, the times spent in different states by an irreducible and aperiodic finite Markov chain, and the maximal points in different regions of a homogeneous Poisson point process.

Keywords: total variation approximation; Stein's method; local dependence; marked point process.
AMS MSC 2010: Primary 60F05, Secondary 60E15; 60G55; 60J27.
Submitted to EJP on July 17, 2018, final version accepted on February 27, 2019.
Supersedes arXiv:1807.06715.

1 Introduction

In this paper, we prove a general theorem that can be used to give bounds in total variation on the accuracy of multivariate discrete normal approximation to the distribution of a random vector W in $\mathbb{Z}^d$, when W is a sum of n random vectors whose dependence structure is local. Our setting is rather similar to that in [Rinott & Rotar (1996)]. In

*Institut für Mathematik, Universität Zürich, Winterthurerstrasse 190, 8057 Zürich, Switzerland; E-mail: a.d.barbour@math.unizh.ch. Work carried out in part at the University of Melbourne and at Monash University, and supported in part by Australian Research Council Grants Nos DP150101459 and DP150103588 and Centre of Excellence for Mathematical and Statistical Frontiers, CE140100049.

†School of Mathematics and Statistics, University of Melbourne, Parkville, VIC 3010, Australia; E-mail: aihuaxia@unimelb.edu.au. Work supported in part by Australian Research Council Grant No. DP150101459.

their paper, Stein's method is used to derive the accuracy, in terms of the convex sets metric, of multivariate normal approximation to suitably normalized sums of *bounded* random vectors; under reasonable conditions, error bounds of order $O(n^{-1/2}\log n)$ are obtained. [Fang (2014)] improves the order of the error to $O(n^{-1/2})$, using slightly different conditions, and also obtains optimal dependence on the dimension d. Here, we are interested in total variation distance bounds, so as to be able to approximate the probabilities of *arbitrary* sets. For random elements of $\mathbb{Z}^d$, this necessitates replacing the multivariate normal distribution by a discretized version. We use the d-dimensional discrete normal distribution $\mathrm{DN}_d(nc, n\Sigma)$ that is obtained from the multivariate normal distribution $\mathcal{N}_d(nc, n\Sigma)$ by assigning the probability of the d-box

$$[i_1 - 1/2, i_1 + 1/2) \times \cdots \times [i_d - 1/2, i_d + 1/2)$$

to the integer vector $(i_1, \ldots, i_d)^T$, for each $(i_1, \ldots, i_d)^T \in \mathbb{Z}^d$. This family of distributions is a natural choice, when approximating a discrete random vector in a central limit setting. We are able to establish discrete normal approximation under conditions broadly analogous to those of [Rinott & Rotar (1996)] and [Fang (2014)], with an error of order $O(n^{-1/2}\log n)$, but without their boundedness assumption; a suitable third moment condition is all that is needed.

For generality, we replace n with an m which is essentially the dimension adjusted trace of the covariance matrix of W. Our approach to establishing approximation in total variation by $\mathrm{DN}_d(mc, m\Sigma)$ is by way of Stein's method. Letting $e^{(i)}$ denote the coordinate vector in the i-direction, we start with a Stein operator $\widetilde{\mathcal{A}}_m$, acting on functions $h\colon \mathbb{Z}^d \to \mathbb{R}$, defined by

$$(\widetilde{\mathcal{A}}_m h)(z) \;:=\; m\mathrm{Tr}(\Sigma\Delta^2 h(z)) - (z - mc)^T \Delta h(z), \quad z \in \mathbb{Z}^d, \tag{1.1}$$

where

$$\Delta_j h(z) \;:=\; h(z + e^{(j)}) - h(z); \quad \Delta^2_{jk} h(z) \;:=\; \Delta_j(\Delta_k h)(z),$$

$\Delta h(z) := (\Delta_1 h(z), \ldots, \Delta_d h(z))^T$, and $\Delta^2 h(z)$ is the matrix whose jk element is $\Delta^2_{jk} h(z)$. For any $z \in \mathbb{Z}^d$ and $0 < r \le \infty$, define

$$\begin{aligned} |\Delta h(z)| \;&:=\; \max_{1\le i\le d} |\Delta_i h(z)|; \quad |\Delta^2 h(z)| \;:=\; \max_{1\le i,k\le d} |\Delta^2_{ik} h(z)|; \\ \|\Delta h\|_{r,\infty} \;&:=\; \max_{z\in\mathbb{Z}^d\cap B_r(mc)} |\Delta h(z)|; \quad \|\Delta^2 h\|_{r,\infty} \;:=\; \max_{z\in\mathbb{Z}^d\cap B_r(mc)} |\Delta^2 h(z)|, \end{aligned} \tag{1.2}$$

where $B_r(x) := \{y \in \mathbb{R}^d \colon |y - x| \le r\}$; note that the centre mc is suppressed in the norm notation. Using the operator $\widetilde{\mathcal{A}}_m$, the following abstract result can be deduced from [Barbour, Luczak & Xia (2018b), Theorem 2.4] and [Barbour, Luczak & Xia (2018a), Remark 4.2]. In its statement, we use $\rho(\Sigma)$ to denote the condition number $\lambda_{\max}(\Sigma)/\lambda_{\min}(\Sigma)$ of a matrix Σ, where $\lambda_{\max}(\Sigma)$ and $\lambda_{\min}(\Sigma)$ denote the largest and smallest eigenvalues of Σ.

Theorem 1.1. *Let W be a random vector in $\mathbb{Z}^d$ with mean $\mu := \mathbb{E}W$ and positive definite covariance matrix $V := \mathbb{E}\{(W-\mu)(W-\mu)^T\}$; define $m := \lceil d^{-1}\mathrm{Tr}V\rceil$, $c := m^{-1}\mu$ and $\Sigma := m^{-1}V$. Set $\delta_0 := \frac{1}{72}\,\rho(\Sigma)^{-3/2}$. Then, for any $0 < \delta \le \delta_0$, there exist $C_{1.1}(\delta), n_{1.1}(\delta) < \infty$, depending continuously on δ and the condition number $\rho(\Sigma)$, but not on d or m, with the following property: if, for some ε_1, ε_{20}, ε_{21} and ε_{22}, and for some $m \ge n_{1.1}(\delta)$,*

(a) $d_{\mathrm{TV}}(\mathcal{L}(W), \mathcal{L}(W + e^{(j)})) \;\le\; \varepsilon_1$, for each $1 \le j \le d$;

(b) $|\mathbb{E}\{\widetilde{\mathcal{A}}_m h(W)\} I[|W-\mu| \le m\delta]|$

$$\le\; \varepsilon_{20}\|h\|_{3m\delta_0/2,\infty} + \varepsilon_{21} m^{1/2}\|\Delta h\|_{3m\delta_0/2,\infty} + \varepsilon_{22} m\|\Delta^2 h\|_{3m\delta_0/2,\infty},$$

for all $h\colon \mathbb{Z}^d \to \mathbb{R}$*, then it follows that*

$$d_{\mathrm{TV}}(\mathcal{L}(W), \mathrm{DN}_d(mc, m\Sigma)) \\ \le\; C_{1.1}(\delta)(d^4(m^{-1/2} + \varepsilon_1) + \varepsilon_{20} + \varepsilon_{21} + \varepsilon_{22}) \log m.$$

The unspecified constants can in principle be deduced from the more detailed information in [Barbour, Luczak & Xia (2018a), Barbour, Luczak & Xia (2018b)].

Applying the theorem in practice may not be easy. Condition (b) is much like the sort of condition that has to be checked to prove multivariate normal approximation using Stein's method [Chen, Goldstein & Shao (2011), p. 337], with differences and derivatives exchanged, except for the indicator $I[|W - mc| \le m\delta]$, which truncates W to the ball $B_{m\delta}(mc)$. The truncation has both good and bad consequences. It introduces an awkward discontinuity inside the expectation, which needs careful treatment in the arguments that follow. On the other hand, it ensures that all the expectations to be considered are finite, and that the function h only has to be evaluated within certain closed balls around mc; this latter feature is important, because the solutions to the Stein equation for this problem may grow large as the distance from mc increases. Condition (a) imposes a certain smoothness on the distribution of W.

In Section 2, we prove a multivariate approximation theorem, Theorem 2.1, with error bounds in the total variation distance, that is much simpler to use than Theorem 1.1. The setting is one of predominately local dependence. The basic elements making up the error bounds are sums of third moments, similar to those that would be expected to quantify the error in the CLT for dissociated summands, together with dependence coefficients analogous to those in [Rinott & Rotar (1996)]. However, there is an extra quantity ε_W appearing in the bound, which quantifies the smoothness of the distribution of W, and which is not as simple to express in concrete terms. We also consider a more general setting, in which W arises from integrating the marks of a marked point process with respect to its ground process on a suitable metric space. For integrals of functionals of a Poisson process, [Schulte & Yukich (2018a), Schulte & Yukich (2018b)] have recently established an order $O(n^{-1/2})$ rate of multivariate approximation with respect to the convex sets metric, using the Malliavin–Stein approach and second order Poincaré inequalities. They require somewhat stronger moment assumptions than ours, but, as in the theorems of [Rinott & Rotar (1996)] and of [Fang (2014)], there is no need to bound an analogue of ε_W.

In Section 3, we introduce a stronger notion of local dependence, that is convenient for many applications. It enables us to give rather simple error bounds, in Corollary 3.1, expressed in terms of an upper bound for the maximum of the third moments of the $|X^{(\alpha)}|$ and the sizes of the neighbourhoods in the dependency graph, both being quantities that typically appear in error bounds in the CLT. It also enables us to give a general result, Theorem 3.2, that is helpful for bounding ε_W. The effectiveness of our bounds is illustrated in a number of examples in Section 4. These also give some insight into why, in addition to the sort of moment conditions that suffice for approximation in metrics weaker than total variation, some smoothness condition is needed.

2 Main theorems

For the ease of use, we present our main results for the accuracy of multivariate discrete normal approximation in two distinct but related settings. We postpone the proofs of the main theorems to Section 5.

In the first setting, we suppose that $W = \sum_{j=1}^n X^{(j)}$ is a sum of n vectors in $\mathbb{R}^d$. As before, we write

$$\mu \;:=\; \mathbb{E}W, \quad V \;:=\; \mathrm{Cov}(W) \quad \text{and} \quad m \;:=\; \lceil d^{-1}\mathrm{Tr}\, V \rceil. \tag{2.1}$$

We assume that there are decompositions of the following form:

(a) For each $1 \le j \le n$, we can write $W = W^{(j)} + Z^{(j)}$, where $W^{(j)} \in \mathbb{Z}^d$ is only weakly dependent on $X^{(j)}$;

(b) For each $1 \le j \le n$, we can write $Z^{(j)} = \sum_{k=1}^{n_j} \widetilde{X}^{(j,k)}$, with $\widetilde{X}^{(j,k)} \in \mathbb{Z}^d$, and then, for each $1 \le k \le n_j$, we can write $W^{(j)} = W^{(j,k)} + Z^{(j,k)}$, where $W^{(j,k)} \in \mathbb{Z}^d$ is only weakly dependent on $(X^{(j)}, \widetilde{X}^{(j,k)})$.

Because of the restrictions to $\mathbb{Z}^d$, centring on the mean is not possible in these decompositions, but it could, for instance, be arranged that each component of $Z^{(j)}$, $\widetilde{X}^{(j,k)}$ and $Z^{(j,k)}$ has mean with modulus at most 1. This makes no difference to the arguments that follow, but the moment sums H_1 and H_2 that appear in the error bounds might otherwise be larger than necessary. We assume that $\mathbb{E}|X^{(j)}|^3 < \infty$ for each $j \in [n] := \{1, 2, \ldots, n\}$, and that $\mathbb{E}|\widetilde{X}^{(j,k)}|^3 < \infty$ and $\mathbb{E}|Z^{(j,k)}|^3 < \infty$ for each $j \in [n]$ and $1 \le k \le n_j$; our bounds are not useful otherwise.

Weak dependence is expressed by the smallness of dependence coefficients analogous to those in [Rinott & Rotar (1996)]. With $\mu^{(j)} := \mathbb{E}X^{(j)}$ and with $|\cdot|_1$ denoting the l_1 norm, we begin by defining

$$\begin{aligned}
\chi_{12j} &:= \mathbb{E}\big|\mathbb{E}(|X^{(j)}| \mid W^{(j)}) - \mathbb{E}|X^{(j)}|\big|;\\
\chi_{13j} &:= \mathbb{E}\big|\mathbb{E}(|X^{(j)}|_1 \mid W^{(j)}) - \mathbb{E}|X^{(j)}|_1\big|; \qquad (2.2)\\
\chi_{2jk} &:= \sum_{i=1}^d \sum_{l=1}^d \mathbb{E}\big|\mathbb{E}\{|X_i^{(j)}|\,|\widetilde{X}_l^{(j,k)}| \mid W^{(j,k)}\} - \mathbb{E}\{|X_i^{(j)}|\,|\widetilde{X}_l^{(j,k)}|\}\big|\\
&\quad + \sum_{i=1}^d \sum_{l=1}^d |\mu_i^{(j)}|\mathbb{E}\big|\mathbb{E}\{|\widetilde{X}_l^{(j,k)}| \mid W^{(j,k)}\} - \mathbb{E}\{|\widetilde{X}_l^{(j,k)}|\}\big|\\
&\quad + \sum_{i=1}^d \sum_{l=1}^d \mathbb{E}\big|\mathbb{E}\{X_i^{(j)}\, \widetilde{X}_l^{(j,k)} \mid W^{(j,k)}\} - \mathbb{E}\{X_i^{(j)}\, \widetilde{X}_l^{(j,k)}\}\big| \qquad (2.3)\\
&\quad + \sum_{i=1}^d \sum_{l=1}^d |\mu_i^{(j)}|\mathbb{E}\big|\mathbb{E}\{\widetilde{X}_l^{(j,k)} \mid W^{(j,k)}\} - \mathbb{E}\{\widetilde{X}_l^{(j,k)}\}\big|,
\end{aligned}$$

and then set

$$\begin{aligned}
\chi_{11} &:= (dm)^{-1/2} \sum_{j=1}^n \mathbb{E}|\mathbb{E}(X^{(j)} \mid W^{(j)}) - \mathbb{E}X^{(j)}|;\\
\chi_{12} &:= (dm)^{-1/2} \sum_{j=1}^n \chi_{12j}; \qquad \chi_{13} = d^{-1}m^{-1/2} \sum_{j=1}^n \chi_{13j};\\
\chi_2 &:= d^{-3}m^{-1} \sum_{j=1}^n \sum_{k=1}^{n_j} \chi_{2jk}; \qquad (2.4)\\
\chi_3 &:= d^{-1}m^{-1} \sum_{j=1}^n \mathbb{E}\{|\mathbb{E}(X^{(j)} \mid W^{(j)}) - \mu^{(j)}|\,|W^{(j)} - \mu|\}.
\end{aligned}$$

We then write $\chi_1 := \max_{1 \le l \le 3} \chi_{1l}$. Note that the m-factors, with m as defined in (2.1), are not present in the quantities in [Rinott & Rotar (1996)] that are directly analogous to χ_{11}, χ_2 and χ_3. This is because, in their formulation, the random variables corresponding to $X^{(j)}$ are normalized to make $\mathrm{Cov}(W)$ close to the identity matrix. Since our sum W is not normalized, to keep its values in $\mathbb{Z}^d$, the elements of its covariance matrix typically grow with n. The quantities χ_{12} and χ_{13} have no direct analogue in [Rinott & Rotar (1996)], and appear only in dealing with the truncation to $B_{n\delta}(\mu)$, something that is not needed in their arguments.

Again recalling m from (2.1), we introduce some moment sums, used in the error estimates, defining

$$\begin{aligned} H_{21} &:= d^{-3/2}m^{-1}\sum_{j=1}^{n}\mathbb{E}\{(|X^{(j)}|+|\mu^{(j)}|)\,|Z^{(j)}|^2\}; \\ H_{22} &:= d^{-3/2}m^{-1}\sum_{j=1}^{n}\sum_{k=1}^{n_j}\mathbb{E}\{(|X^{(j)}|+|\mu^{(j)}|)\,|\widetilde{X}^{(j,k)}|\,|Z^{(j,k)}|\}; \\ H_{23} &:= d^{-3/2}m^{-1}\sum_{j=1}^{n}\sum_{k=1}^{n_j}\mathbb{E}\{(|X^{(j)}|+|\mu^{(j)}|)\,|\widetilde{X}^{(j,k)}|\}\mathbb{E}|Z^{(j,k)}|; \\ H_{24} &:= d^{-3/2}m^{-1}\sum_{j=1}^{n}\sum_{k=1}^{n_j}\mathbb{E}\{(|X^{(j)}|+|\mu^{(j)}|)\,|\widetilde{X}^{(j,k)}|\}\mathbb{E}|Z^{(j)}|, \end{aligned}$$

and then setting

$$\begin{aligned} H_0 &:= d^{-1/2}m^{-1}\sum_{j=1}^{n}\mathbb{E}|X^{(j)}|; \\ H_1 &:= d^{-1}m^{-1}\sum_{j=1}^{n}\sum_{k=1}^{n_j}\mathbb{E}\{(|X^{(j)}|+|\mu^{(j)}|)\,|\widetilde{X}^{(j,k)}|\}; \\ H_2 &:= \max_{1\le l\le 4} H_{2l}. \end{aligned} \tag{2.5}$$

We also assume that

$$\mathbb{E}\{|Z^{(j)}|^2\} \le dm; \qquad \mathbb{E}\{|Z^{(j,k)}|^2\} \le dm, \qquad \text{for all } 1\le j\le n;\ 1\le k\le n_j. \tag{2.6}$$

The various d-factors are designed to offset any automatic dimension dependence in the corresponding quantities, but their choice plays no essential part in the bounds given below.

A rough idea of the magnitudes of the quantities H_0, H_1 and H_2 can be gained as follows. In circumstances in which the weak dependence in the decomposition is sufficiently weak, it is to be expected that

$$\mathrm{Tr}V = \mathbb{E}\{(W-\mu)^T(W-\mu)\} \approx \sum_{j=1}^{n}\sum_{k=1}^{n_j}\mathbb{E}\{(X^{(j)}-\mu^{(j)})^T(\widetilde{X}^{(j,k)}-\mathbb{E}\widetilde{X}^{(j,k)})\} =: \hat{v},$$

say; suppose, for definiteness, that $1/2 \le \hat{v}/\mathrm{Tr}V \le 2$. Using the fact that $\mu^{(j)} = \mathbb{E}X^{(j)}$, and replacing vectors by their moduli in the sum on the right hand side, it follows that dmH_1 is an upper bound for $\hat{v}$. Suppose, as is the case in many applications, that all of the $\{\mathbb{E}|X^{(j)}|^3\}^{1/3}$ and $\{\mathbb{E}|\widetilde{X}^{(j,k)}|^3\}^{1/3}$ are bounded by $c_2\sqrt{d}$ for some $c_2 < \infty$, and also that $\{\mathbb{E}|Z^{(j,k)}|^3\}^{1/3} \le n_* c_2\sqrt{d}$ for all j,k, where $n_* := \max_{1\le j\le n} n_j$. Suppose further that $\hat{v} \ge nn_*c_1^2 d$ for some $0 < c_1 < c_2$, that $n_j \ge 1$ for all j, and that $\mathrm{Tr}V \ge d$. Then it follows that

$$nn_*c_1^2 d \le \hat{v} \le mdH_1 \le 2nn_*c_2^2 d,$$

so that $mdH_1 \le 4(c_2/c_1)^2\,\mathrm{Tr}V$; this implies that $H_1 \le 4(c_2/c_1)^2$, from the definition of m, so that H_1 remains bounded if c_1 and c_2 remain constant as n increases. Similar considerations show also that

$$H_0 \le \frac{4c_2}{n_*c_1^2} \quad \text{and} \quad H_2 \le \frac{2n_*c_2^3}{c_1^2},$$

so that H_0 also remains bounded if c_1 and c_2 remain constant as n increases, whereas H_2 may grow if n_* grows with n. These assumptions are automatically satisfied under the conditions of Corollary 3.1, with the exception of the lower bound $\hat{v} \geq nn_*c_1^2d$ for the variance; this, however, is typically satisfied too, as in Example 4.1. Finally, assumption (2.6) is satisfied if $n_* \leq \frac{1}{2}n(c_1/c_2)^2$, which is true for large enough n if $n_* = o(n)$ as $n \to \infty$. In Example 4.1, this would only require that the underlying graph be sparse, in that the maximal vertex degree should be of smaller order than the number of vertices.

We now make a smoothness assumption on the distributions of $W^{(j)}$ and $W^{(j,k)}$ that is key for approximation in total variation. We assume that, for each $1 \leq j \leq n$, $1 \leq k \leq n_j$ and $1 \leq i \leq d$, we have

$$\begin{aligned} d_{\mathrm{TV}}\big(\mathcal{L}(W^{(j)} + e^{(i)} \,|\, X^{(j)}, Z^{(j)}), \mathcal{L}(W^{(j)} \,|\, X^{(j)}, Z^{(j)})\big) \ &\leq \ \varepsilon_W \text{ a.s.}; \\ d_{\mathrm{TV}}\big(\mathcal{L}(W^{(j,k)} + e^{(i)} \,|\, X^{(j)}, \widetilde{X}^{(j,k)}, Z^{(j,k)}), \mathcal{L}(W^{(j,k)} \,|\, X^{(j)}, \widetilde{X}^{(j,k)}, Z^{(j,k)})\big) \ &\leq \ \varepsilon_W \text{ a.s.}, \end{aligned} \tag{2.7}$$

for some $\varepsilon_W < 1$. Of course, for the bounds that we shall prove, we shall want ε_W to be suitably small. This assumption is clearly useful in establishing Condition (a) of Theorem 1.1, but is also used throughout the treatment of $|\mathbb{E}\{\widetilde{\mathcal{A}}_m h(W)\} I[|W - \mu| \leq m\delta]|$.

Theorem 2.1. *Let $W := \sum_{j=1}^n X^{(j)}$ be decomposed as above, with (2.6) satisfied, and suppose that V is positive definite. Then there exist constants $C_{2.1}$ and $n_{2.1}$, depending continuously on the condition number $\rho(V)$, such that*

$$\begin{aligned} &d_{\mathrm{TV}}(\mathcal{L}(W), \mathrm{DN}_d(\mu, V)) \\ &\quad \leq \ C_{2.1} d^3 \log m \{(d + H_2)\varepsilon_W + (d + H_0 + H_2 + m^{-1/2} H_1) m^{-1/2} + (\chi_1 + \chi_2 + \chi_3)\}, \end{aligned}$$

for all $m \geq n_{2.1}$, where m is as in (2.1), ε_W is as in (2.7), H_0, H_1 and H_2 are as in (2.5), and χ_1, χ_2 and χ_3 are as in (2.4).

Our second setting is somewhat more general. We suppose that W results from integrating the marks of a marked point process with respect to its ground process. We assume that the carrier space Γ of the ground point process Ξ is a locally compact second countable Hausdorff topological space [Kallenberg (1983), p. 11], with Borel σ-field $\mathcal{B}(\Gamma)$. Let $\widetilde{G} := \Gamma \times \mathbb{Z}^d$, and equip it with the product Borel σ-field $\mathcal{B}(\widetilde{G}) = \mathcal{B}(\Gamma) \times \mathcal{B}(\mathbb{Z}^d)$. We use $\mathcal{H}$ to denote the space of all locally finite non-negative integer valued measures ξ on $\widetilde{G}$ such that $\xi(\{\alpha\} \times \mathbb{Z}^d) \leq 1$ for all $\alpha \in \Gamma$. The space $\mathcal{H}$ is endowed with the σ-field $\mathcal{B}(\mathcal{H})$ generated by the vague topology [Kallenberg (1983), p. 169]. A *marked point process* $\tilde{\Xi}$ is a measurable mapping from $(\Omega, \mathcal{F}, \mathbb{P})$ to $(\mathcal{H}, \mathcal{B}(\mathcal{H}))$ [Kallenberg (2017), p. 49]. The induced simple point process $\Xi(\cdot) := \tilde{\Xi}(\cdot \times \mathbb{Z}^d)$ is called the *ground process* [Daley & Vere-Jones (2008), p. 3] or projection [Kallenberg (2017), p. 17] of the marked point process $\tilde{\Xi}$. For $(\alpha, y) \in \Gamma \times \mathbb{Z}^d$ such that $\tilde{\Xi}(\{(\alpha, y)\}) = 1$, we write $X^{(\alpha)} := y$, so that $X^{(\alpha)}$ represents the *mark* of $\tilde{\Xi}$ at α. We assume that the ground process Ξ is locally finite, with mean measure ν.

Let $\{D_\alpha, \alpha \in \Gamma\}$ be a class of neighbourhoods such that, for each $\alpha \in \Gamma$, $D_\alpha \in \mathcal{B}(\Gamma)$ is a Borel set containing α and such that $D = \{(\alpha, \beta) \colon \beta \in D_\alpha, \alpha \in \Gamma\}$ is a measurable subset of the product space $\Gamma^2 := \Gamma \times \Gamma$ with the product Borel σ-field $\mathcal{B}(\Gamma) \times \mathcal{B}(\Gamma)$. For the neighbourhoods $\{D_\alpha, \alpha \in \Gamma\}$, one can easily adapt the proof in [Chen & Xia (2004)] to show that the mapping $(\alpha, \xi) \mapsto (\alpha, \xi|_{D_\alpha \times \mathbb{Z}^d})$ is a measurable mapping from $(\Gamma \times \mathcal{H}, \mathcal{B}(\Gamma) \times \mathcal{B}(\mathcal{H}))$ into itself, where $\xi|_{D_\alpha \times \mathbb{Z}^d}$ is the restriction of $\xi \in \mathcal{H}$ to $D_\alpha \times \mathbb{Z}^d$ [Kallenberg (1983), p. 12].

Our goal is to establish the accuracy of discrete normal approximation to $W = \int_\Gamma X^{(\alpha)} \Xi(d\alpha)$. When $\Gamma = \{1, \ldots, n\}$ and Ξ is the counting measure on Γ, W reduces to the sum in the previous setting, so the bound in Theorem 2.1 is a corollary of that in Theorem 2.2. However, if there is dependence between Ξ and X, then there is significant difference between the two settings. For the latter setting, it is necessary to introduce extra machinery, including the first and second order Palm distributions [Kallenberg (1983), p. 83 and p. 103], to tackle the problem.

For convenience, we use $\mathbb{P}_\alpha$, $\mathbb{E}_\alpha$ and $\mathcal{L}_\alpha$ to stand for the conditional probability, conditional expectation and conditional distribution given $\{\Xi(\{\alpha\}) = 1\}$ respectively. It is a routine exercise [Kallenberg (1983), pp. 83–84] to show that $\mathbb{E}_\alpha$ satisfies

$$\mathbb{E}\int_\Gamma f(X^{(\alpha)}, \alpha)\Xi(d\alpha) = \int_\Gamma \mathbb{E}_\alpha f(X^{(\alpha)}, \alpha)\nu(d\alpha)$$

for all non-negative functions f on $(\mathbb{Z}^d \times \Gamma, \mathcal{B}(\mathbb{Z}^d) \times \mathcal{B}(\Gamma))$. Similarly, for $\alpha \neq \beta$, we use $\mathbb{P}_{\alpha\beta}$, $\mathbb{E}_{\alpha\beta}$ and $\mathcal{L}_{\alpha\beta}$ to stand for the conditional probability, conditional expectation and conditional distribution given $\{\Xi(\{\alpha\}) = 1\} \cap \{\Xi(\{\beta\}) = 1\}$ respectively. Writing $\nu_2(d\alpha, d\beta) = \mathbb{E}(\Xi(d\alpha)\Xi(d\beta))$ for $\alpha \neq \beta$, we can also show that $\mathbb{E}_{\alpha\beta}$ satisfies,

$$\begin{aligned} &\mathbb{E}\int_{\alpha,\beta\in\Gamma,\alpha\neq\beta} f(X^{(\alpha)}, X^{(\beta)}, \alpha, \beta)\Xi(d\beta)\Xi(d\alpha) \\ &= \int_{\alpha,\beta\in\Gamma,\alpha\neq\beta} \mathbb{E}_{\alpha\beta} f(X^{(\alpha)}, X^{(\beta)}, \alpha, \beta)\nu_2(d\alpha, d\beta), \end{aligned} \tag{2.8}$$

for any non-negative measurable function f on $(\mathbb{Z}^d \times \mathbb{Z}^d \times \Gamma \times \Gamma, \mathcal{B}(\mathbb{Z}^d) \times \mathcal{B}(\mathbb{Z}^d) \times \mathcal{B}(\Gamma) \times \mathcal{B}(\Gamma))$. To avoid unnecessary complexity and to keep our notation consistent, we write $\mathbb{P}_{\alpha\alpha} = \mathbb{P}_\alpha$, $\mathbb{E}_{\alpha\alpha} = \mathbb{E}_\alpha$, $\mathcal{L}_{\alpha\alpha} = \mathcal{L}_\alpha$ and $\nu_2(d\alpha, d\alpha) = \nu(d\alpha)$ so that (2.8) can be extended to

$$\mathbb{E}\int_{\alpha,\beta\in\Gamma} f(X^{(\alpha)}, X^{(\beta)}, \alpha, \beta)\Xi(d\beta)\Xi(d\alpha) \;=\; \int_{\alpha,\beta\in\Gamma} \mathbb{E}_{\alpha\beta} f(X^{(\alpha)}, X^{(\beta)}, \alpha, \beta)\nu_2(d\alpha, d\beta).$$

As in the previous setting, we assume that there are decompositions of the following form:

(a') For each $\alpha \in \Gamma$, we can write $W = W^{(\alpha)} + Z^{(\alpha)}$, where, under $\mathbb{P}_\alpha$, $W^{(\alpha)} \in \mathbb{Z}^d$ is only weakly dependent on $X^{(\alpha)}$;

(b') For each $\alpha \in \Gamma$, we can write $Z^{(\alpha)} = \int_{D_\alpha} \widetilde{X}^{(\alpha,\beta)}\Xi(d\beta)$, with $\widetilde{X}^{(\alpha,\beta)} \in \mathbb{Z}^d$, and then, for each $\beta \in D_\alpha$, we can write $W^{(\alpha)} = W^{(\alpha,\beta)} + Z^{(\alpha,\beta)}$, where $W^{(\alpha,\beta)} \in \mathbb{Z}^d$ is only weakly dependent on $\widetilde{X}^{(\alpha,\beta)}$ under $\mathbb{P}_\beta$, and on the pair $(X^{(\alpha)}, \widetilde{X}^{(\alpha,\beta)})$ under $\mathbb{P}_{\alpha\beta}$. In particular, we take $\widetilde{X}^{(\alpha,\alpha)} = X^{(\alpha)}$ if $\Xi(\{\alpha\}) = 1$, $Z^{(\alpha,\alpha)} = Z^{(\alpha)}$ and $W^{(\alpha,\alpha)} = W^{(\alpha)}$.

We assume for each $\alpha \in \Gamma$ that $\mathbb{E}_\alpha|X^{(\alpha)}|^3 < \infty$ ν-a.s., and for each $\alpha \in \Gamma$ and $\beta \in D_\alpha$ that $\mathbb{E}_\beta|\widetilde{X}^{(\alpha,\beta)}|^3 < \infty$ ν-a.s., $\mathbb{E}_{\alpha\beta}|\widetilde{X}^{(\alpha,\beta)}|^3 < \infty$ and $\mathbb{E}_{\alpha\beta}|Z^{(\alpha,\beta)}|^3 < \infty$ ν_2-a.s., since our bounds are not useful otherwise. We set $\mu^{(\alpha)} := \mathbb{E}_\alpha X^{(\alpha)}$ and $\mu := \int_\Gamma \mu^{(\alpha)}\nu(d\alpha) = \mathbb{E}W$, and write

$$V \;:=\; \mathrm{Cov}(W); \quad m \;:=\; \lceil d^{-1}\mathrm{Tr} V\rceil. \tag{2.9}$$

We then define

$$
\begin{aligned}
\chi'_{12\alpha} &:= \mathbb{E}_\alpha\big|\mathbb{E}_\alpha(|X^{(\alpha)}| \,|\, W^{(\alpha)}) - \mathbb{E}_\alpha|X^{(\alpha)}|\big|; \\
\chi'_{13\alpha} &:= \mathbb{E}_\alpha\big|\mathbb{E}_\alpha(|X^{(\alpha)}|_1 \,|\, W^{(\alpha)}) - \mathbb{E}_\alpha|X^{(\alpha)}|_1\big|; \qquad (2.10)\\
\chi'_{2\alpha\beta} &:= \sum_{i=1}^d\sum_{l=1}^d \mathbb{E}_{\alpha\beta}\big|\mathbb{E}_{\alpha\beta}\{|X_i^{(\alpha)}|\,|\widetilde{X}_l^{(\alpha,\beta)}| \,|\, W^{(\alpha,\beta)}\} - \mathbb{E}_{\alpha\beta}\{|X_i^{(\alpha)}|\,|\widetilde{X}_l^{(\alpha,\beta)}|\}\big| \\
&\qquad + \sum_{i=1}^d\sum_{l=1}^d \mathbb{E}_{\alpha\beta}\big|\mathbb{E}_{\alpha\beta}\{X_i^{(\alpha)}\widetilde{X}_l^{(\alpha,\beta)} \,|\, W^{(\alpha,\beta)}\} - \mathbb{E}_{\alpha\beta}\{X_i^{(\alpha)}\widetilde{X}_l^{(\alpha,\beta)}\}\big|; \\
\chi''_{2\alpha\beta} &:= \sum_{i=1}^d\sum_{l=1}^d |\mu_i^{(\alpha)}|\,\mathbb{E}_{\beta}\big|\mathbb{E}_{\beta}\{|\widetilde{X}_l^{(\alpha,\beta)}| \,|\, W^{(\alpha,\beta)}\} - \mathbb{E}_{\beta}\{|\widetilde{X}_l^{(\alpha,\beta)}|\}\big| \\
&\qquad + \sum_{i=1}^d\sum_{l=1}^d |\mu_i^{(\alpha)}|\,\mathbb{E}_{\beta}\big|\mathbb{E}_{\beta}\{\widetilde{X}_l^{(\alpha,\beta)} \,|\, W^{(\alpha,\beta)}\} - \mathbb{E}_{\beta}\{\widetilde{X}_l^{(\alpha,\beta)}\}\big|.
\end{aligned}
$$

Note that the quantities $\chi'_{2\alpha\beta}$ and $\chi''_{2\alpha\beta}$ are more complicated than their counterparts χ_{2jk} in the earlier setting, to allow for possible dependence between the ground process and the marks. Then let

$$
\begin{aligned}
\chi'_{11} &:= (dm)^{-1/2}\int_\Gamma \mathbb{E}_\alpha|\mathbb{E}_\alpha(X^{(\alpha)} \,|\, W^{(\alpha)}) - \mu^{(\alpha)}|\nu(d\alpha); \\
\chi'_{12} &:= (dm)^{-1/2}\int_\Gamma \chi'_{12\alpha}\nu(d\alpha); \qquad \chi'_{13} = d^{-1}m^{-1/2}\int_\Gamma \chi'_{13\alpha}\nu(d\alpha); \\
\chi'_1 &:= \max_{1\le l\le 3}\chi'_{1l} \\
\chi'_2 &:= d^{-3}m^{-1}\left(\int_\Gamma\int_{D_\alpha}\chi'_{2\alpha\beta}\nu_2(d\alpha,d\beta) + \int_\Gamma\int_{D_\alpha}\chi''_{2\alpha\beta}\nu(d\beta)\nu(d\alpha)\right); \qquad (2.11)\\
\chi'_3 &:= d^{-1}m^{-1}\int_\Gamma \mathbb{E}_\alpha\{|\mathbb{E}_\alpha(X^{(\alpha)} \,|\, W^{(\alpha)}) - \mu^{(\alpha)}|\,|W^{(\alpha)} - \mu|\}\nu(d\alpha).
\end{aligned}
$$

We next introduce some moment sums by defining

$$
\begin{aligned}
H'_{21} &:= d^{-3/2}m^{-1}\int_\Gamma\{\mathbb{E}_\alpha\{|X^{(\alpha)}|\,|Z^{(\alpha)}|^2\} + |\mu^{(\alpha)}|\,\mathbb{E}\{|Z^{(\alpha)}|^2\}\}\nu(d\alpha); \\
H'_{22} &:= d^{-3/2}m^{-1}\int_\Gamma\int_{D_\alpha}\mathbb{E}_{\alpha\beta}\{|X^{(\alpha)}|\,|\widetilde{X}^{(\alpha,\beta)}|\,|Z^{(\alpha,\beta)}|\}\nu_2(d\alpha,d\beta) \\
&\qquad + d^{-3/2}m^{-1}\int_\Gamma\int_{D_\alpha}|\mu^{(\alpha)}|\mathbb{E}_\beta\{|\widetilde{X}^{(\alpha,\beta)}|\,|Z^{(\alpha,\beta)}|\}\nu(d\beta)\nu(d\alpha); \\
H'_{23} &:= d^{-3/2}m^{-1}\int_\Gamma\int_{D_\alpha}\mathbb{E}_{\alpha\beta}\{|X^{(\alpha)}|\,|\widetilde{X}^{(\alpha,\beta)}|\,\}\mathbb{E}\{|Z^{(\alpha,\beta)}|\}\nu_2(d\alpha,d\beta) \\
&\qquad + d^{-3/2}m^{-1}\int_\Gamma\int_{D_\alpha}|\mu^{(\alpha)}|\mathbb{E}_\beta\{|\widetilde{X}^{(\alpha,\beta)}|\,\}\mathbb{E}\{|Z^{(\alpha,\beta)}|\}\nu(d\beta)\nu(d\alpha); \\
H'_{24} &:= d^{-3/2}m^{-1}\int_\Gamma\int_{D_\alpha}\mathbb{E}_{\alpha\beta}\{|X^{(\alpha)}|\,|\widetilde{X}^{(\alpha,\beta)}|\,\}\mathbb{E}\{|Z^{(\alpha)}|\}\nu_2(d\alpha,d\beta) \\
&\qquad + d^{-3/2}m^{-1}\int_\Gamma\int_{D_\alpha}|\mu^{(\alpha)}|\mathbb{E}_\beta\{|\widetilde{X}^{(\alpha,\beta)}|\,\}\mathbb{E}\{|Z^{(\alpha)}|\}\nu(d\beta)\nu(d\alpha),
\end{aligned}
\qquad (2.12)
$$

noting the extra complication in H'_{22}, H'_{23} and H'_{24} as compared with H_{22}, H_{23} and H_{24}, and then setting

$$\begin{aligned} H_0' &:= d^{-1/2}m^{-1}\int_\Gamma \mathbb{E}_\alpha|X^{(\alpha)}|\nu(d\alpha); \\ H_1' &:= d^{-1}m^{-1}\int_\Gamma\int_{D_\alpha} \mathbb{E}_{\alpha\beta}\{|X^{(\alpha)}|\,|\widetilde{X}^{(\alpha,\beta)}|\}\nu_2(d\alpha,d\beta) \\ &\quad +d^{-1}m^{-1}\int_\Gamma|\mu^{(\alpha)}|\int_{D_\alpha} \mathbb{E}_\beta\{|\widetilde{X}^{(\alpha,\beta)}|\}\nu(d\beta)\nu(d\alpha); \\ H_2' &:= \max_{1\le l\le 4} H_{2l}'. \end{aligned} \tag{2.13}$$

As a consequence of the dependence between the marks and the ground process, the analogue of (2.6) is more involved: we need to assume that there exists a constant $C \geq 1$ such that

$$\begin{aligned} &\mathbb{E}_\alpha(|W-\mu|^2) \;\le\; Cdm\ \nu-\text{a.s.},\ \mathbb{E}_{\alpha\beta}(|W-\mu|^2) \;\le\; Cdm\ \nu_2-\text{a.s.}, \\ &\mathbb{E}\{|Z^{(\alpha)}|^2\} \;\le\; dm,\ \mathbb{E}_\alpha\{|Z^{(\alpha)}|^2\} \;\le\; dm\ \nu-\text{a.s.}, \\ &\max\{\mathbb{E}_{\alpha\beta}\{|Z^{(\alpha)}|^2\},\mathbb{E}_\beta\{|Z^{(\alpha)}|^2\},\ \mathbb{E}_\beta\{|Z^{(\alpha,\beta)}|^2\},\mathbb{E}_{\alpha\beta}\{|Z^{(\alpha,\beta)}|^2\}\} \;\le\; dm\ \nu_2-\text{a.s.}. \end{aligned} \tag{2.14}$$

The analogue of (2.7) is even more involved. First, for $1 \le i \le d$, ν-a.s. in α and ν_2-a.s. in α, β, we need to find $\varepsilon_W' < 1$ such that

$$\begin{aligned} &d_{\mathrm{TV}}\big(\mathcal{L}(W^{(\alpha)}+e^{(i)}\,|\,X^{(\alpha)},Z^{(\alpha)}),\mathcal{L}(W^{(\alpha)}\,|\,X^{(\alpha)},Z^{(\alpha)})\big) \;\le\; \varepsilon_W'\ \mathbb{P}-\text{a.s.}; \\ &d_{\mathrm{TV}}\big(\mathcal{L}_\alpha(W^{(\alpha)}+e^{(i)}\,|\,X^{(\alpha)},Z^{(\alpha)}),\mathcal{L}_\alpha(W^{(\alpha)}\,|\,X^{(\alpha)},Z^{(\alpha)})\big) \;\le\; \varepsilon_W'\ \mathbb{P}_\alpha-\text{a.s.}; \\ &d_{\mathrm{TV}}\big(\mathcal{L}(W^{(\alpha,\beta)}+e^{(i)}\,|\,X^{(\alpha)},\widetilde{X}^{(\alpha,\beta)},Z^{(\alpha,\beta)}),\mathcal{L}(W^{(\alpha,\beta)}\,|\,X^{(\alpha)},\widetilde{X}^{(\alpha,\beta)},Z^{(\alpha,\beta)})\big) \\ &\qquad\le\; \varepsilon_W'\ \mathbb{P}-\text{a.s.}, \\ &d_{\mathrm{TV}}\big(\mathcal{L}_\beta(W^{(\alpha,\beta)}+e^{(i)}\,|\,X^{(\alpha)},\widetilde{X}^{(\alpha,\beta)},Z^{(\alpha,\beta)}),\mathcal{L}_\beta(W^{(\alpha,\beta)}\,|\,X^{(\alpha)},\widetilde{X}^{(\alpha,\beta)},Z^{(\alpha,\beta)})\big) \\ &\qquad\le\; \varepsilon_W'\ \mathbb{P}_\beta-\text{a.s.}, \\ &d_{\mathrm{TV}}\big(\mathcal{L}_{\alpha\beta}(W^{(\alpha,\beta)}+e^{(i)}\,|\,X^{(\alpha)},\widetilde{X}^{(\alpha,\beta)},Z^{(\alpha,\beta)}),\mathcal{L}_{\alpha\beta}(W^{(\alpha,\beta)}\,|\,X^{(\alpha)},\widetilde{X}^{(\alpha,\beta)},Z^{(\alpha,\beta)})\big) \\ &\qquad\le\; \varepsilon_W'\ \mathbb{P}_{\alpha\beta}-\text{a.s.} \end{aligned} \tag{2.15}$$

We then also need to find $\varepsilon_W'' < 1$ such that

$$\begin{aligned} &d_{\mathrm{TV}}\big(\mathcal{L}_\alpha(W^{(\alpha)}),\mathcal{L}(W^{(\alpha)})\big) \;\le\; \varepsilon_W''\ \nu-\text{a.s.}, \\ &d_{\mathrm{TV}}\big(\mathcal{L}_{\alpha\beta}(W^{(\alpha,\beta)}),\mathcal{L}(W^{(\alpha,\beta)})\big) \;\le\; \varepsilon_W''\ \nu_2-\text{a.s.}, \\ &d_{\mathrm{TV}}\big(\mathcal{L}_\beta(W^{(\alpha,\beta)}),\mathcal{L}(W^{(\alpha,\beta)})\big)\} \;\le\; \varepsilon_W''\ \nu_2-\text{a.s.}. \end{aligned} \tag{2.16}$$

Finally, we need a bound controlling the difference between some conditional and unconditional expectations: we need to find $\varepsilon_W''' < 1$ such that

$$|\mathbb{E}_\alpha(W^{(\alpha)}) - \mathbb{E}(W^{(\alpha)})| \;\le\; \varepsilon_W'''\ \nu-\text{a.s.}\,. \tag{2.17}$$

Fortunately, under many circumstances (see [Barbour & Xia (2006)]), both ε_W'' and ε_W''' can be reduced to 0, as is the case in Example 4.4.

Theorem 2.2. *Let $W := \int_\Gamma X^{(\alpha)}\Xi(d\alpha)$ be decomposed as above, such that* (2.14) *is satisfied, and suppose that V is positive definite. Then there exist constants $C_{2.2}$ and $n_{2.2}$, depending continuously on the condition number $\rho(V)$, such that*

$$\begin{aligned} &d_{\mathrm{TV}}(\mathcal{L}(W),\mathrm{DN}_d(\mu,V)) \\ &\quad\le\; C_{2.2}d^3\log m\big\{(d+H_2')\varepsilon_W' + (d+H_0'+H_2'+m^{-1/2}H_1')m^{-1/2} \\ &\qquad\qquad +\varepsilon_W''(d^{-2}m^{1/2}H_0'+d^{-1}H_1') + d^{-3/2}H_0'\varepsilon_W''' + (\chi_1'+\chi_2'+\chi_3')\big\}, \end{aligned}$$

for all $m \geq n_{2.2}$*, where* m *is as in* (2.9)*,* $\varepsilon'_W, \varepsilon''_W$ *and* ε'''_W *are as in* (2.15)*–*(2.17)*,* H'_0, H'_1 *and* H'_2 *are as in* (2.13)*, and* χ'_1, χ'_2 *and* χ'_3 *are as in* (2.11)*.*

3 Intersection graph dependence

In this section, we consider sums $W := \sum_{j=1}^n X^{(j)}$ of random vectors $X^{(j)}$ that are determined by the values of an underlying collection of independent random elements $(Y_i,\, 1 \leq i \leq M)$; we assume that $X^{(j)} := X^{(j)}((Y_i,\, i \in M_j))$, for some subset $M_j \subset [M] := \{1, 2, \ldots, M\}$. The subsets M_j induce an intersection graph G on $[n]$, in which there is an edge between j and $k \neq j$, $j \sim k$, exactly when $M_j \cap M_k \neq \emptyset$; we denote by $N_j := \{k \in [n] \setminus \{j\} \colon k \sim j\}$ the neighbourhood of j in G. With this definition, $X^{(j)}$ is independent of $(X^{(k)} \colon k \in [n] \setminus (\{j\} \cup N_j))$, and the graph G is a dependency graph in the sense of [Baldi & Rinott (1989)].

Corollary 3.1. *Suppose that the assumptions in the preceding paragraph are satisfied, with* $\mu := \mathbb{E}W$ *and* $V := \mathrm{Cov}(W)$*, and that*

$$1 \;\leq\; \max_{1 \leq j \leq n} d^{-3/2} \mathbb{E}|X^{(j)}|^3 \;=:\; \gamma \;<\; \infty. \tag{3.1}$$

Recalling the notation of Theorem 2.1, with $m := \lceil d^{-1} \mathrm{Tr} V \rceil$*, define*

$$D_j \;:=\; |N_j| \quad \textit{and} \quad \overline{D^2} \;:=\; m^{-1} \sum_{j=1}^n (D_j + 1)^2.$$

Then, if $m \geq n_{2.1}$*,*

$$d_{\mathrm{TV}}(\mathcal{L}(W), \mathrm{DN}_d(\mu, V)) \;\leq\; C_{2.1} d^3 \log m (m^{-1/2} + \varepsilon_W) \{d + 3\gamma \overline{D^2}\},$$

where $C_{2.1}$ *from Theorem 2.1 depends only on the condition number* $\rho(V)$ *of* V*.*

Proof. In this setting, Theorem 2.1 can be employed to prove the corollary. In fact, there is a natural way to define $W^{(j)}$ and $W^{(j,k)}$. For each $j \in [n]$, we define $Z^{(j)} := X^{(j)} + \sum_{k \sim j} X^{(k)}$ and $W^{(j)} := W - Z^{(j)}$, noting that $W^{(j)}$ and $X^{(j)}$ are independent, so that $\widetilde{X}^{(j,k)} = X^{(k)}$, $k \in N_j \cup \{j\}$, and $\chi_1 = \chi_3 = 0$. Then, for $k = j$, $W^{(j,k)} = W^{(j)}$ and $Z^{(j,k)} = 0$; otherwise, for $j \neq k \in [n]$ such that $j \sim k$, we define $W^{(j,k)} := \sum_{l \notin N_j \cup N_k} X^{(l)}$ and $Z^{(j,k)} := W^{(j)} - W^{(j,k)}$; note that $W^{(j,k)}$ and the pair $(X^{(j)}, X^{(k)})$ are independent, so that $\chi_2 = 0$ also. The proof is completed by observing that $H_1 \leq \frac{1}{2}(H_0 + H_2)$, and that $\max\{H_0, H_2\} \leq \gamma \overline{D^2}$. □

The main difficulty in applying the bounds in Theorem 2.1 and Corollary 3.1 is putting a value to ε_W. This can nonetheless often be dealt with, provided that enough of the underlying random variables $(Y_l,\, l \in [M])$ each influence rather few of the $X^{(j)}$. The next theorem gives a way of exploiting this.

Given any $l \in [M]$, define $L_l := \{j \in [n] \colon M_j \ni l\}$, and $S_l := \sum_{j \in L_l} X^{(j)}$; write $\mathcal{G}_l := \sigma(Y_{l'},\, l' \in [M] \setminus \{l\})$, and define

$$d_l^{(i)}(Y) \;:=\; d_{\mathrm{TV}}(\mathcal{L}(S_l \,|\, \mathcal{G}_l), \mathcal{L}(S_l + e^{(i)} \,|\, \mathcal{G}_l)), \quad 1 \leq i \leq d.$$

Given any $j \neq k \in [n]$ such that $j \sim k$, define

$$M^{(j,k)} \;:=\; \bigcup_{j' \in N_j \cup N_k} M_{j'},$$

and find $l_1 < l_2 < \cdots < l_s \in [M] \setminus M^{(j,k)}$ such that $L_{l_r} \cap L_{l_{r'}} = \emptyset$ for all $1 \le r < r' \le s$. Then the vectors $S_{l_1}, \ldots, S_{l_s}$ are conditionally independent, given $\mathcal{F}^{(j,k)} := \sigma\big(Y_l,\, l \notin \{l_1,\ldots,l_s\}\big)$. Write

$$D_i^{(j,k)}(Y) \;:=\; \sum_{r=1}^{s} (1 - d_{l_r}^{(i)}(Y)).$$

Theorem 3.2. *Suppose that, for $j \ne k \in [n]$ such that $j \sim k$, we can find s and $l_1 < l_2 < \cdots < l_s \in [M] \setminus M^{(j,k)}$ such that the sets L_{l_r}, $1 \le r \le s$, are disjoint, and such that*

$$\mathbb{P}[D_i^{(j,k)}(Y) \le T] \;\le\; \eta.$$

Then

$$d_{\mathrm{TV}}\big(\mathcal{L}(W^{(j,k)} + e^{(i)} \,|\, \{X^{(r)},\, r \in N_j \cup N_k\}), \mathcal{L}(W^{(j,k)} \,|\, \{X^{(r)},\, r \in N_j \cup N_k\})\big) \;\le\; \left(\frac{2}{\pi T}\right)^{1/2} + \eta.$$

Proof. Writing $U^{(j,k)} := \sum_{r=1}^{s} S_{l_r}$, we have

$$\begin{aligned} &d_{\mathrm{TV}}\big(\mathcal{L}(W^{(j,k)} + e^{(i)} \,|\, \{X^{(r)},\, r \in N_j \cup N_k\}), \mathcal{L}(W^{(j,k)} \,|\, \{X^{(r)},\, r \in N_j \cup N_k\})\big) \\ &\quad\le\; \mathbb{E}\big\{d_{\mathrm{TV}}\big(\mathcal{L}(W^{(j,k)} + e^{(i)} \,|\, \mathcal{F}^{(j,k)}), \mathcal{L}(W^{(j,k)} \,|\, \mathcal{F}^{(j,k)})\big)\big\} \\ &\quad\le\; \mathbb{E}\big\{d_{\mathrm{TV}}\big(\mathcal{L}(U^{(j,k)} + e^{(i)} \,|\, \mathcal{F}^{(j,k)}), \mathcal{L}(U^{(j,k)} \,|\, \mathcal{F}^{(j,k)})\big)\big\}. \end{aligned}$$

Now, by the Mineka coupling argument [Lindvall (2002), Section II.14],

$$d_{\mathrm{TV}}\big(\mathcal{L}(U^{(j,k)} + e^{(i)} \,|\, \mathcal{F}^{(j,k)}), \mathcal{L}(U^{(j,k)} \,|\, \mathcal{F}^{(j,k)})\big) \;\le\; \left(\frac{2}{\pi D_i^{(j,k)}(Y)}\right)^{1/2},$$

where the constant comes from [Mattner & Roos (2007), Corollary 1.6], and the theorem follows. □

4 Examples

In this section, we demonstrate that Theorems 2.1 and 2.2 can be easily applied in a range of situations. The first three examples are discrete sums, and Theorem 2.1 can be invoked. In the last example, we need Theorem 2.2.

4.1 Graph colouring

As a first example, suppose that the vertices in a graph $G := ([M], E)$ are coloured independently, with colour i being chosen with probability π_i, $1 \le i \le d$. Let Y_l be the colour of vertex l, and let W_i denote the number of edges of G that connect two vertices of colour i; write $W := (W_1, \ldots, W_d)^T$. Then $n := |E|$ is the number of edges in G, and, for $j, k \in E$, $j \sim k$ if j and k share a common vertex. For $l \in [M]$, let δ_l denote the degree of l in G; then, for $j := \{l, l'\} \in E$, $D_j := |N_j| = \delta_l + \delta_{l'} - 2$. Define

$$\widetilde{D} \;:=\; n^{-1} \sum_{j \in E} D_j, \quad \widetilde{D^2} \;:=\; n^{-1} \sum_{j \in E} D_j^2.$$

Then it is easy to compute

$$\begin{aligned} \mu_i &= \mathbb{E} W_i \;=\; n\pi_i^2; \quad V_{ii} \;=\; \mathrm{Var}\, W_i \;=\; n\{\pi_i^2(1-\pi_i^2) + \widetilde{D}\pi_i^3(1-\pi_i)\}; \\ V_{ii'} &= \mathrm{Cov}(W_i, W_{i'}) \;=\; -n\pi_i^2\pi_{i'}^2(1+\widetilde{D}), \quad i \ne i'. \end{aligned}$$

Thus $\mathrm{Tr} V = nd\{c_1 + \widetilde{D} c_2\}$, where

$$c_1 \;:=\; d^{-1}\sum_{i=1}^{d} \pi_i^2(1-\pi_i^2); \quad c_2 \;:=\; d^{-1}\sum_{i=1}^{d} \pi_i^3(1-\pi_i),$$

so that we take $m := \lceil n(c_1 + \widetilde{D} c_2)\rceil$ in Corollary 3.1. We can clearly take $\gamma = 1$ also, and, for fixed d and $\pi_1, \ldots, \pi_d$, this yields a bound

$$d_{\mathrm{TV}}(\mathcal{L}(W), \mathrm{DN}_d(\mu, V)) \;=\; O\bigl\{(m^{-1/2} + \varepsilon_W)\log m\,(1 + \widetilde{D^2}/\widetilde{D})\bigr\},$$

which relies on having a reasonable bound for ε_W.

In order to apply Theorem 3.2, for each $j \sim k \in E$, we want first to find s and $l_1, l_2, \ldots, l_s \in [M] \setminus \{M_j \cup M_k\}$ such that the sets $L_{l_1}, \ldots, L_{l_s}$ are disjoint. Now $L_l = \bigl\{\{l,l'\}\colon l' \in [M], \{l,l'\} \in E\bigr\}$, so that $|L_l| = \delta_l$, and $L_l \cap L_{l'} \neq \emptyset$ exactly when $\{l,l'\} \in E$. Thus we need to find a set of vertices $l_1, \ldots, l_s$ subtending no edges of G (*independent* in the graph theoretical sense). Letting $\delta^* := \max_l \delta_l$, we note that $|[M] \setminus \{M_j \cup M_k\}| \geq M - 3\delta^*$, and that we can thus always take $s \geq s(M, \delta^*) := \lfloor M/(\delta^*+1)\rfloor - 3$.

The next step is to bound $d_l^{(i)}(Y) = d_l^{(i)}(\{Y_{l'},\, \{l,l'\} \in E\})$, for each $1 \leq i \leq d$ and for any l. To do so, let $R_{il} := \sum_{l'\colon \{l,l'\}\in E} I[Y_{l'} = i]$ be the number of neighbours of l in G that have colour i. Then S_l takes one of the values $R_{1l}e^{(1)}, \ldots, R_{dl}e^{(d)} \in \mathbb{Z}^d$, with conditional probabilities $\pi_1, \ldots, \pi_d$. Hence, if $R_{il} = 1$ and $R_{i'l} = 0$ for some $i' \neq i$, then $S_l = e^{(i)}$ with conditional probability π_i, and $S_l = 0$ with conditional probability at least $\pi_{i'}$, giving $d_l^{(i)} \leq 1 - \min\{\pi_i, \pi_{i'}\}$. Hence, for any $i' \neq i$, we have

$$\begin{aligned}\sum_{r=1}^{s}(1 - d_{l_r}^{(i)}) \quad &\geq \quad \sum_{r=1}^{s} I[R_{i,l_r} = 1, R_{i',l_r} = 0]\min\{\pi_i, \pi_{i'}\} \\ &=: \quad \min\{\pi_i, \pi_{i'}\}\widehat{R}(i,i').\end{aligned}$$

Now, if $\delta_l = t$,

$$\mathbb{P}[R_{il} = 1, R_{i'l} = 0] \;=\; h(t,i,i') \;:=\; t\pi_i(1 - \pi_i - \pi_{i'})^{t-1},$$

giving $\mathbb{E}\widehat{R}(i,i') = \sum_{r=1}^{s} h(\delta_{l_r}, i, i') \geq s h_{\min}(i,i')$, where $h_{\min}(i,i') := \min_{1\leq t\leq\delta^*} h(t,i,i')$. Then, since the events $\{R_{i,l_r} = 1, R_{i',l_r} = 0\}$ and $\{R_{i,l_{r'}} = 1, R_{i',l_{r'}} = 0\}$ are independent unless there is a path of length 2 connecting l_r and $l_{r'}$, we have

$$\mathrm{Var}\,(\widehat{R}(i,i')) \;\leq\; \mathbb{E}\widehat{R}(i,i')(1 + \delta^*(\delta^*-1)).$$

Hence, by Chebyshev's inequality,

$$\mathbb{P}[\widehat{R}(i,i') \leq \tfrac{1}{2}s(M,\delta^*)h_{\min}(i,i')] \;\leq\; \frac{4(1+\delta^*(\delta^*-1))}{\mathbb{E}\widehat{R}(i,i')} \;\leq\; \frac{4\delta^{*2}}{s(M,\delta^*)h_{\min}(i,i')}.$$

Thus we can take

$$T \;=\; \tfrac{1}{2}s(M,\delta^*)\min\{\pi_i,\pi_{i'}\}h_{\min}(i,i') \quad \text{and} \quad \eta \;=\; \frac{4\delta^{*2}}{s(M,\delta^*)h_{\min}(i,i')}$$

in Theorem 3.2. If, as $M \to \infty$, $n \geq cM$ for some $c > 0$ and δ^* remains bounded, with the colour probabilities remaining constant, this gives $\varepsilon_W = O(M^{-1/2})$, and so

$$d_{\mathrm{TV}}(\mathcal{L}(W), \mathrm{DN}_d(\mu, V)) \;=\; O\bigl\{M^{-1/2}\log M\bigr\}.$$

The order in M is the same as is obtained, in the context of δ^*-regular graphs and using the convex sets metric, by [Rinott & Rotar (1996)].

Note that, if most of the degrees in G become large as M increases, $h_{\min}(i,i')$ may well converge to zero too fast for the bound on ε_W to be useful, and more sophisticated arguments would be needed. Note also that, for $d=2$, $h(t,i,i')=h(t,1,2)=0$ for all $t\neq 1$, because $\pi_1+\pi_2=1$, and we obtain no bound on ε_W in this way. Indeed, if G is a δ^*-regular graph and $d=2$, ε_W is *not* small, since the distribution of W is concentrated on a sub-lattice of $\mathbb{Z}^2$ if $\delta^*\geq 2$, and $\mathcal{L}(W)$ is no longer close to $DN_d(\mu,V)$ in total variation; see [Barbour, Luczak & Xia (2018b), Section 4.2.1].

The problem can be modified, by only counting a random subset of monochrome edges. Let $(\widetilde{Y}_j,\, j\in E)$ be independent $\mathrm{Be}\,(p)$ random variables, and define $\widetilde{W}_i := \sum_{j\in E}\widetilde{Y}_j X^{(j)}$, where $X^{(j)}$ is as before. Then

$$\begin{aligned}\tilde{\mu} &:= \mathbb{E}\widetilde{W} \;=\; p\mu; \quad \widetilde{V}_{ii} \;:=\; \mathrm{Var}\,(\widetilde{W}_{ii}) \;=\; n\{p\pi_i^2(1-p\pi_i^2)+\widetilde{D}p^2\pi_i^3(1-\pi_i)\};\\ \widetilde{V}_{ii'} &:= \mathrm{Cov}(\widetilde{W}_i,\widetilde{W}_i') \;=\; -np^2\pi_i^2\pi_{i'}^2(1+\widetilde{D}),\end{aligned}$$

giving $m=\lceil nd\{pc_1+p(1-p)c_1'+p^2\widetilde{D}c_2\}\rceil$, where $c_1' := d^{-1}\sum_{i=1}^d \pi_i^4$. As before, for fixed d and $\pi_1,\ldots,\pi_d$, this yields

$$d_{\mathrm{TV}}(\mathcal{L}(\widetilde{W}),\mathrm{DN}_d(\tilde{\mu},\widetilde{V})) \;=\; O\bigl\{(m^{-1/2}+\tilde{\varepsilon}_n)(1+\widetilde{D^2}/\widetilde{D})\log m\bigr\}.$$

However, the quantity $\tilde{\varepsilon}_n$ is rather easier to bound than ε_W, since we can take the independent random variables $(\widetilde{Y}_j,\, j\in E)$ to use in Theorem 3.2, each of which influences only the corresponding $X^{(j)}$. Conditional on the colours, $\mathcal{G}:=\sigma(Y_l,\, l\in[M])$, we have

$$\widetilde{W}_i \;=\; \sum_{j=\{l,l'\}\in E} I[Y_l=Y_{l'}=i]\,\widetilde{Y}_j \;\sim\; \mathrm{Bi}\,(W_i,p),$$

with $\widetilde{W}_1,\ldots,\widetilde{W}_d$ conditionally independent, and hence

$$d_{\mathrm{TV}}(\mathcal{L}(\widetilde{W}\,|\,\mathcal{G}),\mathcal{L}(\widetilde{W}+e^{(i)}\,|\,\mathcal{G}))\leq 1/\sqrt{pW_i}.$$

Using the moments of W calculated above, it follows easily that $\tilde{\varepsilon}_n=O(\{np\}^{-1/2})$, giving

$$d_{\mathrm{TV}}(\mathcal{L}(\widetilde{W}),\mathrm{DN}_d(\tilde{\mu},\widetilde{V})) \;=\; O\bigl\{(Mp\widetilde{D})^{-1/2}(1+\widetilde{D^2}/\widetilde{D})\log M\bigr\}. \tag{4.1}$$

The apparent order in $\widetilde{D}$ is misleading here. If $\widetilde{D}$ is large, the covariance matrix V is ill conditioned, since $\mathrm{Tr}V\asymp n\widetilde{D}\asymp M\widetilde{D}^2$, whereas $\mathrm{Var}\bigl\{\sum_{i=1}^d \pi_i^{-1}W_i\bigr\}=n(d-1)\asymp M\widetilde{D}$. Thus the condition number $\rho(V)$ grows like $\widetilde{D}$, and $\rho(V)$ enters the constant $C_{2.1}$ implied in the order symbol in (4.1). However, if only the joint distribution of, say, $(\widetilde{W}_1,\ldots,\widetilde{W}_{d-1})^T$ is of interest, the corresponding covariance matrix then has condition number that is bounded in $\widetilde{D}$, for fixed d and $\pi_1,\ldots,\pi_d>0$, and the orders in both M and $\widetilde{D}$ are as in (4.1).

A more general modification, in the same spirit, is to choose $(\widetilde{Y}_j^{(i)},\, j\in E, 1\leq i\leq d)$ to be any independent integer valued random variables, with distributions depending only on i, and to set $\widetilde{W}_i := \sum_{j=\{l,l'\}\in E} I[Y_l=Y_{l'}=i]\widetilde{Y}_j^{(i)}$. Then, if the mean and variance of $\widetilde{Y}_1^{(i)}$ are $\tilde{m}^{(i)}$ and $\tilde{v}^{(i)}$, we have

$$\begin{aligned}\tilde{\mu}_i &:= \mathbb{E}\widetilde{W}_i \;=\; n\pi_i^2\tilde{m}^{(i)};\\ \widetilde{V}_{ii} &:= \mathrm{Var}\,(\widetilde{W}_{ii}) \;=\; n\pi_i^2\bigl\{v^{(i)}+(\tilde{m}^{(i)})^2\{1-\pi_i^2+\widetilde{D}\pi_i(1-\pi_i)\}\bigr\};\\ \widetilde{V}_{ii'} &:= \mathrm{Cov}(\widetilde{W}_i,\widetilde{W}_{i'}) \;=\; -n\pi_i^2\pi_{i'}^2\tilde{m}^{(i)}\tilde{m}^{(i')}(1+\widetilde{D}),\end{aligned}$$

from which the corresponding value of m can be deduced. As above, it is not difficult to show that

$$d_{\mathrm{TV}}(\mathcal{L}(\widetilde{W}_i\,|\,\mathcal{G}),\mathcal{L}(\widetilde{W}_i+1\,|\,\mathcal{G})) \;=\; O(1/\sqrt{u^{(i)}W_i}),$$

where $u^{(i)} := 1 - d_{\rm TV}(\widetilde{Y}_1^{(i)}, \widetilde{Y}_1^{(i)} + 1)$, from which it follows that $\tilde{\varepsilon}_n = O((M\widetilde{D})^{-1/2})$. Hence we find from Corollary 3.1 that

$$d_{\rm TV}(\mathcal{L}(\widetilde{W}), {\rm DN}_d(\tilde{\mu}, \widetilde{V})) \; = \; O\bigl\{(Mp\widetilde{D})^{-1/2}\gamma(1 + \widetilde{D^2}) \log M\bigr\}, \tag{4.2}$$

where $\gamma := \max_{1 \le i \le d} \mathbb{E}|\widetilde{Y}_1^{(i)}|^3$.

4.2 Random geometric graphs

Let $M := n^2$ points be distributed uniformly and independently over the torus $T_n := [0, n] \times [0, n]$. For some fixed r, join all pairs of points whose distance apart is less than or equal to r. This yields a particular example of a random geometric graph; the book by [Penrose (2003)] discusses much more general models, and gives a comprehensive treatment of their properties. In this section, we illustrate the application of Theorem 2.1 to counting induced triangles and 2-stars; more complicated examples can be treated in much the same way. If the positions of the points are denoted by $(Y_l,\, 1 \le l \le M)$, we express our statistic as

$$W \; := \; (W_1, W_2)^T : \quad W_i \; := \sum_{j \in [M]_3} I[G_j = G^{(i)}],$$

where $[M]_3$ denotes the set of 3-subsets of $[M]$, $G_j := G(Y_{j_1}, Y_{j_2}, Y_{j_3})$ denotes the induced graph on the points $Y_{j_1}, Y_{j_2}, Y_{j_3}$, $G^{(1)}$ denotes the triangle and $G^{(2)}$ denotes the 2-star.

For any $x \in T_n$, the probability that any given point lies in the circle of radius r around x is $\pi r^2/n^2 =: n^{-2}\hat{p}_r$. Hence $\mathbb{P}[G_j = G^{(i)}] = n^{-4}p_r^{(i)}$ is the same for all $j \in [M]_3$, and $p_r^{(1)}, p_r^{(2)} \le \hat{p}_r^2$. The quantities G_j and G_k are independent unless j and k have at least two of their vertices in common, G_j is independent of the set $(G_k \colon j \cap k = \emptyset)$, and the pair $(G_j, G_{j'})$ is independent of the set $(G_k \colon (j \cup j') \cap k = \emptyset)$. Using these facts, we can make some computations:

$$\begin{aligned} \mu \;\; &:= \;\; \mathbb{E}W \; = \; \binom{M}{3} n^{-4}(p_r^{(1)}, p_r^{(2)})^T \; \sim \; \tfrac{1}{3} n^2 (p_r^{(1)}, p_r^{(2)})^T; \\ V_{ii} \;\; &:= \;\; {\rm Var}\, W_i \; \sim \; c_{ii} n^2; \quad V_{12} \; := \; {\rm Cov}(W_1, W_2) \; \sim \; c_{12} n^2, \end{aligned}$$

and the matrix $\begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}$ is non-singular, with values involving the geometry of intersections of discs in $\mathbb{R}^2$. Thus we can take $m = cn^2$ for some $c > 0$. The quantities H_0 and H_2 are then easily bounded:

$$\begin{aligned} H_0 \;\; &\le \;\; \frac{1}{cn^2\sqrt{2}} \binom{M}{3} 2n^{-4}(p_r^{(1)} + p_r^{(2)}) \; \asymp \; 1; \\ H_2 \;\; &\le \;\; \frac{c'}{cn^2} \binom{M}{3} n^{-4}\hat{p}_r^2\{1 + (Mn^{-2}\hat{p}_r)^4\} \; \asymp \; 1, \end{aligned}$$

giving

$$d_{\rm TV}(\mathcal{L}(W), {\rm DN}_2(\mu, V)) \; = \; O((n^{-1} + \varepsilon_W) \log n).$$

It thus remains to bound ε_W.

To do so, break up $[0, n]^2$ into $9\lfloor n/3r \rfloor^2$ non-overlapping $r \times r$ squares, denoted by $Q_{l,l'} := [(l-1)r, lr) \times [(l'-1), l'r)$. Then there can be no triangles or 2-stars with points in two of the squares in $\mathcal{Q} := (Q_{3l,3l'}, 1 \le l, l' \le \lfloor n/3r \rfloor)$, because points in two of them are more than $2r$ apart. Consider evaluating $d_{\rm TV}(\mathcal{L}(W + e^{(i)} \,|\, \mathcal{G}), \mathcal{L}(W \,|\, \mathcal{G}))$, much as for Theorem 3.2, where $\mathcal{G}$ consists of the positions of all points not in members of $\mathcal{Q}$, together with the *numbers* of points falling in each member of $\mathcal{Q}$. If $S_{l,l'}$ denotes the contribution resulting from assigning positions to the points in $Q_{3l,3l'}$, then the random variables $(S_{l,l'},\, 1 \le l,l' \le \lfloor n/3r \rfloor)$ are conditionally independent, given $\mathcal{G}$. Let $N(A)$ denote the number of points falling in the set $A \subset T_n$. Then the event $E_{l,l'}$ that $N(Q_{3l,3l'}) = 1$, that the rectangle $[(3l+0.25)r, (3l+0.5)r) \times [(3l'-1)r, 3l'r)$ contains two points at a distance between $r/2$ and r from one another, and that $N(U_{l,l'}) = 3$, where $U_{l,l'}$ is the union of $(Q_{r,s},\, l-2 \le r \le l+2, l'-2 \le s \le l'+2)$, is such that $\mathbb{P}[E_{l,l'}] \asymp 1$ as $n \to \infty$, and is the same for all l,l'. Indeed, we have

$$\mathbb{P}[E_{l,l'}] \;=\; \chi\mathbb{P}[N(U_{1,1}) = 3],$$

for a constant $\chi > 0$ that is independent of n also. Conditional on $E_{l,l'}$, we have

$$1 - d_{\mathrm{TV}}\bigl(\mathcal{L}(S_{l,l'} \,|\, E_{l,l'}), \mathcal{L}(S_{l,l'} + e^{(i)} \,|\, E_{l,l'})\bigr) \;=\; u_i, \quad i = 1,2,$$

for $u_1, u_2 > 0$. Now the events $(E_{l,l'},\, 1 \le l,l' \le \lfloor n/3r \rfloor)$ are not independent, but, except for neighbouring pairs of indices, they are only weakly dependent: for r, r' such that $\max\{|r-l|, |r'-l'| \ge 2\}$, we have

$$\begin{aligned}
\mathbb{P}[E_{l,l'}, E_{r,r'}] &= \chi^2\mathbb{P}[\{N(U_{l,l'}) = 3\} \cap \{N(U_{r,r'}) = 3\}] \\
&= (\chi\mathbb{P}[N(U_{1,1}) = 3])^2 \mathrm{Bi}\,(n^2, 25r^2/n^2)\{3\}\,\mathrm{Bi}\,(n^2-3, 25r^2/(n^2-25r^2))\{3\} \\
&= \mathbb{P}[E_{l,l'}]\mathbb{P}[E_{r,r'}]\{1 + O(n^{-2})\}.
\end{aligned}$$

Hence

$$\mathbb{E}\Bigl\{\sum_{l,l'} I[E_{l,l'}]\Bigr\} \;=\; \lfloor n/3r \rfloor^2 \chi\mathbb{P}[N(U_{1,1}) = 3]; \quad \mathrm{Var}\Bigl\{\sum_{l,l'} I[E_{l,l'}]\Bigr\} \;=\; O(n^2),$$

and calculations as for Theorem 3.2 now easily yield $\varepsilon_W = O(n^{-1})$. Hence it follows that

$$d_{\mathrm{TV}}(\mathcal{L}(W), \mathrm{DN}_2(\mu, V)) \;=\; O(n^{-1}\log n).$$

The asymptotics of ε_W are, however, sensitive to the choice of r: if $r = r_n \to \infty$, even logarithmically in n, $\mathbb{P}[N(U_{1,1}) = 3]$ becomes very small, and the bound on ε_W derived in this way is no longer useful.

4.3 Finite Markov chains

Let $(Z_j,\, j \ge 0)$ be an irreducible, aperiodic Markov chain on the finite state space $\{0,1,\ldots,d\}$, and set $X^{(j)} := (I[Z_j = 1], \ldots, I[Z_j = d])^T$. Let $W_n := \sum_{j=1}^n X^{(j)}$ denote the vector of the amounts of time spent in the states $1 \le i \le d$ between times 1 and n. We are interested in the accuracy of approximating the distribution of W_n by $\mathrm{DN}_d(\mu_n, V_n)$, where $\mu_n := \mathbb{E}W_n$ and $V_n := \mathrm{Cov}W_n$; translated Poisson approximation for each component W_{in} separately can be shown to be accurate to order $O(n^{-1/2})$ using the results of [Barbour & Lindvall (2006)]. In a Markov chain, the dependence between the states at different times never completely disappears, so we shall need to make use of the dependence coefficients χ_l, $1 \le l \le 3$. We make the following simplifying assumption:

Assumption A1: $\mathbb{P}[Z_1 = i \,|\, Z_0 = i] \;>\; 0$ for all $0 \le i \le d$.

Clearly, a local decomposition in which

$$Z^{(j)} \;:=\; \sum_{|l-j| \le m_n} X^{(l)} \quad \text{and} \quad Z^{(j,k)} \;:=\; \sum_{\max\{|l-j|,|l-k|\} \le m_n} X^{(l)}$$

is likely to be effective, if m_n is suitably chosen. Because a finite state irreducible aperiodic Markov chain is geometrically ergodic, there exist $0 < \rho < 1$ and $C < \infty$ such that, for all $0 \le i, r \le d$, we have

$$|P_{ir}^{(k)} - \pi_r| \;\le\; C\rho^k, \tag{4.3}$$

where $P_{ir}^{(k)} := \mathbb{P}[Z_k = r \,|\, Z_0 = i]$, and $\pi_r := \lim_{n\to\infty} \mathbb{P}[Z_n = r \,|\, Z_0 = i]$. It is then easy to deduce that, as $n \to \infty$,

$$\begin{aligned} n^{-1}\mu_n &\sim (\pi_1, \ldots, \pi_d)^T \;=:\; \boldsymbol{\pi}; \\ n^{-1}V_{ir;n} &\sim \Big\{\pi_i \sum_{k\ge 1}(P_{ir}^{(k)} - \pi_r) + \pi_r \sum_{k\ge 1}(P_{ri}^{(k)} - \pi_i) + \delta_{ir} - \pi_i\pi_r\Big\} \;=:\; V_{ir}, \end{aligned} \tag{4.4}$$

for $1 \le i, r \le d$. Now, from (4.3), uniformly in i, r, s, q,

$$\mathbb{P}[Z_0 = i, Z_j = r, Z_{j+k} = s \,|\, Z_0 = i, Z_{j+k} = s] \;=\; \frac{\mathbb{P}[Z_0 = i]P_{ir}^{(j)}P_{rq}^{(k)}}{\mathbb{P}[Z_0 = i]P_{iq}^{(j+k)}} \;=\; \pi_r\big(1 + O(\rho^{(j\wedge k)})\big),$$

and

$$\begin{aligned} &\mathbb{P}[Z_0 = i, Z_j = r, Z_{j+l} = s, Z_{j+l+k} = q \,|\, Z_0 = i, Z_{j+l+k} = q] \\ &\quad= \frac{\mathbb{P}[Z_0 = i]P_{ir}^{(j)}P_{rs}^{(l)}P_{sq}^{(k)}}{\mathbb{P}[Z_0 = i]P_{iq}^{(j+l+k)}} \;=\; \pi_r P_{rs}^{(l)}\big(1 + O(\rho^{(j\wedge k)})\big), \end{aligned}$$

so that all the dependence coefficients χ_l, $1 \le l \le 3$, in (2.4) are of order $O(\rho^{m_n})$. It is also immediate, because indicators are bounded random variables, that $H_0 = O(1)$, $H_1 = O(m_n)$ and $H_2 = O(m_n^2)$. It thus remains to consider the quantity ε_W of (2.7).

Assuming that $2m_n \le n/4$, it is enough to bound

$$d_{\mathrm{TV}}\Big(\mathcal{L}_r\Big(\sum_{j=1}^{l} X^{(j)} + e^{(i)}\Big), \mathcal{L}_r\Big(\sum_{j=1}^{l} X^{(j)}\Big)\Big), \tag{4.5}$$

for any $l \ge \lfloor n/4 \rfloor$ and $0 \le r \le d$, where $\mathcal{L}_r$ stands for the distribution given the initial state of the Markov chain is at r. This is because, for $1 \le j \le n/2$ and $k \le j + m_n$, conditioning on the values of Z_j up to time $j + k + m_n$ and using the Markov property, the quantity ε_W in (2.7) is no bigger than any bound for the distance in (4.5), for $l = n - j - k - m_n \ge n/2 - 2m_n \ge n/4$, that is uniform in the initial state r. For $j > n/2$, we note that the same argument works, using the *reversed* Markov chain. We establish (4.5) by using coupling.

Let Z' and Z'' be two copies of the Markov chain Z, both starting in r. We couple them in such a way that the sequence of transitions in the first is the same as that in the second, except that the holding times in 0 and i are allowed to be different. Initially, if $(N'_{0l},\, l \ge 1)$ and $(N''_{0l},\, l \ge 1)$ denote the sequence of successive holding times in 0 of the two chains, then the pair (N'_{0l}, N''_{0l}) is chosen independently of the past according to the Mineka coupling [Lindvall (2002), Section II.14], so that $(N'_{0l} - N''_{0l},\, l \ge 1)$ are the increments of a lazy symmetric random walk with steps in $\{-1, 0, 1\}$. After the first occasion L_0 such that

$$\sum_{l=1}^{L_0}\{N'_{0l} - N''_{0l}\} \;=\; 1,$$

the values of N'_{0l} and N''_{0l} are chosen to be identical. The same strategy is applied to the holding times N'_{il} and N''_{il}, except that they are chosen to be identical after the first occasion L_i on which

$$\sum_{l=1}^{L_i} \{N'_{il} - N''_{il}\} \;=\; -1.$$

Let M_{0i} denote the first time in the underlying Markov chains Z' and Z'' at which both of these occasions have occurred. At this point, both chains have made the same number of steps, because their paths differ only through differences in the partial sums $\sum_l \{N'_{0l} + N'_{il}\}$ and $\sum_l \{N''_{0l} + N''_{il}\}$, and these are equal at all times after M_{0i}. However, at this point, both have spent the same amount of time in states other than i and 0, but Z' has spent one step less in i. By the usual coupling argument, for any set $A \subset \mathbb{Z}^d$,

$$\begin{aligned}
\mathbb{P}_r\Big[\sum_{j=1}^k X^{(j)} + e^{(i)} \in A\Big] \;&=\; \mathbb{P}_r\Big[\sum_{j=1}^k (X^{(j)})' + e^{(i)} \in A\Big] \\
&= \mathbb{P}_r\Big[\Big\{\sum_{j=1}^k (X^{(j)})' + e^{(i)} \in A\Big\} \cap \{M_{0i} \le k\}\Big] \\
&\qquad + \mathbb{P}_r\Big[\Big\{\sum_{j=1}^k (X^{(j)})' + e^{(i)} \in A\Big\} \cap \{M_{0i} > k\}\Big] \\
&= \mathbb{P}_r\Big[\Big\{\sum_{j=1}^k (X^{(j)})'' \in A\Big\} \cap \{M_{0i} \le k\}\Big] + \mathbb{P}_r\Big[\Big\{\sum_{j=1}^k (X^{(j)})' + e^{(i)} \in A\Big\} \cap \{M_{0i} > k\}\Big] \\
&= \mathbb{P}_r\Big[\Big\{\sum_{j=1}^k (X^{(j)})'' \in A\Big\}\Big] + \mathbb{P}_r\Big[\Big\{\sum_{j=1}^k (X^{(j)})' + e^{(i)} \in A\Big\} \cap \{M_{0i} > k\}\Big] \\
&\qquad - \mathbb{P}_r\Big[\Big\{\sum_{j=1}^k (X^{(j)})'' \in A\Big\} \cap \{M_{0i} > k\}\Big] \\
&= \mathbb{P}_r\Big[\Big\{\sum_{j=1}^k X^{(j)} \in A\Big\}\Big] + \mathbb{P}_r\Big[\Big\{\sum_{j=1}^k (X^{(j)})' + e^{(i)} \in A\Big\} \cap \{M_{0i} > k\}\Big] \\
&\qquad - \mathbb{P}_r\Big[\Big\{\sum_{j=1}^k (X^{(j)})'' \in A\Big\} \cap \{M_{0i} > k\}\Big].
\end{aligned}$$

It thus follows that

$$d_{\mathrm{TV}}\Big(\mathcal{L}_r\Big(\sum_{j=1}^k X^{(j)} + e^{(i)}\Big), \mathcal{L}_r\Big(\sum_{j=1}^k X^{(j)}\Big)\Big) \;\le\; \mathbb{P}_r[M_{0i} > k].$$

Now we have $\mathbb{P}[L_0 > l] = O(l^{-1/2})$ and $\mathbb{P}[L_i > l] = O(l^{-1/2})$, by [Lindvall (2002), Section II.14]. Also, because Z has finite state space, the times between visits to 0 and between visits to i have means γ_0 and γ_i and finite variances v_0 and v_i. So, if τ'_{0l} denotes the time at which Z' completes its l-th visit to 0, we have

$$\{M_{0i} > \tfrac{1}{4}n\} \subset \{L_0 > \alpha n\} \cup \{\tau'_{0,\alpha n} > \tfrac{1}{4}n\} \cup \{L_i > \alpha n\} \cup \{\tau'_{i,\alpha n} > \tfrac{1}{4}n\}.$$

Hence it follows by Chebyshev's inequality that, if $\alpha \max\{\gamma_0, \gamma_i\} < 1/8$, then

$$\mathbb{P}_r[M_{0i} > \tfrac{1}{4}n] \;\le\; \mathbb{P}_r[L_0 > \alpha n] + \mathbb{P}_r[L_i > \alpha n] + \frac{\alpha n v_0}{(\frac{1}{4}n - \alpha\gamma_0 n)^2} + \frac{\alpha n v_i}{(\frac{1}{4}n - \alpha\gamma_i n)^2} \;=\; O(n^{-1/2}),$$

where this order follows for the first pair of terms as above, and the second pair are of order $O(n^{-1})$. This shows that $\varepsilon_W = O(n^{-1/2})$.

Theorem 4.1. *Let $(Z_j,\, j \ge 1)$ be an irreducible, aperiodic Markov chain on a finite state space $\{0,1,\ldots,d\}$, that satisfies Assumption A1. Let $W_n := (W_{n1},\ldots,W_{nd})^T$ represent the number of steps spent in the states $1,2,\ldots,d$ up to time n. Then, for any $0 \le r \le d$,*

$$d_{\mathrm{TV}}(\mathcal{L}_r(W_n), \mathrm{DN}_d(n\boldsymbol{\pi}, nV)) \;=\; O(n^{-1/2}\log^3 n),$$

where $\boldsymbol{\pi}$ and V are as given in (4.4).

Proof. We apply Theorem 2.1, taking $m_n = \log n/\log(1/\rho)$, so that $\chi_1 + \chi_2 + \chi_3 = O(n^{-1})$. Then $\log n\, H_2(\varepsilon_W + m^{-1/2}) = O(n^{-1/2}\log^3 n)$ represents the largest order term in the error bound. Finally, it follows from (4.3) that $|\mathbb{E}W_n - n\boldsymbol{\pi}| = O(1)$ and that $|\mathrm{Cov}(W_n)_{ir} - nV_{ir}| = O(1)$ for each i, r also, so that, to the stated accuracy, we can replace the mean and covariance by $n\boldsymbol{\pi}$ and V respectively. □

4.4 Maximal points

Given a configuration Ξ of points in $\mathbb{R}^2$, a point $\boldsymbol{\alpha} = (\alpha_1,\alpha_2)^T \in \Xi$ is called *maximal* if there are no other points $\boldsymbol{\beta} = (\beta_1,\beta_2)^T \in \Xi$ such that $\beta_i \ge \alpha_i$ for $i = 1,2$. In this example, we take Ξ to be a realisation of a Poisson point process with intensity λ on the triangle

$$\Gamma \;:=\; \{\boldsymbol{\alpha} = (\alpha_1,\alpha_2)^T :\; 0 \le \alpha_2 \le 1-\alpha_1,\; 0 \le \alpha_1 \le 1\}.$$

Letting

$$A_{\boldsymbol{\alpha}} \;:=\; \{(x_1,x_2)^T :\; \alpha_1 \le x_1 \le 1-\alpha_2, \alpha_2 \le x_2 \le 1-x_1\} \setminus \{\boldsymbol{\alpha}\},$$

a point $\boldsymbol{\alpha}$ of Ξ is maximal if $\Xi(A_{\boldsymbol{\alpha}}) = 0$. The process of maximal points of Ξ can thus be written as the random point measure $\Upsilon(d\boldsymbol{\alpha}) := \mathbf{1}_{[\Xi(A_{\boldsymbol{\alpha}})=0]}\Xi(d\boldsymbol{\alpha})$, and has mean measure

$$\upsilon(d\boldsymbol{\alpha}) \;:=\; \mathbb{E}\Upsilon(d\boldsymbol{\alpha}) = \lambda e^{-\frac{1}{2}\lambda(1-\alpha_1-\alpha_2)^2} d\alpha_1 d\alpha_2.$$

For $0 \le b_1 < d_1 \le b_2 < d_2 < \infty$, define the strips

$$E_i \;:=\; \big\{\boldsymbol{\alpha} = (\alpha_1,\alpha_2)^T : (1-d_i\lambda^{-1/2}-\alpha_1)\vee 0 \le \alpha_2 < 1-b_i\lambda^{-1/2}-\alpha_1,\; 0 \le \alpha_1 \le 1-b_i\lambda^{-1/2}\big\},$$

parallel to the hypotenuse of Γ and close to it, and define $Y_i = \Upsilon(E_i)$. Our interest is in the approximate joint distribution of $(Y_1,Y_2)^T$.

Proposition 4.2. *Let $\phi(x) = e^{-\frac{x^2}{2}}$ and $\hat m_i = \int_{b_i}^{d_i}\phi(x)dx$, and define*

$$\begin{aligned}
\sigma_{ii} \;&:=\; \hat m_i + 2\hat m_i^2 \int_0^{b_i} \frac{1}{\phi(x)}dx + 2\int_{b_i}^{d_i}\phi(z)dz \int_{b_i}^{z}\frac{1}{\phi(y)}dy\int_y^{d_i}\phi(x)\,dx\\
&\quad -2\hat m_i(\phi(b_i)-\phi(d_i)),\; i = 1,2;\\
\sigma_{12} \;&:=\; 2\hat m_2 \int_{b_1}^{d_1}\phi(z)dz\int_0^z \frac{1}{\phi(y)}\,dy\\
&\quad - \{\hat m_1(\phi(b_2)-\phi(d_2)) + \hat m_2(\phi(b_1)-\phi(d_1))\}\,.
\end{aligned}$$

Then, as $\lambda \to \infty$,

$$\mathbb{E}Y_i \;=\; \upsilon(E_i) \sim \hat m_i\sqrt{\lambda};\quad \mathrm{Var}\,(Y_i) \;\sim\; \sigma_{ii}\sqrt{\lambda},\; i=1,2;\quad \mathrm{Cov}(Y_1,Y_2) \;\sim\; \sigma_{12}\sqrt{\lambda}.$$

Proof. Since $\upsilon(d\boldsymbol{\alpha}) = \mathbb{E}\Upsilon(d\boldsymbol{\alpha}) = \lambda\phi(\sqrt{\lambda}(1-\alpha_1-\alpha_2))d\alpha_1 d\alpha_2$, we have

$$\upsilon(E_i) \;=\; \lambda\int_0^{1-b_i\lambda^{-1/2}} d\alpha_1 \int_{0\vee(1-\alpha_1-d_i\lambda^{-1/2})}^{1-\alpha_1-b_i\lambda^{-1/2}} \phi\left(\sqrt{\lambda}(1-\alpha_1-\alpha_2)\right) d\alpha_2.$$

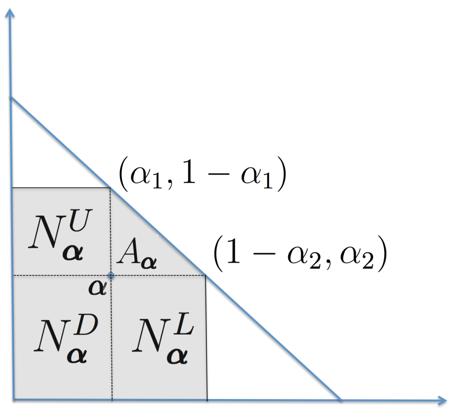

Figure 1: The dependence neighbourhood

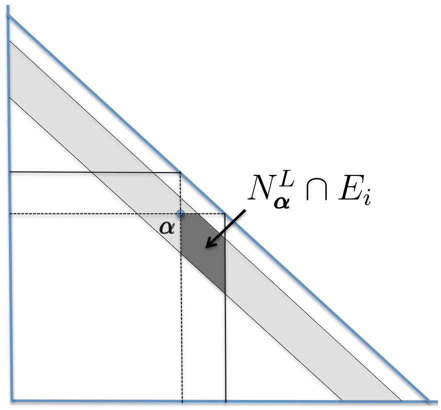

Figure 2: Dark area is $N^L_{\boldsymbol{\alpha}} \cap E_i$

By taking $x = \sqrt{\lambda}(1-\alpha_1-\alpha_2)$ and $y = \alpha_1$, we obtain

$$\upsilon(E_i) \;=\; \sqrt{\lambda}\int_0^{1-b_i\lambda^{-1/2}} dy \int_{b_i}^{d_i\wedge(\sqrt{\lambda}(1-y))} \phi(x)\,dx,$$

from which the first claim follows.

Next, referring to Figure 1, we define

$$\begin{aligned} N^U_{\boldsymbol{\alpha}} &:= \{(x_1,x_2)^T:\ 0\le x_1<\alpha_1, \alpha_2\le x_2\le 1-\alpha_1\};\\ N^L_{\boldsymbol{\alpha}} &:= \{(x_1,x_2)^T:\ \alpha_1\le x_1\le 1-\alpha_2, 0\le x_2<\alpha_2\};\\ N^D_{\boldsymbol{\alpha}} &:= \{(x_1,x_2)^T:\ 0\le x_1<\alpha_1, 0\le x_2<\alpha_2\}, \end{aligned}$$

and then set $N_{\boldsymbol{\alpha}} = A_{\boldsymbol{\alpha}} \cup N^U_{\boldsymbol{\alpha}} \cup N^L_{\boldsymbol{\alpha}} \cup N^D_{\boldsymbol{\alpha}}$. Then, since $I[\Xi(A_{\boldsymbol{\alpha}})=0]$ is independent of $I[\Xi(A_{\boldsymbol{\beta}})=0]$ for $\boldsymbol{\beta}\notin N_{\boldsymbol{\alpha}}\cup\{\boldsymbol{\alpha}\}$, and $\Xi(N^D_{\boldsymbol{\alpha}}) = \Xi(A_{\boldsymbol{\alpha}}) = 0$ if $\Upsilon(\{\boldsymbol{\alpha}\})=1$, we have

$$\begin{aligned} \mathrm{Var}\,(Y_i) \;&=\; \upsilon(E_i) + \int_{E_i} \mathbb{E}\left(\Upsilon((N^L_{\boldsymbol{\alpha}}\cup N^U_{\boldsymbol{\alpha}})\cap E_i)\,|\,\Upsilon(\{\boldsymbol{\alpha}\})=1\right)\upsilon(d\boldsymbol{\alpha})\\ &\quad - \int_{E_i} \upsilon(N_{\boldsymbol{\alpha}}\cap E_i)\upsilon(d\boldsymbol{\alpha}). \end{aligned} \tag{4.6}$$

However, using Figure 2, we obtain

$$\begin{aligned} &\int_{E_i} \mathbb{E}\left(\Upsilon(N^L_{\boldsymbol{\alpha}}\cap E_i)|\Upsilon(\{\boldsymbol{\alpha}\})=1\right)\upsilon(d\boldsymbol{\alpha})\\ &= \lambda\int_{E_i}\upsilon(d\boldsymbol{\alpha})\int_{\alpha_1}^{1-\alpha_2} d\beta_1 \int_{(1-d_i\lambda^{-1/2}-\beta_1)\vee 0}^{\alpha_2\wedge(1-b_i\lambda^{-1/2}-\beta_1)} e^{-\frac{\lambda}{2}((1-\beta_1-\beta_2)^2-(1-\beta_1-\alpha_2)^2)}\,d\beta_2\\ &= \sqrt{\lambda}\int_0^{1-b_i\lambda^{-1/2}} d\alpha_1 \int_{b_i}^{d_i\wedge((1-\alpha_1)\sqrt{\lambda})} \phi(z)\,dz \int_0^z \frac{1}{\phi(y)}dy \int_{y\vee b_i}^{d_i\wedge(y-z+\sqrt{\lambda}(1-\alpha_1))} \phi(x)\,dx\\ &\sim \sqrt{\lambda}\hat{m}_i^2\int_0^{b_i}\frac{1}{\phi(y)}dy + \sqrt{\lambda}\int_{b_i}^{d_i}\phi(z)\,dz\int_{b_i}^{z}\frac{1}{\phi(y)}dy\int_y^{d_i}\phi(x)\,dx, \end{aligned} \tag{4.7}$$

where the last equality is from the change of variables

$$1-\alpha_2 \;=\; \alpha_1 + z\lambda^{-1/2}, \quad x \;=\; (1-\beta_1-\beta_2)\sqrt{\lambda} \;\;\text{and}\;\; y \;=\; (1-\beta_1-\alpha_2)\sqrt{\lambda}. \tag{4.8}$$

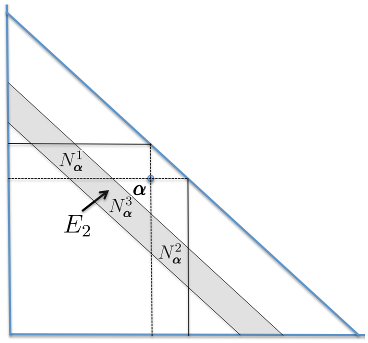

Figure 3: $N^1_{\boldsymbol{\alpha}},\ N^2_{\boldsymbol{\alpha}},\ N^3_{\boldsymbol{\alpha}}$

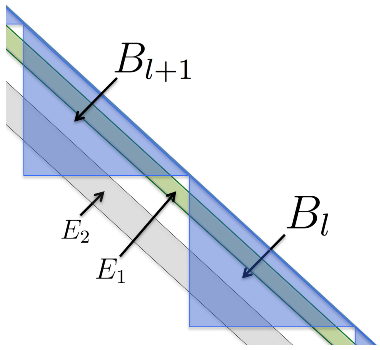

Figure 4: B_l

By symmetry, the calculation for $N^D_{\boldsymbol{\alpha}} \cap E_i$ gives an identical result. Similarly, by taking $1-\alpha_2 = \alpha_1 + z\lambda^{-1/2}$, $y = \sqrt{\lambda}(\alpha_1 - \beta_1)$ and $x = \sqrt{\lambda}(1-\beta_1-\beta_2)$ in the second equality below, we get

$$\begin{aligned} \upsilon(N_{\boldsymbol{\alpha}} \cap E_i) &= \int_{(\alpha_1 - d_i\lambda^{-1/2})\vee 0}^{1-\alpha_2} d\beta_1 \int_{(1-d_i\lambda^{-1/2}-\beta_1)\vee 0}^{(1-\alpha_1)\wedge(1-b_i\lambda^{-1/2}-\beta_1)} \lambda\phi(\sqrt{\lambda}(1-\beta_1-\beta_2))d\beta_2 \\ &= \int_{-z}^{d_i\wedge(\alpha_1\sqrt{\lambda})} dy \int_{b_i\vee y}^{d_i\wedge(y+\sqrt{\lambda}(1-\alpha_1))} \phi(x)dx, \end{aligned}$$

which implies that

$$\begin{aligned} &\int_{E_i} \upsilon(N_{\boldsymbol{\alpha}} \cap E_i)\upsilon(d\boldsymbol{\alpha}) \\ &= \sqrt{\lambda}\int_0^{1-b_i\lambda^{-1/2}} d\alpha_1 \int_{b_i}^{d_i\wedge((1-\alpha_1)\sqrt{\lambda})} \phi(z)\,dz \int_{-z}^{d_i\wedge(\alpha_1\sqrt{\lambda})} dy \int_{b_i\vee y}^{d_i\wedge(y+\sqrt{\lambda}(1-\alpha_1))} \phi(x)\,dx \\ &\sim \sqrt{\lambda}\int_{b_i}^{d_i} \phi(z)dz \int_{-z}^{d_i} dy \int_{b_i\vee y}^{d_i} \phi(x)\,dx \\ &= 2\sqrt{\lambda}\hat{m}_i(\phi(b_i)-\phi(d_i)). \end{aligned} \tag{4.9}$$

Combining (4.7) and (4.9) with (4.6) gives the second claim.

Finally we estimate $\mathrm{Cov}(Y_1, Y_2)$. For $\boldsymbol{\alpha} \in E_1$, we refer to Figure 3 and define $N^1_{\boldsymbol{\alpha}} := E_2 \cap N^U_{\boldsymbol{\alpha}}$, $N^2_{\boldsymbol{\alpha}} := E_2 \cap N^L_{\boldsymbol{\alpha}}$ and $N^3_{\boldsymbol{\alpha}} := E_2 \cap N^D_{\boldsymbol{\alpha}}$. Then we can express the covariance as

$$\mathrm{Cov}(Y_1,Y_2) = 2\int_{E_1} \mathbb{E}[\Upsilon(N^2_{\boldsymbol{\alpha}})|\Upsilon(\{\boldsymbol{\alpha}\})=1)]\upsilon(d\boldsymbol{\alpha}) - \int_{E_1} \upsilon(N^1_{\boldsymbol{\alpha}} \cup N^2_{\boldsymbol{\alpha}} \cup N^3_{\boldsymbol{\alpha}})\upsilon(d\boldsymbol{\alpha}). \tag{4.10}$$

For the first term, we have

$$\begin{aligned} &\mathbb{E}[\Upsilon(N^2_{\boldsymbol{\alpha}})|\Upsilon(\{\boldsymbol{\alpha}\})=1)] \\ &= \lambda\int_{\alpha_1}^{1-\alpha_2} d\beta_1 \int_{(1-d_2\lambda^{-1/2}-\beta_1)\vee 0}^{1-b_2\lambda^{-1/2}-\beta_1} \frac{\phi(\sqrt{\lambda}(1-\beta_1-\beta_2))}{\phi(\sqrt{\lambda}(1-\beta_1-\alpha_2))}d\beta_2 \\ &= \int_0^z \phi^{-1}(y)dy \int_{b_2}^{d_2\wedge(y-z+\sqrt{\lambda}(1-\alpha_1))} \phi(x)dx \\ &\sim \hat{m}_2\int_0^z \phi^{-1}(y)dy, \end{aligned} \tag{4.11}$$

for $\alpha_1 < 1$, where the last equality is from the change of variables specified in (4.8). It thus follows from (4.11) that

$$\begin{aligned}
&\int_{E_1} \mathbb{E}[\Upsilon(N^2_{\boldsymbol\alpha})|\Upsilon(\{\boldsymbol\alpha\})=1)]\upsilon(d\boldsymbol\alpha)\\
&\sim\ \hat m_2\sqrt{\lambda}\int_0^{1-b_1\lambda^{-1/2}} d\alpha_1 \int_{b_1}^{d_1\wedge((1-\alpha_1)\sqrt{\lambda})}\phi(z)\,dz\int_0^z \phi^{-1}(y)\,dy\\
&\sim\ \hat m_2\sqrt{\lambda}\int_{b_1}^{d_1}\phi(z)\,dz\int_0^z\phi^{-1}(y)\,dy. \qquad (4.12)
\end{aligned}$$

Likewise, using the convention that $\int_{c_1}^{c_2} f(x)dx = 0$ for $c_1 > c_2$, we have

$$\begin{aligned}
&\int_{E_1}\upsilon(N^1_{\boldsymbol\alpha}\cup N^2_{\boldsymbol\alpha}\cup N^3_{\boldsymbol\alpha})\upsilon(d\boldsymbol\alpha)\\
&=\ \lambda\int_{E_1}\upsilon(d\boldsymbol\alpha)\int_{(\alpha_1-d_2\lambda^{-1/2})\vee 0}^{1-\alpha_2} d\beta_1\int_{(1-d_2\lambda^{-1/2}-\beta_1)\vee 0}^{(1-b_2\lambda^{-1/2}-\beta_1)\wedge(1-\alpha_1)}\phi(\sqrt{\lambda}(1-\beta_1-\beta_2))d\beta_2\\
&=\ \sqrt{\lambda}\int_0^{1-b_1\lambda^{-1/2}} d\alpha_1\int_{b_1}^{d_1\wedge(\sqrt{\lambda}(1-\alpha_1))}\phi(z)\,dz\\
&\qquad\qquad\times\int_0^{(z+d_2)\wedge(z+\sqrt{\lambda}\alpha_1)} dy\int_{b_2\vee(y-z)}^{d_2\wedge(y-z+(1-\alpha_1)\sqrt{\lambda})}\phi(x)\,dx\\
&\sim\ \sqrt{\lambda}\int_{b_1}^{d_1}\phi(z)dz\int_0^{z+d_2}dy\int_{b_2\vee(y-z)}^{d_2}\phi(x)\,dx\\
&=\ \sqrt{\lambda}\hat m_1(\phi(b_2)-\phi(d_2))+\sqrt{\lambda}\hat m_2(\phi(b_1)-\phi(d_1)), \qquad (4.13)
\end{aligned}$$

where, again, we used the the change of variables in (4.8) for the penultimate equality. Combining (4.12) and (4.13) with (4.10) completes the proof. □

Theorem 4.3. *Let $W = (Y_1, Y_2)^T$, $\mu = \mathbb{E}W$ and $V = \mathrm{Cov}(W)$ be as in Proposition 4.2. Then, as $\lambda\to\infty$,*

$$d_{\mathrm{TV}}(\mathcal{L}(W), \mathrm{DN}_2(\mu, V))\ =\ O\left(\lambda^{-1/4}\ln(\lambda)\right).$$

Proof. In order to apply Theorem 2.2 to the maximal points in $E' := E_1\cup E_2$, we need to establish suitable decompositions. As neighbourhoods, we take $D_{\boldsymbol\alpha} := N_{\boldsymbol\alpha}\cup\{\boldsymbol\alpha\}$, with $N_{\boldsymbol\alpha}$ as defined in the proof of Proposition 4.2 (see Figure 1). Proposition 4.2 ensures that $\mathrm{Tr}\,V\asymp\lambda^{1/2}$, and so $m\asymp\lambda^{1/2}$ also. We assign a mark

$$X^{(\boldsymbol\alpha)}\ :=\ \mathbf{1}_{[\Xi(A_{\boldsymbol\alpha})=0]}(\mathbf{1}_{[\boldsymbol\alpha\in E_1]}, \mathbf{1}_{[\boldsymbol\alpha\in E_2]})^T$$

if $\Xi(\{\boldsymbol\alpha\}) = 1$, so that $\mu^{(\boldsymbol\alpha)} = \mathbb{E}X^{(\boldsymbol\alpha)}$, and define

$$\tilde X^{(\boldsymbol\alpha,\boldsymbol\beta)}\ :=\ X^{(\boldsymbol\beta)}, \qquad \boldsymbol\beta\in D_{\boldsymbol\alpha}.$$

Then

$$W\ =\ \int_{\boldsymbol\alpha\in E'} X^{(\boldsymbol\alpha)}\Xi(d\boldsymbol\alpha), \quad\text{and}\quad \nu(d\boldsymbol\alpha)\ =\ \lambda d\alpha_1 d\alpha_2, \quad \nu_2(d\boldsymbol\alpha, d\boldsymbol\beta)\ =\ \nu(d\boldsymbol\alpha)\nu(d\boldsymbol\beta).$$

We now decompose the integral as follows. For each $\boldsymbol\alpha\in E'$, define

$$\begin{aligned}
Z^{(\boldsymbol\alpha)}\ &:=\ \int_{\boldsymbol\beta\in D_{\boldsymbol\alpha}\cap E'} X^{(\boldsymbol\beta)}\Xi(d\boldsymbol\beta); \quad W^{(\boldsymbol\alpha)}\ :=\ \int_{\boldsymbol\gamma\in D^c_{\boldsymbol\alpha}\cap E'} X^{(\boldsymbol\gamma)}\Xi(d\boldsymbol\gamma);\\
Z^{(\boldsymbol\alpha,\boldsymbol\beta)}\ &:=\ \int_{\boldsymbol\gamma\in D^c_{\boldsymbol\alpha}\cap D_{\boldsymbol\beta}\cap E'} X^{(\boldsymbol\gamma)}\Xi(d\boldsymbol\gamma); \quad W^{(\boldsymbol\alpha,\boldsymbol\beta)}\ :=\ \int_{\boldsymbol\gamma\in D^c_{\boldsymbol\alpha}\cap D^c_{\boldsymbol\beta}\cap E'} X^{(\boldsymbol\gamma)}\Xi(d\boldsymbol\gamma).
\end{aligned}$$

This decomposition ensures that $X^{(\boldsymbol{\alpha})}$ is independent of $W^{(\boldsymbol{\alpha})}$ with respect to $\mathbb{P}$ and $\mathbb{P}_{\boldsymbol{\alpha}}$, and that $W^{(\boldsymbol{\alpha},\boldsymbol{\beta})}$ is independent of $(X^{(\boldsymbol{\alpha})}, X^{(\boldsymbol{\beta})})$ with respect to $\mathbb{P}$, $\mathbb{P}_{\boldsymbol{\alpha}}$, $\mathbb{P}_{\boldsymbol{\beta}}$ and $\mathbb{P}_{\boldsymbol{\alpha\beta}}$. This immediately implies that

$$\chi'_{11} \;=\; \chi'_{12} \;=\; \chi'_{13} \;=\; \chi'_2 \;=\; \chi'_3 \;=\; \varepsilon''_W \;=\; \varepsilon'''_W \;=\; 0.$$

Hence, it suffices to show that $H'_0,\ H'_1,\ H'_{2i},\ i = 1,2,3,4$, are all of order $O(1)$, that (2.14) holds and that $\varepsilon'_W = O(\lambda^{-1/4})$.

For brevity, we write $\varsigma_{\boldsymbol{\alpha}} = \mathbf{1}_{[\Xi(A_{\boldsymbol{\alpha}})=0]}$ and $\theta_{\boldsymbol{\alpha}} = \mathbb{E}(\varsigma_{\boldsymbol{\alpha}})$. Clearly, $\mathbb{E}_{\boldsymbol{\alpha}}(\varsigma_{\boldsymbol{\alpha}}) = \theta_{\boldsymbol{\alpha}}$ also, and

$$\mathbb{E}_{\boldsymbol{\alpha\beta}}(\varsigma_{\boldsymbol{\alpha}}\varsigma_{\boldsymbol{\beta}}) \;=\; \left\{ \begin{array}{ll} \mathbb{E}(\varsigma_{\boldsymbol{\alpha}}\varsigma_{\boldsymbol{\beta}}) & \text{for } \boldsymbol{\beta} \in N^U_{\boldsymbol{\alpha}} \cup N^L_{\boldsymbol{\alpha}} \\ 0 & \text{for } \boldsymbol{\beta} \in N^D_{\boldsymbol{\alpha}} \cup A_{\boldsymbol{\alpha}} \end{array} \right\} \;\le\; \mathbb{E}(\varsigma_{\boldsymbol{\alpha}}\varsigma_{\boldsymbol{\beta}}),$$

where $N^U_{\boldsymbol{\alpha}}$, $N^L_{\boldsymbol{\alpha}}$ and $N^D_{\boldsymbol{\alpha}}$ are defined in Figure 1. Noting that $|X^{(\boldsymbol{\alpha})}| = \varsigma_{\boldsymbol{\alpha}}\mathbf{1}_{[\boldsymbol{\alpha}\in E']}$, we have

$$\mathbb{E}\int_{\boldsymbol{\alpha}\in E'} \varsigma_{\boldsymbol{\alpha}}\Xi(d\boldsymbol{\alpha}) \;=\; \int_{\boldsymbol{\alpha}\in E'} \mathbb{E}_{\boldsymbol{\alpha}}(\varsigma_{\boldsymbol{\alpha}})\nu(d\boldsymbol{\alpha}) \;=\; \int_{\boldsymbol{\alpha}\in E'} \theta_{\boldsymbol{\alpha}}\nu(d\boldsymbol{\alpha}) \;=\; |\mu|_1 \;=\; O(\lambda^{1/2}). \tag{4.14}$$

It thus follows that

$$H'_0 \;=\; d^{-1/2}m^{-1}\int_{\boldsymbol{\alpha}\in E'} \mathbb{E}_{\boldsymbol{\alpha}}(\varsigma_{\boldsymbol{\alpha}})\nu(d\boldsymbol{\alpha}) \;=\; d^{-1/2}m^{-1}\int_{\boldsymbol{\alpha}\in E'} \theta_{\boldsymbol{\alpha}}\nu(d\boldsymbol{\alpha}) \;=\; O(1).$$

For H'_1 and H'_{2i}, we repeatedly need to apply the estimate

$$\nu(D_{\boldsymbol{\alpha}} \cap E') \;=\; \lambda O(\lambda^{-1}) \;=\; O(1), \tag{4.15}$$

which follows because the area of $D_{\boldsymbol{\alpha}} \cap E'$ is of order $O(\lambda^{-1})$ and the intensity of Ξ is of order $O(\lambda)$. Note that the bound (4.15) and all the upper bounds below are uniform in $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$.

First, with (4.15) in mind, we obtain

$$\begin{aligned} \int_{\boldsymbol{\beta}\in D_{\boldsymbol{\alpha}}\cap E'} \mathbb{E}\{\varsigma_{\boldsymbol{\beta}} \,|\, \varsigma_{\boldsymbol{\alpha}} = 1\}\nu(d\boldsymbol{\beta}) \;&\le\; \nu(D_{\boldsymbol{\alpha}} \cap E') \;=\; O(1), \\ \int_{\boldsymbol{\beta}\in D_{\boldsymbol{\alpha}}\cap E'} \theta_{\boldsymbol{\beta}}\nu(d\boldsymbol{\beta}) \;&\le\; \nu(D_{\boldsymbol{\alpha}} \cap E') \;=\; O(1), \end{aligned}$$

which, together with (4.14), imply that

$$\begin{aligned} &\int_{\boldsymbol{\alpha}\in E'}\int_{\boldsymbol{\beta}\in D_{\boldsymbol{\alpha}}\cap E'} \{\mathbb{E}(\varsigma_{\boldsymbol{\alpha}}\varsigma_{\boldsymbol{\beta}}) + \theta_{\boldsymbol{\alpha}}\theta_{\boldsymbol{\beta}}\}\nu(d\boldsymbol{\beta})\nu(d\boldsymbol{\alpha}) \\ &= \int_{\boldsymbol{\alpha}\in E'}\int_{\boldsymbol{\beta}\in D_{\boldsymbol{\alpha}}\cap E'} \{\mathbb{E}\left[\varsigma_{\boldsymbol{\beta}} \,|\, \varsigma_{\boldsymbol{\alpha}} = 1\right] + \theta_{\boldsymbol{\beta}}\}\,\theta_{\boldsymbol{\alpha}}\nu(d\boldsymbol{\beta})\nu(d\boldsymbol{\alpha}) \\ &= O(1)\int_{\boldsymbol{\alpha}\in E'} \theta_{\boldsymbol{\alpha}}\nu(\boldsymbol{\alpha}) \;=\; O(\lambda^{1/2}). \end{aligned} \tag{4.16}$$

It therefore follows from (4.16) that

$$\begin{aligned} H'_1 \;&\le\; d^{-1}m^{-1}\int_{\boldsymbol{\alpha}\in E'}\int_{\boldsymbol{\beta}\in D_{\boldsymbol{\alpha}}\cap E'} \{\mathbb{E}_{\boldsymbol{\alpha\beta}}(\varsigma_{\boldsymbol{\alpha}}\varsigma_{\boldsymbol{\beta}}) + \theta_{\boldsymbol{\alpha}}\mathbb{E}_{\boldsymbol{\beta}}(\varsigma_{\boldsymbol{\beta}})\}\nu(d\boldsymbol{\beta})\nu(d\boldsymbol{\alpha}) \\ &\le\; d^{-1}m^{-1}\int_{\boldsymbol{\alpha}\in E'}\int_{\boldsymbol{\beta}\in D_{\boldsymbol{\alpha}}\cap E'} \{\mathbb{E}(\varsigma_{\boldsymbol{\alpha}}\varsigma_{\boldsymbol{\beta}}) + \theta_{\boldsymbol{\alpha}}\theta_{\boldsymbol{\beta}}\}\,\nu(d\boldsymbol{\beta})\nu(d\boldsymbol{\alpha}) \;=\; O(1). \end{aligned}$$

To show that $H'_{21} = O(1)$, we proceed as follows. With $\psi_{\boldsymbol{\alpha}} := \Xi(D_{\boldsymbol{\alpha}} \cap E')$, we obtain from (4.15) that

$$\mathbb{E}\left\{|Z^{(\boldsymbol{\alpha})}|^2\right\} \;\le\; \mathbb{E}\{\psi^2_{\boldsymbol{\alpha}}\} \;=\; \nu(D_{\boldsymbol{\alpha}} \cap E') + \nu(D_{\boldsymbol{\alpha}} \cap E')^2 \;=\; O(1). \tag{4.17}$$

Since $\varsigma_{\boldsymbol\alpha}$ is independent of $\psi'_{\boldsymbol\alpha} := \Xi(D_{\boldsymbol\alpha} \cap A^c_{\boldsymbol\alpha} \cap E')$, it follows from (4.15) and (4.17) that

$$\mathbb{E}_{\boldsymbol\alpha}\{|Z^{(\boldsymbol\alpha)}|^2 \,|\, \varsigma_{\boldsymbol\alpha} = 1\} \;\le\; \mathbb{E}_{\boldsymbol\alpha}\left\{ (\psi'_{\boldsymbol\alpha})^2 \middle| \varsigma_{\boldsymbol\alpha} = 1 \right\} \;=\; \mathbb{E}\left\{ (1+\psi'_{\boldsymbol\alpha})^2 \right\} \;\le\; \mathbb{E}\{(1+\psi_{\boldsymbol\alpha})^2\} \;=\; O(1). \tag{4.18}$$

Combining (4.17) and (4.18) with (4.14) then ensures that

$$\begin{aligned} H'_{21} &\le d^{-3/2}m^{-1}\int_{\boldsymbol\alpha\in E'} \left\{ \mathbb{E}_{\boldsymbol\alpha}(\varsigma_{\boldsymbol\alpha}|Z^{(\boldsymbol\alpha)}|^2) + \theta_{\boldsymbol\alpha}\mathbb{E}(|Z^{(\boldsymbol\alpha)}|^2) \right\} \nu(d\boldsymbol\alpha) \\ &= d^{-3/2}m^{-1}\int_{\boldsymbol\alpha\in E'} \left\{ \mathbb{E}_{\boldsymbol\alpha}[|Z^{(\boldsymbol\alpha)}|^2 \,|\, \varsigma_{\boldsymbol\alpha} = 1] + \mathbb{E}[|Z^{(\boldsymbol\alpha)}|^2] \right\} \theta_{\boldsymbol\alpha}\nu(d\boldsymbol\alpha) \\ &= O(\lambda^{-1/2})\int_{\boldsymbol\alpha\in E'} \theta_{\boldsymbol\alpha}\nu(d\boldsymbol\alpha) \;=\; O(1). \end{aligned}$$

In order to bound H'_{22}, H'_{23} and H'_{24}, we apply (4.15) again to get the estimates

$$\begin{aligned} \mathbb{E}_{\boldsymbol\beta}\{|Z^{(\boldsymbol\alpha,\boldsymbol\beta)}| \,|\, \varsigma_{\boldsymbol\beta} = 1\} &\le \mathbb{E}_{\boldsymbol\beta}\Xi(D^c_{\boldsymbol\alpha} \cap D_{\boldsymbol\beta} \cap A^c_{\boldsymbol\beta} \cap E') \;\le\; 1 + \nu(D_{\boldsymbol\beta} \cap E') \;=\; O(1), \\ \mathbb{E}_{\boldsymbol\alpha\boldsymbol\beta}\{|Z^{(\boldsymbol\alpha,\boldsymbol\beta)}| \,|\, \varsigma_{\boldsymbol\alpha} = \varsigma_{\boldsymbol\beta} = 1\} &\le \mathbb{E}_{\boldsymbol\alpha\boldsymbol\beta}\Xi(D^c_{\boldsymbol\alpha} \cap D_{\boldsymbol\beta} \cap A^c_{\boldsymbol\alpha} \cap A^c_{\boldsymbol\beta} \cap E') \\ &\le 2 + \nu(D_{\boldsymbol\beta} \cap E') \;=\; O(1), \\ \mathbb{E}|Z^{(\boldsymbol\alpha,\boldsymbol\beta)}| &\le \mathbb{E}\Xi(D^c_{\boldsymbol\alpha} \cap D_{\boldsymbol\beta} \cap E') \;\le\; \nu(D_{\boldsymbol\beta} \cap E') \;=\; O(1), \\ \mathbb{E}|Z^{(\boldsymbol\alpha)}| &\le \mathbb{E}\Xi(D_{\boldsymbol\alpha} \cap E') \;=\; \nu(D_{\boldsymbol\alpha} \cap E') \;=\; O(1). \end{aligned}$$

These in turn show that

$$\begin{aligned} H'_{22} &\le d^{-3/2}m^{-1}\int_{\boldsymbol\alpha\in E'}\int_{\boldsymbol\beta\in D_{\boldsymbol\alpha}\cap E'} \left\{ \mathbb{E}_{\boldsymbol\alpha\boldsymbol\beta}(\varsigma_{\boldsymbol\alpha}\varsigma_{\boldsymbol\beta}|Z^{(\boldsymbol\alpha,\boldsymbol\beta)}|) + \theta_{\boldsymbol\alpha}\mathbb{E}_{\boldsymbol\beta}(\varsigma_{\boldsymbol\beta}|Z^{(\boldsymbol\alpha,\boldsymbol\beta)}|) \right\} \nu(d\boldsymbol\beta)\nu(d\boldsymbol\alpha) \\ &= d^{-3/2}m^{-1}\int_{\boldsymbol\alpha\in E'}\int_{\boldsymbol\beta\in D_{\boldsymbol\alpha}\cap E'} \mathbb{E}_{\boldsymbol\alpha\boldsymbol\beta}(|Z^{(\boldsymbol\alpha,\boldsymbol\beta)}| \,|\, \varsigma_{\boldsymbol\alpha} = \varsigma_{\boldsymbol\beta} = 1) \\ &\qquad\qquad\qquad\qquad \times\mathbb{P}_{\boldsymbol\alpha\boldsymbol\beta}(\varsigma_{\boldsymbol\alpha} = \varsigma_{\boldsymbol\beta} = 1)\nu(d\boldsymbol\beta)\nu(d\boldsymbol\alpha) \\ &\quad + d^{-3/2}m^{-1}\int_{\boldsymbol\alpha\in E'}\int_{\boldsymbol\beta\in D_{\boldsymbol\alpha}\cap E'} \mathbb{E}_{\boldsymbol\beta}(|Z^{(\boldsymbol\alpha,\boldsymbol\beta)}| \,|\, \varsigma_{\boldsymbol\beta} = 1)\theta_{\boldsymbol\alpha}\theta_{\boldsymbol\beta}\nu(d\boldsymbol\beta)\nu(d\boldsymbol\alpha) \\ &= O(\lambda^{-1/2})\int_{\boldsymbol\alpha\in E'}\int_{\boldsymbol\beta\in D_{\boldsymbol\alpha}\cap E'} \{\mathbb{E}(\varsigma_{\boldsymbol\alpha}\varsigma_{\boldsymbol\beta}) + \theta_{\boldsymbol\alpha}\theta_{\boldsymbol\beta}\} \;=\; O(1), \qquad (4.19) \\ H'_{23} &= d^{-3/2}m^{-1}\int_{\boldsymbol\alpha\in E'}\int_{\boldsymbol\beta\in D_{\boldsymbol\alpha}\cap E'} \{\mathbb{E}_{\boldsymbol\alpha\boldsymbol\beta}(\varsigma_{\boldsymbol\alpha}\varsigma_{\boldsymbol\beta}) + \theta_{\boldsymbol\alpha}\mathbb{E}_{\boldsymbol\beta}(\varsigma_{\boldsymbol\beta})\}\, \mathbb{E}|Z^{(\boldsymbol\alpha,\boldsymbol\beta)}|\nu(d\boldsymbol\beta)\nu(d\boldsymbol\alpha) \\ &= O(\lambda^{-1/2})\int_{\boldsymbol\alpha\in E'}\int_{\boldsymbol\beta\in D_{\boldsymbol\alpha}\cap E'} \{\mathbb{E}(\varsigma_{\boldsymbol\alpha}\varsigma_{\boldsymbol\beta}) + \theta_{\boldsymbol\alpha}\theta_{\boldsymbol\beta}\}\nu(d\boldsymbol\beta)\nu(d\boldsymbol\alpha) = O(1), \qquad (4.20) \end{aligned}$$

and

$$\begin{aligned} H'_{24} &= d^{-3/2}m^{-1}\int_{\boldsymbol\alpha\in E'}\int_{\boldsymbol\beta\in D_{\boldsymbol\alpha}\cap E'} \{\mathbb{E}_{\boldsymbol\alpha\boldsymbol\beta}(\varsigma_{\boldsymbol\alpha}\varsigma_{\boldsymbol\beta}) + \theta_{\boldsymbol\alpha}\mathbb{E}_{\boldsymbol\beta}(\varsigma_{\boldsymbol\beta})\}\mathbb{E}|Z^{(\boldsymbol\alpha)}|\nu(d\boldsymbol\beta)\nu(d\boldsymbol\alpha) \\ &= O(\lambda^{-1/2})\int_{\boldsymbol\alpha\in E'}\int_{\boldsymbol\beta\in D_{\boldsymbol\alpha}\cap E'} \{\mathbb{E}(\varsigma_{\boldsymbol\alpha}\varsigma_{\boldsymbol\beta}) + \theta_{\boldsymbol\alpha}\theta_{\boldsymbol\beta}\}\nu(d\boldsymbol\beta)\nu(d\boldsymbol\alpha) \;=\; O(1), \qquad (4.21) \end{aligned}$$

where the last equalities in (4.19)–(4.21) are from (4.16).

Next, we turn to (2.14). In view of (4.15) and (4.17), we have the bounds

$$\begin{aligned} \mathbb{E}_{\boldsymbol\alpha}(|Z^{(\boldsymbol\alpha)}|^2) &\le \mathbb{E}\{(1+\psi_{\boldsymbol\alpha})^2\} \;=\; O(1), \quad & \mathbb{E}_{\boldsymbol\beta}(|Z^{(\boldsymbol\alpha)}|^2) \;\le\; \mathbb{E}\{(1+\psi_{\boldsymbol\alpha})^2\} \;=\; O(1), \\ \mathbb{E}(|Z^{(\boldsymbol\alpha,\boldsymbol\beta)}|^2) &\le \mathbb{E}(\psi^2_{\boldsymbol\beta}) \;=\; O(1), & \mathbb{E}_{\boldsymbol\beta}(|Z^{(\boldsymbol\alpha,\boldsymbol\beta)}|^2) \;\le\; \mathbb{E}\{(1+\psi_{\boldsymbol\beta})^2\} \;=\; O(1), \\ \mathbb{E}_{\boldsymbol\alpha\boldsymbol\beta}(|Z^{(\boldsymbol\alpha)}|^2) &\le \mathbb{E}\{(2+\psi_{\boldsymbol\alpha})^2\} \;=\; O(1), & \mathbb{E}_{\boldsymbol\alpha\boldsymbol\beta}(|Z^{(\boldsymbol\alpha,\boldsymbol\beta)}|^2) \;\le\; \mathbb{E}\{(1+\psi_{\boldsymbol\beta})^2\} \;=\; O(1). \end{aligned}$$

To show that both $\mathbb{E}_{\boldsymbol{\alpha}}(|W-\mu|^2)$ and $\mathbb{E}_{\boldsymbol{\alpha\beta}}(|W-\mu|^2)$ are bounded by $Cdm = O(\lambda^{1/2})$, for a suitably chosen C, we use the following crude estimates, which are adequate under local dependence conditions:

$$\begin{aligned}\mathbb{E}_{\boldsymbol{\alpha}}(|W-\mu|^2) &\le 2\mathbb{E}_{\boldsymbol{\alpha}}(|W^{(\boldsymbol{\alpha})}-\mu|^2) + 2\mathbb{E}_{\boldsymbol{\alpha}}(|Z^{(\boldsymbol{\alpha})}|^2) \;=\; 2\mathbb{E}(|W^{(\boldsymbol{\alpha})}-\mu|^2) + 2\mathbb{E}_{\boldsymbol{\alpha}}(|Z^{(\boldsymbol{\alpha})}|^2)\\ &\le 4\mathbb{E}(|W-\mu|^2) + 4\mathbb{E}(|Z^{(\boldsymbol{\alpha})}|^2) + 2\mathbb{E}_{\boldsymbol{\alpha}}(|Z^{(\boldsymbol{\alpha})}|^2) \;=\; O(\lambda^{1/2}),\end{aligned}$$

and

$$\begin{aligned}\mathbb{E}_{\boldsymbol{\alpha\beta}}(|W-\mu|^2) &\le 2\mathbb{E}_{\boldsymbol{\alpha\beta}}(|W^{(\boldsymbol{\alpha},\boldsymbol{\beta})}-\mu|^2) + 2\mathbb{E}_{\boldsymbol{\alpha\beta}}(|Z^{(\boldsymbol{\alpha})}+Z^{(\boldsymbol{\alpha},\boldsymbol{\beta})}|^2)\\ &= 2\mathbb{E}(|W^{(\boldsymbol{\alpha},\boldsymbol{\beta})}-\mu|^2) + 2\mathbb{E}_{\boldsymbol{\alpha\beta}}(|Z^{(\boldsymbol{\alpha})}+Z^{(\boldsymbol{\alpha},\boldsymbol{\beta})}|^2)\\ &\le 4\mathbb{E}(|W-\mu|^2) + 4\mathbb{E}(|Z^{(\boldsymbol{\alpha})}+Z^{(\boldsymbol{\alpha},\boldsymbol{\beta})}|^2) + 2\mathbb{E}_{\boldsymbol{\alpha\beta}}(|Z^{(\boldsymbol{\alpha})}+Z^{(\boldsymbol{\alpha},\boldsymbol{\beta})}|^2)\\ &= O(\lambda^{1/2});\end{aligned}$$

hence (2.14) holds.

Finally, we show that $\varepsilon'_W = O(\lambda^{-1/4})$. Referring to Figure 4, we fix $\theta \ge d_2$ as a constant, and set $\kappa := \lfloor(\sqrt{\lambda}/\theta)-1\rfloor$. We then define

$$B_l \;:=\; A_{(1-(l+1)\theta\lambda^{-1/2},\, l\theta\lambda^{-1/2})}; \quad \boldsymbol{\eta}_l \;:=\; \int_{\boldsymbol{\alpha}\in B_l\cap E'} X^{(\boldsymbol{\alpha})}\Xi(d\boldsymbol{\alpha}), \quad l=0,1,\ldots,\kappa.$$

Then the $\boldsymbol{\eta}_l$'s are independent and identically distributed random vectors. For any $\boldsymbol{\alpha}\in E'$ and $\boldsymbol{\beta}\in D_{\boldsymbol{\alpha}}\cap E'$, there are at most three of B_l's such that $B_l\cap(D_{\boldsymbol{\alpha}}\cup D_{\boldsymbol{\beta}})\neq\emptyset$, so we eliminate such $\boldsymbol{\eta}_l$'s and define $W'_{\boldsymbol{\alpha},\boldsymbol{\beta}} := \sum_{l:B_l\cap(D_{\boldsymbol{\alpha}}\cup D_{\boldsymbol{\beta}})=\emptyset}\boldsymbol{\eta}_l$. We use $W'_{\boldsymbol{\alpha},\boldsymbol{\beta}}$ to estimate ε'_W. To this end, let $\mathcal{F}_{\boldsymbol{\alpha},\boldsymbol{\beta}}$ be the σ-algebra generated by the configurations of points of Ξ in $\Gamma\setminus\left(\cup_{l:\,B_l\cap(D_{\boldsymbol{\alpha}}\cup D_{\boldsymbol{\beta}})=\emptyset}B_l\right)$, and let $d_{TV}\left(W^{(\boldsymbol{\alpha})}, W^{(\boldsymbol{\alpha})}+e^{(i)}\middle|\,\mathcal{F}_{\boldsymbol{\alpha},\boldsymbol{\beta}}\right)$ denote the total variation distance between $W^{(\boldsymbol{\alpha})}$ and $W^{(\boldsymbol{\alpha})}+e^{(i)}$ given configurations in $\mathcal{F}_{\boldsymbol{\alpha},\boldsymbol{\beta}}$ under $\mathbb{P}$. Then it follows that

$$\begin{aligned}&d_{\mathrm{TV}}\big(\mathcal{L}(W^{(\boldsymbol{\alpha})}+e^{(i)}\,|\,X^{(\boldsymbol{\alpha})},Z^{(\boldsymbol{\alpha})}),\mathcal{L}(W^{(\boldsymbol{\alpha})}\,|\,X^{(\boldsymbol{\alpha})},Z^{(\boldsymbol{\alpha})})\big)\\ &\quad\le \operatorname{ess\,sup}\left\{d_{TV}\left(W^{(\boldsymbol{\alpha})},W^{(\boldsymbol{\alpha})}+e^{(i)}\middle|\,\mathcal{F}_{\boldsymbol{\alpha},\boldsymbol{\beta}}\right)\right\} \;=\; d_{TV}(W'_{\boldsymbol{\alpha},\boldsymbol{\beta}}, W'_{\boldsymbol{\alpha},\boldsymbol{\beta}}+e^{(i)}),\end{aligned}\tag{4.22}$$

where $\operatorname{ess\,sup}$ stands for the essential supremum. Likewise,

$$\begin{aligned}&d_{\mathrm{TV}}\big(\mathcal{L}_\alpha(W^{(\boldsymbol{\alpha})}+e^{(i)}\,|\,X^{(\boldsymbol{\alpha})},Z^{(\boldsymbol{\alpha})}),\mathcal{L}_\alpha(W^{(\boldsymbol{\alpha})}\,|\,X^{(\boldsymbol{\alpha})},Z^{(\boldsymbol{\alpha})})\big)\\ &\quad\le \operatorname{ess\,sup}\left\{d_{TV}\left(W^{(\boldsymbol{\alpha})},W^{(\boldsymbol{\alpha})}+e^{(i)}\middle|\,\mathcal{F}_{\boldsymbol{\alpha},\boldsymbol{\beta}}\right)\right\} \;=\; d_{TV}(W'_{\boldsymbol{\alpha},\boldsymbol{\beta}}, W'_{\boldsymbol{\alpha},\boldsymbol{\beta}}+e^{(i)}),\\ &d_{\mathrm{TV}}\big(\mathcal{L}(W^{(\boldsymbol{\alpha},\boldsymbol{\beta})}+e^{(i)}\,|\,X^{(\boldsymbol{\alpha})},\widetilde{X}^{(\boldsymbol{\alpha},\boldsymbol{\beta})},Z^{(\boldsymbol{\alpha},\boldsymbol{\beta})}),\mathcal{L}(W^{(\boldsymbol{\alpha},\boldsymbol{\beta})}\,|\,X^{(\boldsymbol{\alpha})},\widetilde{X}^{(\boldsymbol{\alpha},\boldsymbol{\beta})},Z^{(\boldsymbol{\alpha},\boldsymbol{\beta})})\big)\\ &\quad\le \operatorname{ess\,sup}\left\{d_{TV}\left(W^{(\boldsymbol{\alpha},\boldsymbol{\beta})},W^{(\boldsymbol{\alpha},\boldsymbol{\beta})}+e^{(i)}\middle|\,\mathcal{F}_{\boldsymbol{\alpha},\boldsymbol{\beta}}\right)\right\} \;=\; d_{TV}(W'_{\boldsymbol{\alpha},\boldsymbol{\beta}}, W'_{\boldsymbol{\alpha},\boldsymbol{\beta}}+e^{(i)}),\\ &d_{\mathrm{TV}}\big(\mathcal{L}_\beta(W^{(\boldsymbol{\alpha},\boldsymbol{\beta})}+e^{(i)}\,|\,X^{(\boldsymbol{\alpha})},\widetilde{X}^{(\boldsymbol{\alpha},\boldsymbol{\beta})},Z^{(\boldsymbol{\alpha},\boldsymbol{\beta})}),\mathcal{L}_\beta(W^{(\boldsymbol{\alpha},\boldsymbol{\beta})}\,|\,X^{(\boldsymbol{\alpha})},\widetilde{X}^{(\boldsymbol{\alpha},\boldsymbol{\beta})},Z^{(\boldsymbol{\alpha},\boldsymbol{\beta})})\big)\\ &\quad\le \operatorname{ess\,sup}\left\{d_{TV}\left(W^{(\boldsymbol{\alpha},\boldsymbol{\beta})},W^{(\boldsymbol{\alpha},\boldsymbol{\beta})}+e^{(i)}\middle|\,\mathcal{F}_{\boldsymbol{\alpha},\boldsymbol{\beta}}\right)\right\} \;=\; d_{TV}(W'_{\boldsymbol{\alpha},\boldsymbol{\beta}}, W'_{\boldsymbol{\alpha},\boldsymbol{\beta}}+e^{(i)}),\\ &d_{\mathrm{TV}}\big(\mathcal{L}_{\alpha\beta}(W^{(\boldsymbol{\alpha},\boldsymbol{\beta})}+e^{(i)}\,|\,X^{(\boldsymbol{\alpha})},\widetilde{X}^{(\boldsymbol{\alpha},\boldsymbol{\beta})},Z^{(\boldsymbol{\alpha},\boldsymbol{\beta})}),\mathcal{L}_{\alpha\beta}(W^{(\boldsymbol{\alpha},\boldsymbol{\beta})}\,|\,X^{(\boldsymbol{\alpha})},\widetilde{X}^{(\boldsymbol{\alpha},\boldsymbol{\beta})},Z^{(\boldsymbol{\alpha},\boldsymbol{\beta})})\big)\\ &\quad\le \operatorname{ess\,sup}\left\{d_{TV}\left(W^{(\boldsymbol{\alpha},\boldsymbol{\beta})},W^{(\boldsymbol{\alpha},\boldsymbol{\beta})}+e^{(i)}\middle|\,\mathcal{F}_{\boldsymbol{\alpha},\boldsymbol{\beta}}\right)\right\} \;=\; d_{TV}(W'_{\boldsymbol{\alpha},\boldsymbol{\beta}}, W'_{\boldsymbol{\alpha},\boldsymbol{\beta}}+e^{(i)}).\end{aligned}\tag{4.23}$$

On the other hand,

$$d_{TV}(\boldsymbol{\eta}_1,\boldsymbol{\eta}_1+e^{(i)}) \;\le\; 1-\mathbb{P}(\boldsymbol{\eta}_1=0)\wedge\mathbb{P}(\boldsymbol{\eta}_1=e^{(i)}).$$

Noting that B_1 and $B_1\cap E_i$ satisfy

$$\nu(B_1) \;=\; \frac{\theta^2}{2}, \quad\text{and}\quad \nu(B_1\cap E_i) \;=\; \frac{1}{2}(d_i-b_i)(2\theta-(d_i+b_i)),$$

we obtain

$$\begin{aligned}
\mathbb{P}(\boldsymbol{\eta}_1 = 0) &\geq \mathbb{P}(\Upsilon(B_1) = 0) \;=\; e^{-\frac{1}{2}\theta^2}, \\
\mathbb{P}(\boldsymbol{\eta}_1 = e^{(i)}) &\geq \mathbb{P}(\Upsilon(B_1 \cap E_i) = 1, \Upsilon(B_1 \setminus E_i) = 0) \\
&= \frac{1}{2}(d_i - b_i)(2\theta - (d_i + b_i))e^{-\frac{1}{2}\theta^2},
\end{aligned}$$

which together imply that

$$d_{TV}(\boldsymbol{\eta}_1, \boldsymbol{\eta}_1 + e^{(i)}) \;\leq\; 1 - \left\{1 \wedge \left(\frac{1}{2}(d_i - b_i)(2\theta - (d_i + b_i))\right)\right\} e^{-\frac{1}{2}\theta^2}.$$

Hence it follows from Lemma 4.1 of [Barbour, Luczak & Xia (2018b)] that

$$d_{TV}(W'_{\boldsymbol{\alpha},\boldsymbol{\beta}}, W'_{\boldsymbol{\alpha},\boldsymbol{\beta}} + e^{(i)}) \;=\; O\left((\kappa - 3)^{-1/2}\right) \;=\; O(\lambda^{-1/4}),$$

since $\kappa \geq (\sqrt{\lambda}/\theta) - 2$. This, together with (4.22) and (4.23), ensures that $\varepsilon'_W = O(\lambda^{-1/4})$, and completes the proof of the theorem. □

5 The proofs of Theorems 2.1 and 2.2

Before proving our main theorems, we establish an auxiliary lemma. It is useful in what follows to be able to extend the definition of a function h from a ball $B_{m\delta}(x) \cap \mathbb{Z}^d$ to the whole of $\mathbb{Z}^d$ in such a way that, in the notation of (1.2), $\|\Delta h\|_\infty := \sup_{z \in \mathbb{Z}^d} |\Delta h(z)|$ can be bounded in terms of $\|\Delta h\|_{3m\delta/2,\infty}$. That this can be done, if $m\delta \geq 2\sqrt{d}$, is proved using the following lemma.

Lemma 5.1. *Let $h\colon \mathbb{Z}^d \to \mathbb{R}$ be given. Then, for any $x \in \mathbb{R}^d$ and $r > 0$, it is possible to modify h outside the set $\mathbb{Z}^d \cap B_r(x)$ in such a way that the resulting function $\tilde{h}$ satisfies $\|\Delta\tilde{h}\|_\infty \leq \sqrt{d}\|\Delta h\|_{r+\sqrt{d},\infty}$.*

Proof. First, for all $y = (y_1, \ldots, y_d) \in B_r(x)$, we have

$$Z(y) \;:=\; \lfloor y \rfloor + \{0,1\}^d \;\subset\; B_{r+\sqrt{d}}(x),$$

where $\lfloor y \rfloor := (\lfloor y_1 \rfloor, \ldots, \lfloor y_d \rfloor)$, because, for each $z \in Z(y)$, $|z - y| \leq \sqrt{d}$. Extend the definition of h to all $y \in B_r(x)$ by averaging over the values at the points $Z(y)$:

$$h(y) \;:=\; \sum_{q \in \{0,1\}^d} \left\{\prod_{i=1}^{d}(1 - \{y_i\} + q_i(2\{y_i\} - 1))\right\} h(\lfloor y \rfloor + q),$$

where $\{y_i\} := y_i - \lfloor y_i \rfloor$. It is immediate that h is continuous in $B_r(x)$, and that, for y in the interior of any unit cube,

$$\begin{aligned}
|D_j h(y)| &= \Bigl| \sum_{q \in \{0,1\}^d} (2q_j - 1)\Bigl\{\prod_{i \neq j}(1 - \{y_i\} + q_i(2\{y_i\} - 1))\Bigr\} h(\lfloor y \rfloor + q)\Bigr| \\
&\leq \sum_{q' \in \{0,1\}^{j-1} \times \{0\} \times \{0,1\}^{d-j}} \Bigl\{\prod_{i \neq j}(1 - \{y_i\} + q'_i(2\{y_i\} - 1))\Bigr\} |\Delta_j h(\lfloor y \rfloor + q')| \\
&\leq \|\Delta h\|_{r+\sqrt{d},\infty}.
\end{aligned}$$

Hence it follows that $|h(y) - h(y')| \leq \sqrt{d}\|\Delta h\|_{r+\sqrt{d},\infty}|y - y'|$ for any $y, y' \in B_r(x)$.

Now define $\tilde{h}$ on $\mathbb{R}^d$ by setting $\tilde{h}(y) = h(y)$ on $B_r(x)$, and $\tilde{h}(y) = h(\pi_x y)$ for $y \notin B_r(x)$, where $\pi_x y := x + r(y - x)/|y - x|$ is the projection of y onto the surface of $B_r(x)$. Then, since

$$|a - b| \;\geq\; \left|\frac{a}{|a|} - \frac{b}{|b|}\right| \quad \text{if} \quad |a|, |b| \geq 1,$$

it follows that

$$|\tilde{h}(y)-\tilde{h}(y')| \;=\; |h(\pi_x y)-h(\pi_x y')| \;\le\; \sqrt{d}\|\Delta h\|_{r+\sqrt{d},\infty}|\pi_x y-\pi_x y'| \;\le\; \sqrt{d}\|\Delta h\|_{r+\sqrt{d},\infty}|y-y'|,$$

and so $\|\Delta\tilde{h}\|_\infty \le \sqrt{d}\|\Delta h\|_{r+\sqrt{d},\infty}$. □

We are now in a position to prove our main theorems.

Proofs of Theorems 2.1 and 2.2. We first prove Theorem 2.2. Condition (a) of Theorem 1.1 follows directly from (2.15), with ε'_W for ε_1. We thus turn to Condition (b), using the Stein operator $\widetilde{\mathcal{A}}_m$, as in (1.1), with m as defined in (2.9), and with

$$c \;:=\; m^{-1}\mu; \quad \Sigma \;:=\; m^{-1}V. \tag{5.1}$$

As a first step, choose some $\delta > 0$ such that $2\delta \le \delta_0$, where δ_0 is as in Theorem 1.1. Given any function h to be used in Theorem 1.1(b), use Lemma 5.1 to continue it outside $B_{3m\delta/2}(\mu)$ in such a way that

$$\|\Delta h\|_\infty \;\le\; \sqrt{d}\|\Delta h\|_{2m\delta,\infty} \;\le\; \sqrt{d}\|\Delta h\|_{m\delta_0,\infty}, \tag{5.2}$$

possible provided that $\sqrt{d} \le m\delta/2$; since the bound given in the theorem is trivial (taking $C_{2.1} \ge 1$ if necessary) if $m \le d^8$, it is enough for this to suppose that $\delta\sqrt{m} \ge 2$. We now observe by Cauchy–Schwarz and Chebyshev's inequality that

$$\begin{aligned}&|\mathbb{E}\{(W-\mu)^T\Delta h(W)I[|W-\mu| > m\delta]\}| \\ &\quad\le\; \|\Delta h\|_\infty\{\mathbb{E}|W-\mu|^2\mathbb{P}[|W-\mu| > m\delta]\}^{1/2} \;\le\; \|\Delta h\|_\infty \mathrm{Tr}(V)/(m\delta) \;\le\; d/\delta\|\Delta h\|_\infty.\end{aligned}$$

This allows the second part of $\mathbb{E}\{\widetilde{\mathcal{A}}_m h(W)I[|W-\mu| \le m\delta]\}$ to be computed without the indicator, at little cost:

$$\begin{aligned}&|\mathbb{E}\{(W-\mu)^T\Delta h(W)I[|W-\mu| \le m\delta]\} - \mathbb{E}\{(W-\mu)^T\Delta h(W)\}| \\ &\quad\le\; (d/\delta)\,\|\Delta h\|_\infty.\end{aligned} \tag{5.3}$$

Then, expanding W as an integral and using $\mathbb{E}_\alpha X^{(\alpha)} = \mu^{(\alpha)}$, we have

$$\mathbb{E}\{(W-\mu)^T\Delta h(W)\} \tag{5.4}$$

$$\begin{aligned}&= \int_\Gamma \{\mathbb{E}_\alpha\{(X^{(\alpha)})^T\Delta h(Z^{(\alpha)}+W^{(\alpha)})\} - \mathbb{E}\{(\mu^{(\alpha)})^T\Delta h(Z^{(\alpha)}+W^{(\alpha)})\}\}\nu(d\alpha) \\ &= \int_\Gamma \mathbb{E}_\alpha\{(X^{(\alpha)})^T\left(\Delta h(Z^{(\alpha)}+W^{(\alpha)}) - \Delta h(W^{(\alpha)})\right)\}\nu(d\alpha) \\ &\quad -\int_\Gamma \mathbb{E}\{(\mu^{(\alpha)})^T\left(\Delta h(Z^{(\alpha)}+W^{(\alpha)}) - \Delta h(W^{(\alpha)})\right)\}\nu(d\alpha) + \eta_1 + \eta_2,\end{aligned} \tag{5.5}$$

where

$$\begin{aligned}|\eta_1| \;&=\; \left|\int_\Gamma \mathbb{E}_\alpha\left(\left(\mathbb{E}_\alpha\left((X^{(\alpha)})^T\middle|W^{(\alpha)}\right) - (\mu^{(\alpha)})^T\right)\Delta h(W^{(\alpha)})\right)\nu(d\alpha)\right| \\ &\le\; (dm)^{1/2}\|\Delta h\|_\infty\chi'_{11}, \\ |\eta_2| \;&=\; \left|\int_\Gamma (\mu^{(\alpha)})^T\left(\mathbb{E}_\alpha\Delta h(W^{(\alpha)}) - \mathbb{E}\Delta h(W^{(\alpha)})\right)\nu(d\alpha)\right| \\ &\le\; 2d^{1/2}m\|\Delta h\|_\infty H'_0\varepsilon''_W.\end{aligned} \tag{5.6}$$

The next step is to approximate $\Delta h(Z^{(\alpha)} + W^{(\alpha)}) - \Delta h(W^{(\alpha)})$ by $\Delta^2 h(W^{(\alpha)}) Z^{(\alpha)}$ in (5.5), and to take care of the error. This is accomplished in a number of steps. First, in view of Condition (b) of Theorem 1.1, we need to express bounds on the second differences of h in terms of their supremum in some $m\eta$-ball around $\mu = mc$; we do not have an analogue of Lemma 5.1 for the second differences. Thus we re-introduce truncation, to ensure that both $W^{(\alpha)}$ and W are close enough to μ. From (2.10) and (2.14), and by Chebyshev's inequality, we have

$$\begin{aligned}
&\mathbb{E}_\alpha\bigl\{|(X^{(\alpha)})^T\{\Delta h(Z^{(\alpha)} + W^{(\alpha)}) - \Delta h(W^{(\alpha)})\}|(I[|W^{(\alpha)} - \mu| > m\delta] + I[|Z^{(\alpha)}| > \sqrt{m}])\bigr\} \\
&\quad\le\ 2\|\Delta h\|_\infty\bigl\{\mathbb{E}_\alpha\{|X^{(\alpha)}|(\mathbb{P}_\alpha[|W^{(\alpha)} - \mu| > m\delta] + m^{-1}|Z^{(\alpha)}|^2)\} + \chi'_{12\alpha}\bigr\} \\
&\quad\le\ 2\|\Delta h\|_\infty\bigl\{\mathbb{E}_\alpha\{|X^{(\alpha)}|(\mathbb{P}_\alpha[|W - \mu| > m\delta/2] \\
&\qquad\quad +\mathbb{P}_\alpha[|Z^{(\alpha)}| > m\delta/2] + m^{-1}|Z^{(\alpha)}|^2)\} + \chi'_{12\alpha}\bigr\} \\
&\quad\le\ 2\|\Delta h\|_\infty\bigl\{m^{-1}\mathbb{E}_\alpha\{|X^{(\alpha)}|\bigl(8C\delta^{-2}d + |Z^{(\alpha)}|^2\bigr)\} + \chi'_{12\alpha}\bigr\}
\end{aligned} \tag{5.7}$$

and

$$\begin{aligned}
&\mathbb{E}\bigl\{|(\mu^{(\alpha)})^T\{\Delta h(Z^{(\alpha)} + W^{(\alpha)}) - \Delta h(W^{(\alpha)})\}|(I[|W^{(\alpha)} - \mu| > m\delta] + I[|Z^{(\alpha)}| > \sqrt{m}])\bigr\} \\
&\quad\le\ 2\|\Delta h\|_\infty\bigl\{\{|\mu^{(\alpha)}|(\mathbb{P}[|W^{(\alpha)} - \mu| > m\delta] + m^{-1}\mathbb{E}(|Z^{(\alpha)}|^2))\}\bigr\} \\
&\quad\le\ 2\|\Delta h\|_\infty\bigl\{\{|\mu^{(\alpha)}|(\mathbb{P}[|W - \mu| > m\delta/2] + \mathbb{P}[|Z^{(\alpha)}| > m\delta/2] + m^{-1}\mathbb{E}(|Z^{(\alpha)}|^2))\}\bigr\} \\
&\quad\le\ 2\|\Delta h\|_\infty m^{-1}|\mu^{(\alpha)}|\bigl(8\delta^{-2}d + \mathbb{E}(|Z^{(\alpha)}|^2)\bigr).
\end{aligned} \tag{5.8}$$

Integrating over α with respect to ν, it thus follows from (2.13) that

$$\begin{aligned}
&\int_\Gamma \mathbb{E}_\alpha\Bigl|(X^{(\alpha)})^T\{\Delta h(Z^{(\alpha)} + W^{(\alpha)}) - \Delta h(W^{(\alpha)})\} \\
&\qquad - (X^{(\alpha)})^T\{\Delta h(Z^{(\alpha)} + W^{(\alpha)}) - \Delta h(W^{(\alpha)})\}I[|W^{(\alpha)} - \mu| \le m\delta]I[|Z^{(\alpha)}| \le \sqrt{m}]\Bigr|\nu(d\alpha) \\
&\le\ 2d^{3/2}\|\Delta h\|_\infty(8C\delta^{-2}H'_0 + H'_{21}) + 2\|\Delta h\|_\infty(dm)^{1/2}\chi'_{12},
\end{aligned} \tag{5.9}$$

and

$$\begin{aligned}
&\int_\Gamma \mathbb{E}\Bigl|(\mu^{(\alpha)})^T\{\Delta h(Z^{(\alpha)} + W^{(\alpha)}) - \Delta h(W^{(\alpha)})\} \\
&\qquad - (\mu^{(\alpha)})^T\{\Delta h(Z^{(\alpha)} + W^{(\alpha)}) - \Delta h(W^{(\alpha)})\}I[|W^{(\alpha)} - \mu| \le m\delta]I[|Z^{(\alpha)}| \le \sqrt{m}]\Bigr|\nu(d\alpha) \\
&\le\ 2d^{3/2}\|\Delta h\|_\infty(8\delta^{-2}H'_0 + H'_{21}).
\end{aligned} \tag{5.10}$$

The integrals on the right hand side of (5.5) can thus be replaced by

$$\begin{aligned}
&\int_\Gamma \mathbb{E}_\alpha\bigl\{(X^{(\alpha)})^T\,(\Delta h(Z^{(\alpha)} + W^{(\alpha)}) - \Delta h(W^{(\alpha)}))I[|W^{(\alpha)} - \mu| \le m\delta]I[|Z^{(\alpha)}| \le \sqrt{m}]\bigr\}\nu(d\alpha) \\
&-\int_\Gamma \mathbb{E}\bigl\{(\mu^{(\alpha)})^T\,(\Delta h(Z^{(\alpha)} + W^{(\alpha)}) - \Delta h(W^{(\alpha)}))I[|W^{(\alpha)} - \mu| \le m\delta]I[|Z^{(\alpha)}| \le \sqrt{m}]\bigr\}\nu(d\alpha) \\
&\quad=\ \int_\Gamma \mathbb{E}_\alpha\bigl\{(X^{(\alpha)})^T\,\Delta^2 h(W^{(\alpha)})Z^{(\alpha)}I[|W^{(\alpha)} - \mu| \le m\delta]I[|Z^{(\alpha)}| \le \sqrt{m}]\bigr\}\nu(d\alpha) \qquad (5.11)\\
&\qquad + \int_\Gamma \mathbb{E}_\alpha\bigl\{(X^{(\alpha)})^T\,(\Delta h(Z^{(\alpha)} + W^{(\alpha)}) - \Delta h(W^{(\alpha)}) - \Delta^2 h(W^{(\alpha)})Z^{(\alpha)}) \\
&\qquad\qquad I[|W^{(\alpha)} - \mu| \le m\delta]I[|Z^{(\alpha)}| \le \sqrt{m}]\bigr\}\nu(d\alpha) \qquad (5.12)\\
&\qquad -\int_\Gamma \mathbb{E}\bigl\{(\mu^{(\alpha)})^T\,\Delta^2 h(W^{(\alpha)})Z^{(\alpha)}I[|W^{(\alpha)} - \mu| \le m\delta]I[|Z^{(\alpha)}| \le \sqrt{m}]\bigr\}\nu(d\alpha) \qquad (5.13)\\
&\qquad - \int_\Gamma \mathbb{E}\bigl\{(\mu^{(\alpha)})^T\,(\Delta h(Z^{(\alpha)} + W^{(\alpha)}) - \Delta h(W^{(\alpha)}) - \Delta^2 h(W^{(\alpha)})Z^{(\alpha)}) \\
&\qquad\qquad I[|W^{(\alpha)} - \mu| \le m\delta]I[|Z^{(\alpha)}| \le \sqrt{m}]\bigr\}\nu(d\alpha), \qquad (5.14)
\end{aligned}$$

having truncation in both $W^{(\alpha)}$ and $Z^{(\alpha)}$, with errors bounded by (5.9) and (5.10).

Now (5.12) and (5.14) can be represented in terms of sums of second differences of h. Defining $\widehat{Z}^{[\alpha,l]} := \sum_{t=1}^{l} Z_t^{(\alpha)} e^{(t)}$, $1 \le l \le d$, we have

$$\begin{aligned}
&(e^{(i)})^T\{\Delta h(Z^{(\alpha)}+W^{(\alpha)}) - \Delta h(W^{(\alpha)}) - \Delta^2 h(W^{(\alpha)})Z^{(\alpha)}\} \\
&\quad = \sum_{l=1}^{d} \left\{ \sum_{s=0}^{Z_l^{(\alpha)}-1} \{\Delta_{il}^2 h(W^{(\alpha)} + \widehat{Z}^{[\alpha,l-1]} + se^{(l)}) - \Delta_{il}^2 h(W^{(\alpha)})\} I[Z_l^{(\alpha)} \ge 1] \right. \\
&\qquad\qquad \left. - \sum_{s=Z_l^{(\alpha)}}^{-1} \{\Delta_{il}^2 h(W^{(\alpha)} + \widehat{Z}^{[\alpha,l-1]} + se^{(l)}) - \Delta_{il}^2 h(W^{(\alpha)})\} I[Z_l^{(\alpha)} \le -1] \right\}.
\end{aligned}$$

Writing

$$h_{il}(w,\delta) \;:=\; \Delta_{il}^2 h(w) I[|w-\mu| \le m\delta],$$

it then follows that

$$\begin{aligned}
&\{\Delta_{il}^2 h(W^{(\alpha)} + \widehat{Z}^{[\alpha,l-1]} + se^{(l)}) - \Delta_{il}^2 h(W^{(\alpha)})\} I[|W^{(\alpha)} - \mu| \le m\delta] \\
&\quad = \; h_{il}(W^{(\alpha)} + \widehat{Z}^{[\alpha,l-1]} + se^{(l)}, \delta) - h_{il}(W^{(\alpha)}, \delta) \qquad (5.15)\\
&\qquad - \Delta_{il}^2 h(W^{(\alpha)} + \widehat{Z}^{[\alpha,l-1]} + se^{(l)}) \\
&\qquad\qquad \{I[|W^{(\alpha)} + \widehat{Z}^{[\alpha,l-1]} + se^{(l)} - \mu| \le m\delta] - I[|W^{(\alpha)} - \mu| \le m\delta]\}. \qquad (5.16)
\end{aligned}$$

The contribution from (5.15) to (5.12) and (5.14) can be respectively bounded by first taking the expectation conditional on $X^{(\alpha)}$ and $Z^{(\alpha)}$, and using (2.15); this gives

$$\begin{aligned}
&\big|\mathbb{E}_\alpha\big\{(X_i^{(\alpha)})\{h_{il}(W^{(\alpha)} + \widehat{Z}^{[\alpha,l-1]} + se^{(l)}, \delta) - h_{il}(W^{(\alpha)}, \delta)\} I[|Z^{(\alpha)}| \le \sqrt{m}] \,|\, X^{(\alpha)}, Z^{(\alpha)}\big\}\big| \\
&\quad \le \; |X_i^{(\alpha)}|\, 2\|\Delta^2 h\|_{m\delta,\infty}\, \varepsilon'_W\, (|s| + |\widehat{Z}^{[\alpha,l-1]}|_1), \qquad (5.17)
\end{aligned}$$

and

$$\begin{aligned}
&\big|\mathbb{E}\big\{(\ \mu_i^{(\alpha)})\{h_{il}(W^{(\alpha)} + \widehat{Z}^{[\alpha,l-1]} + se^{(l)}, \delta) - h_{il}(W^{(\alpha)}, \delta)\} I[|Z^{(\alpha)}| \le \sqrt{m}] \,|\, X^{(\alpha)}, Z^{(\alpha)}\big\}\big| \\
&\quad \le \; |\mu_i^{(\alpha)}|\, 2\|\Delta^2 h\|_{m\delta,\infty}\, \varepsilon'_W\, (|s| + |\widehat{Z}^{[\alpha,l-1]}|_1). \qquad (5.18)
\end{aligned}$$

Adding over s and over $1 \le i \le d$, integrating over $\alpha \in \Gamma$ with respect to ν and taking expectations with respect to $\mathbb{E}_\alpha$ and $\mathbb{E}$ respectively, we get error bounds of at most

$$\varepsilon'_W \int_\Gamma \mathbb{E}_\alpha\{|X^{(\alpha)}|_1\, |Z^{(\alpha)}|_1 (|Z^{(\alpha)}|_1 + 1)\} \|\Delta^2 h\|_{3m\delta/2,\infty}\, \nu(d\alpha) \;\le\; 2d^3 m \|\Delta^2 h\|_{3m\delta/2,\infty} H'_{21}\, \varepsilon'_W \qquad (5.19)$$

and

$$\varepsilon'_W \int_\Gamma \mathbb{E}\{|\mu^{(\alpha)}|_1\, |Z^{(\alpha)}|_1 (|Z^{(\alpha)}|_1 + 1)\} \|\Delta^2 h\|_{3m\delta/2,\infty}\, \nu(d\alpha) \;\le\; 2d^3 m \|\Delta^2 h\|_{3m\delta/2,\infty} H'_{21}\, \varepsilon'_W. \qquad (5.20)$$

For the contribution from (5.16) to (5.12) and (5.14), recalling that $\sqrt{m} \ge 2/\delta$, we have

$$\begin{aligned}
&\big|I[|W^{(\alpha)} + \widehat{Z}^{[\alpha,l-1]} + se^{(l)} - \mu| \le m\delta] - I[|W^{(\alpha)} - \mu| \le m\delta]\big|\, I[|Z^{(\alpha)}| \le \sqrt{m}] \\
&\quad \le \; I[|W^{(\alpha)} - \mu| > \tfrac{1}{2} m\delta]
\end{aligned}$$

for $0 \le s < Z_l^{(\alpha)}$ if $Z_l^{(\alpha)} \ge 1$, and for $Z_l^{(\alpha)} \le s < 0$ if $Z_l^{(\alpha)} < 0$. Arguing for $Z_l^{(\alpha)} \ge 1$, we thus have

$$\Big|\sum_{s=0}^{Z_l^{(\alpha)}-1} \Delta_{il}^2 h(W^{(\alpha)} + \widehat{Z}^{[\alpha,l-1]} + se^{(l)})$$
$$\big\{I[|W^{(\alpha)} + \widehat{Z}^{[\alpha,l-1]} + se^{(l)} - \mu| \le m\delta] - I[|W^{(\alpha)} - \mu| \le m\delta]\big\} I[|Z^{(\alpha)}| \le \sqrt{m}]\Big|$$
$$\le Z_l^{(\alpha)} I[|W^{(\alpha)} - \mu| > \tfrac{1}{2}m\delta] \|\Delta^2 h\|_{3m\delta/2,\infty},$$

from which it follows that

$$\Big|\sum_{i=1}^d \mathbb{E}_\alpha\Big\{X_i^{(\alpha)} \sum_{l=1}^d \sum_{s=0}^{Z_l^{(\alpha)}-1} \Delta_{il}^2 h(W^{(\alpha)} + \widehat{Z}^{[\alpha,l-1]} + se^{(l)})$$
$$\big\{I[|W^{(\alpha)} + \widehat{Z}^{[\alpha,l-1]} + se^{(l)} - \mu| \le m\delta] - I[|W^{(\alpha)} - \mu| \le m\delta]\big\} I[|Z^{(\alpha)}| \le \sqrt{m}]\Big\}\Big|$$
$$\le \|\Delta^2 h\|_{3m\delta/2,\infty} \big\{\mathbb{E}_\alpha |X^{(\alpha)}|_1\, \mathbb{P}_\alpha[|W^{(\alpha)} - \mu| > \tfrac{1}{2}m\delta] + \chi'_{13\alpha}\big\}\sqrt{dm} \tag{5.21}$$

and that

$$\Big|\sum_{i=1}^d \mathbb{E}\Big\{\mu_i^{(\alpha)} \sum_{l=1}^d \sum_{s=0}^{Z_l^{(\alpha)}-1} \Delta_{il}^2 h(W^{(\alpha)} + \widehat{Z}^{[\alpha,l-1]} + se^{(l)})$$
$$\{I[|W^{(\alpha)} + \widehat{Z}^{[\alpha,l-1]} + se^{(l)} - \mu| \le m\delta] - I[|W^{(\alpha)} - \mu| \le m\delta]\} I[|Z^{(\alpha)}| \le \sqrt{m}]\Big\}\Big|$$
$$\le \|\Delta^2 h\|_{3m\delta/2,\infty} |\mu^{(\alpha)}|_1\, \mathbb{P}[|W^{(\alpha)} - \mu| > \tfrac{1}{2}m\delta]\sqrt{dm}. \tag{5.22}$$

The argument for $Z_l^{(\alpha)} < 0$ is almost exactly the same.

The first part of (5.21) yields at most

$$\sqrt{dm}\|\Delta^2 h\|_{3m\delta/2,\infty} d^{1/2}\mathbb{E}_\alpha|X^{(\alpha)}|\,\mathbb{P}_\alpha[|W^{(\alpha)} - \mu| > \tfrac{1}{2}m\delta]$$
$$\le d\sqrt{m}\|\Delta^2 h\|_{3m\delta/2,\infty}\mathbb{E}_\alpha|X^{(\alpha)}|\,\{\mathbb{P}_\alpha[|W - \mu| > \tfrac{1}{4}m\delta] + \mathbb{P}_\alpha[|Z^{(\alpha)}| > \tfrac{1}{4}m\delta]\}$$
$$\le 32Cd^2\delta^{-2}m^{-1/2}\|\Delta^2 h\|_{3m\delta/2,\infty}\mathbb{E}_\alpha|X^{(\alpha)}|, \tag{5.23}$$

and (5.22) generates at most

$$\sqrt{dm}\|\Delta^2 h\|_{3m\delta/2,\infty} d^{1/2}|\mu^{(\alpha)}|\,\mathbb{P}[|W^{(\alpha)} - \mu| > \tfrac{1}{2}m\delta]$$
$$\le d\sqrt{m}\|\Delta^2 h\|_{3m\delta/2,\infty}|\mu^{(\alpha)}|\,\{\mathbb{P}[|W - \mu| > \tfrac{1}{4}m\delta] + \mathbb{P}[|Z^{(\alpha)}| > \tfrac{1}{4}m\delta]\}$$
$$\le 32d^2\delta^{-2}m^{-1/2}\|\Delta^2 h\|_{3m\delta/2,\infty}|\mu^{(\alpha)}|, \tag{5.24}$$

using Assumption (2.14) and Chebyshev's inequality in the last steps. Integrating over α with respect to ν, we deduce that the contribution from (5.16) to (5.12) is bounded by

$$\Big(32Cd^2\delta^{-2}m^{-1/2}\int_\Gamma \mathbb{E}_\alpha|X^{(\alpha)}|\nu(d\alpha) + d^{3/2}m\chi'_{13}\Big)\|\Delta^2 h\|_{3m\delta/2,\infty}$$
$$\le (32Cd^{5/2}\delta^{-2}m^{-1/2}H'_0 + d^{3/2}\chi'_{13})\, m\|\Delta^2 h\|_{3m\delta/2,\infty} \tag{5.25}$$

and to (5.14) is bounded by

$$\Big(32d^2\delta^{-2}m^{-1/2}\int_\Gamma |\mu^{(\alpha)}|\nu(d\alpha)\Big)\|\Delta^2 h\|_{3m\delta/2,\infty}$$
$$\le (32d^{5/2}\delta^{-2}m^{-1/2}H'_0)\, m\|\Delta^2 h\|_{3m\delta/2,\infty}. \tag{5.26}$$

This leaves the quantities in (5.11) and (5.13). First, we easily have

$$\begin{aligned}&\bigl|\mathbb{E}_\alpha\bigl\{(X^{(\alpha)})^T\Delta^2 h(W^{(\alpha)})Z^{(\alpha)}I[|W^{(\alpha)}-\mu|\le m\delta]I[|Z^{(\alpha)}|>\sqrt{m}]\bigr\}\bigr|\\ &\qquad\le\ m^{-1/2}\|\Delta^2 h\|_{m\delta,\infty}\mathbb{E}_\alpha\{|X^{(\alpha)}|_1\,|Z^{(\alpha)}|_1\,|Z^{(\alpha)}|\},\end{aligned}$$

and

$$\begin{aligned}&\bigl|\mathbb{E}\bigl\{(\mu^{(\alpha)})^T\Delta^2 h(W^{(\alpha)})Z^{(\alpha)}I[|W^{(\alpha)}-\mu|\le m\delta]I[|Z^{(\alpha)}|>\sqrt{m}]\bigr\}\bigr|\\ &\qquad\le\ m^{-1/2}\|\Delta^2 h\|_{m\delta,\infty}|\mu^{(\alpha)}|_1\,\mathbb{E}\{|Z^{(\alpha)}|_1\,|Z^{(\alpha)}|\},\end{aligned}$$

so that $I[|Z^{(\alpha)}|\le\sqrt{m}]$ can be dispensed with by incurring an extra error of at most

$$2m\|\Delta^2 h\|_{m\delta,\infty}\,d^{5/2}m^{-1/2}H'_{21}. \tag{5.27}$$

Then we can expand $Z^{(\alpha)}$, giving

$$\begin{aligned}&\mathbb{E}_\alpha\bigl\{(X^{(\alpha)})^T\Delta^2 h(W^{(\alpha)})Z^{(\alpha)}I[|W^{(\alpha)}-\mu|\le m\delta]\bigr\}\\ &\qquad=\ \int_{D_\alpha}\mathbb{E}_{\alpha\beta}\bigl\{(X^{(\alpha)})^T\Delta^2 h(W^{(\alpha)})\widetilde{X}^{(\alpha,\beta)}I[|W^{(\alpha)}-\mu|\le m\delta]\bigr\}\nu_\alpha(d\beta),\end{aligned} \tag{5.28}$$

where $\nu_\alpha(d\beta):=\nu_2(d\alpha,d\beta)/\nu(d\alpha)$, and

$$\begin{aligned}&\mathbb{E}\bigl\{(\mu^{(\alpha)})^T\Delta^2 h(W^{(\alpha)})Z^{(\alpha)}I[|W^{(\alpha)}-\mu|\le m\delta]\bigr\}\\ &\qquad=\ \int_{D_\alpha}\mathbb{E}_{\beta}\bigl\{(\mu^{(\alpha)})^T\Delta^2 h(W^{(\alpha)})\widetilde{X}^{(\alpha,\beta)}I[|W^{(\alpha)}-\mu|\le m\delta]\bigr\}\nu(d\beta),\end{aligned} \tag{5.29}$$

and then introduce the indicator $I[|Z^{(\alpha,\beta)}|\le\frac{1}{2}m\delta]$ in exchange for an error of at most

$$\begin{aligned}&\|\Delta^2 h\|_{m\delta,\infty}\int_\Gamma\int_{D_\alpha}\mathbb{E}_{\alpha\beta}\{|X^{(\alpha)}|_1\,|\widetilde{X}^{(\alpha,\beta)}|_1|Z^{(\alpha,\beta)}|/\tfrac{1}{2}m\delta\}\nu_2(d\alpha,d\beta)\\ &\qquad\le\ 2\delta^{-1}d^{5/2}H'_{22}\|\Delta^2 h\|_{m\delta,\infty}.\end{aligned} \tag{5.30}$$

$$\begin{aligned}&\|\Delta^2 h\|_{m\delta,\infty}\int_\Gamma\int_{D_\alpha}|\mu^{(\alpha)}|_1\mathbb{E}_{\beta}\{|\widetilde{X}^{(\alpha,\beta)}|_1|Z^{(\alpha,\beta)}|/\tfrac{1}{2}m\delta\}\nu(d\beta)\nu(d\alpha)\\ &\qquad\le\ 2\delta^{-1}d^{5/2}H'_{22}\|\Delta^2 h\|_{m\delta,\infty}.\end{aligned} \tag{5.31}$$

The next step is to split $\Delta^2 h(W^{(\alpha)})$ in (5.28) and (5.29), for $\beta\in D_\alpha$, giving

$$\Delta^2 h(W^{(\alpha)})\ =\ (\Delta^2 h(W^{(\alpha)})-\Delta^2 h(W^{(\alpha,\beta)}))+\Delta^2 h(W^{(\alpha,\beta)}). \tag{5.32}$$

Much as for (5.25), we write

$$\begin{aligned}&(\Delta^2 h(W^{(\alpha)})-\Delta^2 h(W^{(\alpha,\beta)}))I[|W^{(\alpha)}-\mu|\le m\delta]\\ &\quad=\ (\Delta^2 h(W^{(\alpha)})I[|W^{(\alpha)}-\mu|\le m\delta]-\Delta^2 h(W^{(\alpha,\beta)})I[|W^{(\alpha,\beta)}-\mu|\le m\delta])\\ &\qquad+\Delta^2 h(W^{(\alpha,\beta)})(I[|W^{(\alpha,\beta)}-\mu|\le m\delta]-I[|W^{(\alpha)}-\mu|\le m\delta]).\end{aligned}$$

Now, using (2.15), we deduce that

$$\begin{aligned}&\bigl|\mathbb{E}_{\alpha\beta}\bigl\{(X^{(\alpha)})^T(\Delta^2 h(W^{(\alpha)})I[|W^{(\alpha)}-\mu|\le m\delta]-\Delta^2 h(W^{(\alpha,\beta)})I[|W^{(\alpha,\beta)}-\mu|\le m\delta])\widetilde{X}^{(\alpha,\beta)}\\ &\qquad I[|Z^{(\alpha,\beta)}|\le\tfrac{1}{2}m\delta]\,|\,X^{(\alpha)},\widetilde{X}^{(\alpha,\beta)},Z^{(\alpha,\beta)}\bigr\}\bigr|\\ &\quad\le\ |X^{(\alpha)}|_1\,2\|\Delta^2 h\|_{3m\delta/2,\infty}\,|\widetilde{X}^{(\alpha,\beta)}|_1|Z^{(\alpha,\beta)}|_1\,\varepsilon'_W,\end{aligned} \tag{5.33}$$

and

$$\begin{aligned}&\bigl|\mathbb{E}_{\beta}\bigl\{(\mu^{(\alpha)})^T(\Delta^2 h(W^{(\alpha)})I[|W^{(\alpha)}-\mu|\le m\delta]-\Delta^2 h(W^{(\alpha,\beta)})I[|W^{(\alpha,\beta)}-\mu|\le m\delta])\widetilde{X}^{(\alpha,\beta)}\\ &\qquad I[|Z^{(\alpha,\beta)}|\le\tfrac{1}{2}m\delta]\,|\,X^{(\alpha)},\widetilde{X}^{(\alpha,\beta)},Z^{(\alpha,\beta)}\bigr\}\bigr|\\ &\quad\le\ |\mu^{(\alpha)}|_1\,2\|\Delta^2 h\|_{3m\delta/2,\infty}\,|\widetilde{X}^{(\alpha,\beta)}|_1|Z^{(\alpha,\beta)}|_1\,\varepsilon'_W,\end{aligned} \tag{5.34}$$

giving a first contribution to the errors incurred in (5.11) and (5.13) by splitting $\Delta^2 h(W^{(\alpha)})$ in (5.28) and (5.29) of

$$2m\|\Delta^2 h\|_{3m\delta/2,\infty}\, d^3 H'_{22}\varepsilon'_W. \tag{5.35}$$

For the remaining contribution, because

$$|I[|W^{(\alpha,\beta)}-\mu|\le m\delta]-I[|W^{(\alpha)}-\mu|\le m\delta]|\,I[|Z^{(\alpha,\beta)}|\le \tfrac12 m\delta] \;\le\; I[|W^{(\alpha,\beta)}-\mu|>\tfrac12 m\delta],$$

we have

$$\begin{aligned}
&\mathbb{E}_{\alpha\beta}\big\{\big|(X^{(\alpha)})^T\Delta^2 h(W^{(\alpha,\beta)})(I[|W^{(\alpha,\beta)}-\mu|\le m\delta]-I[|W^{(\alpha)}-\mu|\le m\delta])\widetilde{X}^{(\alpha,\beta)}\big| \\
&\qquad\qquad \times I[|Z^{(\alpha,\beta)}|\le \tfrac12 m\delta]\big\} \\
&\quad\le\quad \|\Delta^2 h\|_{3m\delta/2,\infty}\big(\mathbb{E}_{\alpha\beta}\{|X^{(\alpha)}|_1\,|\widetilde{X}^{(\alpha,\beta)}|_1\}\mathbb{P}_{\alpha\beta}[|W^{(\alpha,\beta)}-\mu|>\tfrac12 m\delta]+\chi'_{2\alpha\beta}\big),
\end{aligned}\tag{5.36}$$

and

$$\begin{aligned}
&\mathbb{E}_{\beta}\big\{\big|(\mu^{(\alpha)})^T\Delta^2 h(W^{(\alpha,\beta)})(I[|W^{(\alpha,\beta)}-\mu|\le m\delta]-I[|W^{(\alpha)}-\mu|\le m\delta])\widetilde{X}^{(\alpha,\beta)}\big| \\
&\qquad\qquad \times I[|Z^{(\alpha,\beta)}|\le \tfrac12 m\delta]\big\} \\
&\quad\le\quad \|\Delta^2 h\|_{3m\delta/2,\infty}\big(|\mu^{(\alpha)}|_1\,\mathbb{E}_{\beta}\{|\widetilde{X}^{(\alpha,\beta)}|_1\}\mathbb{P}_{\beta}[|W^{(\alpha,\beta)}-\mu|>\tfrac12 m\delta]+\chi''_{2\alpha\beta}\big).
\end{aligned}\tag{5.37}$$

Integrating over $\beta\in D_\alpha$ and then $\alpha\in\Gamma$, and using Assumption (2.14), the first part of (5.36) gives at most

$$\begin{aligned}
&d\|\Delta^2 h\|_{3m\delta/2,\infty}\int_\Gamma\int_{D_\alpha}\mathbb{E}_{\alpha\beta}\{|X^{(\alpha)}|\,|\widetilde{X}^{(\alpha,\beta)}|\} \\
&\qquad\qquad \big\{\mathbb{P}_{\alpha\beta}[|W-\mu|>\tfrac14 m\delta]+\mathbb{P}_{\alpha\beta}[|Z^{(\alpha)}|>\tfrac18 m\delta]+\mathbb{P}_{\alpha\beta}[|Z^{(\alpha,\beta)}|>\tfrac18 m\delta]\big\}\nu_2(d\alpha,d\beta) \\
&\le\quad \frac{144Cd^2}{m\delta^2}\|\Delta^2 h\|_{3m\delta/2,\infty}\int_\Gamma\int_{D_\alpha}\mathbb{E}_{\alpha\beta}\{|X^{(\alpha)}|\,|\widetilde{X}^{(\alpha,\beta)}|\}\nu_2(d\alpha,d\beta) \\
&=\quad \frac{144Cd^3}{\delta^2}\,H'_1\|\Delta^2 h\|_{3m\delta/2,\infty}
\end{aligned}\tag{5.38}$$

and the first part of (5.37) produces at most

$$\begin{aligned}
&d\|\Delta^2 h\|_{3m\delta/2,\infty}\int_\Gamma\int_{D_\alpha}|\mu^{(\alpha)}|\,\mathbb{E}_{\beta}\{|\widetilde{X}^{(\alpha,\beta)}|\} \\
&\qquad\qquad \big\{\mathbb{P}_{\beta}[|W-\mu|>\tfrac14 m\delta]+\mathbb{P}_{\beta}[|Z^{(\alpha)}|>\tfrac18 m\delta]+\mathbb{P}_{\beta}[|Z^{(\alpha,\beta)}|>\tfrac18 m\delta]\big\}\nu(d\beta)\nu(d\alpha) \\
&\le\quad \frac{144Cd^2}{m\delta^2}\|\Delta^2 h\|_{3m\delta/2,\infty}\int_\Gamma\int_{D_\alpha}|\mu^{(\alpha)}|\,\mathbb{E}_{\beta}\{|\widetilde{X}^{(\alpha,\beta)}|\}\nu(d\beta)\nu(d\alpha) \\
&=\quad \frac{144Cd^3}{\delta^2}\,H'_1\|\Delta^2 h\|_{3m\delta/2,\infty}.
\end{aligned}\tag{5.39}$$

The second parts of (5.36) and (5.37) give at most $d^3 m\|\Delta^2 h\|_{3m\delta/2,\infty}\chi'_2$. Thus (5.35), (5.36), (5.37), (5.38) and (5.39) together give a contribution to the error of at most

$$m\|\Delta^2 h\|_{3m\delta/2,\infty}\,\{2d^3 H'_{22}\,\varepsilon'_W+288Cd^3 m^{-1}H'_1\delta^{-2}+d^3\chi'_2\}. \tag{5.40}$$

Thus, having used (5.32) to replace $\Delta^2 h(W^{(\alpha)})$ by $\Delta^2 h(W^{(\alpha,\beta)})$ in (5.28) and (5.29), with the error being bounded by the sum of (5.30), (5.31) and (5.40), we are left with

$$\mathbb{E}_{\alpha\beta}\big\{(X^{(\alpha)})^T\Delta^2 h(W^{(\alpha,\beta)})\widetilde{X}^{(\alpha,\beta)}I[|W^{(\alpha)}-\mu|\le m\delta]I[|Z^{(\alpha,\beta)}|\le\tfrac12 m\delta]\big\} \tag{5.41}$$

and

$$\mathbb{E}_{\beta}\big\{(\mu^{(\alpha)})^T\Delta^2 h(W^{(\alpha,\beta)})\widetilde{X}^{(\alpha,\beta)}I[|W^{(\alpha)}-\mu|\le m\delta]I[|Z^{(\alpha,\beta)}|\le\tfrac12 m\delta]\big\}. \tag{5.42}$$

Exactly as above, we can replace $I[|W^{(\alpha)} - \mu| \le m\delta]$ by $I[|W^{(\alpha,\beta)} - \mu| \le m\delta]$, adding a second contribution as in (5.38), (5.39) and $d^3 m \|\Delta^2 h\|_{3m\delta/2,\infty}\chi_2'$ to the error. Then, to remove the factor $I[|Z^{(\alpha,\beta)}| \le \frac{1}{2}m\delta]$, note that

$$\begin{aligned}
&\mathbb{E}_{\alpha\beta}\big\{|(X^{(\alpha)})^T \Delta^2 h(W^{(\alpha,\beta)})\widetilde{X}^{(\alpha,\beta)}|\, I[|W^{(\alpha,\beta)} - \mu| \le m\delta]\, I[|Z^{(\alpha,\beta)}| > \tfrac{1}{2}m\delta]\big\} \\
&\quad\le\ \|\Delta^2 h\|_{m\delta,\infty}\mathbb{E}_{\alpha\beta}\{|X^{(\alpha)}|_1\, |\widetilde{X}^{(\alpha,\beta)}|_1 |Z^{(\alpha,\beta)}|\}/\tfrac{1}{2}m\delta, \qquad (5.43)
\end{aligned}$$

and that

$$\begin{aligned}
&\mathbb{E}_{\beta}\big\{|(\mu^{(\alpha)})^T \Delta^2 h(W^{(\alpha,\beta)})\widetilde{X}^{(\alpha,\beta)}|\, I[|W^{(\alpha,\beta)} - \mu| \le m\delta]\, I[|Z^{(\alpha,\beta)}| > \tfrac{1}{2}m\delta]\big\} \\
&\quad\le\ \|\Delta^2 h\|_{m\delta,\infty}\mathbb{E}_{\beta}\{|\mu^{(\alpha)}|_1\, |\widetilde{X}^{(\alpha,\beta)}|_1 |Z^{(\alpha,\beta)}|\}/\tfrac{1}{2}m\delta. \qquad (5.44)
\end{aligned}$$

Integrating over $\beta \in D_\alpha$ and $\alpha \in \Gamma$ thus gives a contribution to the error of at most

$$2d^{5/2}\delta^{-1}H_{22}'\|\Delta^2 h\|_{m\delta,\infty}. \qquad (5.45)$$

After these adjustments, we are left with

$$\begin{aligned}
&\int_\Gamma\int_{D_\alpha} \mathbb{E}_{\alpha\beta}\{(X^{(\alpha)})^T\Delta^2 h(W^{(\alpha,\beta)})\widetilde{X}^{(\alpha,\beta)} I[|W^{(\alpha,\beta)} - \mu| \le m\delta]\}\nu_2(d\alpha, d\beta) \qquad (5.46)\\
&\quad=\ \int_\Gamma\int_{D_\alpha} \mathrm{Tr}\big(\mathbb{E}_{\alpha\beta}((X^{(\alpha)})(\widetilde{X}^{(\alpha,\beta)})^T) \\
&\qquad\qquad \times\mathbb{E}_{\alpha\beta}\{\Delta^2 h(W^{(\alpha,\beta)}) I[|W^{(\alpha,\beta)} - \mu| \le m\delta]\}\big)\nu_2(d\alpha, d\beta) + \eta_3 \\
&\quad=\ \int_\Gamma\int_{D_\alpha} \mathrm{Tr}\big(\mathbb{E}_{\alpha\beta}((X^{(\alpha)})(\widetilde{X}^{(\alpha,\beta)})^T) \\
&\qquad\qquad \times\mathbb{E}\{\Delta^2 h(W^{(\alpha,\beta)}) I[|W^{(\alpha,\beta)} - \mu| \le m\delta]\}\big)\nu_2(d\alpha, d\beta) + \eta_3 + \eta_4
\end{aligned}$$

and

$$\begin{aligned}
&\int_\Gamma\int_{D_\alpha} \mathbb{E}_{\beta}\{(\mu^{(\alpha)})^T\Delta^2 h(W^{(\alpha,\beta)})\widetilde{X}^{(\alpha,\beta)} I[|W^{(\alpha,\beta)} - \mu| \le m\delta]\}\nu(d\beta)\nu(d\alpha) \qquad (5.47)\\
&\quad=\ \int_\Gamma\int_{D_\alpha} \mathrm{Tr}\big(\mathbb{E}_{\beta}((\mu^{(\alpha)})(\widetilde{X}^{(\alpha,\beta)})^T) \\
&\qquad\qquad \times\mathbb{E}_{\beta}\{\Delta^2 h(W^{(\alpha,\beta)}) I[|W^{(\alpha,\beta)} - \mu| \le m\delta]\}\big)\nu(d\beta)\nu(d\alpha) + \eta_5 \\
&\quad=\ \int_\Gamma\int_{D_\alpha} \mathrm{Tr}\big(\mathbb{E}_{\beta}((\mu^{(\alpha)})(\widetilde{X}^{(\alpha,\beta)})^T) \\
&\qquad\qquad \times\mathbb{E}\{\Delta^2 h(W^{(\alpha,\beta)}) I[|W^{(\alpha,\beta)} - \mu| \le m\delta]\}\big)\nu(d\beta)\nu(d\alpha) + \eta_5 + \eta_6.
\end{aligned}$$

One can bound η_3 and η_5 by

$$|\eta_3| + |\eta_5|\ \le\ m\|\Delta^2 h\|_{m\delta,\infty}\, d^3\chi_2', \qquad (5.48)$$

and each of η_4 and η_6 by

$$\max\{|\eta_4|, |\eta_6|\}\ \le\ 2d^2 m\|\Delta^2 h\|_{m\delta,\infty} H_1'\varepsilon_W''. \qquad (5.49)$$

Since, from (2.15), for any $1 \le l, m \le d$, we have

$$\begin{aligned}
&|\mathbb{E}\{\Delta^2_{lm} h(W) I[|W - \mu| \le m\delta]\} - \mathbb{E}\{\Delta^2_{lm} h(W^{(\alpha,\beta)}) I[|W^{(\alpha,\beta)} - \mu| \le m\delta]\}| \\
&\quad\le\ \|\Delta^2 h\|_{m\delta,\infty}(\mathbb{E}|Z^{(\alpha,\beta)}|_1 + \mathbb{E}|Z^{(\alpha)}|_1)\varepsilon_W', \qquad (5.50)
\end{aligned}$$

we can replace $\mathbb{E}\{\Delta^2 h(W^{(\alpha,\beta)}) I[|W^{(\alpha,\beta)} - \mu| \le m\delta]\}$ by $\mathbb{E}\{\Delta^2 h(W) I[|W - \mu| \le m\delta]\}$ in (5.46) and (5.47), introducing further errors of at most

$$\begin{aligned}
&\int_\Gamma \int_{D_\alpha} \sum_{i=1}^d \sum_{l=1}^d |\mathbb{E}_{\alpha\beta}\{X_i^{(\alpha)} \widetilde{X}_l^{(\alpha,\beta)}\}| (\mathbb{E}|Z^{(\alpha,\beta)}|_1 + \mathbb{E}|Z^{(\alpha)}|_1) \nu_2(d\alpha, d\beta) \|\Delta^2 h\|_{m\delta,\infty} \varepsilon'_W \\
&\quad \le\ \|\Delta^2 h\|_{m\delta,\infty} \varepsilon'_W \int_\Gamma \int_{D_\alpha} \mathbb{E}_{\alpha\beta}\{|X^{(\alpha)}|_1\, |\widetilde{X}^{(\alpha,\beta)}|_1\} (\mathbb{E}|Z^{(\alpha,\beta)}|_1 + \mathbb{E}|Z^{(\alpha)}|_1) \nu_2(d\alpha, d\beta) \\
&\quad \le\ d^3 m \|\Delta^2 h\|_{m\delta,\infty} (H'_{23} + H'_{24}) \varepsilon'_W, \qquad (5.51)
\end{aligned}$$

and

$$\begin{aligned}
&\int_\Gamma \int_{D_\alpha} \sum_{i=1}^d \sum_{l=1}^d |\mathbb{E}_{\beta}\{\mu_i^{(\alpha)} \widetilde{X}_l^{(\alpha,\beta)}\}| (\mathbb{E}|Z^{(\alpha,\beta)}|_1 + \mathbb{E}|Z^{(\alpha)}|_1) \nu(d\beta)\nu(d\alpha) \|\Delta^2 h\|_{m\delta,\infty} \varepsilon'_W \\
&\quad \le\ \|\Delta^2 h\|_{m\delta,\infty} \varepsilon'_W \int_\Gamma \int_{D_\alpha} |\mu^{(\alpha)}|_1\, \mathbb{E}_{\beta}\{|\widetilde{X}^{(\alpha,\beta)}|_1\} (\mathbb{E}|Z^{(\alpha,\beta)}|_1 + \mathbb{E}|Z^{(\alpha)}|_1) \nu(d\beta)\nu(d\alpha) \\
&\quad \le\ d^3 m \|\Delta^2 h\|_{m\delta,\infty} (H'_{23} + H'_{24}) \varepsilon'_W, \qquad (5.52)
\end{aligned}$$

and leaving the principal term of

$$\mathbb{E}\{\mathrm{Tr}(\widehat{V}\Delta^2 h(W))\, I[|W-\mu| \le m\delta]\}, \qquad (5.53)$$

where

$$\begin{aligned}
\widehat{V} \ &:=\ \int_\Gamma \int_{D_\alpha} \mathbb{E}_{\alpha\beta}\big(X^{(\alpha)} (\widetilde{X}^{(\alpha,\beta)})^T\big) \nu_2(d\alpha, d\beta) - \int_\Gamma \int_{D_\alpha} \mathbb{E}_{\beta}\big(\mu^{(\alpha)} (\widetilde{X}^{(\alpha,\beta)})^T\big) \nu(d\beta)\nu(d\alpha) \\
&=\ \mathbb{E}\int_\Gamma \int_{D_\alpha} X^{(\alpha)} (\widetilde{X}^{(\alpha,\beta)})^T \Xi(d\beta)\Xi(d\alpha) - \mathbb{E}\int_\Gamma \int_{D_\alpha} \mu^{(\alpha)} (\widetilde{X}^{(\alpha,\beta)})^T \Xi(d\beta)\nu(d\alpha) \\
&=\ \mathbb{E}\int_\Gamma X^{(\alpha)} (Z^{(\alpha)})^T \Xi(d\alpha) - \mathbb{E}\int_\Gamma \mu^{(\alpha)} (Z^{(\alpha)})^T \nu(d\alpha). \qquad (5.54)
\end{aligned}$$

We now recall the first term in $\mathbb{E}\{\widetilde{\mathcal{A}}_m h(W) I[|W-\mu| \le m\delta]\}$, which is

$$\mathbb{E}\{\mathrm{Tr}(V\Delta^2 h(W)) I[|W-\mu| \le m\delta]\}, \qquad (5.55)$$

differing from that in (5.53) only because the matrix $V = \mathrm{Cov}(W)$ replaces $\widehat{V}$. If approximation by $\mathrm{DN}_d(\mu, \widehat{V})$ is required, it is now enough to collect the various errors. If not, we can write

$$V \ =\ \mathrm{Cov}(W) \ =\ \mathbb{E}\int_\Gamma \{X^{(\alpha)} W^T\} \Xi(d\alpha) - \mathbb{E}\int_\Gamma \{\mu^{(\alpha)} W^T\} \nu(d\alpha),$$

so that, recalling $W = W^{(\alpha)} + Z^{(\alpha)}$, we have

$$\begin{aligned}
V - \widehat{V} \ &=\ \mathbb{E}\int_\Gamma \{X^{(\alpha)} (W^{(\alpha)})^T\} \Xi(d\alpha) - \mathbb{E}\int_\Gamma \{\mu^{(\alpha)} (W^{(\alpha)})^T\} \nu(d\alpha) \\
&=\ \mathbb{E}\int_\Gamma \{X^{(\alpha)} (W^{(\alpha)} - \mu)^T\} \Xi(d\alpha) - \mathbb{E}\int_\Gamma \{\mu^{(\alpha)} (W^{(\alpha)} - \mu)^T\} \nu(d\alpha).
\end{aligned}$$

Defining

$$\begin{aligned}
V' \ &:=\ \int_\Gamma \mathbb{E}_\alpha\big\{(\mathbb{E}_\alpha(X^{(\alpha)} \,|\, W^{(\alpha)}) - \mu^{(\alpha)})(W^{(\alpha)} - \mu)^T\big\} \nu(d\alpha), \\
V'' \ &:=\ \int_\Gamma \mu^{(\alpha)} \big\{\mathbb{E}_\alpha\left((W^{(\alpha)})^T\right) - \mathbb{E}\left((W^{(\alpha)})^T\right)\big\} \nu(d\alpha),
\end{aligned}$$

we thus have

$$V - \widehat{V} \;=\; V' + V''.$$

Hence the difference between (5.53) and (5.55) can be bounded by

$$\begin{aligned}
&\|\Delta^2 h\|_{m\delta,\infty} \sum_{i=1}^{d}\sum_{l=1}^{d}(|V'_{il}| + |V''_{il}|) \\
&\quad\le\; \|\Delta^2 h\|_{m\delta,\infty}\int_\Gamma \mathbb{E}_\alpha\{|\mathbb{E}_\alpha(X^{(\alpha)} \mid W^{(\alpha)}) - \mu^{(\alpha)}|_1\, |W^{(\alpha)} - \mu|_1\}\nu(d\alpha) \\
&\qquad + \|\Delta^2 h\|_{m\delta,\infty}\int_\Gamma |\mu^{(\alpha)}|_1 \Big|\mathbb{E}_\alpha\left((W^{(\alpha)})^T\right) - \mathbb{E}\left((W^{(\alpha)})^T\right)\Big|_1 \nu(d\alpha) \\
&\quad\le\; d^2 m\|\Delta^2 h\|_{m\delta,\infty}\chi'_3 + d^{3/2} m\|\Delta^2 h\|_{m\delta,\infty} H'_0\varepsilon'''_W, \qquad (5.56)
\end{aligned}$$

where the second element in (5.56) is from (2.17).

Adding the error bounds in (5.3), (5.6), (5.9), (5.10), (5.19), (5.20), (5.25), (5.26) (5.27), (5.30), (5.31), (5.40), (5.45), (5.48), (5.49), (5.51), (5.52) and (5.56), using (5.2) and with $\sqrt{m} \ge 2/\delta$, gives

$$\begin{aligned}
&|\mathbb{E}\{\widetilde{\mathcal{A}}_m h(W) I[|W-\mu| \le m\delta]\}| \\
&\quad\le\; C_1(\delta)\{m^{-1/2}d^{3/2}(1 + H'_0 + H'_2) + d^{1/2}\chi'_1 + (dm)^{1/2}H'_0\varepsilon''_W\}\, m^{1/2}\|\Delta h\|_{m\delta_0,\infty} \\
&\qquad + C_2(\delta)\big\{\varepsilon'_W d^3 H'_2 + d^3 m^{-1} H'_1 + m^{-1/2} d^{5/2}(H'_0 + H'_2) \\
&\qquad\qquad + d^{3/2}\chi'_1 + d^3\chi'_2 + d^2\chi'_3 + d^2 H'_1\varepsilon''_W + d^{3/2}H'_0\varepsilon'''_W\big\}\, m\|\Delta^2 h\|_{3m\delta/2,\infty}.
\end{aligned}$$

Recalling Theorem 1.1, Theorem 2.2 follows.

Theorem 2.1 can be deduced from Theorem 2.2 directly by taking $\Gamma = \{1, \ldots, n\}$, Ξ as the counting measure on Γ so that $\Xi(\{i\}) = \nu(\{i\}) = 1$ for all $i \in \Gamma$ and $\nu_2(\{i\}, \{j\}) = 1$ for all $i, j \in \Gamma$; replacing $\int$ with $\sum$; α with j, β with k; $\mathbb{E}_\alpha$, $\mathbb{E}_\beta$, $\mathbb{E}_{\alpha\beta}$ with $\mathbb{E}$; $\mathbb{P}_\alpha$, $\mathbb{P}_\beta$, $\mathbb{P}_{\alpha\beta}$ with $\mathbb{P}$ so that $\varepsilon''_W = \varepsilon'''_W = 0$; χ', H', ε'_W with χ, H, ε_W. $\square$

References

[Baldi & Rinott (1989)] P. Baldi & Y. Rinott (1989) On normal approximations of distributions in terms of dependency graphs. *Ann. Probab.* **17**, 1646–1650. MR-1048950

[Barbour & Lindvall (2006)] A. D. Barbour & T. Lindvall (2006) Translated Poisson approximation for Markov chains. *J. Theor. Probab.* **19**, 609–630. MR-2280512

[Barbour, Luczak & Xia (2018a)] A. D. Barbour, M. J. Luczak & A. Xia (2018a) Multivariate approximation in total variation, I: equilibrium distributions of Markov jump processes. *Ann. Probab.* **46**, 1351–1404. MR-3785590

[Barbour, Luczak & Xia (2018b)] A. D. Barbour, M. J. Luczak & A. Xia (2018b) Multivariate approximation in total variation, II: discrete normal approximation. *Ann. Probab.* **46**, 1405–1440. MR-3785591

[Barbour & Xia (2006)] A. D. Barbour & A. Xia (2006) Normal approximation for random sums. *Adv. Appl. Probab.* **38**, 693–728. MR-2256874

[Chen, Goldstein & Shao (2011)] L. H. Y. Chen, L. Goldstein & Q.-M. Shao (2011) *Normal approximation by Stein's method.* Springer–Verlag, Berlin. MR-2732624

[Chen & Xia (2004)] L. H. Y. Chen & A. Xia (2004) Stein's method, Palm theory and Poisson process approximation. *Ann. Probab.* **32**, 2545–2569. MR-2078550

[Daley & Vere-Jones (2008)] D. J. Daley & D. Vere-Jones (2008) *An Introduction to the Theory of Point Processes.* Vol. 2, Springer, New York. MR-2371524

[Fang (2014)] X. Fang (2014) Discretized normal approximation by Stein's method. *Bernoulli* **20**, 1404–1431. MR-3217448

[Kallenberg (1983)] O. Kallenberg (1983) *Random Measures.* Academic Press, New York. MR-0818219

[Kallenberg (2017)] O. Kallenberg (2017) *Random Measures, Theory and Applications.* Springer, Switzerland. MR-3642325

[Lindvall (2002)] T. Lindvall (2002) *Lectures on the coupling method.* 2nd Edn, Dover Publications, Mineola NY. MR-1924231

[Mattner & Roos (2007)] L. Mattner & B. Roos (2007) A shorter proof of Kanter's Bessel function concentration bound. *Probab. Theory Relat. Fields,* **139**, 191–205. MR-2322695

[Penrose (2003)] M. Penrose (2003) *Random geometric graphs.* Oxford University Press. MR-1986198

[Rinott & Rotar (1996)] Y. Rinott & V. Rotar (1996) A multivariate CLT for local dependence with $n^{-1/2}\log n$ rate and applications to multivariate graph statistics. *J. Multivariate Anal.* **56**, 333–350. MR-1379533

[Schulte & Yukich (2018a)] M. Schulte & J. E. Yukich (2018a) Multivariate second order Poincaré inequalities for Poisson functionals. arXiv:1803.11059

[Schulte & Yukich (2018b)] M. Schulte & J. E. Yukich (2018b) Rates of multivariate normal approximation for statistics in geometric probability. In preparation.