The Annals of Applied Probability
2021, Vol. 31, No. 6, 2881–2923
https://doi.org/10.1214/21-AAP1666

PALM THEORY, RANDOM MEASURES AND STEIN COUPLINGS

BY LOUIS H. Y. CHEN[1], ADRIAN RÖLLIN[2] AND AIHUA XIA[3]

[1]*Department of Mathematics, National University of Singapore, matchyl@nus.edu.sg*
[2]*Department of Statistics and Applied Probability, National University of Singapore, adrian.roellin@nus.edu.sg*
[3]*School of Mathematics and Statistics, University of Melbourne, aihuaxia@unimelb.edu.au*

We establish a general Berry–Esseen type bound which gives optimal bounds in many situations under suitable moment assumptions. By combining the general bound with Palm theory, we deduce a new error bound for assessing the accuracy of normal approximation to statistics arising from random measures, including stochastic geometry. We illustrate the use of the bound in four examples: completely random measures, excursion random measure of a locally dependent random process, and the total edge length of Ginibre–Voronoi tessellations and of Poisson–Voronoi tessellations. Moreover, we apply the general bound to Stein couplings and discuss the special cases of local dependence and additive functionals in occupancy problems.

1. Introduction. The pioneering work of Stein (1972), well known as Stein's method, provides a set of tools to estimate the error in the approximation of the distributions of random variables by a specific distribution, and it has proven to be particularly powerful in the presence of dependence. Indeed, many forms of Stein's method have been developed to study a variety of random phenomena, and the comprehensive monographs by Barbour, Holst and Janson (1992) and Chen, Goldstein and Shao (2011) give accounts to that diversity.

It has become clear over the past decades that Stein's method is naturally related to size biasing and its point process counterpart Palm theory; see, for example, the results of Goldstein and Rinott (1996), Chen and Xia (2004) and Goldstein and Xia (2006). While Goldstein and Rinott (1996) and Goldstein and Xia (2006) considered size-bias couplings, Chen and Xia (2004) studied Poisson process approximation for point processes using Palm theory. The work of Chen and Xia (2004) suggests that for normal approximation for statistics resulting from a random measure including those in stochastic geometry, it may be fruitful to combine Stein's method with Palm theory as well. Thus, in this article, we study normal approximation for statistics associated with random measures through their Palm distributions.

To this end, we first prove a general result, Theorem 2.1, which can be thought of as an extension of Theorem 2.1 of Chen and Shao (2004) to settings that are not restricted to local dependence. We then connect our result with Palm theory in Section 3 to bound the errors of normal approximation for statistics arising from random measures. In order to illustrate the approach, we then estimate in Section 4 the errors in the normal approximation for completely random measures, the excursion random measure of a locally dependent random process and the total edge length of Ginibre–Voronoi tessellations as well as Poisson–Voronoi tessellations. The first three examples do not assume the Poisson process as an underlying point process. Theorem 2.1 can also be easily combined with Stein couplings, giving rise to Theorem 5.1 in Section 5, with applications to local dependence and problems from random occupancy.

Received April 2020; revised August 2020.
MSC2020 subject classifications. Primary 60F05; secondary 60G55, 60G57.
Key words and phrases. Stein's method, normal approximation, Palm distribution, random measure, stochastic geometry, Stein coupling.

Our main theorems are formulated in such a way so as to give optimal rates of convergence in many applications. The cost we have to pay are higher moment requirements, but in many applications these are naturally satisfied.

2. A general theorem. Let W be such that $\mathbb{E}W = 0$ and $\operatorname{Var} W = 1$. Theorem 2.1 of Chen and Shao (2004) shows that if W is a sum of LD1 (see Section 5.1 for more details) locally dependent random variables then there exists a random function $\hat{K}(t)$ such that

$$\mathbb{E}\{Wf(W)\} = \mathbb{E}\int_{-\infty}^{\infty} f'(W+t)\hat{K}(t)\,dt \tag{2.1}$$

for all absolutely continuous functions f for which the expectations exist. A bound on the Kolmogorov distance $d_{\mathrm{K}}(\mathscr{L}(W), \mathcal{N}(0,1))$ is then obtained without further dependence assumption. A crucial step in the proof is the use of a concentration inequality (Proposition 3.1 of Chen and Shao (2004)) established under the LD1 local dependence. A careful examination of the proof of the proposition reveals that the concentration inequality actually holds if W only satisfies (2.1), in which case the bound is expressed in terms of $\hat{K}(t)$ instead of the locally dependent random variables. Consequently, Theorem 2.1 of Chen and Shao (2004) holds for any W if $\operatorname{Var}(W) = 1$ and there exists a random function $\hat{K}(t)$ such that (2.1) holds.

It was observed by Chen and Röllin (2010) that the proof of Theorem 2.1 of Chen and Shao (2004) can be simplified if the concentration inequality is replaced by a recursive inequality, which was inspired by Raič (2003) and which is (2.14) in this paper. Using this approach, they obtain a bound for W satisfying a Stein coupling assumption. In this paper we use the recursive inequality approach to obtain a simpler bound for W assuming that $\operatorname{Var} W = 1$ and that (2.1) holds for some random function $\hat{K}(t)$. As in the proof of Theorem 2.1 of Chen and Shao (2004), Young's inequality ((2.11) in this paper) is used to separate the product of two random variables. A crucial step in the proof of Chen and Shao (2004) is to use Young's inequality together with the concentration inequality, whereas in this paper, it is used together with the recursive inequality. Also in this paper, the random function $\hat{K}(t)$ is decomposed as $\hat{K}^{\mathrm{in}}(t) + \hat{K}^{\mathrm{out}}(t)$ to allow greater flexibility in applications. We now state and prove the general theorem.

THEOREM 2.1. *Let W be such that $\mathbb{E}W = 0$ and $\operatorname{Var} W = 1$. Suppose there is a random function $\hat{K}(t)$ such that (2.1) holds for all absolutely continuous functions f for which the expectations exist, and assume we can write $\hat{K}(t) = \hat{K}^{\mathrm{in}}(t) + \hat{K}^{\mathrm{out}}(t)$, where $\hat{K}^{\mathrm{in}}(t) = 0$ for $|t| > 1$. Define $K(t) = \mathbb{E}\hat{K}(t)$, $K^{\mathrm{in}}(t) = \mathbb{E}\hat{K}^{\mathrm{in}}(t)$, and $K^{\mathrm{out}}(t) = \mathbb{E}\hat{K}^{\mathrm{out}}(t)$. Then*

$$d_{\mathrm{K}}\big(\mathscr{L}(W), \mathcal{N}(0,1)\big) \le 2r_1 + 11r_2 + 5r_3 + 10r_4 + 7r_5, \tag{2.2}$$

where

$$r_1 = \mathbb{E}\left|\int_{|t|\le 1}\big(\hat{K}^{\mathrm{in}}(t) - K^{\mathrm{in}}(t)\big)\,dt\right|, \qquad r_2 = \int_{|t|\le 1}\big|tK^{\mathrm{in}}(t)\big|\,dt,$$

$$r_3 = \mathbb{E}\int_{-\infty}^{\infty}\big|\hat{K}^{\mathrm{out}}(t)\big|\,dt, \qquad r_4 = \mathbb{E}\int_{|t|\le 1}\big(\hat{K}^{\mathrm{in}}(t) - K^{\mathrm{in}}(t)\big)^2\,dt,$$

$$r_5 = \left(\mathbb{E}\int_{|t|\le 1}|t|\big(\hat{K}^{\mathrm{in}}(t) - K^{\mathrm{in}}(t)\big)^2\,dt\right)^{1/2}.$$

PROOF. From (2.1), by letting $f(w) = w$, we obtain $\int_{-\infty}^{\infty} K(t)\,dt = 1$. For $x \in \mathbb{R}$ and $\varepsilon > 0$, define

$$h_{x,\varepsilon}(w) := \begin{cases} 1 & \text{if } w \le x, \\ 0 & \text{if } w \ge x+\varepsilon, \\ 1 + \varepsilon^{-1}(x-w) & \text{if } x < w < x+\varepsilon. \end{cases}$$

Let $f_{x,\varepsilon}$ be the bounded solution of the Stein equation

$$f'_{x,\varepsilon}(w) - wf_{x,\varepsilon}(w) = h_{x,\varepsilon}(w) - \mathbb{E}h_{x,\varepsilon}(Z), \tag{2.3}$$

where $Z \sim \mathcal{N}(0,1)$. The bounded solution $f_{x,\varepsilon}$ of (2.3) is unique and is given by

$$f_{x,\varepsilon}(w) = -e^{\frac{1}{2}w^2}\int_w^\infty e^{-\frac{1}{2}t^2}\big[h_{x,\varepsilon}(t) - \mathbb{E}h_{x,\varepsilon}(Z)\big]\,dt$$

(see Chen, Goldstein and Shao (2011), page 15). We have for all $w, v \in \mathbb{R}$,

$$0 \le f_{x,\varepsilon}(w) \le 1, \qquad \big|f'_{x,\varepsilon}(w)\big| \le 1, \qquad \big|f'_{x,\varepsilon}(w) - f'_{x,\varepsilon}(v)\big| \le 1 \tag{2.4}$$

and

$$\begin{aligned}\big|f'_{x,\varepsilon}(w+t) - f'_{x,\varepsilon}(w)\big| &\le \big(|w|+1\big)|t| + \frac{1}{\varepsilon}\int_{t\wedge 0}^{t\vee 0}\mathbf{1}[x \le w+u \le x+\varepsilon]\,du \\ &\le \big(|w|+1\big)|t| + \mathbf{1}[x-0\vee t \le w \le x-0\wedge t+\varepsilon].\end{aligned} \tag{2.5}$$

The bounds (2.4) and (2.5) were obtained by Chen and Shao (2004), page 2010. Bounds for all cases of h and their proofs were given by Chen, Goldstein and Shao (2011), Section 2.2. Now write

$$\begin{aligned}&\mathbb{E}h_{x,\varepsilon}(W) - \mathbb{E}h_{x,\varepsilon}(Z)\\ &\quad= \mathbb{E}\int_{-\infty}^\infty f'_{x,\varepsilon}(W)K(t)\,dt - \mathbb{E}\int_{-\infty}^\infty f'_{x,\varepsilon}(W+t)\hat{K}(t)\,dt\\ &\quad= \mathbb{E}\int_{|t|\le 1} f'_{x,\varepsilon}(W)\big(K^{\mathrm{in}}(t) - \hat{K}^{\mathrm{in}}(t)\big)\,dt\\ &\qquad+ \mathbb{E}\int_{-\infty}^\infty f'_{x,\varepsilon}(W)\big(K^{\mathrm{out}}(t) - \hat{K}^{\mathrm{out}}(t)\big)\,dt\\ &\qquad+ \mathbb{E}\int_{-\infty}^\infty \big(f'_{x,\varepsilon}(W) - f'_{x,\varepsilon}(W+t)\big)\hat{K}^{\mathrm{out}}(t)\,dt\\ &\qquad+ \mathbb{E}\int_{|t|\le 1} \big(f'_{x,\varepsilon}(W) - f'_{x,\varepsilon}(W+t)\big)\big(\hat{K}^{\mathrm{in}}(t) - K^{\mathrm{in}}(t)\big)\,dt\\ &\qquad+ \mathbb{E}\int_{|t|\le 1} \big(f'_{x,\varepsilon}(W) - f'_{x,\varepsilon}(W+t)\big)K^{\mathrm{in}}(t)\,dt\\ &\quad=: R_1 + R_2 + R_3 + R_4 + R_5.\end{aligned} \tag{2.6}$$

By (2.4), we obtain bounds

$$\begin{aligned}|R_1| &= \left|\mathbb{E}\left\{f'_{x,\varepsilon}(W)\int_{|t|\le 1}\big(\hat{K}^{\mathrm{in}}(t) - K^{\mathrm{in}}(t)\big)\,dt\right\}\right|\\ &\le \mathbb{E}\left|\int_{|t|\le 1}\big(\hat{K}^{\mathrm{in}}(t) - K^{\mathrm{in}}(t)\big)\,dt\right| = r_1,\end{aligned} \tag{2.7}$$

$$\begin{aligned}|R_2| &= \left|\mathbb{E}\left\{f'_{x,\varepsilon}(W)\int_{-\infty}^\infty\big(K^{\mathrm{out}}(t) - \hat{K}^{\mathrm{out}}(t)\big)\,dt\right\}\right|\\ &\le \mathbb{E}\int_{-\infty}^\infty\big|K^{\mathrm{out}}(t)\big|\,dt + \mathbb{E}\int_{-\infty}^\infty\big|\hat{K}^{\mathrm{out}}(t)\big|\,dt \le 2\mathbb{E}\int_{-\infty}^\infty\big|\hat{K}^{\mathrm{out}}(t)\big|\,dt = 2r_3\end{aligned} \tag{2.8}$$

and

$$|R_3| \le \mathbb{E}\int_{-\infty}^\infty\big|\hat{K}^{\mathrm{out}}(t)\big|\,dt = r_3. \tag{2.9}$$

By (2.5),

$$
\begin{aligned}
|R_4| &\le \mathbb{E}\biggl\{(|W|+1)\int_{|t|\le 1}|t|\bigl|\hat{K}^{\mathrm{in}}(t)-K^{\mathrm{in}}(t)\bigr|\,dt\biggr\}\\
&\quad + \mathbb{E}\int_{|t|\le 1}\mathbf{1}[x-0\vee t\le W\le x-0\wedge t+\varepsilon]\bigl|\hat{K}^{\mathrm{in}}(t)-K^{\mathrm{in}}(t)\bigr|\,dt \qquad (2.10)\\
&=: R_{4,1}+R_{4,2}.
\end{aligned}
$$

Recall Young's inequality: For $a, b, c > 0$, we have

$$
ab \le \frac{ca^2}{2}+\frac{b^2}{2c}. \qquad (2.11)
$$

Using this inequality with $c=\alpha>0$, $a=(|W|+1)\sqrt{|t|}$ and $b=\sqrt{|t|}|\hat{K}^{\mathrm{in}}(t)-K^{\mathrm{in}}(t)|$, we have

$$
\begin{aligned}
R_{4,1} &\le \frac{\alpha}{2}\mathbb{E}\int_{|t|\le 1}\bigl(|W|+1\bigr)^2|t|\,dt+\frac{1}{2\alpha}\mathbb{E}\int_{|t|\le 1}|t|\bigl(\hat{K}^{\mathrm{in}}(t)-K^{\mathrm{in}}(t)\bigr)^2\,dt\\
&\le 2\alpha+\frac{1}{2\alpha}\mathbb{E}\int_{|t|\le 1}|t|\bigl(\hat{K}^{\mathrm{in}}(t)-K^{\mathrm{in}}(t)\bigr)^2\,dt=2\alpha+\frac{r_5^2}{2\alpha}.
\end{aligned}
$$

By letting $\alpha=r_5$,

$$
R_{4,1}\le 2.5r_5. \qquad (2.12)
$$

Using the inequality (2.11) again, but with $c=(2d+0.4|t|+0.4\varepsilon)/(\theta\beta)$ for $\theta,\beta>0$, $b=\mathbf{1}[x-0\vee t\le W\le x-0\wedge t+\varepsilon]$ and $a=|\hat{K}^{\mathrm{in}}(t)-K^{\mathrm{in}}(t)|$, we obtain

$$
\begin{aligned}
R_{4,2} &\le \frac{\theta\beta}{2}\mathbb{E}\int_{|t|\le 1}\bigl(2d+0.4|t|+0.4\varepsilon\bigr)^{-1}\mathbf{1}[x-0\vee t\le W\le x-0\wedge t+\varepsilon]\,dt\\
&\quad +\frac{1}{2\theta\beta}\mathbb{E}\int_{|t|\le 1}\bigl(2d+0.4|t|+0.4\varepsilon\bigr)\bigl(\hat{K}^{\mathrm{in}}(t)-K^{\mathrm{in}}(t)\bigr)^2\,dt. \qquad (2.13)
\end{aligned}
$$

Let

$$
d=d_{\mathrm{K}}\bigl(\mathscr{L}(W),\mathcal{N}(0,1)\bigr),\qquad d_\varepsilon=\sup_{x\in\mathbb{R}}\bigl|\mathbb{E}h_{x,\varepsilon}(W)-h_{x,\varepsilon}(Z)\bigr|;
$$

then it is not difficult to see that, for $a\le b$,

$$
\mathbb{P}[a\le W\le b]\le 2d+\frac{1}{\sqrt{2\pi}}(b-a)\le 2d+0.4(b-a),\quad d\le d_\varepsilon+0.4\varepsilon. \qquad (2.14)
$$

By (2.14),

$$
\mathbb{P}[x-0\vee t\le W\le x-0\wedge t+\varepsilon]\le 2d+0.4|t|+0.4\varepsilon.
$$

Using this, (2.13) yields

$$
\begin{aligned}
R_{4,2} &\le \frac{\theta\beta}{2}\int_{|t|\le 1}dt+\frac{2d+0.4\varepsilon}{2\theta\beta}\mathbb{E}\int_{|t|\le 1}\bigl(\hat{K}^{\mathrm{in}}(t)-K^{\mathrm{in}}(t)\bigr)^2\,dt\\
&\quad +\frac{0.4}{2\theta\beta}\mathbb{E}\int_{|t|\le 1}|t|\bigl(\hat{K}^{\mathrm{in}}(t)-K^{\mathrm{in}}(t)\bigr)^2\,dt=\theta\beta+\frac{d+0.2\varepsilon}{\theta\beta}r_4+\frac{0.2}{\theta\beta}r_5^2.
\end{aligned}
$$

By letting $\beta=d+0.2\varepsilon+r_5$, we obtain

$$
R_{4,2}\le\theta(d+0.2\varepsilon+r_5)+\frac{1}{\theta}r_4+\frac{0.2}{\theta}r_5=\theta d+0.2\theta\varepsilon+\frac{1}{\theta}r_4+\Bigl(\theta+\frac{0.2}{\theta}\Bigr)r_5. \qquad (2.15)
$$

By (2.5) again, we have

$$
\begin{aligned}
|R_5| &\le \mathbb{E}\int_{|t|\le 1}\bigl(|W|+1\bigr)\bigl|tK^{\mathrm{in}}(t)\bigr|\,dt \\
&\quad + \frac{1}{\varepsilon}\mathbb{E}\int_{|t|\le 1}\int_{0\wedge t}^{0\vee t}\mathbf{1}[x\le W+u\le x+\varepsilon]\bigl|K^{\mathrm{in}}(t)\bigr|\,du\,dt \\
&\le 2\int_{|t|\le 1}\bigl|tK^{\mathrm{in}}(t)\bigr|\,dt + \frac{1}{\varepsilon}\int_{|t|\le 1}\int_{0\wedge t}^{0\vee t}\mathbb{P}[x\le W+u\le x+\varepsilon]\bigl|K^{\mathrm{in}}(t)\bigr|\,du\,dt \\
&\le 2\int_{|t|\le 1}\bigl|tK^{\mathrm{in}}(t)\bigr|\,dt + \frac{1}{\varepsilon}\int_{|t|\le 1}(2d+0.4\varepsilon)\bigl|tK^{\mathrm{in}}(t)\bigr|\,dt = 2r_2 + \frac{2d+0.4\varepsilon}{\varepsilon}r_2.
\end{aligned}
\tag{2.16}
$$

Letting $\varepsilon = \frac{1}{2}d$ and combining (2.6), (2.7), (2.8), (2.9), (2.10), (2.12), (2.15) and (2.16), we obtain

$$
d_\varepsilon \le 1.1\theta d + r_1 + 6.4r_2 + 3r_3 + \frac{1}{\theta}r_4 + \biggl(2.5+\theta+\frac{0.2}{\theta}\biggr)r_5.
$$

This, together with (2.14), yields

$$
d \le 1.1\theta d + 0.2d + r_1 + 6.4r_2 + 3r_3 + \frac{1}{\theta}r_4 + \biggl(2.5+\theta+\frac{0.2}{\theta}\biggr)r_5,
$$

which implies

$$
d \le (0.8-1.1\theta)^{-1}\biggl\{r_1 + 6.4r_2 + 3r_3 + \frac{1}{\theta}r_4 + \biggl(2.5+\theta+\frac{0.2}{\theta}\biggr)r_5\biggr\}.
$$

Letting $\theta = 0.18$, we obtain

$$
d \le 2r_1 + 11r_2 + 5r_3 + 10r_4 + 7r_5,
$$

and this proves Theorem 2.1. □

REMARK 2.2. We have introduced $\hat{K}^{\mathrm{in}}$ and $\hat{K}^{\mathrm{out}}$ mainly to allow for truncation. Since we have kept the theorem general, different types of truncation are possible, and we will show this in various applications in this article.

EXAMPLE 2.3. We will check the optimality of the bounds in Theorem 2.1 by taking W as a sum of independent random variables. Let $\xi_1,\dots,\xi_n$ be independent with $\mathbb{E}\xi_i = 0$ and $\operatorname{Var}(\xi_i) = \sigma_i^2$, $i = 1,\dots,n$. Define $B^2 = \sum_{i=1}^n \sigma_i^2$, $X_i = \xi_i/B$, $i = 1,\dots,n$, and $W = \sum_{i=1}^n X_i$. Then $\mathbb{E}W = 0$ and $\operatorname{Var}(W) = 1$ and W satisfy the Stein identity (2.1) with

$$
\hat{K}(t) = \sum_{i=1}^n X_i\bigl(\mathbf{1}[-X_i < t \le 0] - \mathbf{1}[-X_i > t > 0]\bigr).
$$

Define

$$
\hat{K}^{\mathrm{in}}(t) = \sum_{i=1}^n X_i\mathbf{1}\bigl[|X_i|\le 1\bigr]\bigl(\mathbf{1}[-X_i < t \le 0] - \mathbf{1}[-X_i > t > 0]\bigr),
$$

$$
\hat{K}^{\mathrm{out}}(t) = \sum_{i=1}^n X_i\mathbf{1}\bigl[|X_i|> 1\bigr]\bigl(\mathbf{1}[-X_i < t \le 0] - \mathbf{1}[-X_i > t > 0]\bigr).
$$

Clearly $\hat{K}(t) = \hat{K}^{\mathrm{in}}(t) + \hat{K}^{\mathrm{out}}(t)$ and $\hat{K}^{\mathrm{in}}(t) = 0$ for $|t| > 1$. Straightforward calculations yield

$$r_1 \le \left(\sum_{i=1}^{n} \mathrm{Var}(X_i^2 \mathbf{1}[|X_i| \le 1])\right)^{1/2} \le \frac{\sqrt{\sum_{i=1}^{n} \mathbb{E}\{\xi_i^4 \mathbf{1}[|\xi_i| \le B]\}}}{B^2};$$

$$r_2 \le \frac{1}{2}\sum_{i=1}^{n} \mathbb{E}\{|X_i|^3 \mathbf{1}[|X_i| \le 1]\} = \frac{\sum_{i=1}^{n} \mathbb{E}\{|\xi_i|^3 \mathbf{1}[|\xi_i| \le B]\}}{2B^3};$$

$$r_3 \le \sum_{i=1}^{n} \mathbb{E}\{X_i^2 \mathbf{1}[|X_i| > 1]\} = \frac{\sum_{i=1}^{n} \mathbb{E}\{\xi_i^2 \mathbf{1}[|\xi_i| > B]\}}{B^2};$$

$$r_4 \le \sum_{i=1}^{n} \mathbb{E}\{|X_i|^3 \mathbf{1}[|X_i| \le 1]\} = \frac{\sum_{i=1}^{n} \mathbb{E}\{|\xi_i|^3 \mathbf{1}[|\xi_i| \le B]\}}{B^3};$$

$$r_5 \le \frac{1}{\sqrt{2}}\left(\sum_{i=1}^{n} \mathbb{E}\{X_i^4 \mathbf{1}[|X_i| \le 1]\}\right)^{1/2} = \frac{\sqrt{\sum_{i=1}^{n} \mathbb{E}\{\xi_i^4 \mathbf{1}[|\xi_i| \le B]\}}}{\sqrt{2}B^2}.$$

Assuming that $\mathbb{E}|\xi_i|^3 < \infty$ for $i = 1, \dots, n$, we obtain

$$d_{\mathrm{K}}(\mathcal{L}(W), \mathcal{N}(0,1)) \le \frac{7\sqrt{\sum_{i=1}^{n} \mathbb{E}\{\xi_i^4 \mathbf{1}[|\xi_i| \le B]\}}}{B^2} + \frac{15.5\sum_{i=1}^{n} \mathbb{E}|\xi_i|^3}{B^3}. \tag{2.17}$$

If both $\sum_{i=1}^{n} \mathbb{E}\{\xi_i^4 \mathbf{1}[|\xi_i| \le B]\}$ and $\sum_{i=1}^{n} \mathbb{E}|\xi_i|^3$ are $\mathrm{O}(B^2)$, such as in the i.i.d. case, then the bound in (2.17) is $\mathrm{O}(B^{-1})$, which agrees with the order of the Berry–Esseen bound $C\sum_{i=1}^{n} \mathbb{E}|\xi_i|^3/B^3$. We expect that in most applications, the bound on the Kolmogorov distance in (2.2) should give the optimal or near optimal order. In particular, for integer-valued random variables, the bounds can never be better than the scaling factor (see the general argument of Englund (1981)), so that, for instance, the bounds in Corollary 5.5 on occupancy problems are optimal whenever the additive functional is integer-valued.

3. Random measures. Let Γ be a locally compact separable metric space. Let Ξ be a random measure on Γ with finite intensity measure Λ, and let Ξ_α be the Palm measure associated with Ξ at $\alpha \in \Gamma$ (see Kallenberg (1983), pages 83, 103). We have

$$\mathbb{E}\left\{\int_\Gamma f(\alpha, \Xi)\Xi(d\alpha)\right\} = \mathbb{E}\left\{\int_\Gamma f(\alpha, \Xi_\alpha)\Lambda(d\alpha)\right\} \tag{3.1}$$

for real-valued functions $f(\cdot,\cdot)$ for which the expectations exist (see Kallenberg (1983), page 84). If Ξ is a simple point process, the distribution of Ξ_α can be interpreted as the conditional distribution of Ξ given that a point of Ξ at α has occurred. On the other hand, if $\Lambda(\{\alpha\}) > 0$, then $\Xi(\{\alpha\})$ is a nonnegative random variable with positive mean and $\Xi_\alpha(\{\alpha\})$ is a $\Xi(\{\alpha\})$-size-biased random variable. Therefore, in general, we may interpret the Palm measure as a "size-biased random measure". For the special case where f is absolutely continuous from $\mathbb{R}$ to $\mathbb{R}$, we obtain

$$\mathbb{E}\{|\Xi| f(|\Xi|)\} = \mathbb{E}\int_\Gamma f(|\Xi_\alpha|)\Lambda(d\alpha), \tag{3.2}$$

provided the expectations and integral exist, where $|\Xi| = \Xi(\Gamma)$. Let $\lambda = \Lambda(\Gamma) = \mathbb{E}|\Xi|$, $B^2 = \mathrm{Var}(|\Xi|)$ and define

$$W = \frac{|\Xi| - \lambda}{B}, \qquad W_\alpha = \frac{|\Xi_\alpha| - \lambda}{B}. \tag{3.3}$$

Assume that Ξ and Ξ_α, $\alpha \in \Gamma$, are defined on the same probability space, and define

$$\Delta_\alpha = W_\alpha - W, \qquad Y_\alpha = |\Xi_\alpha| - |\Xi|.$$

From (3.2), we have

$$\begin{aligned}\mathbb{E}\{Wf(W)\} &= \frac{1}{B}\mathbb{E}\int_\Gamma \big(f(W_\alpha) - f(W)\big)\Lambda(d\alpha) = \frac{1}{B}\mathbb{E}\int_\Gamma \int_0^{\Delta_\alpha} f'(W+t)\,dt\,\Lambda(d\alpha)\\ &= \frac{1}{B}\mathbb{E}\int_\Gamma \int_{-\infty}^{\infty} f'(W+t)\big(\mathbf{1}[\Delta_\alpha > t > 0] - \mathbf{1}[\Delta_\alpha < t \le 0]\big)\,dt\,\Lambda(d\alpha)\\ &= \mathbb{E}\int_{-\infty}^{\infty} f'(W+t)\hat{K}(t)\,dt,\end{aligned} \tag{3.4}$$

where

$$\hat{K}(t) = \frac{1}{B}\int_\Gamma \big(\mathbf{1}[\Delta_\alpha > t > 0] - \mathbf{1}[\Delta_\alpha < t \le 0]\big)\Lambda(d\alpha). \tag{3.5}$$

We now apply Theorem 2.1 to (3.4) to obtain the following theorem.

THEOREM 3.1. *Let W and W_α, $\alpha \in \Gamma$, be as defined in (3.3), and assume that Ξ and Ξ_α are defined on the same probability space. Define $\hat{K}(t)$ as in (3.5), and let*

$$\hat{K}^{\mathrm{in}}(t) = \frac{1}{B}\int_\Gamma \big(\mathbf{1}[\Delta_\alpha > t > 0] - \mathbf{1}[\Delta_\alpha < t \le 0]\big)\mathbf{1}\big[|\Delta_\alpha| \le 1\big]\Lambda(d\alpha),$$

$$\hat{K}^{\mathrm{out}}(t) = \frac{1}{B}\int_\Gamma \big(\mathbf{1}[\Delta_\alpha > t > 0] - \mathbf{1}[\Delta_\alpha < t \le 0]\big)\mathbf{1}\big[|\Delta_\alpha| > 1\big]\Lambda(d\alpha).$$

Moreover, let

$$K(t) = \mathbb{E}\hat{K}(t), \qquad K^{\mathrm{in}}(t) = \mathbb{E}\hat{K}^{\mathrm{in}}(t), \qquad K^{\mathrm{out}}(t) = \mathbb{E}\hat{K}^{\mathrm{out}}(t).$$

Then

$$d_{\mathrm{K}}\big(\mathscr{L}(W), \mathcal{N}(0,1)\big) \le 2r_1' + 5.5r_2' + 5r_3' + 10r_4' + 7r_5',$$

where r_1', r_2', r_3', r_4' and r_5' are given by (3.6), (3.7), (3.8), (3.9) and (3.10) respectively.

PROOF. The proof of this theorem is reduced to calculating the error terms in Theorem 2.1, which yields

$$\begin{aligned} r_1' := r_1 &= \frac{1}{B}\mathbb{E}\bigg|\int_\Gamma \big(\Delta_\alpha \mathbf{1}\big[|\Delta_\alpha| \le 1\big] - \mathbb{E}\big\{\Delta_\alpha \mathbf{1}\big[|\Delta_\alpha| \le 1\big]\big\}\big)\Lambda(d\alpha)\bigg|\\ &= \frac{1}{B^2}\mathbb{E}\bigg|\int_\Gamma \big(Y_\alpha \mathbf{1}\big[|Y_\alpha| \le B\big] - \mathbb{E}\big\{Y_\alpha \mathbf{1}\big[|Y_\alpha| \le B\big]\big\}\big)\Lambda(d\alpha)\bigg|; \end{aligned} \tag{3.6}$$

$$\begin{aligned} 2r_2 &= \frac{1}{B}\int_\Gamma \mathbb{E}\big\{\Delta_\alpha^2 \mathbf{1}\big[|\Delta_\alpha| \le 1\big]\big\}\Lambda(d\alpha)\\ &= \frac{1}{B^3}\int_\Gamma \mathbb{E}\big\{Y_\alpha^2 \mathbf{1}\big[|Y_\alpha| \le B\big]\big\}\Lambda(d\alpha) =: r_2'; \end{aligned} \tag{3.7}$$

$$\begin{aligned} r_3 &= \frac{1}{B}\int_\Gamma \mathbb{E}\big\{|\Delta_\alpha| \mathbf{1}\big[|\Delta_\alpha| > 1\big]\big\}\Lambda(d\alpha)\\ &= \frac{1}{B^2}\int_\Gamma \mathbb{E}\big\{|Y_\alpha| \mathbf{1}\big[|Y_\alpha| > B\big]\big\}\Lambda(d\alpha) =: r_3'; \end{aligned} \tag{3.8}$$

$$
\begin{aligned}
r_4' := r_4 &= \frac{1}{B^2}\int_{|t|\le 1}\int_\Gamma\int_\Gamma \operatorname{Cov}\bigl(\mathbf{1}[1 \ge \Delta_\alpha > t > 0] - \mathbf{1}[-1 \le \Delta_\alpha < t \le 0],\\
&\qquad \mathbf{1}[1 \ge \Delta_\beta > t > 0] - \mathbf{1}[-1 \le \Delta_\beta < t \le 0]\bigr)\Lambda(d\alpha)\Lambda(d\beta)\,dt\\
&= \frac{1}{B^2}\int_0^1\int_\Gamma\int_\Gamma \operatorname{Cov}\bigl(\mathbf{1}[1 \ge \Delta_\alpha > t > 0], \mathbf{1}[1 \ge \Delta_\beta > t > 0]\bigr)\\
&\quad \times \Lambda(d\alpha)\Lambda(d\beta)\,dt\\
&\quad + \frac{1}{B^2}\int_{-1}^0\int_\Gamma\int_\Gamma \operatorname{Cov}\bigl(\mathbf{1}[-1 \le \Delta_\alpha < t < 0], \mathbf{1}[-1 \le \Delta_\beta < t < 0]\bigr)\\
&\quad \times \Lambda(d\alpha)\Lambda(d\beta)\,dt; \qquad (3.9)\\
r_5' := r_5&\\
&= \frac{1}{B}\biggl(\int_0^1\int_\Gamma\int_\Gamma t\operatorname{Cov}\bigl(\mathbf{1}[1 \ge \Delta_\alpha > t > 0], \mathbf{1}[1 \ge \Delta_\beta > t > 0]\bigr)\\
&\quad \times \Lambda(d\alpha)\Lambda(d\beta)\,dt\\
&\quad - \int_{-1}^0\int_\Gamma\int_\Gamma t\operatorname{Cov}\bigl(\mathbf{1}[-1 \le \Delta_\alpha < t < 0], \mathbf{1}[-1 \le \Delta_\beta < t < 0]\bigr)\\
&\quad \times \Lambda(d\alpha)\Lambda(d\beta)\,dt\biggr)^{1/2}. \qquad (3.10)
\end{aligned}
$$

This completes the proof of Theorem 3.1. □

Using the fact that independence implies uncorrelatedness, we have the following corollary.

COROLLARY 3.2. *Let Ξ be a random measure on Γ with finite mean measure Λ such that $\mathbb{E}|\Xi|^4 < \infty$, and set $B^2 = \operatorname{Var}(|\Xi|)$. Assume that Ξ and Ξ_α, $\alpha \in \Gamma$, are defined on the same probability space. Define*

$$W = \frac{|\Xi| - \mathbb{E}|\Xi|}{B}, \qquad W_\alpha = \frac{|\Xi_\alpha| - \mathbb{E}|\Xi|}{B}, \qquad \Delta_\alpha = W_\alpha - W.$$

Assume that there is a set $D \in \mathcal{B}(\Gamma \times \Gamma)$ such that D is symmetric, that is, $\{(x, y) : (y, x) \in D\} = D$, and for all $(\alpha, \beta) \ne D$, Δ_α and Δ_β are independent. Then

$$d_{\mathrm{K}}\bigl(\mathcal{L}(W), \mathcal{N}(0, 1)\bigr) \le 7s_1 + 5.5s_2 + 10s_3,$$

where

$$
\begin{aligned}
s_1 &= \frac{1}{B^2}\biggl(\int_{(\alpha,\beta)\in D}\mathbb{E}\bigl\{Y_\alpha^2\mathbf{1}\bigl[|Y_\alpha| \le B\bigr]\bigr\}\Lambda(d\alpha)\Lambda(d\beta)\biggr)^{\frac{1}{2}};\\
s_2 &= \frac{1}{B^3}\int_\Gamma \mathbb{E}Y_\alpha^2\Lambda(d\alpha);\\
s_3 &= \frac{1}{B^3}\int_{(\alpha,\beta)\in D}\mathbb{E}\bigl\{|Y_\alpha|\mathbf{1}\bigl[|Y_\alpha| \le B\bigr]\bigr\}\Lambda(d\alpha)\Lambda(d\beta).
\end{aligned}
$$

PROOF. By Theorem 3.1, we have

$$
\begin{aligned}
r_1' &\le \frac{1}{B}\biggl(\mathbb{E}\biggl\{\biggl[\int_\Gamma \bigl[\Delta_\alpha \mathbf{1}\bigl[|\Delta_\alpha| \le 1\bigr] - \mathbb{E}\bigl\{\Delta_\alpha \mathbf{1}\bigl[|\Delta_\alpha| \le 1\bigr]\bigr\}\bigr]\Lambda(d\alpha)\biggr]^2\biggr\}\biggr)^{1/2} \\
&= \frac{1}{B}\biggl(\int_{(\alpha,\beta)\in D} \operatorname{Cov}\bigl(\Delta_\alpha \mathbf{1}\bigl[|\Delta_\alpha| \le 1\bigr], \Delta_\beta \mathbf{1}\bigl[|\Delta_\beta| \le 1\bigr]\bigr)\Lambda(d\alpha)\Lambda(d\beta)\biggr)^{\frac{1}{2}} \\
&\le \frac{1}{B}\Bigl(\int_{(\alpha,\beta)\in D} \frac{1}{2}\bigl(\operatorname{Var}\bigl(\Delta_\alpha \mathbf{1}\bigl[|\Delta_\alpha| \le 1\bigr]\bigr) + \operatorname{Var}\bigl(\Delta_\beta \mathbf{1}\bigl[|\Delta_\beta| \le 1\bigr]\bigr)\Lambda(d\alpha)\Lambda(d\beta)\Bigr)^{\frac{1}{2}} \qquad (3.11) \\
&= \frac{1}{B}\biggl(\int_{(\alpha,\beta)\in D} \operatorname{Var}\bigl(\Delta_\alpha \mathbf{1}\bigl[|\Delta_\alpha| \le 1\bigr]\bigr)\Lambda(d\alpha)\Lambda(d\beta)\biggr)^{\frac{1}{2}} \\
&\le \frac{1}{B^2}\biggl(\int_{(\alpha,\beta)\in D} \mathbb{E}\bigl\{Y_\alpha^2 \mathbf{1}\bigl[|Y_\alpha| \le B\bigr]\bigr\}\Lambda(d\alpha)\Lambda(d\beta)\biggr)^{\frac{1}{2}} =: s_1,
\end{aligned}
$$

where the second equality is due to the symmetry of the set D. Next,

$$
\begin{aligned}
r_2' + r_3' &\le \frac{1}{B^3}\int_\Gamma \mathbb{E}Y_\alpha^2 \Lambda(d\alpha) =: s_2, \\
r_4' &= \frac{1}{B^2}\int_0^1 \int_{(\alpha,\beta)\in D} \operatorname{Cov}\bigl(\mathbf{1}[1 \ge \Delta_\alpha > t > 0], \mathbf{1}[1 \ge \Delta_\beta > t > 0]\bigr) \\
&\quad \times \Lambda(d\alpha)\Lambda(d\beta)\,dt \\
&\quad + \frac{1}{B^2}\int_{-1}^0 \int_{(\alpha,\beta)\in D} \operatorname{Cov}\bigl(\mathbf{1}[-1 \le \Delta_\alpha < t < 0], \mathbf{1}[-1 \le \Delta_\beta < t < 0]\bigr) \\
&\quad \times \Lambda(d\alpha)\Lambda(d\beta)\,dt \\
&\le \frac{1}{B^2}\int_{(\alpha,\beta)\in D} \mathbb{E}\bigl\{\min\bigl(|\Delta_\alpha|\mathbf{1}\bigl[|\Delta_\alpha| \le 1\bigr], |\Delta_\beta|\mathbf{1}\bigl[|\Delta_\beta| \le 1\bigr]\bigr)\bigr\}\Lambda(d\alpha)\Lambda(d\beta) \\
&\le \frac{1}{B^3}\int_{(\alpha,\beta)\in D} \mathbb{E}\bigl\{|Y_\alpha|\mathbf{1}\bigl[|Y_\alpha| \le B\bigr]\bigr\}\Lambda(d\alpha)\Lambda(d\beta) =: s_3;
\end{aligned}
$$

and finally,

$$
\begin{aligned}
r_5' &= \frac{1}{B}\biggl(\int_0^1 \int_{(\alpha,\beta)\in D} t \operatorname{Cov}\bigl(\mathbf{1}[1 \ge \Delta_\alpha > t > 0], \mathbf{1}[1 \ge \Delta_\beta > t > 0]\bigr) \\
&\quad \times \Lambda(d\alpha)\Lambda(d\beta)\,dt \\
&\quad - \int_{-1}^0 \int_{(\alpha,\beta)\in D} t \operatorname{Cov}\bigl(\mathbf{1}[-1 \le \Delta_\alpha < t < 0], \mathbf{1}[-1 \le \Delta_\beta < t < 0]\bigr) \\
&\quad \times \Lambda(d\alpha)\Lambda(d\beta)\,dt\biggr)^{1/2} \\
&\le \frac{1}{B}\biggl(\int_0^1 \int_{(\alpha,\beta)\in D} t\mathbb{E}\bigl\{\mathbf{1}[1 \ge \Delta_\alpha > t > 0]\mathbf{1}[1 \ge \Delta_\beta > t > 0]\bigr\} \qquad (3.12) \\
&\quad \times \Lambda(d\alpha)\Lambda(d\beta)\,dt \\
&\quad - \int_{-1}^0 \int_{(\alpha,\beta)\in D} t\mathbb{E}\bigl\{\mathbf{1}[-1 \le \Delta_\alpha < t < 0]\mathbf{1}[-1 \le \Delta_\beta < t < 0]\bigr\} \\
&\quad \times \Lambda(d\alpha)\Lambda(d\beta)\,dt\biggr)^{1/2}
\end{aligned}
$$

$$= \frac{1}{\sqrt{2}B}\Big(\int_{(\alpha,\beta)\in D} \mathbb{E}\{\min(\Delta_\alpha^2 \mathbf{1}[|\Delta_\alpha| \le 1], \Delta_\beta^2 \mathbf{1}[|\Delta_\beta| \le 1])\}\Lambda(d\alpha)\Lambda(\beta)\Big)^{\frac{1}{2}}$$

$$\le \frac{1}{\sqrt{2}B^2}\Big(\int_{(\alpha,\beta)\in D} \mathbb{E}\{Y_\alpha^2 \mathbf{1}[|Y_\alpha| \le B]\}\Lambda(d\alpha)\Lambda(\beta)\Big)^{\frac{1}{2}} = \frac{1}{\sqrt{2}} s_1.$$

The proof of the corollary is completed by combining (3.11) to (3.12). □

4. Applications.

4.1. *Completely random measures.* A random measure Ξ on the carrier space $(\Gamma, \mathcal{B}(\Gamma))$ is said to be *completely random* (see Kingman (1967)) if for any $k \ge 1$ and any pairwise disjoint sets $A_1, \dots, A_k \in \mathcal{B}(\Gamma)$, $\Xi(A_i)$, $1 \le i \le k$, are independent. Well-known examples include the compound Poisson process with cluster distributions on $\mathbb{R}_+ := [0, \infty)$ (see Daley and Vere-Jones (2003), page 198), the Gamma process (see Daley and Vere-Jones (2008), page 11) and the Pólya sum process (see Zessin (2009) and Rafler (2011)). The former two processes cannot in general be represented as an integral of a random field with respect to a point process with finite mean measure, hence they are not covered by the general theory of Barbour and Xia (2006).

THEOREM 4.1. *Let Ξ be a completely random measure with mean measure Λ and finite fourth moment $\mathbb{E}|\Xi|^4 < \infty$. Let $\mu := \mu_\Xi := \Lambda(\Gamma)$, $B^2 := \mathrm{Var}(|\Xi|)$ and $W = (|\Xi| - \mathbb{E}|\Xi|)/B$, then*

$$\begin{aligned} &d_{\mathrm{K}}(\mathscr{L}(W), \mathcal{N}(0,1)) \\ &\quad\le \frac{10}{B^2}\Big(\sum_{\alpha\in\Gamma} \mathbb{E}\{\Xi(\{\alpha\})^3\}\Lambda(\{\alpha\})\Big)^{1/2} + \frac{5.5}{B^3}\mathbb{E}\sum_{\alpha\in\Gamma}\Xi(\{\alpha\})^3 \\ &\qquad + \frac{25.5}{B^3}\sum_{\alpha\in\Gamma}\mathbb{E}\{\Xi(\{\alpha\})^2\}\Lambda(\{\alpha\}) \qquad (4.1) \\ &\quad\le \frac{10}{B^2}\Big(\sum_{\alpha\in\Gamma} \mathbb{E}\{\Xi(\{\alpha\})^3\}\Lambda(\{\alpha\})\Big)^{1/2} + \frac{31}{B^3}\mathbb{E}\sum_{\alpha\in\Gamma}\Xi(\{\alpha\})^3. \qquad (4.2) \end{aligned}$$

REMARK 4.2. (1) If Λ is diffuse at α (i.e., $\Lambda(\{\alpha\}) = 0$), then $\Xi(\{\alpha\}) = 0$ a.s. Hence, if Λ is a diffuse measure, then the bound (4.1) is reduced to

$$d_{\mathrm{K}}(\mathcal{L}(W), \mathcal{N}(0,1)) \le \frac{5.5}{B^3}\mathbb{E}\sum_{\alpha\in\Gamma}\Xi(\{\alpha\})^3 = \frac{5.5}{B^3}\int_\Gamma \mathbb{E}\{\Xi_\alpha(\{\alpha\})^2\}\Lambda(d\alpha).$$

(2) For a simple Poisson point process with $\Lambda(\Gamma) = \lambda$, the bound in (4.2) becomes $31\lambda^{-1/2}$, which compares favourably with those in the literature; see, for example, Lachièze-Rey, Schulte and Yukich (2019).

PROOF OF THEOREM 4.1. Using similar notation as in Corollary 3.2, we have $Y_\alpha = \Xi_\alpha(\{\alpha\}) - \Xi(\{\alpha\})$ and Y_α is independent of Y_β unless $\alpha = \beta$. Hölder's inequality ensures that

$$\mathbb{E}\{\Xi(\{\alpha\})^2\}\Lambda(\{\alpha\}) \le \mathbb{E}\{\Xi(\{\alpha\})^3\}. \qquad (4.3)$$

Hence, direct computation gives

$$\begin{aligned} s_1 &\le \frac{1}{B^2}\Big(\sum_\Gamma \mathbb{E}\{Y_\alpha^2\}\Lambda(\{\alpha\})^2\Big)^{1/2} \\ &\le \frac{1}{B^2}\Big(\sum_\Gamma (\mathbb{E}\{\Xi(\{\alpha\})^2\} + \mathbb{E}\{\Xi_\alpha(\{\alpha\})^2\})\Lambda(\{\alpha\})^2\Big)^{1/2} \\ &= \frac{1}{B^2}\Big(\sum_\Gamma (\mathbb{E}\{\Xi(\{\alpha\})^2\}\Lambda(\{\alpha\})^2 + \mathbb{E}\{\Xi(\{\alpha\})^3\}\Lambda(\{\alpha\}))\Big)^{1/2} \\ &\le \frac{\sqrt{2}}{B^2}\Big(\sum_\Gamma \mathbb{E}\{\Xi(\{\alpha\})^3\}\Lambda(\{\alpha\})\Big)^{1/2}, \end{aligned}$$

where the equality is due to (3.1) and the last inequality follows from (4.3). The same reasoning gives

$$\begin{aligned} s_2 &\le \frac{1}{B^3}\int_\Gamma (\mathbb{E}\{\Xi(\{\alpha\})^2\} + \mathbb{E}\{\Xi_\alpha(\{\alpha\})^2\})\Lambda(d\alpha) \\ &= \frac{1}{B^3}\Big(\sum_{\alpha\in\Gamma} \mathbb{E}\{\Xi(\{\alpha\})^2\}\Lambda(\{\alpha\}) + \mathbb{E}\int_\Gamma \Xi(\{\alpha\})^2\Xi(d\alpha)\Big) \\ &= \frac{1}{B^3}\Big(\sum_{\alpha\in\Gamma} \mathbb{E}\{\Xi(\{\alpha\})^2\}\Lambda(\{\alpha\}) + \mathbb{E}\sum_{\alpha\in\Gamma} \Xi(\{\alpha\})^3\Big), \end{aligned}$$

where the first equality follows from the fact that $\Lambda(\{\alpha\}) = 0$ implies $\Xi(\{\alpha\}) = 0$ a.s. and (3.1). Finally,

$$\begin{aligned} s_3 &\le \frac{1}{B^3}\sum_\Gamma (\mathbb{E}\{\Xi(\{\alpha\})\} + \mathbb{E}\{\Xi_\alpha(\{\alpha\})\})\Lambda(\{\alpha\})^2 \\ &= \frac{1}{B^3}\sum_\Gamma (\mathbb{E}\{\Xi(\{\alpha\})\}\Lambda(\{\alpha\})^2 + \mathbb{E}\{\Xi(\{\alpha\})^2\}\Lambda(\{\alpha\})) \\ &\le \frac{2}{B^3}\sum_\Gamma \mathbb{E}\{\Xi(\{\alpha\})^2\}\Lambda(\{\alpha\}). \end{aligned}$$

Combining these estimates and Corollary 3.2 gives (4.1). (4.2) is an immediate consequence of (4.1) and (4.3). □

COROLLARY 4.3. *Let $\Xi^{(i)}$ for $1 \le i \le n$ be independent random measures on the carrier space $(S, \mathcal{S})$. Define $\Xi = \sum_{i=1}^n \Xi^{(i)}$, $B^2 = \operatorname{Var}(|\Xi|)$ and $W = (|\Xi| - \mathbb{E}|\Xi|)/B$. Then*

$$d_{\mathrm{K}}(\mathscr{L}(W), \mathcal{N}(0,1)) \le \frac{10}{B^2}\Big(\sum_{i=1}^n \mathbb{E}|\Xi^{(i)}|\mathbb{E}|\Xi^{(i)}|^3\Big)^{1/2} + \frac{31}{B^3}\sum_{i=1}^n \mathbb{E}|\Xi^{(i)}|^3.$$

PROOF. Define $\Xi' = \sum_{i=1}^n |\Xi^{(i)}|\delta_i$, where δ_i is the Dirac measure at i, then Ξ' is a completely random measure on the carrier space $\Gamma := \{1, \dots, n\}$ with mean measure $\Lambda'(\{i\}) = \mathbb{E}|\Xi^{(i)}|$, $i \in S'$. We have $B^2 = \operatorname{Var}(|\Xi'|)$, $W = (|\Xi'| - \mathbb{E}|\Xi'|)/B$, hence the claim follows from (4.2). □

4.2. *Excursion random measure.* Let $(S, \mathcal{B}(S))$ be a metric space and $\{X_t, 0 \le t \le T\}$ be an S-valued random process. Define $\mathcal{F}_{a,b} = \sigma\{X_t : a \le t \le b\}$, for $0 \le a \le b \le T$. We say that $\{X_t, 0 \le t \le T\}$ is *l-dependent* with $l > 0$ if $\mathcal{F}_{0,b}$ is independent of $\mathcal{F}_{b+l,T}$ for all $0 \le b < b + l \le T$. We define the excursion random measure

$$\Xi(dt) = \mathbf{1}_E(t, X_t)\,dt, \qquad E \in \mathcal{B}\big([0,T] \times S\big).$$

Define $\mu = \mathbb{E}\Xi([0,T])$, $B = \sqrt{\operatorname{Var}(|\Xi|)}$ and $W = (|\Xi| - \mu)/B$.

The excursion random measure of a stationary process was defined by Hsing and Leadbetter (1998). It was shown by Hsing and Leadbetter (1998) that the asymptotic distribution of the excursion random measure at high levels of exceedances gives a range of useful information about the extremal behavior of the stationary process. Under very general conditions, Hsing and Leadbetter (1998) demonstrated that various asymptotic properties of the excursion random measures can be established. Our Theorem 3.1 can be used to prove the following normal approximation error bound for the total excursion time of l-dependent random processes.

THEOREM 4.4. *For the l-dependent process $\{X_t, 0 \le t \le T\}$, we have*

$$d_{\mathrm{K}}\big(\mathscr{L}(W), \mathcal{N}(0,1)\big) \le \frac{(14\sqrt{2}+8)l^{3/2}\mu^{1/2}}{B^2} + \frac{102l^2\mu}{B^3}. \tag{4.4}$$

PROOF. Write $\Lambda(dt) = \mathbb{E}\Xi(dt)$ and $N_\alpha = [0,T] \cap [\alpha - l, \alpha + l]$ for $\alpha \in [0,T]$. Since $\{X_t, 0 \le t \le T\}$ is l-dependent, we can take Ξ_α such that $Y_\alpha = \Xi_\alpha(N_\alpha) - \Xi(N_\alpha)$ for all $\alpha \in [0,T]$, and Y_α is independent of Y_β for all $|\alpha - \beta| > 2l$. Moreover, we have $|Y_\alpha| \le 2l$. Hence,

$$\begin{aligned}
r_1' &\le \frac{1}{B^2}\bigg(\iint_{|\beta-\alpha|\le 2l} \operatorname{Cov}\big(Y_\alpha \mathbf{1}\big[|Y_\alpha| \le B\big], Y_\beta \mathbf{1}\big[|Y_\beta| \le B\big]\big)\Lambda(d\alpha)\Lambda(d\beta)\bigg)^{1/2} \\
&\le \frac{1}{B^2}\bigg(\iint_{|\beta-\alpha|\le 2l} \frac{1}{2}\big[\operatorname{Var}\big(Y_\alpha \mathbf{1}\big[|Y_\alpha| \le B\big]\big) + \operatorname{Var}\big(Y_\beta \mathbf{1}\big[|Y_\beta| \le B\big]\big)\big]\Lambda(d\alpha)\Lambda(d\beta)\bigg)^{1/2} \\
&= \frac{1}{B^2}\bigg(\iint_{|\beta-\alpha|\le 2l} \operatorname{Var}\big(Y_\alpha \mathbf{1}\big[|Y_\alpha| \le B\big]\big)\Lambda(d\alpha)\Lambda(d\beta)\bigg)^{1/2} \\
&\le \frac{1}{B^2}\bigg(\iint_{|\beta-\alpha|\le 2l} \mathbb{E}\big(Y_\alpha^2\big)\Lambda(d\alpha)\Lambda(d\beta)\bigg)^{1/2} \le \frac{4}{B^2}\sqrt{l^3\mu}.
\end{aligned} \tag{4.5}$$

Similarly,

$$\begin{aligned}
r_2' + r_3' &= \frac{1}{B^3}\int_0^T \mathbb{E}\big(Y_\alpha^2\big)\Lambda(d\alpha) \le 4l^2\mu/B^3, \\
r_4' &= \frac{1}{B^2}\int_0^1 \iint_{|\beta-\alpha|\le 2l} \operatorname{Cov}\big(\mathbf{1}[1 \ge \Delta_\alpha > t > 0], \mathbf{1}[1 \ge \Delta_\beta > t > 0]\big) \\
&\qquad \times \Lambda(d\alpha)\Lambda(d\beta)\,dt \\
&\qquad + \frac{1}{B^2}\int_{-1}^0 \iint_{|\beta-\alpha|\le 2l} \operatorname{Cov}\big(\mathbf{1}[-1 \le \Delta_\alpha < t < 0], \mathbf{1}[-1 \le \Delta_\beta < t < 0]\big) \\
&\qquad \times \Lambda(d\alpha)\Lambda(d\beta)\,dt \\
&\le \frac{1}{B^2}\int_0^1 \iint_{|\beta-\alpha|\le 2l} \mathbb{E}\mathbf{1}[1 \ge \Delta_\alpha > t > 0]\Lambda(d\alpha)\Lambda(d\beta)\,dt
\end{aligned}$$

$$+\frac{1}{B^2}\int_{-1}^{0}\iint_{|\beta-\alpha|\le 2l}\mathbb{E}\mathbf{1}[-1\le\Delta_\alpha<t<0]\Lambda(d\alpha)\Lambda(d\beta)\,dt$$
$$=\frac{1}{B^3}\iint_{|\beta-\alpha|\le 2l}\mathbb{E}|Y_\alpha|\Lambda(d\alpha)\Lambda(d\beta)\le\frac{8l^2}{B^3}\mu$$

and

$$\begin{aligned}r_5'=\frac{1}{B}\Big(&\int_0^1\iint_{|\beta-\alpha|\le 2l}t\operatorname{Cov}\big(\mathbf{1}[1\ge\Delta_\alpha>t>0],\mathbf{1}[1\ge\Delta_\beta>t>0]\big)\\&\times\Lambda(d\alpha)\Lambda(d\beta)\,dt\\&-\int_{-1}^0\iint_{|\beta-\alpha|\le 2l}t\operatorname{Cov}\big(\mathbf{1}[-1\le\Delta_\alpha<t<0],\mathbf{1}[-1\le\Delta_\beta<t<0]\big)\\&\times\Lambda(d\alpha)\Lambda(d\beta)\,dt\Big)^{1/2}\\\le\frac{1}{B}\Big\{&\int_0^1\iint_{|\beta-\alpha|\le 2l}t\mathbb{E}\mathbf{1}[1\ge\Delta_\alpha>t>0]\Lambda(d\alpha)\Lambda(d\beta)\,dt\\&-\int_{-1}^0\iint_{|\beta-\alpha|\le 2l}t\mathbb{E}\mathbf{1}[-1\le\Delta_\alpha<t<0]\Lambda(d\alpha)\Lambda(d\beta)\,dt\Big\}^{1/2}\\\le\frac{1}{B}\Big\{&\frac{1}{2}\iint_{|\beta-\alpha|\le 2l}\mathbb{E}\{\Delta_\alpha^2\}\Lambda(d\alpha)\Lambda(d\beta)\,dt\Big\}^{1/2}\le\frac{2\sqrt{2}l^{3/2}\mu^{1/2}}{B^2}.\end{aligned}\tag{4.6}$$

Finally, we have from Theorem 3.1 that

$$d_{\mathrm{K}}\big(\mathscr{L}(W),\mathcal{N}(0,1)\big)\le 2r_1'+5.5\big(r_2'+r_3'\big)+10r_4'+7r_5',$$

so collecting (4.5) to (4.6), we obtain (4.4). □

COROLLARY 4.5. *Let I_i, $1\le i\le n$, be independent indicator random variables such that $\mathbb{P}[I_i=1]=p_i$. Let $S_n=\sum_{i=1}^{n-k+1}\prod_{j=i}^{i+k-1}I_j$, the number of k-runs in the sequence. Define $\mu_n=\mathbb{E}S_n=\sum_{i=1}^{n-k+1}\prod_{j=i}^{i+k-1}p_j$, $B_n=\sqrt{\operatorname{Var}(S_n)}$, $W=(S_n-\mu_n)/B_n$, then*

$$d_{\mathrm{K}}\big(\mathscr{L}(W),\mathcal{N}(0,1)\big)\le\frac{(14\sqrt{2}+8)k^{3/2}\mu_n^{1/2}}{B_n^2}+\frac{102k^2\mu_n}{B_n^3}.$$

In particular, if $p_i=p\in(0,1)$ for all i, then

$$d_{\mathrm{K}}\big(\mathscr{L}(W),\mathcal{N}(0,1)\big)=\mathrm{O}\big(n^{-1/2}\big).$$

PROOF. We define $X_t:=X_{\lfloor t\rfloor}=\prod_{j=\lfloor t\rfloor}^{\lfloor t\rfloor+k-1}I_j$, $0\le t<n-k+2$. Let $E=[1,n-k+2)\times\{1\}$, $\Xi(dt)=\mathbf{1}_E(t,X_t)\,dt$, then the claim follows from (4.4) with $l=k$. □

4.3. *The total edge length of Ginibre–Voronoi tessellations.* The Ginibre point process (see Ginibre (1965), Soshnikov (2000), Mehta (1991) and Goldman (2010)) has attracted considerable attention recently because of its wide use in modeling mobile networks (see Torrisi and Leonardi (2014), Miyoshi and Shirai (2014a), Miyoshi and Shirai (2014b) and Keeler, Ross and Xia (2018)). The Ginibre point process is a special class of the Gibbs point process family and it exhibits a repulsion between the points. The repulsive character makes the cells more regular than those coming from a Poisson point process, hence in applications

the Ginibre–Voronoi tessellation often fits better than the Poisson–Voronoi tessellation (see Rider (2004), Le Caër and Ho (1990) and Goldman (2010)).

The Ginibre point process is defined through the factorial moment measures. For a locally finite process $\boldsymbol{Y}$ on a Polish space S, the nth order *factorial moment measure* $\nu^{(n)}$ of $\boldsymbol{Y}$ is defined by the relation (see Kallenberg (1983), pages 109–110)

$$\mathbb{E}\left[\int_{S^n} f(y_1,\dots,y_n)\boldsymbol{Y}(dy_1)(\boldsymbol{Y}-\delta_{y_1})(dy_2)\dots\left(\boldsymbol{Y}-\sum_{i=1}^{n-1}\delta_{y_i}\right)(dy_n)\right]$$
$$=\int_{S^n} f(y_1,\dots,y_n)\nu^{(n)}(dy_1,\dots,dy_n),$$

where f ranges over all Borel measurable functions $h: S^n \to [0,\infty)$. The Ginibre point process on the complex plane $\mathbb{C}$ is defined as follows.

DEFINITION 4.6. We say the point process $\boldsymbol{X}$ on the complex plane $\mathbb{C}$ ($\cong \mathbb{R}^2$) is the Ginibre point process if its factorial moment measures are given by

$$\nu^{(n)}(dx_1,\dots,dx_n)=\rho^{(n)}(x_1,\dots,x_n)\,dx_1\dots dx_n,\quad n\ge 1,$$

where $\rho^{(n)}(x_1,\dots,x_n)$ is the determinant of the $n\times n$ matrix with (i,j)th entry

$$K(x_i,x_j)=\frac{1}{\pi}e^{-\frac{1}{2}(|x_i|^2+|x_j|^2)}e^{x_i\bar{x}_j}.$$

Here and in the sequel, $\bar{x}$ and $|x|$ are the complex conjugate and modulus of x.

The Ginibre point process has the mean measure $\boldsymbol{\mu}(dx)=\frac{1}{\pi}dx$. Goldman (2010) stated that the Palm process $\boldsymbol{X}_x$ of the Ginibre point process $\boldsymbol{X}$ at the location x satisfies

$$\boldsymbol{X}\stackrel{d}{=}(\boldsymbol{X}_x\setminus\{x\})\cup\{x+Z\}, \tag{4.7}$$

where $\stackrel{d}{=}$ stands for "equals in distribution", $Z=(Z_1,Z_2\sqrt{-1})$ with (Z_1,Z_2) having bivariate normal on $\mathbb{R}^2$ with mean $(0,0)$ and covariance matrix $\left[\begin{smallmatrix}1/2 & 0\\ 0 & 1/2\end{smallmatrix}\right]$. That is, the Palm process $\boldsymbol{X}_x$ can be obtained by removing a point from the process which is Gaussian distributed from x and then adding x to $\boldsymbol{X}$. It is still an open problem to know how Z is correlated with $\boldsymbol{X}_x\setminus\{x\}$ (see Goldman (2010), Problem 2, page 27).

As Schreiber and Yukich (2013) (see also Xia and Yukich (2015)), we consider the window $Q_\lambda:=\{(s_1,s_2\sqrt{-1}): -0.5\sqrt{\lambda}\le s_1,s_2\le 0.5\sqrt{\lambda}\}\subset\mathbb{C}$. For a realization $\boldsymbol{x}$ of $\boldsymbol{X}$ and $x\in\boldsymbol{x}$, let $\mathcal{C}(x,\boldsymbol{x})$ be the set of every point in $\mathbb{C}$ whose (Euclidean) distance to x is less than or equal to its distance to any other point of $\boldsymbol{x}$. The set $\mathcal{C}(x,\boldsymbol{x})$ is called the *Voronoi cell* centered at x and the collection of $\mathcal{C}(x,\boldsymbol{x})$, $x\in\boldsymbol{x}$, is called the *Voronoi tessellation* induced by $\boldsymbol{x}$.

Note that when the Voronoi cell centers are close to the boundary of Q_λ, our definition of Voronoi cells is slightly different from that of Penrose and Yukich (2001) (see also Baryshnikov and Yukich (2005), Schreiber and Yukich (2013) and Xia and Yukich (2015)). This is because the Voronoi cells defined by Penrose and Yukich (2001) do not satisfy the translation invariant property which is a crucial condition for obtaining the central limit theorems of the Voronoi tessellation statistics.[1]

If we define the random measure

$$\Xi(dx)=L(x,\boldsymbol{X})\boldsymbol{X}(dx),$$

[1]This minor issue was noted by Penrose (2007), and we thank J. Yukich for bringing this to our attention.

where $L(x, \boldsymbol{X}) := L_\lambda(x, \boldsymbol{X})$ is one half the total edge length of the *finite* length edges (hence we exclude all infinite edges) in the cell $\mathcal{C}(x, \boldsymbol{X})$, then the total edge length of the Ginibre–Voronoi tessellation induced by $\boldsymbol{X}$ with centers in $\boldsymbol{X} \cap Q_\lambda$ can be written as

$$\mathcal{L}(\lambda) := |\Xi| = \int_{Q_\lambda} L(x, \boldsymbol{X}) \boldsymbol{X}(dx).$$

THEOREM 4.7. *Let* $B^2 = \operatorname{Var}(\mathcal{L}(\lambda))$ *and* $W = (\mathcal{L}(\lambda) - \mathbb{E}\mathcal{L}(\lambda))/B$. *We have*

$$\lim_{\lambda\to\infty} \lambda^{-1}\mathbb{E}\mathcal{L}(\lambda) \in (0, \infty), \qquad \lim_{\lambda\to\infty} \lambda^{-1} B^2 \in (0, \infty) \tag{4.8}$$

and

$$d_{\mathrm{K}}\big(\mathscr{L}(W), \mathcal{N}(0, 1)\big) = \mathrm{O}\big(\lambda^{-1/2} \ln \lambda\big). \tag{4.9}$$

REMARK 4.8. The Ginibre–Voronoi tessellation is a special case of the Gibbs–Voronoi tessellations studied by Xia and Yukich (2015). Theorem 2.3 of Xia and Yukich (2015) gives

$$d_{\mathrm{K}}\big(\mathscr{L}(W), \mathcal{N}(0, 1)\big) = \mathrm{O}\big(\lambda^{-1/2} (\ln \lambda)^4\big),$$

which is slightly worse than (4.9).

REMARK 4.9. The error estimate for the total edge length of the Ginibre–Voronoi tessellation is also valid for the more general class of α-Ginibre point processes with $0 < \alpha < 1$. As a matter of fact, an α-Ginibre point process can be constructed by "deleting, independently and with probability $1 - \alpha$, each point of the Ginibre point process and then applying the homothety of ratio $\sqrt{\alpha}$ to the remaining points in order to restore the intensity of the process" (Goldman (2010)). Hence, for α-Ginibre Voronoi tessellations, except notational complexity, our proof goes through without any difficulty.

REMARK 4.10. We do not know if the bound of (4.9) is of the correct order.

To prove Theorem 4.7, we need the following lemmas. We note that the estimate of the void probability (4.10), albeit very simple, is not new and better estimates were given by Błaszczyszyn, Yogeshwaran and Yukich (2019), Lemma 1.7 in the Supplement.

LEMMA 4.11. *For* $A \subset \mathbb{C}$ *with the area* $|A|$, *we have*

$$\mathbb{P}\big[\boldsymbol{X}(A) = 0\big] \le e^{-|A|/\pi} \tag{4.10}$$

and

$$\mathbb{E}\big[\boldsymbol{X}(A)^l\big] \le 2^{\frac{1}{2}l(l-1)} \big[1 \vee \big(|A|/\pi\big)\big]^l \tag{4.11}$$

for all $l \in \mathbb{N} := \{1, 2, \dots\}$.

PROOF. For $\theta \in (0, 1)$, let $g(\theta) = \mathbb{E}\{(1-\theta)^{\mathbf{X}(A)}\}$, then

$$g'(\theta) = -\mathbb{E}\big\{(1-\theta)^{\mathbf{X}(A)-1}\mathbf{X}(A)\big\} = -\int_A \mathbb{E}\big\{(1-\theta)^{\mathbf{X}_x(A)-1}\big\}\boldsymbol{\mu}(dx). \tag{4.12}$$

Using (4.7), we can construct $\boldsymbol{X}$ and $\boldsymbol{X}_x$ together such that $\boldsymbol{X}_x(A) \le \boldsymbol{X}(A) + 1$ a.s. Hence, it follows from (4.12) that

$$g'(\theta) \le -\int_A \mathbb{E}\big\{(1-\theta)^{\mathbf{X}(A)}\big\}\boldsymbol{\mu}(dx) = -g(\theta)\boldsymbol{\mu}(A).$$

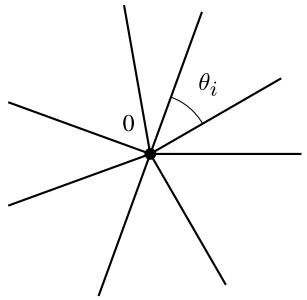

FIG. 1. *Isosceles triangles.*

However, $g(0) = 1$, we obtain $g(\theta) \le e^{-\theta \boldsymbol{\mu}(A)}$, which implies

$$\mathbb{P}[\boldsymbol{X}(A) = 0] = g(1) \le e^{-\boldsymbol{\mu}(A)} = e^{-|A|/\pi},$$

as claimed in (4.10). In terms of (4.11), we can use the construction $\boldsymbol{X}_x(A) \le \boldsymbol{X}(A) + 1$ a.s. again and the inequality $(a+b)^{l-1} \le 2^{l-2}(a^{l-1} + b^{l-1})$ for all $a, b \ge 0$ to obtain

$$\begin{aligned}\mathbb{E}[\boldsymbol{X}(A)^l] &= \mathbb{E}\int_A \boldsymbol{X}(A)^{l-1}\boldsymbol{X}(dx) = \int_A \mathbb{E}[\boldsymbol{X}_x(A)^{l-1}]\boldsymbol{\mu}(dx)\\ &\le \int_A \mathbb{E}[(\boldsymbol{X}(A)+1)^{l-1}]\boldsymbol{\mu}(dx) = \boldsymbol{\mu}(A)\mathbb{E}[(\boldsymbol{X}(A)+1)^{l-1}]\\ &\le 2^{l-2}\boldsymbol{\mu}(A)\{\mathbb{E}[\boldsymbol{X}(A)^{l-1}]+1\} \le 2^{l-1}\boldsymbol{\mu}(A)[1 \vee \mathbb{E}(\boldsymbol{X}(A)^{l-1})].\end{aligned}$$

Hence, (4.11) follows by induction. □

LEMMA 4.12. *Suppose that rays emanating from the center* 0 *of* $\mathbb{C}$ *divide* $\mathbb{C}$ *into disjoint congruent isosceles triangles* A_i *with angles* θ_i, $i = 1, \dots, k$ (*see Figure* 1), *where* k *may be finite or infinity. If* $\boldsymbol{x} \subset \mathbb{C}$ *satisfies* $\boldsymbol{x} \cap A_i \ne \varnothing$ *and* $\theta_i \le \pi/3$ *for all* $i = 1, \dots, k$, *then the Voronoi cell* $\mathcal{C}(0, \boldsymbol{x})$ *is contained in the disk* $B(0, d(\boldsymbol{x}))$, *where* $B(x, r) := \{u \in \mathbb{C} : |u - x| \le r\}$ *and* $d(\boldsymbol{x}) := \sup\{|u| : u \in \boldsymbol{x}\}$.

PROOF. For each $y \in \mathcal{C}(0, \boldsymbol{x})$, there exists a triangle A_i such that $y \in A_i$. Since $\boldsymbol{x} \cap A_i \ne \varnothing$, there exists a point $v \in \boldsymbol{x} \cap A_i$ (see Figure 2) and it follows from $y \in \mathcal{C}(0, \boldsymbol{x})$ that $|y| \le |y - v|$. This in turn implies $|v| \ge 2|y|\cos(\theta) \ge |y|$, that is, $y \in B(0, d(\boldsymbol{x})) := \{u : |u| \le d(\boldsymbol{x})\}$. This completes the proof. □

PROOF OF THEOREM 4.7. From the definition of the Ginibre point process, it is a Gibbs point process with a pair potential function (Osada and Shirai (2016)), hence (4.8) are direct corollaries of Theorem 2.1 of Schreiber and Yukich (2013) and Theorem 1.1 of Xia and Yukich (2015). Hence, it remains to show (4.9).

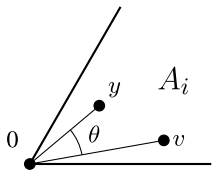

FIG. 2. $\theta \le \pi/3$.

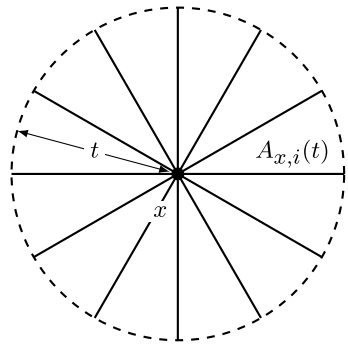

FIG. 3. *Twelve disjoint congruent equal sectors.*

We divide the disk $B^o(x,t) := \{u : 0 < |u-x| < t\}$ into twelve disjoint congruent equal sectors $A_{x,i}(t)$, $i = 1, \dots, 12$, (see Figure 3) and define

$$T_x := \inf\{t : A_{x,i}(t) \cap \boldsymbol{X} \neq \varnothing, i = 1, \dots, 12\},$$

(see McGivney and Yukich (1999) and Penrose and Yukich (2001)). The area of $A_{x,i}(t)$ is $|A_{x,i}(t)| = \pi t^2/12$, so it follows from Lemma 4.11 that

$$\begin{aligned} \mathbb{P}[T_x > t] &= \mathbb{P}\left(\bigcup_{i=1}^{12}\{\boldsymbol{X}(A_{x,i}(t)) = 0\}\right) \\ &\le 12\mathbb{P}[\boldsymbol{X}(A_{x,1}(t)) = 0] \le 12e^{-t^2/12}. \end{aligned} \tag{4.13}$$

Write $Y_x := |\Xi_x| - |\Xi|$, we establish that, for $k \in \mathbb{N}$,

$$\mathbb{E}|Y_x|^k \le C(k), \tag{4.14}$$

where $C(k)$ is a constant dependent on k. In fact, for a bounded measurable function f on the space of all locally finite measures on $\mathbb{C}$, a routine exercise of random measures gives

$$\begin{aligned} \mathbb{E}f(\Xi_x) &= \frac{\mathbb{E}\{f(\sum_{y\in\boldsymbol{X}\cap Q_\lambda} L(y,\boldsymbol{X}))L(x,\boldsymbol{X})\boldsymbol{X}(dx)\}}{\mathbb{E}\{L(x,\boldsymbol{X})\boldsymbol{X}(dx)\}} \\ &= \frac{\mathbb{E}\{f(\sum_{y\in\boldsymbol{X}_x\cap Q_\lambda} L(y,\boldsymbol{X}_x))L(x,\boldsymbol{X}_x)\}}{\mathbb{E}L(x,\boldsymbol{X}_x)}. \end{aligned}$$

That is, if the edges of the Voronoi tessellations are not affected by moving the point from $x+Z$ to x, then their distribution is not affected either. On the other hand, adding a point at x does not affect the Voronoi cells centered at points outside $B(x,3T_x)$ and deleting a point at $x+Z$ does not affect the Voronoi cells centered at points outside $B(x+Z,3T_{x+Z})$ (see McGivney and Yukich (1999), Section 4). Therefore, by Lemma 4.12, the change of Voronoi edge lengths due to shifting a point at $x+Z$ to x is bounded by $2\pi[T_x\boldsymbol{X}(B(x,3T_x)) + T_{Z+x}\boldsymbol{X}(B(x+Z,3T_{x+Z}))]$, so

$$\begin{aligned} \mathbb{E}|Y_x|^k &\le (2\pi)^k\mathbb{E}\{\big(T_x\mathbf{X}\big(B(x,3T_x)\big) + T_{Z+x}\mathbf{X}\big(B(x+Z,3T_{x+Z})\big)\big)^k\} \\ &\le 0.5(4\pi)^k\mathbb{E}\{T_x^k\boldsymbol{X}\big(B(x,3T_x)\big)^k + T_{Z+x}^k\boldsymbol{X}\big(B(x+Z,3T_{x+Z})\big)^k\} \\ &= (4\pi)^k\mathbb{E}\{T_0^k\boldsymbol{X}\big(B(0,3T_0)\big)^k\}, \end{aligned} \tag{4.15}$$

where the second inequality follows from the fact that $(a+b)^k \le 2^{k-1}(a^k+b^k)$ for all $a, b \ge 0$ and the equality holds because $T_{Z+x}\boldsymbol{X}(B(x+Z,3T_{x+Z}))$ and $T_x\boldsymbol{X}(B(x,3T_x))$ have

the same distribution as that of $T_0\boldsymbol{X}(B(0,3T_0))$. However,

$$\begin{aligned}
&\mathbb{E}\{T_0^k\boldsymbol{X}\big(B(0,3T_0)\big)^k\}\\
&\quad\le 12\mathbb{E}\int_{\mathbb{C}}|y|^k\boldsymbol{X}\big(B\big(0,3|y|\big)\big)^k\\
&\qquad\times\mathbf{1}\Bigg[\{\boldsymbol{X}\big(A_{0,1}\big(|y|\big)\big)=0\}\cap\bigcap_{i=2}^{12}\{\boldsymbol{X}\big(A_{0,i}\big(|y|\big)\big)\ge 1\}\Bigg]\boldsymbol{X}(dy)\\
&\quad\le 12\mathbb{E}\int_{\mathbb{C}}|y|^k\boldsymbol{X}\big(B\big(0,3|y|\big)\big)^k\mathbf{1}\big[\boldsymbol{X}\big(A_{0,1}\big(|y|\big)\big)=0\big]\boldsymbol{X}(dy).
\end{aligned}\tag{4.16}$$

We apply the Georgii–Nguyen–Zessin integral characterization of Gibbs point processes (see Møller and Waagepetersen (2004), Chapter 6.4) to obtain that the conditional probability of observing an extra point of $\boldsymbol{X}$ in the volume element dy, given that configuration without that point, equals $\pi^{-1}\exp(-\beta\Delta^{\Psi}(\{y\},\boldsymbol{X}))\,dy\le\pi^{-1}\,dy$, where $\Delta^{\Psi}(\{y\},\boldsymbol{X})\ge 0$ is the local energy function and $1/\beta\ge 0$ is the temperature (see Xia and Yukich (2015), Section 1.1). Hence, it follows from (4.16) that

$$\begin{aligned}
&\mathbb{E}\{T_0^k\boldsymbol{X}\big(B(0,3T_0)\big)^k\}\\
&\quad\le\frac{12}{\pi}\mathbb{E}\int_{\mathbb{C}}|y|^k(\boldsymbol{X}+\delta_y)\big(B\big(0,3|y|\big)\big)^k\\
&\qquad\times\mathbf{1}\big[(\boldsymbol{X}+\delta_y)\big(A_{0,1}\big(|y|\big)\big)=0\big]\exp\big(-\beta\Delta^{\Psi}\big(\{y\},\boldsymbol{X}\big)\big)\,dy\\
&\quad\le\frac{12}{\pi}\mathbb{E}\int_{\mathbb{C}}|y|^k(\boldsymbol{X}+\delta_y)\big(B\big(0,3|y|\big)\big)^k\mathbf{1}\big[(\boldsymbol{X}+\delta_y)\big(A_{0,1}\big(|y|\big)\big)=0\big]\,dy\\
&\quad=24\mathbb{E}\int_0^{\infty}t^k(\boldsymbol{X}+\delta_y)\big(B(0,3t)\big)^k\mathbf{1}\big[\boldsymbol{X}\big(A_{0,1}(t)\big)=0\big]\,dt\\
&\quad\le 24\cdot 2^{k-1}\mathbb{E}\int_0^{\infty}t^k\{\boldsymbol{X}\big(B(0,3t)\big)^k+1\}\mathbf{1}\big[\boldsymbol{X}\big(A_{0,1}(t)\big)=0\big]\,dt\\
&\quad\le 24\cdot 2^{k-1}\int_0^{\infty}t^k\sqrt{\mathbb{E}\{\mathbf{X}\big(B(0,3t)\big)^{2k}\}\mathbb{P}\big[\boldsymbol{X}\big(A_{0,1}(t)\big)=0\big]}\,dt\\
&\qquad+24\cdot 2^{k-1}\int_0^{\infty}t^k\mathbb{P}\big[\boldsymbol{X}\big(A_{0,1}(t)\big)=0\big]\,dt\\
&\quad\le 12\times 2^{0.5k(2k+1)}\int_0^{\infty}\max\{1,9^kt^{2k}\}t^ke^{-t^2/24}\,dt+24\cdot 2^{k-1}\int_0^{\infty}t^ke^{-t^2/12}\,dt,
\end{aligned}$$

where the last inequality follows from (4.10). This, together with (4.15), yields the bound in (4.14).

We now apply Theorem 3.1 to establish (4.9).

The estimate of r_1'. To simplify the notation, we write $Y_x'=Y_x\mathbf{1}[|Y_x|\le B]$. Set

$$U_x=\{T_x\le 4\sqrt{3\ln\lambda}\}\cap\{T_{Z+x}\le 4\sqrt{3\ln\lambda}\}\cap\{|Z|\le 2\sqrt{3\ln\lambda}\}.$$

Using (4.13), we have

$$\begin{aligned}
\mathbb{P}\big[U_x^c\big]=\mathbb{P}\big[U_y^c\big]&\le\mathbb{P}[T_x>4\sqrt{3\ln\lambda}]+\mathbb{P}[T_{x+Z}>4\sqrt{3\ln\lambda}]+\mathbb{P}\big[|Z|>2\sqrt{3\ln\lambda}\big]\\
&=\mathrm{O}(\lambda^{-4}).
\end{aligned}\tag{4.17}$$

This, together with (4.14), ensures

$$\begin{aligned}\operatorname{Cov}(Y'_x\mathbf{1}[U_x^c], Y'_y) &\le \sqrt{\operatorname{Var}(Y'_x\mathbf{1}[U_x^c])\operatorname{Var}(Y'_y)} \le \sqrt{\mathbb{E}\{(Y'_x)^2\mathbf{1}[U_x^c]\}\mathbb{E}\{(Y'_y)^2\}}\\ &\le \sqrt{\mathbb{E}\{Y_x^2\mathbf{1}[U_x^c]\}\mathbb{E}Y_y^2} \le \sqrt{\sqrt{\mathbb{E}Y_x^4\mathbb{P}[U_x^c]}\mathbb{E}Y_y^2} = \mathrm{O}(\lambda^{-1}).\end{aligned} \tag{4.18}$$

Similarly, we can derive

$$\operatorname{Cov}(Y'_x\mathbf{1}[U_x], Y'_y\mathbf{1}[U_y^c]) = \mathrm{O}(\lambda^{-1}) \tag{4.19}$$

and

$$\begin{aligned}\operatorname{Cov}(Y'_x\mathbf{1}[U_x], Y'_y\mathbf{1}[U_y]) &\le \sqrt{\operatorname{Var}(Y'_x\mathbf{1}[U_x])\operatorname{Var}(Y'_y\mathbf{1}[U_y])}\\ &\le \sqrt{\mathbb{E}Y_x^2\mathbb{E}Y_y^2} = \mathrm{O}(1).\end{aligned} \tag{4.20}$$

Assume $|x-y| > 20\sqrt{3\ln\lambda}$. Conditional on U_xU_y, Y_x is independent of Y_y, hence

$$\mathbb{E}(Y'_xY'_y|U_xU_y) = \mathbb{E}(Y'_x|U_xU_y)\mathbb{E}(Y'_y|U_xU_y).$$

Using (4.14) and (4.17), we obtain

$$|\mathbb{E}(Y'_x\mathbf{1}[U_xU_y^c])| \le \{\mathbb{E}[(Y'_x\mathbf{1}[U_x])^2]\mathbb{P}[U_y^c]\}^{1/2} \le \{\mathbb{E}[Y_x^2]\mathbb{P}[U_y^c]\}^{1/2} = \mathrm{O}(\lambda^{-2}).$$

The same argument gives that all $\mathbb{E}(Y'_x\mathbf{1}[U_xU_y])$, $\mathbb{E}(Y'_y\mathbf{1}[U_xU_y])$ and $\mathbb{E}(Y'_x\mathbf{1}[U_x])$ are of order $\mathrm{O}(1)$ and $\mathbb{E}(Y'_y\mathbf{1}[U_x^cU_y]) = \mathrm{O}(\lambda^{-2})$, hence

$$\begin{aligned}&\operatorname{Cov}(Y'_x\mathbf{1}[U_x], Y'_y\mathbf{1}[U_y])\\ &\quad= \mathbb{E}(Y'_x|U_xU_y)\mathbb{E}(Y'_y|U_xU_y)\mathbb{P}[U_xU_y] - \mathbb{E}(Y'_x\mathbf{1}[U_xU_y])\mathbb{E}(Y'_y\mathbf{1}[U_xU_y])\\ &\qquad- \mathbb{E}(Y'_x\mathbf{1}[U_xU_y^c])\mathbb{E}(Y'_y\mathbf{1}[U_xU_y]) - \mathbb{E}(Y'_x\mathbf{1}[U_x])\mathbb{E}(Y'_y\mathbf{1}[U_x^cU_y])\\ &\quad= \mathbb{E}(Y'_x\mathbf{1}[U_xU_y])\mathbb{E}(Y'_y\mathbf{1}[U_xU_y])\big[\mathbb{P}[U_xU_y]\big]^{-1}(1-\mathbb{P}[U_xU_y])\\ &\qquad+ \mathrm{O}(\lambda^{-2}) + \mathrm{O}(\lambda^{-2})\\ &\quad= \mathrm{O}(\lambda^{-2}),\end{aligned} \tag{4.21}$$

where the last equation follows from (4.17) since $1-\mathbb{P}[U_xU_y] \le \mathbb{P}[U_x^c] + \mathbb{P}[U_y^c] = o(\lambda^{-4})$. Now,

$$\begin{aligned}\iint_{\Gamma^2}\operatorname{Cov}(Y'_x, Y'_y)\Lambda(dx)\Lambda(dy) &= \iint_{\Gamma^2}\operatorname{Cov}(Y'_x\mathbf{1}[U_x^c], Y'_y)\Lambda(dx)\Lambda(dy)\\ &\quad+ \iint_{\Gamma^2}\operatorname{Cov}(Y'_x\mathbf{1}[U_x], Y'_y\mathbf{1}[U_y^c])\Lambda(dx)\Lambda(dy)\\ &\quad+ \iint_{|x-y|>20\sqrt{3\ln\lambda}}\operatorname{Cov}(Y'_x\mathbf{1}[U_x], Y'_y\mathbf{1}[U_y])\Lambda(dx)\Lambda(dy)\\ &\quad+ \iint_{|x-y|\le 20\sqrt{3\ln\lambda}}\operatorname{Cov}(Y'_x\mathbf{1}[U_x], Y'_y\mathbf{1}[U_y])\Lambda(dx)\Lambda(dy).\end{aligned}$$

Using (4.18) for the first term, (4.19) for the second term, (4.21) for the third term and (4.20) for the last term, we have

$$\iint_{\Gamma^2}\operatorname{Cov}(Y'_x, Y'_y)\Lambda(dx)\Lambda(dy) = \mathrm{O}(\lambda\ln\lambda).$$

This gives the estimate of r_1' as

$$r_1' = \mathrm{O}(\lambda^{-1})\Big[\iint_{\Gamma^2} \mathrm{Cov}(Y_x', Y_y')\Lambda(dx)\Lambda(dy)\Big]^{1/2} = \mathrm{O}(\lambda^{-1/2}\sqrt{\ln\lambda}). \tag{4.22}$$

The estimate of $r_2' + r_3'$. Applying (4.14) gives

$$r_2' + r_3' \le B^{-3}\int_\Gamma \mathbb{E}Y_x^2\Lambda(dx) = \mathrm{O}(\lambda^{-1/2}). \tag{4.23}$$

The estimate of r_4'. To simplify the notation, we write

$$\zeta_{x,t} = \begin{cases} \mathbf{1}[1 \ge \Delta_x > t > 0] & \text{for } t > 0, \\ \mathbf{1}[-1 \le \Delta_x < t < 0] & \text{for } t < 0. \end{cases}$$

If $|x-y| > 20\sqrt{3\ln\lambda}$, we have

$$\begin{aligned} \mathrm{Cov}(\zeta_{x,t}, \zeta_{y,t}) = {} & \mathrm{Cov}\big(\zeta_{x,t}\mathbf{1}\big[U_x^c\big], \zeta_{y,t}\big) + \mathrm{Cov}\big(\zeta_{x,t}\mathbf{1}[U_x], \zeta_{y,t}\mathbf{1}\big[U_y^c\big]\big) \\ & + \mathrm{Cov}\big(\zeta_{x,t}\mathbf{1}[U_x], \zeta_{y,t}\mathbf{1}[U_y]\big). \end{aligned} \tag{4.24}$$

We apply (4.17) to obtain

$$\mathrm{Cov}\big(\zeta_{x,t}\mathbf{1}\big[U_x^c\big], \zeta_{y,t}\big) \le \mathbb{P}\big[U_x^c\big] = \mathrm{O}(\lambda^{-4}). \tag{4.25}$$

Likewise,

$$\mathrm{Cov}\big(\zeta_{x,t}\mathbf{1}[U_x], \zeta_{y,t}\mathbf{1}\big[U_y^c\big]\big) = \mathrm{O}(\lambda^{-4}).$$

Given U_xU_y, $\zeta_{x,t}$ is independent of $\zeta_{y,t}$, hence

$$\mathbb{E}(\zeta_{x,t}\zeta_{y,t}|U_xU_y) = \mathbb{E}(\zeta_{x,t}|U_xU_y)\mathbb{E}(\zeta_{y,t}|U_xU_y).$$

This ensures that we can expand the last term of (4.24) into

$$\begin{aligned} \mathrm{Cov}\big(\zeta_{x,t}\mathbf{1}[U_x], \zeta_{y,t}\mathbf{1}[U_y]\big) = {} & \mathbb{E}(\zeta_{x,t}|U_xU_y)\mathbb{E}(\zeta_{y,t}|U_xU_y)\mathbb{P}[U_xU_y]\big(1-\mathbb{P}[U_xU_y]\big) \\ & - \mathbb{E}(\zeta_{x,t}|U_x)\mathbb{P}[U_x]\mathbb{E}\big(\zeta_{y,t}|U_x^cU_y\big)\mathbb{P}\big[U_x^cU_y\big] \\ & - \mathbb{E}\big(\zeta_{x,t}|U_xU_y^c\big)\mathbb{P}\big[U_xU_y^c\big]\mathbb{E}(\zeta_{y,t}|U_xU_y)\mathbb{P}[U_xU_y] \\ = {} & \mathrm{O}\big(\mathbb{P}\big[U_x^c\big] + \mathbb{P}\big[U_y^c\big]\big) = \mathrm{O}(\lambda^{-4}), \end{aligned} \tag{4.26}$$

again, by (4.17). Combining estimates (4.25)–(4.26), we obtain from (4.24) that, when $|x-y| > 20\sqrt{3\ln\lambda}$,

$$\mathrm{Cov}(\zeta_{x,t}, \zeta_{y,t}) = \mathrm{O}(\lambda^{-4}). \tag{4.27}$$

This implies

$$\begin{aligned} & \int_0^1 \iint_{\Gamma^2} \mathrm{Cov}(\zeta_{x,t}, \zeta_{y,t})\Lambda(dx)\Lambda(dy)\,dt \\ & \quad = \int_0^1 \iint_{|x-y|>20\sqrt{3\ln\lambda}} \mathrm{Cov}(\zeta_{x,t}, \zeta_{y,t})\Lambda(dx)\Lambda(dy)\,dt \\ & \qquad + \int_0^1 \iint_{|x-y|\le 20\sqrt{3\ln\lambda}} \mathrm{Cov}(\zeta_{x,t}, \zeta_{y,t})\Lambda(dx)\Lambda(dy)\,dt \\ & \quad \le \mathrm{O}(\lambda^{-2}) + \int_0^1 \iint_{|x-y|\le 20\sqrt{3\ln\lambda}} \mathbb{E}\zeta_{x,t}\Lambda(dx)\Lambda(dy)\,dt \\ & \quad = \mathrm{O}(\lambda^{-2}) + B^{-1}\iint_{|x-y|\le 20\sqrt{3\ln\lambda}} \mathbb{E}|Y_x|\Lambda(dx)\Lambda(dy) \le \mathrm{O}(\lambda^{1/2}\ln\lambda), \end{aligned} \tag{4.28}$$

where the last inequality is due to (4.14). Similarly, we can also establish

$$\int_{-1}^{0}\iint_{\Gamma^2}\operatorname{Cov}(\zeta_{x,t},\zeta_{y,t})\Lambda(dx)\Lambda(dy)\,dt=\mathrm{O}\big(\lambda^{1/2}\ln\lambda\big). \tag{4.29}$$

Adding (4.28) and (4.29) gives

$$r_4'=\mathrm{O}\big(\lambda^{-1/2}\ln\lambda\big). \tag{4.30}$$

The estimate of r_5'. We make use of (4.27) to get

$$\begin{aligned}
&\int_0^1\iint_{\Gamma^2} t\operatorname{Cov}(\zeta_{x,t},\zeta_{y,t})\Lambda(dx)\Lambda(dy)\,dt\\
&\quad=\int_0^1\iint_{|x-y|>20\sqrt{3\ln\lambda}} t\operatorname{Cov}(\zeta_{x,t},\zeta_{y,t})\Lambda(dx)\Lambda(dy)\,dt\\
&\quad\quad+\int_0^1\iint_{|x-y|\le 20\sqrt{3\ln\lambda}} t\operatorname{Cov}(\zeta_{x,t},\zeta_{y,t})\Lambda(dx)\Lambda(dy)\,dt\\
&\quad\le\mathrm{O}\big(\lambda^{-2}\big)+\int_0^1\iint_{|x-y|\le 20\sqrt{3\ln\lambda}} t\mathbb{E}\zeta_{x,t}\Lambda(dx)\Lambda(dy)\,dt\\
&\quad=\mathrm{O}\big(\lambda^{-2}\big)+\frac{B^{-2}}{2}\iint_{|x-y|\le 20\sqrt{3\ln\lambda}}\mathbb{E}|Y_x|^2\Lambda(dx)\Lambda(dy)\le\mathrm{O}(\ln\lambda),
\end{aligned} \tag{4.31}$$

where, again, the last inequality follows from (4.14). Correspondingly, we can deduce the following bound:

$$\int_{-1}^{0}\iint_{\Gamma^2} t\operatorname{Cov}(\zeta_{x,t},\zeta_{y,t})\Lambda(dx)\Lambda(dy)\,dt=\mathrm{O}(\ln\lambda). \tag{4.32}$$

Combining (4.31) and (4.32) yields

$$r_5'=\mathrm{O}\big(\lambda^{-1/2}\sqrt{\ln\lambda}\big). \tag{4.33}$$

Finally, we collect all the estimates in (4.22), (4.23), (4.30) and (4.33) to achieve the bound (4.9), as claimed. □

4.4. *The total edge length of Poisson–Voronoi tessellations.* The Poisson–Voronoi tessellations have been studied extensively since Avram and Bertsimas (1993). For normal approximation of the total edge length of Poisson–Voronoi tessellations, an error bound of the optimal order was established by Lachièze-Rey, Schulte and Yukich (2019) using the Malliavin–Stein approach. In this subsection, we demonstrate that Theorem 3.1 can be utilized to derive an error bound of the same order.

Similar to the previous subsection, we consider $\boldsymbol{X}$ as a Poisson point process on $\mathbb{R}^2$ with mean measure $\boldsymbol{\mu}(dx)=dx$ and set $Q_\lambda=\{(x_1,x_2):-\sqrt{\lambda}/2\le x,y\le\sqrt{\lambda}/2\}$. For the ease of reading, we briefly recap a few essential terminologies. For a realization $\boldsymbol{x}$ of $\boldsymbol{X}$ and $x\in\boldsymbol{x}$, we define $\mathcal{C}(x,\boldsymbol{x})$ as the set of every point in $\mathbb{R}^2$ whose (Euclidean) distance to x is less than or equal to its distance to any other point of $\boldsymbol{x}$. The collection of $\mathcal{C}(x,\boldsymbol{x})$, $x\in\boldsymbol{x}$, is called the *Poisson–Voronoi tessellation* induced by the realization $\boldsymbol{x}$ of $\boldsymbol{X}$. Again, we write $L(x,\boldsymbol{X}):=L_\lambda(x,\boldsymbol{X})$ as one half the total edge length of the *finite* length edges in the cell $\mathcal{C}(x,\boldsymbol{X})$, then the total edge length of the Poisson–Voronoi tessellation induced by $\boldsymbol{X}$ with centers in $\boldsymbol{X}\cap Q_\lambda$ can be summarized as

$$\mathcal{L}(\lambda):=\int_{Q_\lambda}L(x,\boldsymbol{X})\boldsymbol{X}(dx).$$

THEOREM 4.13. *Let* $B^2 = \operatorname{Var}(\mathcal{L}(\lambda))$ *and* $W = (\mathcal{L}(\lambda) - \mathbb{E}\mathcal{L}(\lambda))/B$. *We have*

$$\lim_{\lambda\to\infty} \lambda^{-1}\mathbb{E}\mathcal{L}(\lambda) \in (0,\infty), \qquad \lim_{\lambda\to\infty} \lambda^{-1}B^2 \in (0,\infty) \tag{4.34}$$

and

$$d_{\mathrm{K}}\big(\mathscr{L}(W), \mathcal{N}(0,1)\big) = \mathrm{O}(\lambda^{-1/2}). \tag{4.35}$$

PROOF. The claim (4.34) can be found in Avram and Bertsimas (1993), hence it suffices to show (4.35). To this end, we write $\Xi(dx) = L(x, \boldsymbol{X})\boldsymbol{X}(dx)$ and apply Theorem 3.1. As observed by McGivney and Yukich (1999), Section 4, adding a point at α does not affect the Voronoi cells centered at points outside $B(\alpha, 3T_\alpha)$, while Y_α (respectively Y_β) is determined by the configuration $B(\alpha, 3T_\alpha) \cap \boldsymbol{X}$ (respectively $B(\beta, 3T_\beta) \cap \boldsymbol{X}$). Thus Y_α and Y_β are conditionally independent given T_α and T_β with $|\alpha - \beta| > 3(T_\alpha + T_\beta)$, which implies

$$\begin{aligned}
&\operatorname{Cov}(Y'_\alpha, Y'_\beta)\\
&= \int_0^\infty \int_0^\infty \mathbb{E}(Y'_\alpha Y'_\beta | T_\alpha = s_1, T_\beta = s_2)\mathbb{P}(T_\alpha \in ds_1, T_\beta \in ds_2)\\
&\quad - \int_0^\infty \mathbb{E}(Y'_\alpha | T_\alpha = s_1)\mathbb{P}(T_\alpha \in ds_1) \int_0^\infty \mathbb{E}(Y'_\beta | T_\beta = s_2)\mathbb{P}(T_\beta \in ds_2)\\
&= \iint_{3(s_1+s_2)\ge|\alpha-\beta|} \mathbb{E}(Y'_\alpha Y'_\beta | T_\alpha = s_1, T_\beta = s_2)\mathbb{P}(T_\alpha \in ds_1, T_\beta \in ds_2)\\
&\quad - \iint_{3(s_1+s_2)\ge|\alpha-\beta|} \mathbb{E}(Y'_\alpha | T_\alpha = s_1)\mathbb{E}(Y'_\beta | T_\beta = s_2)\mathbb{P}(T_\alpha \in ds_1)\mathbb{P}(T_\beta \in ds_2).
\end{aligned} \tag{4.36}$$

However, direct verification ensures

$$\mathbb{P}(T_\alpha \in ds) = 2\pi\big(1 - e^{-\frac{\pi s^2}{12}}\big)^{11} e^{-\frac{\pi s^2}{12}} s\, ds. \tag{4.37}$$

By checking the relationship of the event $\{T_\beta \in ds_2\}$ and various possible cases of the event $\{T_\alpha = s_1\}$, we obtain

$$\mathbb{P}(T_\beta \in ds_2 | T_\alpha = s_1) \le \frac{e^{-\frac{\pi s_2^2}{12}} \frac{2\pi s_2}{12}\, ds_2}{1 - e^{-\frac{\pi s_1^2}{12}}},$$

which, together with (4.37), implies

$$\begin{aligned}
\mathbb{P}(T_\alpha \in ds_1, T_\beta \in ds_2) &= \mathbb{P}(T_\beta \in ds_2 | T_\alpha = s_1)\mathbb{P}(T_\alpha \in ds_1)\\
&\le \frac{\pi^2}{3} s_1 s_2 e^{-\frac{\pi(s_1^2+s_2^2)}{12}}\, ds_1\, ds_2.
\end{aligned} \tag{4.38}$$

Since the change of the total edge lengths as the result of adding a point at α (respectively β) can be bounded by $2\pi T_\alpha$ (respectively $2\pi T_\beta$), if we use C to represent a constant independent of λ, α and β, whose value may vary from one line to another, then we have the following crude estimates for Y' and the same estimates also hold for Y in place of Y':

$$\begin{aligned}
\big|\mathbb{E}(Y'_\alpha | T_\alpha = s_1, T_\beta = s_2)\big| &\le C(s_1^3 + 1),\\
\big|\mathbb{E}(Y'_\alpha | T_\alpha = s_1)\big| &\le C(s_1^3 + 1),\\
\big|\mathbb{E}(Y'_\alpha Y'_\beta | T_\alpha = s_1, T_\beta = s_2)\big| &\le C(s_1^3 s_2^3 + 1),\\
\big|\mathbb{E}(Y'^2_\alpha | T_\alpha = s_1, T_\beta = s_2)\big| &\le C(s_1^4 + 1),\\
\big|\mathbb{E}(Y'^2_\alpha | T_\alpha = s_1)\big| &\le C(s_1^4 + 1).
\end{aligned} \tag{4.39}$$

Combining (4.36), (4.37), (4.38) and (4.39) gives

$$
\begin{aligned}
(4.40)\qquad |\mathrm{Cov}(Y'_\alpha, Y'_\beta)| &\le C \iint_{3(s_1+s_2)\ge|\alpha-\beta|} (s_1^4 s_2^4 + 1) e^{-\frac{\pi(s_1^2+s_2^2)}{12}}\, ds_1\, ds_2 \\
&\le C\big(|\alpha-\beta|^3+1\big) e^{-\frac{\pi|\alpha-\beta|^2}{432}}.
\end{aligned}
$$

Using (4.40), we have

$$
\begin{aligned}
(4.41)\qquad r'_1 &\le \mathrm{O}(\lambda^{-1}) \bigg(\iint_{Q_\lambda\times Q_\lambda} \mathrm{Cov}(Y'_\alpha, Y'_\beta) \Lambda(d\alpha)\Lambda(d\beta) \bigg)^{1/2} \\
&\le \mathrm{O}(\lambda^{-1}) \bigg(\iint_{Q_\lambda\times Q_\lambda} \big(|\alpha-\beta|^3+1\big) e^{-\frac{\pi|\alpha-\beta|^2}{432}} \Lambda(d\alpha)\Lambda(d\beta) \bigg)^{1/2} = \mathrm{O}\big(\lambda^{-1/2}\big).
\end{aligned}
$$

For r'_2 and r'_3, we use (4.37) and (4.39) to obtain

$$
\begin{aligned}
\mathbb{E}\big(Y_\alpha^2\big) &= \int_0^\infty \mathbb{E}\big(Y_\alpha^2 | T_\alpha = s\big) \mathbb{P}(T_\alpha \in ds) \le \int_0^\infty C\big(s^4+1\big) e^{-\frac{\pi s^2}{12}} s\, ds = C, \\
\mathbb{E}\big(Y_\alpha'^2\big) &= \int_0^\infty \mathbb{E}\big(Y_\alpha'^2 | T_\alpha = s\big) \mathbb{P}(T_\alpha \in ds) \le \int_0^\infty C\big(s^4+1\big) e^{-\frac{\pi s^2}{12}} s\, ds = C,
\end{aligned}
$$

which implies

$$
r'_2 \le \mathrm{O}\big(\lambda^{-3/2}\big) \int_{Q_\lambda} \mathbb{E}\big(Y_\alpha'^2\big) \Lambda(d\alpha) = \mathrm{O}\big(\lambda^{-1/2}\big)
$$

and

$$
r'_3 \le \mathrm{O}\big(\lambda^{-3/2}\big) \int_{Q_\lambda} \mathbb{E}\big(Y_\alpha^2\big) \Lambda(d\alpha) = \mathrm{O}\big(\lambda^{-1/2}\big).
$$

For r'_4, when $3(s_1+s_2) < |\alpha-\beta|$, we have

$$
\begin{aligned}
&\mathbb{P}(1 \ge \Delta_\alpha > t > 0, 1 \ge \Delta_\beta > t > 0 | T_\alpha = s_1, T_\beta = s_2) \\
&\quad = \mathbb{P}(1 \ge \Delta_\alpha > t > 0 | T_\alpha = s_1) \mathbb{P}(1 \ge \Delta_\beta > t > 0 | T_\beta = s_2).
\end{aligned}
$$

Hence

$$
\begin{aligned}
&\mathrm{Cov}\big(\mathbf{1}[1 \ge \Delta_\alpha > t > 0], \mathbf{1}[1 \ge \Delta_\beta > t > 0]\big) \\
&= \int_0^\infty \int_0^\infty \mathbb{P}(1 \ge \Delta_\alpha > t > 0, 1 \ge \Delta_\beta > t > 0 | T_\alpha = s_1, T_\beta = s_2) \\
&\quad \times \mathbb{P}(T_\alpha \in ds_1, T_\beta \in ds_2) \\
&\quad - \bigg(\int_0^\infty \mathbb{P}(1 \ge \Delta_\alpha > t > 0 | T_\alpha = s) \mathbb{P}(T_\alpha \in ds) \bigg)^2 \\
&= \iint_{3(s_1+s_2)\ge|\alpha-\beta|} \mathbb{P}(1 \ge \Delta_\alpha > t > 0, 1 \ge \Delta_\beta > t > 0 | T_\alpha = s_1, T_\beta = s_2) \\
&\quad \times \mathbb{P}(T_\alpha \in ds_1, T_\beta \in ds_2) \\
&\quad - \iint_{3(s_1+s_2)\ge|\alpha-\beta|} \mathbb{P}(1 \ge \Delta_\alpha > t > 0 | T_\alpha = s_1) \mathbb{P}(1 \ge \Delta_\beta > t > 0 | T_\beta = s_2) \\
&\quad \times \mathbb{P}(T_\alpha \in ds_1) \mathbb{P}(T_\beta \in ds_2),
\end{aligned}
$$

which implies

$$
\begin{aligned}
&\int_0^1 \bigl|\mathrm{Cov}\bigl(\mathbf{1}[1 \geq \Delta_\alpha > t > 0], \mathbf{1}[1 \geq \Delta_\beta > t > 0]\bigr)\bigr|\,dt \\
&\quad \leq \iint_{3(s_1+s_2)\geq|\alpha-\beta|} \int_0^1 \mathbb{P}(1 \geq \Delta_\alpha > t > 0, 1 \geq \Delta_\beta > t > 0 | T_\alpha = s_1, T_\beta = s_2)\,dt \\
&\qquad \times \mathbb{P}(T_\alpha \in ds_1, T_\beta \in ds_2) \\
&\qquad + \iint_{3(s_1+s_2)\geq|\alpha-\beta|} \int_0^1 \mathbb{P}(1 \geq \Delta_\alpha > t > 0 | T_\alpha = s_1)\mathbb{P}(1 \geq \Delta_\beta > t > 0 | T_\beta = s_2)\,dt \\
&\qquad \times \mathbb{P}(T_\alpha \in ds_1)\mathbb{P}(T_\beta \in ds_2) \\
&\quad \leq 2 \iint_{3(s_1+s_2)\geq|\alpha-\beta|, s_1<s_2} \mathbb{E}\bigl(|\Delta_\alpha| | T_\alpha = s_1, T_\beta = s_2\bigr)\mathbb{P}(T_\alpha \in ds_1, T_\beta \in ds_2) \\
&\qquad + 2 \iint_{3(s_1+s_2)\geq|\alpha-\beta|, s_1<s_2} \mathbb{E}\bigl(|\Delta_\alpha| | T_\alpha = s_1\bigr)\mathbb{P}(T_\alpha \in ds_1)\mathbb{P}(T_\beta \in ds_2) \\
&\quad \leq \mathrm{O}\bigl(\lambda^{-1/2}\bigr) \iint_{3(s_1+s_2)\geq|\alpha-\beta|, s_1<s_2} \bigl(s_1^3+1\bigr)\mathbb{P}(T_\alpha \in ds_1, T_\beta \in ds_2) \\
&\qquad + \mathrm{O}\bigl(\lambda^{-1/2}\bigr) \iint_{3(s_1+s_2)\geq|\alpha-\beta|, s_1<s_2} \bigl(s_1^3+1\bigr)\mathbb{P}(T_\alpha \in ds_1)\mathbb{P}(T_\beta \in ds_2) \\
&\quad \leq \mathrm{O}\bigl(\lambda^{-1/2}\bigr) e^{-\frac{\pi|\alpha-\beta|^2}{432}},
\end{aligned}
\tag{4.42}
$$

where the second last inequality is from (4.39) and the last inequality is obtained as in (4.40). Thus, it follows from (4.42) and the same argument for (4.41) that

$$
\begin{aligned}
&\biggl|\iint_{Q_\lambda \times Q_\lambda} \int_0^1 \mathrm{Cov}\bigl(\mathbf{1}[1 \geq \Delta_\alpha > t > 0], \mathbf{1}[1 \geq \Delta_\beta > t > 0]\bigr)\,dt\,\Lambda(d\alpha)\Lambda(d\beta)\biggr| \\
&\qquad \leq \mathrm{O}\bigl(\lambda^{-1/2}\bigr) \iint_{Q_\lambda \times Q_\lambda} e^{-\frac{\pi|\alpha-\beta|^2}{432}} \Lambda(d\alpha)\Lambda(d\beta) = \mathrm{O}\bigl(\lambda^{1/2}\bigr).
\end{aligned}
\tag{4.43}
$$

Likewise, we can show that

$$
\begin{aligned}
&\biggl|\iint_{Q_\lambda \times Q_\lambda} \int_{-1}^0 \mathrm{Cov}\bigl(\mathbf{1}[-1 \leq \Delta_\alpha < t < 0], \mathbf{1}[-1 \leq \Delta_\beta < t < 0]\bigr)\,dt\,\Lambda(d\alpha)\Lambda(d\beta)\biggr| \\
&\qquad = \mathrm{O}\bigl(\lambda^{1/2}\bigr).
\end{aligned}
\tag{4.44}
$$

Combining (4.43) and (4.44) gives $r_4' = \mathrm{O}(\lambda^{-1/2})$. For r_5', we can bound it in the same way as for r_4'. In fact, we replace (4.42) with

$$
\begin{aligned}
&\int_0^1 t\bigl|\mathrm{Cov}\bigl(\mathbf{1}[1 \geq \Delta_\alpha > t > 0], \mathbf{1}[1 \geq \Delta_\beta > t > 0]\bigr)\bigr|\,dt \\
&\quad \leq \iint_{3(s_1+s_2)\geq|\alpha-\beta|} \int_0^1 t\mathbb{P}(1 \geq \Delta_\alpha > t > 0, 1 \geq \Delta_\beta > t > 0 | T_\alpha = s_1, T_\beta = s_2)\,dt \\
&\qquad \times \mathbb{P}(T_\alpha \in ds_1, T_\beta \in ds_2) \\
&\qquad + \iint_{3(s_1+s_2)\geq|\alpha-\beta|} \int_0^1 t\mathbb{P}(1 \geq \Delta_\alpha > t > 0 | T_\alpha = s_1)\mathbb{P}(1 \geq \Delta_\beta > t > 0 | T_\beta = s_2)\,dt \\
&\qquad \times \mathbb{P}(T_\alpha \in ds_1)\mathbb{P}(T_\beta \in ds_2)
\end{aligned}
$$

$$
\begin{aligned}
&\le \iint_{3(s_1+s_2)\ge|\alpha-\beta|,s_1<s_2} \mathbb{E}\big(\Delta_\alpha^2|T_\alpha=s_1,T_\beta=s_2\big)\mathbb{P}(T_\alpha\in ds_1,T_\beta\in ds_2)\\
&\quad+\iint_{3(s_1+s_2)\ge|\alpha-\beta|,s_1<s_2} \mathbb{E}\big(\Delta_\alpha^2|T_\alpha=s_1\big)\mathbb{P}(T_\alpha\in ds_1)\mathbb{P}(T_\beta\in ds_2)\\
&\le \mathrm{O}\big(\lambda^{-1}\big)\iint_{3(s_1+s_2)\ge|\alpha-\beta|,s_1<s_2}(s_1^4+1)\mathbb{P}(T_\alpha\in ds_1,T_\beta\in ds_2)\\
&\quad+\mathrm{O}\big(\lambda^{-1}\big)\iint_{3(s_1+s_2)\ge|\alpha-\beta|,s_1<s_2}(s_1^4+1)\mathbb{P}(T_\alpha\in ds_1)\mathbb{P}(T_\beta\in ds_2)\\
&=\mathrm{O}\big(\lambda^{-1}\big)e^{-\frac{\pi|\alpha-\beta|^2}{432}},
\end{aligned}
$$

where the second last inequality is from (4.39) and the last inequality is from a similar argument leading to (4.40). Therefore,

$$
\begin{aligned}
&\left|\int_0^1\iint_{Q_\lambda\times Q_\lambda} t\,\mathrm{Cov}\big(\mathbf{1}[1\ge\Delta_\alpha>t>0],\mathbf{1}[1\ge\Delta_\beta>t>0]\big)\,dt\,\Lambda(d\alpha)\Lambda(d\beta)\right|\\
&\qquad\le \mathrm{O}\big(\lambda^{-1}\big)\iint_{Q_\lambda\times Q_\lambda} e^{-\frac{\pi|\alpha-\beta|^2}{432}}\Lambda(d\alpha)\Lambda(d\beta)=\mathrm{O}(1).
\end{aligned}
\tag{4.45}
$$

The same reasoning can be adjusted to show that

$$
\begin{aligned}
&\left|\int_{-1}^0\iint_{Q_\lambda\times Q_\lambda} t\,\mathrm{Cov}\big(\mathbf{1}[-1\le\Delta_\alpha<t<0],\mathbf{1}[-1\le\Delta_\beta<t<0]\big)\,dt\,\Lambda(d\alpha)\Lambda(d\beta)\right|\\
&\qquad=\mathrm{O}(1).
\end{aligned}
\tag{4.46}
$$

The anticipated order of $r_5'=\mathrm{O}(\lambda^{-1/2})$ can be observed from (4.45) and (4.46). The proof is completed by observing that r_i', $1\le i\le 5$, are all of the order $\mathrm{O}(\lambda^{-1/2})$. □

5. Stein couplings. The following notion was introduced by Chen and Röllin (2010). A triple of random variables (W,W',G) defined on the same probability space is called a *Stein coupling* if

$$
\mathbb{E}\big\{Gf(W')-Gf(W)\big\}=\mathbb{E}\big\{Wf(W)\big\}
\tag{5.1}
$$

for all absolutely continuous functions f with $f(x)=\mathrm{O}(1+|x|)$. By taking $f(x)=1$, (5.1) implies that $\mathbb{E}W=0$, and by taking $f(w)=w$, (5.1) implies that $\mathrm{Var}\,W=\mathbb{E}\{G(W'-W)\}$; as usual, we assume that $\mathrm{Var}\,W=1$.

Now suppose (W,W',G) is a Stein coupling. Let $\Delta=W'-W$, and let $\mathcal{F}$ be a σ-algebra with respect to which W is measurable. Then

$$
\mathbb{E}\big\{Wf(W)\big\}=\mathbb{E}\int_{-\infty}^{\infty}f'(W+t)\hat{K}(t)\,dt,
$$

where

$$
\hat{K}(t)=\mathbb{E}\big\{G\big(\mathbf{1}[\Delta>t>0]-\mathbf{1}[\Delta<t\le 0]\big)\big|\mathcal{F}\big\}.
$$

One particular choice of K^{in} and K^{out} that turns out to be useful is

$$
\begin{aligned}
\hat{K}^{\mathrm{in}}(t)&=\mathbb{E}\big\{G\big(\mathbf{1}[\Delta\wedge 1\ge t>0]-\mathbf{1}\big[\Delta\vee(-1)\le t\le 0\big]\big)\big|\mathcal{F}\big\},\\
\hat{K}^{\mathrm{out}}(t)&=\mathbb{E}\big\{G\big[\mathbf{1}\big(\Delta\ge t>\Delta\wedge 1\big]-\mathbf{1}\big[\Delta<t<\Delta\vee(-1)\big]\big)\big|\mathcal{F}\big\},
\end{aligned}
$$

from which, via Theorem 2.1, we can immediately deduce a bound on the normal approximation of W in terms of quantities involving only the Stein coupling. Whereas r_2 and r_3

are typically straightforward to bound, r_1, r_4 and r_5 can be difficult to handle without further assumptions. The following result shows, however, that, by introducing additional auxiliary random variables, we can obtain upper bounds on these quantities that are far more tractable.

THEOREM 5.1. *Let (W, W', G) be a Stein coupling; set $\Delta = W' - W$. Let (G', Δ') be a conditionally independent copy of (G, Δ) given $\mathcal{F}$, and let (G^*, Δ^*) be an unconditionally independent copy of (G, Δ). Then*

$$d_{\mathrm{K}}(\mathscr{L}(W), \mathcal{N}(0,1)) \le 9s_1 + 11s_2 + 5s_3 + 10s_4,$$

where

$$\begin{aligned}
s_1 &= \bigl(\mathbb{E}\bigl\{|GG'|(|\Delta| \wedge 1)(|\Delta' - \Delta^*| \wedge 2) + |G(G' - G^*)|(|\Delta| \wedge 1)(|\Delta^*| \wedge 1)\bigr\}\bigr)^{1/2},\\
s_2 &= \mathbb{E}\bigl\{|G|(|\Delta| \wedge 1)^2\bigr\}, \qquad s_3 = \mathbb{E}\bigl\{|G|(|\Delta| - 1)\mathbf{1}[|\Delta| > 1]\bigr\},\\
s_4 &= \mathbb{E}\bigl\{|GG'|(|\Delta' - \Delta^*| \wedge 1) + |G(G' - G^*)|(|\Delta'| \wedge |\Delta^*| \wedge 1)\bigr\}.
\end{aligned}$$

PROOF. Let $\mathbb{E}^{\mathcal{F}}$ denote conditional expectation with respect to $\mathcal{F}$, and let $\bar{x} = (x \wedge 1) \vee (-1)$. We apply Theorem 2.1; using Lemma 5.9(i) in the last inequality,

$$\begin{aligned}
r_1^2 &= \bigl(\mathbb{E}|\mathbb{E}^{\mathcal{F}}(G\bar{\Delta}) - \mathbb{E}(G\bar{\Delta})|\bigr)^2 \le \operatorname{Var}\mathbb{E}^{\mathcal{F}}(G\bar{\Delta}) = \mathbb{E}\bigl\{\bigl(\mathbb{E}^{\mathcal{F}}(G\bar{\Delta})\bigr)^2 - \bigl(\mathbb{E}\{G\bar{\Delta}\}\bigr)^2\bigr\}\\
&= \mathbb{E}\bigl\{GG'\bar{\Delta}\bar{\Delta}' - GG^*\bar{\Delta}\bar{\Delta}^*\bigr\} \le \mathbb{E}\bigl\{|G\bar{\Delta}||G'\bar{\Delta}' - G^*\bar{\Delta}^*|\bigr\}\\
&\le \mathbb{E}\bigl\{|G\bar{\Delta}|\bigl(|G'|(|\Delta' - \Delta^*| \wedge 2) + |G' - G^*|(|\Delta^*| \wedge 1)\bigr)\bigr\} = s_1^2.
\end{aligned}$$

The expressions for s_2 and s_3 are easily obtained from r_2 and r_3. Letting $I_x^{\pm}(t) = \mathbf{1}[x \wedge 1 \ge t > 0] - \mathbf{1}[x \vee (-1) \le t \le 0]$ and using Lemma 5.9(ii),

$$\begin{aligned}
r_4 &= \mathbb{E}\int_{|t|\le 1}\bigl(\hat{K}^{\mathrm{in}}(t) - K^{\mathrm{in}}(t)\bigr)^2\,dt = \mathbb{E}\int_{|t|\le 1}\bigl(\mathbb{E}^{\mathcal{F}}\bigl\{GI_t^{\pm}(\Delta)\bigr\} - \mathbb{E}\bigl\{GI_t^{\pm}(\Delta)\bigr\}\bigr)^2\,dt\\
&= \mathbb{E}\int_{|t|\le 1}\bigl(\mathbb{E}^{\mathcal{F}}\bigl\{GI_t^{\pm}(\Delta)\bigr\}\bigr)^2 - \bigl(\mathbb{E}\bigl\{GI_t^{\pm}(\Delta)\bigr\}\bigr)^2\,dt\\
&= \mathbb{E}\int_{|t|\le 1}\bigl(\mathbb{E}^{\mathcal{F}}\bigl\{GG'I_t^{\pm}(\Delta)I_t^{\pm}(\Delta')\bigr\} - \mathbb{E}\bigl\{GG^*I_t^{\pm}(\Delta)I_t^{\pm}(\Delta^*)\bigr\}\bigr)\,dt\\
&= \mathbb{E}\int_{|t|\le 1}\bigl(GG'I_t^{\pm}(\Delta)I_t^{\pm}(\Delta') - GG^*I_t^{\pm}(\Delta)I_t^{\pm}(\Delta^*)\bigr)\,dt\\
&= \mathbb{E}\bigl\{GG'(|\Delta| \wedge |\Delta'| \wedge 1)\mathbf{1}[\Delta\Delta' > 0] - GG^*(|\Delta| \wedge |\Delta^*| \wedge 1)\mathbf{1}[\Delta\Delta^* > 0]\bigr\}\\
&\le \mathbb{E}\bigl\{|GG'|(|\Delta' - \Delta^*| \wedge 1)\bigr\} + \mathbb{E}\bigl\{|G||G' - G^*|(|\Delta| \wedge |\Delta^*| \wedge 1)\bigr\} = s_4.
\end{aligned}$$

Finally, similar to the estimate of r_4, using Lemma 5.9(iii),

$$\begin{aligned}
r_5^2 &= \mathbb{E}\int_{|t|\le 1}|t|\bigl(\hat{K}^{\mathrm{in}}(t) - K^{\mathrm{in}}(t)\bigr)^2\,dt\\
&= \mathbb{E}\int_{|t|\le 1}|t|\bigl(GG'I_t^{\pm}(\Delta)I_t^{\pm}(\Delta') - GG^*I_t^{\pm}(\Delta)I_t^{\pm}(\Delta^*)\bigr)\,dt\\
&= \frac{1}{2}\mathbb{E}\bigl\{GG'(|\Delta| \wedge |\Delta'| \wedge 1)^2\mathbf{1}[\Delta\Delta' > 0] - GG^*(|\Delta| \wedge |\Delta^*| \wedge 1)^2\mathbf{1}[\Delta\Delta^* > 0]\bigr\}\\
&\le \mathbb{E}\biggl\{|GG'|(|\Delta| \wedge 1)(|\Delta' - \Delta^*| \wedge 1) + \frac{1}{2}|G||G' - G^*|(|\Delta| \wedge |\Delta^*| \wedge 1)^2\biggr\} \le s_1^2.
\end{aligned}$$

This concludes the proof. □

Many couplings in the literature, such as the exchangeable pair and the size-bias coupling, can be formulated as Stein couplings, to which Theorem 5.1 can be applied. In what follows, we construct Stein couplings for local dependence and additive functionals in classical occupancy problems and apply Theorem 5.1 to obtain new error bounds, of which the former improves a result of Chen and Shao (2004).

5.1. *Local dependence.* Consider a sequence of centered random variables $X_1, \dots, X_n$ which are locally dependent in the following sense. For each $1 \le i \le n$, there is a set $A_i \subset \{1, \dots, n\}$ such that X_i and $(X_j)_{j \in A_i^c}$ are independent of each other. Moreover, for each $1 \le i \le n$, there is a set $B_i \subset \{1, \dots, n\}$ such that $A_i \subset B_i$ and such that $(X_j)_{j \in A_i}$ and $(X_j)_{j \in B_i^c}$ are independent of each other. Let $W = \sum_{i=1}^n X_i$, $Y_i = \sum_{j \in A_i} X_j$ and assume that $\operatorname{Var} W = 1$.

This and related types of dependency structures were extensively studied by Chen and Shao (2004); in particular, it is shown that under the above assumptions and if there is $2 < \rho \le 4$ such that $\mathbb{E}|X_i|^\rho + \mathbb{E}|Y_i|^\rho \le \theta^p$ for some $\theta > 0$ and all $1 \le i \le n$, then

$$d_{\mathrm{K}}\big(\mathscr{L}(W), \mathcal{N}(0,1)\big) \le (13 + 11\kappa) n \theta^{3 \wedge \rho} + 2.5 \theta^{p/2} \sqrt{\kappa n}, \tag{5.2}$$

where

$$\kappa = \sup_{1 \le i \le n} \big|\{1 \le j \le n : B_j \cap B_i \ne \varnothing\}\big|; \tag{5.3}$$

see Chen and Shao (2004), Theorem 2.2.

We can easily reproduce (5.2) (up to constants) and further improve it by means of Theorem 5.1.

THEOREM 5.2. *Let $X_1, \dots, X_n$ be centered random variables with dependency neighborhoods A_i and B_i as described above, set $W = \sum_{i=1}^n X_i$, and assume that $\operatorname{Var}(W) = 1$, and assume that there is $\theta > 0$ such that $\mathbb{E}\{|X_i|^\rho\} \vee \mathbb{E}\{|Y_i|^\rho\} \le \theta^\rho$ for some $2 < \rho \le 4$ and each $1 \le i \le n$. Then*

$$d_{\mathrm{K}}\big(\mathscr{L}(W), \mathcal{N}(0,1)\big) \le (16 + 67\lambda) n \theta^{3 \wedge \rho} + 28 \theta^{\rho/2} \sqrt{\lambda n},$$

where

$$\lambda = \sup_{1 \le i \le n} \big|\{1 \le j \le n : A_j \cap B_i \ne \varnothing\}\big|. \tag{5.4}$$

REMARK 5.3. Note that λ in (5.4) is upper bounded by κ in (5.3), and in fact, λ can be substantially smaller than κ.

PROOF OF THEOREM 5.2. Let I and J be independent random variables, uniformly distributed on set of indices $\{1, \dots, n\}$ independently of all else. Let $W' = W - Y_I$ and $G = -nX_I$; then (W, W', G) is a Stein coupling, and we have $\Delta = -Y_I$. Let $\mathcal{F} = \sigma(X_1, \dots, X_n)$, and let $(X_1^*, \dots, X_n^*)$ be an independent copy of $(X_1, \dots, X_n)$. Let $G' = -nX_J$ and $\Delta' = -Y_J$; clearly, (G', Δ') is an independent copy of (G, Δ) given $\mathcal{F}$. Moreover, with

$$\big(G^*, \Delta^*\big) = \begin{cases} (-nX_J, -Y_J) & \text{if } A_J \cap B_I = \varnothing, \\ (-nX_J^*, -Y_J^*) & \text{if } A_J \cap B_I \ne \varnothing; \end{cases}$$

it is easy to see that (G^*, Δ^*) is an unconditionally independent copy of (G, Δ). We now apply Theorem 5.1. To this end, recall Young's inequality; for nonnegative real numbers a and b we have $ab \le a^p/p + b^q/q$ whenever p and q are Hölder conjugates.

We start by bounding s_2; we have

$$s_2 = \sum_{i=1}^{n} \mathbb{E}\{|X_i|(|Y_i| \wedge 1)^2\}.$$

If $3 \le \rho \le 4$, Young's inequality yields

$$\mathbb{E}\{|X_i|(|Y_i| \wedge 1)^2\} \le \frac{\mathbb{E}|X_i|^3}{3} + \frac{2\mathbb{E}|Y_i|^3}{3} \le \theta^3,$$

and if $2 < \rho < 3$, Young's inequality yields

$$\mathbb{E}\{|X_i|(|Y_i| \wedge 1)^2\} \le \frac{\mathbb{E}|X_i|^\rho}{\rho} + \frac{(\rho-1)\mathbb{E}(|Y_i| \wedge 1)^{2\rho/(\rho-1)}}{\rho} \le \biggl(\frac{1}{\rho} + \frac{\rho-1}{\rho}\biggr)\theta^\rho = \theta^\rho,$$

so that $s_2 \le n\theta^{3\wedge\rho}$. We continue to bound s_3; we have

$$s_3 = \sum_{i=1}^{n} \mathbb{E}\{|X_i|(|Y_i| - 1)\mathbf{1}[|Y_i| > 1]\}.$$

If $3 \le \rho \le 4$, Young's inequality yields

$$\mathbb{E}\{|X_i|(|Y_i| - 1)\mathbf{1}[|Y_i| > 1]\} \le \frac{\mathbb{E}|X_i|^3}{3} + \frac{2\mathbb{E}|Y_i|^3}{3} \le \theta^3,$$

and if $2 < \rho < 3$, Young's inequality yields

$$\begin{aligned}\mathbb{E}\{|X_i|(|Y_i| - 1)\mathbf{1}[|Y_i| > 1]\} &\le \frac{\mathbb{E}|X_i|^\rho}{\rho} + \frac{(\rho-1)\mathbb{E}(|Y_i|\mathbf{1}[|Y_i| > 1])^{\rho/(\rho-1)}}{\rho} \\ &\le \frac{\mathbb{E}|X_i|^\rho}{\rho} + \frac{(\rho-1)\mathbb{E}|Y_i|^\rho}{\rho} \le \theta^\rho,\end{aligned}$$

so that $s_3 \le n\theta^{3\wedge\rho}$. Now, in order to bound s_4, note that

$$\mathbb{E}\{|GG'|(|\Delta' - \Delta^*| \wedge 1)\} = \sum_{i=1}^{n} \sum_{j:A_j \cap B_i \neq \varnothing}^{n} \mathbb{E}\{|X_i X_j|(|Y_j - Y_j^*| \wedge 1)\}.$$

Using again Young's inequality, we have for $3 \le \rho \le 4$,

$$\mathbb{E}\{|X_i X_j|(|Y_j - Y_j^*| \wedge 1)\} \le \frac{1}{3}\mathbb{E}\{|X_i|^3 + |X_j|^3 + 4|Y_j|^3 + 4|Y_j^*|^3\} \le \frac{10}{3}\theta^3,$$

and for $2 < \rho < 3$,

$$\begin{aligned}\mathbb{E}\{|X_i X_j|(|Y_j - Y_j^*| \wedge 1)\} &\le \frac{1}{\rho}\mathbb{E}|X_i|^\rho + \frac{1}{\rho}\mathbb{E}|X_j|^\rho + \frac{\rho-2}{\rho}\mathbb{E}(|Y_j - Y_j^*| \wedge 1)^{\rho/(\rho-2)} \\ &\le \frac{1}{\rho}\mathbb{E}|X_i|^\rho + \frac{1}{\rho}\mathbb{E}|X_j|^\rho + \frac{\rho-2}{\rho}\mathbb{E}(|Y_j - Y_j^*|)^\rho \\ &\le \frac{1}{\rho}\mathbb{E}|X_i|^\rho + \frac{1}{\rho}\mathbb{E}|X_j|^\rho + 2^{\rho-1}\frac{\rho-2}{\rho}(\mathbb{E}|Y_j|^\rho + \mathbb{E}|Y_j^*|^\rho) \\ &\le \biggl(\frac{2}{\rho} + 2^\rho\frac{\rho-2}{\rho}\biggr)\theta^\rho \le \frac{10}{3}\theta^\rho.\end{aligned}$$

Similarly,

$$\begin{aligned}\mathbb{E}\{|G(G' - G^*)|(|\Delta'| \wedge |\Delta^*| \wedge 1)\} &\le \mathbb{E}\{|G(G' - G^*)|(|\Delta'| \wedge 1)\} \\ &= \sum_{i=1}^{n} \sum_{j:A_j \cap B_i \neq \varnothing}^{n} \mathbb{E}\{|X_i(X_j - X_j^*)|(|Y_j| \wedge 1)\}.\end{aligned}$$

Using again Young's inequality, we have for $3 \le \rho \le 4$,

$$\mathbb{E}\{|X_i(X_j - X_j^*)|(|Y_j| \wedge 1)\} \le \frac{1}{3}\mathbb{E}\{|X_i|^3 + 4|X_j|^3 + 4|X_j^*|^3 + |Y_j|^3\} \le \frac{10}{3}\theta^3,$$

and for $2 < \rho < 3$,

$$\begin{aligned}
&\mathbb{E}\{|X_i(X_j - X_j^*)|(|Y_j| \wedge 1)\} \\
&\quad \le \frac{1}{\rho}\mathbb{E}|X_i|^\rho + \frac{1}{\rho}\mathbb{E}|X_j - X_j^*|^\rho + \frac{\rho-2}{\rho}\mathbb{E}(|Y_j| \wedge 1)^{\rho/(\rho-2)} \\
&\quad \le \frac{1}{\rho}\mathbb{E}|X_i|^\rho + \frac{1}{\rho}\mathbb{E}|X_j - X_j^*|^\rho + \frac{\rho-2}{\rho}\mathbb{E}|Y_j|^\rho \\
&\quad \le \frac{1}{\rho}\mathbb{E}|X_i|^\rho + \frac{2^{\rho-1}}{\rho}(\mathbb{E}|X_j|^\rho + \mathbb{E}|X_j^*|^\rho) + \frac{\rho-2}{\rho}\mathbb{E}|Y_j|^\rho \\
&\quad \le \Big(\frac{1+2^\rho}{\rho} + \frac{\rho-2}{\rho}\Big)\theta^\rho \le \frac{10}{3}\theta^\rho.
\end{aligned}$$

Hence, $s_4 \le \frac{20}{3}n\lambda\theta^{3\wedge\rho}$. Finally, to bound s_1, note that

$$\begin{aligned}
&\mathbb{E}\{|GG'|(|\Delta| \wedge 1)(|\Delta' - \Delta^*| \wedge 2)\} \\
&\quad = \sum_{i=1}^n \sum_{j:A_j \cap B_i \neq \varnothing}^n \mathbb{E}\{|X_i X_j|(|Y_i| \wedge 1)(|Y_j - Y_j^*| \wedge 2)\}.
\end{aligned}$$

Using Young's inequality, we have

$$\begin{aligned}
&\mathbb{E}\Big\{|2X_i X_j|(|Y_i| \wedge 1)\Big(\Big(\frac{1}{2}|Y_j - Y_j^*|\Big) \wedge 1\Big)\Big\} \\
&\quad \le \frac{\mathbb{E}|2X_i|^\rho}{\rho} + \frac{\mathbb{E}|X_j|^\rho}{\rho} \\
&\qquad + \frac{(\rho-2)(\mathbb{E}(|Y_i| \wedge 1)^{2\rho/(\rho-2)} + \mathbb{E}((\frac{1}{2}|Y_j - Y_j^*|) \wedge 1)^{2\rho/(\rho-2)})}{2\rho} \\
&\quad \le \frac{\mathbb{E}|2X_i|^\rho}{\rho} + \frac{\mathbb{E}|X_j|^\rho}{\rho} + \frac{(\rho-2)(\mathbb{E}|Y_i|^\rho + 2^{-\rho}\mathbb{E}|Y_j - Y_j^*|^\rho)}{2\rho} \\
&\quad \le \Big(\frac{1+2^\rho}{\rho} + \frac{(\rho-2)(1+2^{-\rho+\rho-1+1})}{2\rho}\Big)\theta^\rho \le \frac{19}{4}\theta^\rho.
\end{aligned}$$

Similarly,

$$\begin{aligned}
&\mathbb{E}\{|G(G' - G^*)|(|\Delta| \wedge 1)(|\Delta^*| \wedge 1)\} \\
&\quad = \sum_{i=1}^n \sum_{j:A_j \cap B_i \neq \varnothing}^n \mathbb{E}\{|X_i(X_j - X_j^*)|(|Y_i| \wedge 1)(|Y_j^*| \wedge 1)\}.
\end{aligned}$$

Using Young's inequality another time, we have

$$\begin{aligned}
&\mathbb{E}\{|X_i(X_j - X_j^*)|(|Y_i| \wedge 1)(|Y_j^*| \wedge 1)\} \\
&\quad \le \frac{\mathbb{E}|X_i|^\rho}{\rho} + \frac{\mathbb{E}|X_j - X_j^*|^\rho}{\rho} + \frac{(\rho-2)(\mathbb{E}(|Y_i| \wedge 1)^{2\rho/(\rho-2)} + \mathbb{E}(|Y_j^*| \wedge 1)^{2\rho/(\rho-2)})}{2\rho} \\
&\quad \le \frac{\mathbb{E}|X_i|^\rho}{\rho} + \frac{\mathbb{E}|X_j - X_j^*|^\rho}{\rho} + \frac{(\rho-2)(\mathbb{E}|Y_i|^\rho + \mathbb{E}|Y_j^*|^\rho)}{2\rho} \\
&\quad \le \biggl(\frac{1+2^\rho}{\rho} + \frac{\rho-2}{\rho}\biggr)\theta^\rho \le \frac{19}{4}\theta^\rho.
\end{aligned}$$

Thus, $s_1^2 \le \frac{19}{2} n\lambda\theta^\rho$. Combining the bounds and applying Theorem 5.1, the final bound follows. □

5.2. *Additive functionals in the classical occupancy scheme.* Consider the following multinomial urn model. A total of m balls are independently distributed among n urns in such a way that a ball is placed in urn i with probability p_i, where $\sum_{i=1}^m p_i = 1$. Let ξ_i be the number of balls urn i contains after the balls have been distributed. For each $1 \le i \le n$, let φ_i be a real-valued function on the nonnegative integers. We are interested in the statistic

$$V = \sum_{i=1}^n \varphi_i(\xi_i),$$

respectively, the centered and normalized version

$$W = \frac{1}{\sigma}\sum_{i=1}^n (\varphi_i(\xi_i) - \mu_i),$$

where $\mu_i = \mathbb{E}\varphi_i(\xi_i)$ and $\sigma^2 = \operatorname{Var} V$. In the special case where $p_i = 1/n$ and $\varphi_i = \varphi$, this statistic has been studied by various authors; in particular, Dembo and Rinott (1996) used Stein's method and size-biasing to obtain error bounds, but only for smooth probability metrics (we refer to their paper for general references). We give the corresponding result for the Kolmogorov distance for general p_i and φ_i, which is, to the best of our knowledge, new in this generality.

THEOREM 5.4. *Let m, n and W be as in the preceding paragraph. Assume there are positive constants K_1, K_2 and K_3 such that*

$$|\varphi_i(x)| \le K_1 e^{K_1 x}, \quad x \ge 0, 1 \le i \le n,$$

and such that

$$\sup_{1\le i\le n} p_i \le \frac{K_2}{m} \quad \text{and} \quad n \le K_3 m. \tag{5.5}$$

Then there is a constant $C := C(K_1, K_2, K_3)$ such that

$$d_{\mathrm{K}}(\mathscr{L}(W), \mathcal{N}(0,1)) \le C\biggl(\frac{n^{1/2}}{\sigma^2} + \frac{n}{\sigma^3}\biggr). \tag{5.6}$$

Specializing to the case originally considered by Dembo and Rinott (1996), we have that $\sigma^2 \asymp n$ as long as φ is not a linear function (see Dembo and Rinott (1996), Remark 3.1), and we have the following corollary.

COROLLARY 5.5. *If $\varphi_i = \varphi$ for all i for some nonlinear function φ, if $p_i = 1/n$, and if $0 < \lim_{n\to\infty} n/m < \infty$, then $\sigma^2 \asymp n$ and thus,*

$$d_{\mathrm{K}}\big(\mathscr{L}(W), \mathcal{N}(0,1)\big) = \mathrm{O}\Big(\frac{1}{n^{1/2}}\Big).$$

PROOF OF THEOREM 5.4. In what follows the reference set in expressions like "$i \neq I$" or "$i \notin A$" is $\{1,\dots,n\}$, so that these expressions have to be read as "$i \in \{1,\dots,n\} \setminus \{I\}$" or "$i \in \{1,\dots,n\} \setminus A$", respectively. Moreover, we use the convention that sums of the form "$\sum_{i\neq j}$" stand for single sums over the first variable, not double sums over both variables.

Stein coupling. Denote by $(\xi_i)_{1\le i\le n}$ the ball counts in the respective urns, and let $\mathcal{F} = \sigma(\xi_1,\dots,\xi_n)$. Let I be uniformly distributed on $\{1,\dots,n\}$, independently of all else. Given I, let $\iota_1, \iota_2, \dots$ be an i.i.d. sequence, where $\mathbb{P}[\iota_1 = i|I] = p_i/(1-p_I)$ for all $i \neq I$, let $\eta_i = \xi_i + \sum_{k=1}^{\xi_I} \mathbf{1}[\iota_k = i]$ for $i \neq I$, and let $N := \{\iota_1,\dots,\iota_{\xi_I}\}$. The family of random variables $(\eta_i)_{i\neq I}$ represents the configuration of balls in the urns if the balls from the Ith urn are redistributed among the other urns, and N is the set of urns having received at least one ball during that redistribution.

Let W be defined as before, and let

$$G = -\frac{n}{\sigma}\big(\varphi_I(\xi_I) - \mu_I\big), \qquad W' = \frac{1}{\sigma}\Big(-\mu_I + \sum_{i\neq I}\big(\varphi_i(\eta_i) - \mu_i\big)\Big).$$

It is not difficult to see that (W, W', G) is a Stein coupling and that

$$\Delta = -\frac{1}{\sigma}\Big(\varphi_I(\xi_I) + \sum_{i\in N}\big(\varphi_i(\xi_i) - \varphi_i(\eta_i)\big)\Big);$$

see, for example, Chen and Röllin (2010), Construction 2A.

Construction of (G', Δ'). Let $\mathcal{F} = \sigma(\xi_1,\dots,\xi_n)$, and let J be uniformly distributed on $\{1,\dots,n\}$, independently of all else. Given J, let $\iota'_1, \iota'_2, \dots$ be an i.i.d. sequence, where $\mathbb{P}[\iota'_1 = i|J] = p_i/(1-p_J)$ for all $i \neq J$, and let $\eta'_i = \xi_i + \sum_{k=1}^{\xi_J} \mathbf{1}[\iota'_k = i]$ for $i \neq J$, and let $N' := \{\iota'_1,\dots,\iota'_{\xi_J}\}$. Define

$$G' = -\frac{n}{\sigma}\big(\varphi_J(\xi_J) - \mu_J\big), \qquad \Delta' = -\frac{1}{\sigma}\Big(\varphi_J(\xi_J) + \sum_{i\in N'}\big(\varphi_i(\xi_i) - \varphi_i(\eta'_i)\big)\Big).$$

Clearly, (G', Δ') is an independent copy of (G, Δ) conditionally on $\mathcal{F}$.

Construction of (G^*, Δ^*). We first construct a realization $(\xi^*_i)_{1\le i\le n}$ of the urn process which is independent of (G, Δ) but which is still closely coupled to $(\xi_i)_{1\le i\le n}$. To this end, let $(\xi^\bullet_i)_{1\le i\le n}$ be an independent copy of $(\xi_i)_{1\le i\le n}$. Set

$$\xi^*_i = \xi^\bullet_i \quad \text{for all } i \in N \cup \{I\}.$$

Now, let $\chi = \sum_{i\in N\cup\{I\}} \xi_i$ and $\chi^\bullet = \sum_{i\in N\cup\{I\}} \xi^\bullet_i$; we distinguish three cases.

(i) If $\chi^\bullet = \chi$, set

$$\xi^*_i = \xi_i \quad \text{for all } i \notin N \cup \{I\}.$$

(ii) If $\chi^\bullet < \chi$, let $\iota^*_1, \dots, \iota^*_{\chi-\chi^\bullet}$ be i.i.d. random variables on $\{1,\dots,n\} \setminus (N \cup \{I\})$ with distribution given by $\mathbb{P}[\iota^*_1 = k|I, N] = p_k/(1 - \sum_{j\in N\cup\{I\}} p_j)$ for $k \notin N \cup \{I\}$, and set

$$\xi^*_i = \xi_i + \sum_{j=1}^{\chi-\chi^\bullet} \mathbf{1}\big[\iota^*_j = i\big] \quad \text{for all } i \notin N \cup \{I\}.$$

(iii) If $\chi^\bullet > \chi$, let $\iota_1^*, \dots, \iota_{\chi^\bullet - \chi}$ be constructed recursively as follows. Let ι_1^* have distribution given by $\mathbb{P}[\iota_1^* = k | I, (\xi_i)_{i \notin N \cup \{I\}}] = \xi_k/(m - \chi)$ for $k \notin N \cup \{I\}$. For $1 < l < \chi^\bullet - \chi$, assume $\iota_1^*, \dots, \iota_l^*$ have been sampled, and let ι_{l+1}^* have distribution given by $\mathbb{P}[\iota_{l+1}^* = k | I, (\xi_i)_{i \notin N \cup \{I\}}, (\iota_i^*)_{1 \le i \le l}] = (\xi_k - \sum_{j=1}^l \mathbf{1}[\iota_j^* = k])/(m - \chi - l)$ for $k \notin N \cup \{I\}$. Finally, set

$$\xi_i^* = \xi_i - \sum_{j=1}^{\chi^\bullet - \chi} \mathbf{1}\big[\iota_j^* = i\big] \quad \text{for all } i \notin N \cup \{I\}.$$

It is not difficult to check that the distribution of $(\xi^*)_{1 \le i \le n}$ is the same regardless of $(\xi_i)_{i \in N \cup \{I\}}$ (the key to this is the observation that $(\xi_i)_{i \notin N \cup \{I\}}$ given $(\xi_i)_{i \notin N \cup \{I\}}$ is like distributing $n - \chi$ balls independently among the urns $i \notin N \cup \{I\}$ proportionally to their respective probabilities $(p_i)_{i \notin N \cup \{I\}}$), so that $(\xi^*)_{1 \le i \le n}$ is independent of $(\xi_i)_{i \in N \cup \{I\}}$.

With $\iota_1', \iota_2', \dots$ as before, we set $\eta_i^* = \xi_i^* + \sum_{k=1}^{\xi_J^*} \mathbf{1}[\iota_k' = i]$ and $N^* = \{\iota_1', \dots, \iota_{\xi_J^*}'\}$ and define

$$G^* = -\frac{n}{\sigma}\big(\varphi_J\big(\xi_J^*\big) - \mu_J\big), \qquad \Delta^* = -\frac{1}{\sigma}\Big(\varphi_J\big(\xi_J^*\big) + \sum_{i \in N^*} \big(\varphi_i\big(\xi_i^*\big) - \varphi_i\big(\eta_i^*\big)\big)\Big).$$

Since $(\xi_i^*)_{1 \le i \le n}$, J and $(\iota_1', \iota_2', \dots)$ are independent of $(\xi_i)_{i \in N \cup \{I\}}$, it follows that (G^*, Δ^*) is an independent copy of (G, Δ).

Bounding the error terms—preliminaries. We first show that we can assume without loss of generality that

$$\sup_{1 \le i \le n} p_i \le \frac{K_2}{m} \wedge \frac{1}{2}. \tag{5.7}$$

Indeed, there are only finitely many m such that $K_2/m \ge 1/2$, that is, such that $m \le 2K_2$, and hence, using the second condition of (5.5), there are only finitely many pairs (m, n) such that $K_2/m \ge 1/2$. For these finitely many cases, we can choose C large enough to make (5.6) true whenever $\sigma^2 > 0$.

We will use (5.7) repeatedly to conclude that, for example,

$$\frac{p_j}{1 - p_i} \le \frac{2K_2}{m} \le \frac{C}{m}.$$

Throughout the proof we will use C to denote a constant that can change from expression to expression but that only depends on K_1, K_2 and K_3.

Bounding the error terms—an event of small probability. Define the event

$$B_1 = \big\{\xi_J = \xi_J^*\big\} \cap \big\{\xi_{\iota_j'} = \xi_{\iota_j'}^* \text{ for all } 1 \le j \le \xi_J\big\}$$

and note that

$$B_1 \subset \big\{G' = G^*\big\}, \qquad B_1 \subset \big\{\Delta' = \Delta^*\big\}.$$

Let $N^\bullet = \{\iota_1^*, \dots, \iota_{|\chi^\bullet - \chi|}^*\} = \{i \notin N \cup \{I\} : \xi_i \ne \xi_i^*\}$, let $M = N^\bullet \cup N \cup \{I\}$ (disjoint union!), and define the event

$$B_2 = \big\{M \cap \{J\} \cap N' = \varnothing\big\}.$$

Clearly, $B_2 \subset B_1$; thus, setting $A = B_2^c$, we have

$$\big\{G' \ne G^*\big\} \subset A, \qquad \big\{\Delta' \ne \Delta^*\big\} \subset A.$$

We proceed to bound $\mathbb{P}[A]$. Let $\mathcal{G} = \sigma(I, (\iota_i)_{i\geq 1}, (\iota_i^*)_{i\geq 1}, (\xi_i)_{1\leq i\leq n}, (\xi_i^\bullet)_{1\leq i\leq n})$; we have

$$\begin{aligned}\mathbb{P}[A|\mathcal{G}] &= \frac{1}{n}\sum_{j=1}^{n}\mathbb{P}[A|J=j,\mathcal{G}] \leq \frac{|M|}{n} + \frac{1}{n}\sum_{j\notin M}\mathbb{E}\{|M\cap N'||J=j,\mathcal{G}\}\\ &\leq \frac{|M|}{n} + \frac{1}{n}\sum_{j\notin M}\mathbb{E}\Bigg\{\sum_{k=1}^{\xi_J}\mathbf{1}[\iota'_k\in M]\Bigg|J=j,\mathcal{G}\Bigg\} = \frac{|M|}{n} + \frac{1}{n}\sum_{j\notin M}\xi_j\sum_{k\in M}\frac{p_k}{1-p_j}\\ &\leq \frac{|M|}{n} + \frac{2K_2|M|}{nm}\sum_{j\notin M}\xi_j.\end{aligned}$$

Since $\sum_{j\notin M}\xi_j \leq m$, and since

$$|M| = 1 + |N| + |\chi - \chi^\bullet| \leq C(1+\chi+\chi^\bullet),$$

we obtain

$$\mathbb{P}\big[A|I,(\iota_i)_{i\geq 1},(\iota_i^*)_{i\geq 1},(\xi_i)_{1\leq i\leq n},(\xi_i^\bullet)_{1\leq i\leq n}\big] \leq \frac{C}{n}(1+\chi+\chi^\bullet). \tag{5.8}$$

Let now $\mathcal{G} = \sigma(I,(\iota_i)_{i\geq 1},\xi_I)$. We have

$$\begin{aligned}&\mathbb{E}\Bigg\{\sum_{i\in N\cup\{I\}}(\xi_i+\xi_i^\bullet)\Bigg|\mathcal{G}\Bigg\}\\ &\leq \mathbb{E}\Bigg\{\sum_{i\neq I}(\xi_i+\xi_i^\bullet)\sum_{k=1}^{\xi_I}\mathbf{1}[\iota_k=i]\Bigg|\mathcal{G}\Bigg\}\\ &= \mathbb{E}\Bigg\{\sum_{i\neq I}(\xi_i+\xi_i^\bullet)\sum_{k=1}^{\xi_I}\frac{p_i}{1-p_I}\Bigg|\mathcal{G}\Bigg\} \leq \frac{2K_2\xi_I}{m}\mathbb{E}\Bigg\{\sum_{i\neq I}(\xi_i+\xi_i^\bullet)\Bigg|\mathcal{G}\Bigg\} \leq 4K_2\xi_I,\end{aligned}$$

which, together with (5.8), implies that

$$\mathbb{P}\big[A|I,(\iota_i)_{i\geq 1},\xi_I\big] \leq \frac{C(\xi_I+1)}{n}.$$

Taking expectation and applying Lemma 5.7, we obtain

$$\mathbb{P}[A] \leq \frac{C}{n}.$$

Bounding the error terms—G and Δ. Note first that

$$|\Delta|^3 \leq \frac{C}{\sigma^3}\Bigg(e^{K_1\xi_I} + 2\sum_{i\in N}e^{K_1(\xi_i+\xi_I)}\Bigg)^3 \leq \frac{C}{\sigma^3}\Bigg(e^{3K_1\xi_I} + |N|^2\sum_{i\in N}e^{3K_1(\xi_i+\xi_I)}\Bigg),$$

and since $|N| \leq \xi_I \leq e^{\xi_I}$, we can further bound this by

$$|\Delta|^3 \leq \frac{C}{\sigma^3}\sum_{i\in N\cup\{I\}}e^{C(\xi_i+\xi_I)}.$$

Now, given I and ξ_I, the family of random variables $(\xi_i)_{i\neq I}$ is again an urn model; in particular, we have

$$\mathscr{L}(\xi_i|I,\xi_I) = \mathrm{Bi}\Bigg(m-\xi_I, \frac{p_i}{1-p_I}\Bigg).$$

Since the set N is chosen independently of $(\xi_i)_{i\neq I}$, we can apply Lemma 5.7 consecutively, and we obtain

$$\mathbb{E}\sum_{i\in N\cup\{I\}} e^{C(\xi_i+\xi_I)} \le C\mathbb{E}\{(1+|N|)e^{C\xi_I}\} \le C\mathbb{E}\{e^{C\xi_I}\} \le C.$$

Thus, $\mathbb{E}|\Delta|^3 \le C/\sigma^3$. Now,

$$|G|^3 \le \frac{Cn^3}{\sigma^3}(e^{3K_1\xi_I} + \mu_I^3),$$

and since $\mu_i^3 \le K_1\mathbb{E}e^{3K_1\xi_i}$, we have

$$\mathbb{E}|G|^3 \le \frac{Cn^3}{\sigma^3}\mathbb{E}e^{C\xi_I} \le \frac{Cn^3}{\sigma^3}.$$

Bounding the error terms—$G\mathbf{1}_A$ and $\Delta\mathbf{1}_A$. Note first that

$$\Delta^4 \le \frac{C}{\sigma^4}\Big(e^{K_1\xi_I} + 2\sum_{i\in N} e^{K_1(\xi_i+\xi_I)}\Big)^4 \le \frac{C}{\sigma^4}\Big(e^{4K_1\xi_I} + |N|^3\sum_{i\in N} e^{4K_1(\xi_i+\xi_I)}\Big),$$

and since $|N| \le \xi_I$, we can further bound this by

$$\Delta^4 \le \frac{C}{\sigma^4}\sum_{i\in N\cup\{I\}} e^{C(\xi_i+\xi_I)}.$$

Thus, using (5.8), we obtain that

$$\begin{aligned}\mathbb{E}\{\Delta^4\mathbf{1}_A\} &\le \frac{C}{n\sigma^4}\mathbb{E}\Big\{\Big(\sum_{i\in N\cup\{I\}} e^{C(\xi_i+\xi_I)}\Big)\Big(1+\sum_{i\in N\cup\{I\}}(\xi_i+\xi_i^\bullet)\Big)\Big\}\\ &\le \frac{C}{n\sigma^4}\mathbb{E}\Big\{\Big(\sum_{i\in N\cup\{I\}} e^{C(\xi_i+\xi_I)}\Big)\Big(\sum_{i\in N\cup\{I\}} e^{C(\xi_i+\xi_i^\bullet)}\Big)\Big\}\\ &\le \frac{C}{n\sigma^4}\mathbb{E}\Big(\sum_{i\in N\cup\{I\}} e^{C(\xi_i+\xi_i^\bullet+\xi_I)}\Big)^2 \le \frac{C}{n\sigma^4}\mathbb{E}\sum_{i\in N\cup\{I\}} e^{C(\xi_i+\xi_i^\bullet+\xi_I)}.\end{aligned}$$

Similarly as before,

$$\mathbb{E}\sum_{i\in N\cup\{I\}} e^{C(\xi_i+\xi_i^\bullet+\xi_I)} \le C\mathbb{E}\{(1+|N|)e^{C\xi_I}\} \le C. \tag{5.9}$$

Thus, $\mathbb{E}\{\Delta^4\mathbf{1}_A\} \le \frac{C}{n\sigma^4}$. In very much the same way, we deduce that $\mathbb{E}\{G^4\mathbf{1}_A\} \le \frac{Cn^3}{\sigma^4}$.

Bounding the error terms—$G'\mathbf{1}_A$ and $\Delta'\mathbf{1}_A$. Similarly as for Δ^4,

$$\Delta'^4 \le \frac{C}{\sigma^4}\sum_{i\in N'\cup\{J\}} e^{C(\xi_i+\xi_J)}.$$

Let us now prove first that

$$\begin{aligned}&\mathbb{P}\big[A|J,(\iota_i')_{1\le i\le n},(\xi_i)_{1\le i\le n},(\xi_i^\bullet)_{1\le i\le n}\big]\\ &\quad\le \frac{C}{n}\Big(1+\xi_J+\frac{1}{m}\Big(1+\sum_{i\in N'\cup\{J\}}\xi_i\Big)\sum_{i\notin N'\cup\{J\}}(\xi_i+\xi_i^\bullet)^2\Big).\end{aligned} \tag{5.10}$$

To this end, let $\mathcal{G} = \sigma(J, (\iota_i')_{1\le i\le n}, (\xi_i)_{1\le i\le n}, (\xi_i^\bullet)_{1\le i\le n})$. We have

$$\begin{aligned}\mathbb{P}[A|\mathcal{G}] &= \frac{1}{n}\sum_{i=1}^n \mathbb{P}[A|I=i,\mathcal{G}] = \frac{1+|N'|}{n} + \frac{1}{n}\sum_{i\notin N'\cup\{J\}} \mathbb{P}[A|I=i,\mathcal{G}]\\ &\le \frac{1+\xi_J}{n} + \frac{1}{n}\sum_{i\notin N'\cup\{J\}} \mathbb{E}\{|M\cap(N'\cup\{J\})||I=i,\mathcal{G}\}\\ &\le \frac{1+\xi_J}{n} + \frac{1}{n}\sum_{i\notin N'\cup\{J\}} \mathbb{E}\{|N\cap(N'\cup\{J\})| + |N^\bullet\cap(N'\cup\{J\})||I=i,\mathcal{G}\}.\end{aligned}$$

Now, for $i\notin N'\cup\{J\}$,

$$\begin{aligned}&\mathbb{E}\{|N\cap(N'\cup\{J\})||I=i,\mathcal{G}\}\\ &\quad\le \mathbb{E}\Biggl\{\sum_{k=1}^{\xi_I}\mathbf{1}[\iota_k\in N'\cup\{J\}]\Biggm| I=i,\mathcal{G}\Biggr\} = \xi_i\sum_{k\in N'\cup\{J\}}\frac{p_k}{1-p_i} \le \frac{C\xi_i(1+\xi_J)}{m},\end{aligned}$$

and

$$\begin{aligned}&\mathbb{E}\{|N^\bullet\cap(N'\cup\{J\})||I=i,\mathcal{G}\}\\ &\quad\le \mathbb{E}\Biggl\{\sum_{k=1}^{|\chi-\chi^\bullet|}\mathbf{1}[\iota_k^*\in N'\cup\{J\}]\Biggm| I=i,\mathcal{G}\Biggr\}\\ &\quad\le \mathbb{E}\Biggl\{\sum_{k=1}^{|\chi-\chi^\bullet|}\Biggl(\sum_{l\in N'\cup\{J\}}\frac{p_l}{1-\sum_{j\in N\cup\{I\}}p_j} + \frac{\sum_{l\in N'\cup\{J\}}\xi_l}{m-\chi-k+1}\Biggr)\Biggm| I=i,\mathcal{G}\Biggr\}\\ &\quad\le \mathbb{E}\Biggl\{\sum_{k=1}^{|\chi-\chi^\bullet|}\Biggl((1+\xi_J)\mathbf{1}\Biggl[\xi_i+1\le\frac{m}{2K_2}\Biggr]\frac{2K_2}{m}\\ &\qquad + \mathbf{1}\Biggl[\xi_i+1>\frac{m}{2K_2}\Biggr] + \frac{\sum_{l\in N'\cup\{J\}}\xi_l}{m-\chi-k+1}\Biggr)\Biggm| I=i,\mathcal{G}\Biggr\}\\ &\quad\le \frac{C(\xi_i+\xi_J+1)}{m}\mathbb{E}\{|\chi-\chi^\bullet||I=i,\mathcal{G}\}\\ &\qquad + \Biggl(\sum_{k\in N'\cup\{J\}}\xi_k\Biggr)\mathbb{E}\Biggl\{\sum_{k=1}^{|\chi-\chi^\bullet|}\frac{1}{m-\chi-k+1}\Biggm| I=i,\mathcal{G}\Biggr\}.\end{aligned}$$

First observe that

$$\begin{aligned}&\mathbb{E}\{|\chi-\chi^\bullet||I=i,\mathcal{G}\}\\ &\quad\le \mathbb{E}\{\chi+\chi^\bullet|I=i,\mathcal{G}\} = \mathbb{E}\Biggl\{\sum_{k\in N\cup\{I\}}(\xi_k+\xi_k^\bullet)\Biggm| I=i,\mathcal{G}\Biggr\}\\ &\quad= \xi_i+\xi_i^\bullet+\sum_{k\ne i}(\xi_k+\xi_k^\bullet)\mathbb{E}\Biggl\{\sum_{l=1}^{\xi_i}\mathbf{1}[\iota_l=k]\Biggm| I=i,\mathcal{G}\Biggr\}\\ &\quad= \xi_i+\xi_i^\bullet+\sum_{k\ne i}(\xi_k+\xi_k^\bullet)\sum_{l=1}^{\xi_i}\frac{p_k}{1-p_i} \le \xi_i+\xi_i^\bullet+\sum_{k\ne i}(\xi_k+\xi_k^\bullet)\sum_{l=1}^{\xi_i}\frac{2K_2}{m} \le C(\xi_i+\xi_i^\bullet)\end{aligned}$$

and

$$\begin{aligned}
&\mathbb{E}\{(\chi+\chi^\bullet)^2|I=i,\mathcal{G}\}\\
&\quad\le \mathbb{E}\Big\{\sum_{j\in N\cup\{I\}}\sum_{k\in N\cup\{I\}}(\xi_j+\xi_j^\bullet)(\xi_k+\xi_k^\bullet)\Big|I=i,\mathcal{G}\Big\}\\
&\quad=(\xi_i+\xi_i^\bullet)^2+2(\xi_i+\xi_i^\bullet)\mathbb{E}\Big\{\sum_{j\in N}(\xi_j+\xi_j^\bullet)+\sum_{j\in N}\sum_{k\in N}(\xi_j+\xi_j^\bullet)(\xi_k+\xi_k^\bullet)\Big|I=i,\mathcal{G}\Big\}.
\end{aligned}$$

Now,

$$\begin{aligned}
&\mathbb{E}\Big\{\sum_{j\in N}(\xi_j+\xi_j^\bullet)\Big|I=i,\mathcal{G}\Big\}\\
&\quad\le \mathbb{E}\Big\{\sum_{j\ne I}(\xi_j+\xi_j^\bullet)\Big(\sum_{l=1}^{\xi_I}\mathbf{1}[\iota_l=j]\Big)\Big|I=i,\mathcal{G}\Big\}\\
&\quad=\sum_{j\ne i}(\xi_j+\xi_j^\bullet)\mathbb{E}\Big\{\sum_{l=1}^{\xi_I}\mathbf{1}[\iota_l=j]\Big|I=i,\mathcal{G}\Big\}=\sum_{j\ne i}(\xi_j+\xi_j^\bullet)\sum_{l=1}^{\xi_i}\frac{p_j}{1-p_i}\\
&\quad=\sum_{j\ne i}(\xi_j+\xi_j^\bullet)\sum_{l=1}^{\xi_i}\frac{2K_2}{m}\le 2K_2\xi_i,
\end{aligned}$$

and

$$\begin{aligned}
&\mathbb{E}\Big\{\sum_{j\in N}\sum_{k\in N}(\xi_j+\xi_j^\bullet)(\xi_k+\xi_k^\bullet)\Big|I=i,\mathcal{G}\Big\}\\
&\quad\le \mathbb{E}\Big\{\sum_{j\ne I}\sum_{k\ne I}(\xi_j+\xi_j^\bullet)(\xi_k+\xi_k^\bullet)\sum_{l=1}^{\xi_I}\mathbf{1}[\iota_l=j]\sum_{u=1}^{\xi_I}\mathbf{1}[\iota_u=k]\Big|I=i,\mathcal{G}\Big\}\\
&\quad=\sum_{j\ne i}\sum_{k\ne i}(\xi_j+\xi_j^\bullet)(\xi_k+\xi_k^\bullet)\mathbb{E}\Big\{\sum_{l=1}^{\xi_I}\mathbf{1}[\iota_l=j]\sum_{u=1}^{\xi_I}\mathbf{1}[\iota_u=k]\Big|I=i,\mathcal{G}\Big\}\\
&\quad\le \sum_{j\ne i}\sum_{k\ne i}(\xi_j+\xi_j^\bullet)(\xi_k+\xi_k^\bullet)\mathbb{E}\Big\{\sum_{l=1}^{\xi_I}\mathbf{1}[\iota_l=j]\Big|I=i,\mathcal{G}\Big\}\\
&\qquad\times\mathbb{E}\Big\{\sum_{u=1}^{\xi_I}\mathbf{1}[\iota_u=k]\Big|I=i,\mathcal{G}\Big\}\\
&\quad\le \sum_{j\ne i}\sum_{k\ne i}(\xi_j+\xi_j^\bullet)(\xi_k+\xi_k^\bullet)\Big(\xi_i\frac{p_j}{1-p_i}\Big)^2\\
&\quad\le \sum_{j\ne i}\sum_{k\ne i}(\xi_j+\xi_j^\bullet)(\xi_k+\xi_k^\bullet)\Big(\xi_i\frac{2K_2}{m}\Big)^2\le 16K_2^2\xi_i^2,
\end{aligned}$$

so that

$$\mathbb{E}\{(\chi+\chi^\bullet)^2|I=i,\mathcal{G}\}\le(\xi_i+\xi_i^\bullet)^2+2(\xi_i+\xi_i^\bullet)(2K_2\xi_i+16K_2^2\xi_i^2)\le C(\xi_i+\xi_i^\bullet)^2.$$

Now, since $a + |a - b| \le 2(a + b)$ for nonnegative numbers a and b, we have

$$\begin{aligned}
&\mathbb{E}\biggl\{\sum_{k=1}^{|\chi-\chi^\bullet|}\frac{1}{m-\chi-k+1}\Big|I=i,\mathcal{G}\biggr\}\\
&\quad\le \mathbb{E}\biggl\{\frac{|\chi-\chi^\bullet|}{m-m/2}+2\log(m+1)\mathbf{1}[2(\chi+\chi^\bullet)>m/2]\Big|I=i,\mathcal{G}\biggr\}\\
&\quad\le \frac{C}{m}\bigl(\mathbb{E}\{|\chi-\chi^\bullet||I=i,\mathcal{G}\}\mathbb{E}\{(\chi+\chi^\bullet)^2|I=i,\mathcal{G}\}\bigr)\le\frac{C(\xi_i+\xi_i^\bullet)^2}{m}.
\end{aligned}$$

This leads to

$$\begin{aligned}
&\mathbb{E}\{|N^\bullet\cap(N'\cup\{J\})||I=i,\mathcal{G}\}\\
&\quad\le C\biggl(\frac{\xi_J+1}{m}+\frac{\xi_i+1}{m}\biggr)(\xi_i+\xi_i^\bullet)+C\biggl(\sum_{k\in N'\cup\{J\}}\xi_k\biggr)\frac{1}{m}(\xi_i+\xi_i^\bullet)^2\\
&\quad\le \frac{C}{m}\biggl[(\xi_J+\xi_i+1)(\xi_i+\xi_i^\bullet)+\biggl(\sum_{k\in N'\cup\{J\}}\xi_k\biggr)(\xi_i+\xi_i^\bullet)^2\biggr]\\
&\quad\le \frac{C}{m}\biggl[\xi_J(\xi_i+\xi_i^\bullet)+(\xi_i+\xi_i^\bullet)^2+\biggl(\sum_{k\in N'\cup\{J\}}\xi_k\biggr)(\xi_i+\xi_i^\bullet)^2\biggr]\\
&\quad\le \frac{C}{m}\biggl(1+\sum_{k\in N'\cup\{J\}}\xi_k\biggr)(\xi_i+\xi_i^\bullet)^2.
\end{aligned}$$

Putting all together, (5.10) follows. Hence,

$$\begin{aligned}
&\mathbb{E}\{\Delta'^4\mathbf{1}_A\}\\
&\quad\le\frac{C}{n\sigma^4}\mathbb{E}\biggl\{\biggl(\sum_{i\in N'\cup\{J\}}e^{C(\xi_i+\xi_J)}\biggr)\times\biggl(1+\xi_J+\frac{1}{m}\biggl(1+\sum_{i\in N'\cup\{J\}}\xi_i\biggr)\sum_{i\notin N'\cup\{J\}}(\xi_i+\xi_i^\bullet)^2\biggr)\biggr\}\\
&\quad\le\frac{C}{n\sigma^4}\mathbb{E}\biggl\{\biggl(\sum_{i\in N'\cup\{J\}}e^{C(\xi_i+\xi_J)}\biggr)\times\biggl(e^{\xi_J}+\frac{1}{m}\biggl(\sum_{i\in N'\cup\{J\}}e^{\xi_i}\biggr)\sum_{i\notin N'\cup\{J\}}e^{2(\xi_i+\xi_i^\bullet)}\biggr)\biggr\}.
\end{aligned}$$

Now, similarly as in (5.9),

$$\mathbb{E}\biggl\{e^{\xi_J}\sum_{i\in N'\cup\{J\}}e^{C(\xi_i+\xi_J)}\biggr\}\le C.$$

Moreover

$$\begin{aligned}
&\frac{1}{m}\mathbb{E}\biggl\{\biggl(\sum_{i\in N'\cup\{J\}}e^{C(\xi_i+\xi_J)}\biggr)\biggl(\sum_{i\in N'\cup\{J\}}e^{\xi_i}\biggr)\biggl(\sum_{i\notin N'\cup\{J\}}e^{2(\xi_i+\xi_i^\bullet)}\biggr)\biggr\}\\
&\quad\le\frac{1}{m}\mathbb{E}\biggl\{\biggl(\sum_{i\in N'\cup\{J\}}e^{C'(\xi_i+\xi_J)}\biggr)\biggl(\sum_{i\notin N'\cup\{J\}}e^{2(\xi_i+\xi_i^\bullet)}\biggr)\biggr\}\\
&\quad\le\frac{1}{m}\mathbb{E}\biggl\{\mathbb{E}\biggl(\sum_{i\in N'\cup\{J\}}e^{C'(\xi_i+\xi_J)}\Big|\xi_J\biggr)\mathbb{E}\biggl(\sum_{i\notin N'\cup\{J\}}e^{2(\xi_i+\xi_i^\bullet)}\Big|\xi_J\biggr)\biggr\}\le\frac{Cn}{m}\le C,
\end{aligned}$$

where for the second last inequality we used the fact that conditionally on J and ξ_J, the family $(\xi_i)_{i\ne J}$ is a urn model and hence negatively associated and we can therefore apply Lemma 5.8, and where for the last inequality we proceeded in the same way as in (5.9). So,

putting all together we obtain $\mathbb{E}\{\Delta'^4\mathbf{1}_A\} \le C/(\sigma^4 n)$. In very much the same way we deduce that $\mathbb{E}\{G'^4\mathbf{1}_A\} \le Cn^3/\sigma^4$.

Bounding the error terms—$G^*\mathbf{1}_A$ and $\Delta^*\mathbf{1}_A$. Similarly as for Δ^4, and using that $\xi_i^* \le \xi_i + \xi_i^\bullet + \chi$, we have

$$\Delta^{*4} \le \frac{C}{\sigma^4} \sum_{i \in N^* \cup \{J\}} e^{C(\xi_i^* + \xi_J^*)} \le e^{C\chi} \sum_{i \in N^* \cup \{J\}} e^{C(\xi_i + \xi_i^\bullet)}. \tag{5.11}$$

Let $\mathcal{G} = \sigma(I, (\iota_i)_{i \ge 1}, (\iota_i^*)_{i \ge 1}, (\xi_i)_{1 \le i \le n}, (\xi_i^\bullet)_{1 \le i \le n})$; we have

$$\mathbb{E}\biggl\{ \sum_{i \in N^* \cup \{J\}} e^{C(\xi_i + \xi_i^\bullet)} \mathbf{1}_A \biggm| \mathcal{G} \biggr\} = \frac{1}{n} \sum_{j=1}^n \mathbb{E}\biggl\{ \sum_{i \in N^* \cup \{J\}} e^{C(\xi_i + \xi_i^\bullet)} \mathbf{1}_A \biggm| J = j, \mathcal{G} \biggr\}. \tag{5.12}$$

For $j \in M$ we have

$$\begin{aligned}
&\mathbb{E}\biggl\{ \sum_{i \in N^* \cup \{J\}} e^{C(\xi_i + \xi_i^\bullet)} \mathbf{1}_A \biggm| J = j, \mathcal{G} \biggr\} \\
&\quad \le e^{C(\xi_j + \xi_j^\bullet)} + \mathbb{E}\biggl\{ \sum_{i \ne J} e^{C(\xi_i + \xi_i^\bullet)} \sum_{k=1}^{\xi_J^*} \mathbf{1}[\iota_k' = i] \biggm| J = j, \mathcal{G} \biggr\} \\
&\quad \le e^{C(\xi_j + \xi_j^\bullet)} + \mathbb{E}\biggl\{ \sum_{i \ne J} e^{C(\xi_i + \xi_i^\bullet)} \sum_{k=1}^{\xi_J^*} \frac{p_i}{1 - p_J} \biggm| J = j, \mathcal{G} \biggr\} \\
&\quad \le e^{C(\xi_j + \xi_j^\bullet)} + \frac{2K_2}{m} \sum_{i \ne j} e^{C(\xi_i + \xi_i^\bullet)} \xi_j^*.
\end{aligned} \tag{5.13}$$

For $j \notin M$, we have

$$\begin{aligned}
&\mathbb{E}\biggl\{ \sum_{i \in N^* \cup \{J\}} e^{C(\xi_i + \xi_i^\bullet)} \mathbf{1}_A \biggm| J = j, \mathcal{G} \biggr\} \\
&\quad \le \mathbb{E}\biggl\{ |M \cap N^*| \sum_{i \in N^* \cup \{J\}} e^{C(\xi_i + \xi_i^\bullet)} \biggm| J = j, \mathcal{G} \biggr\} \\
&\quad \le \mathbb{E}\biggl\{ \sum_{k=1}^{\xi_J^*} \mathbf{1}[\iota_k' \in M] \Bigl(e^{C(\xi_J + \xi_J^\bullet)} + \sum_{i \ne J} e^{C(\xi_i + \xi_i^\bullet)} \sum_{k=1}^{\xi_J^*} \mathbf{1}[\iota_k' = i] \Bigr) \biggm| J = j, \mathcal{G} \biggr\} \\
&\quad \le \frac{2K_2|M|}{m} \xi_j^* e^{C(\xi_j + \xi_j^\bullet)} + \mathbb{E}\biggl\{ \xi_J^* \sum_{i \in M} e^{C(\xi_i + \xi_i^\bullet)} \sum_{k=1}^{\xi_J^*} \mathbf{1}[\iota_k' = i] \biggm| J = j, \mathcal{G} \biggr\} \\
&\qquad + \mathbb{E}\biggl\{ \sum_{i \ne M \cup \{J\}} e^{C(\xi_i + \xi_i^\bullet)} \sum_{k=1}^{\xi_J^*} \mathbf{1}[\iota_k' \in M] \sum_{k=1}^{\xi_J^*} \mathbf{1}[\iota_k' = i] \biggm| J = j, \mathcal{G} \biggr\} \\
&\quad \le \frac{2K_2|M|}{m} \xi_j^* e^{C(\xi_j + \xi_j^\bullet)} + \xi_j^* \sum_{i \in M} e^{C(\xi_i + \xi_i^\bullet)} \xi_j^* \frac{2K_2}{m} \\
&\qquad + \sum_{i \ne M \cup \{j\}} e^{C(\xi_i + \xi_i^\bullet)} \frac{2K_2|M|\xi_j^*}{m} \frac{2K_2 \xi_j^*}{m},
\end{aligned} \tag{5.14}$$

where in the last inequality we used that the indicators $(\mathbf{1}[\iota'_k \in M], (\mathbf{1}[\iota'_k = i])$ are negatively associated whenever $i \notin M$. Combining (5.13) and (5.14) with (5.12), we obtain

$$\begin{aligned}
&\mathbb{E}\Big\{\sum_{i\in N^*\cup\{J\}} e^{C(\xi_i+\xi_i^\bullet)}\mathbf{1}_A\Big|\mathcal{G}\Big\}\\
&\le \frac{C}{n}\sum_{j=1}^{n}\Big(e^{C(\xi_j+\xi_j^\bullet)} + \frac{1}{m}\sum_{i\neq j}\xi_j^* e^{C(\xi_i+\xi_i^\bullet)} + \frac{|M|}{m}\xi_j^* e^{C(\xi_j+\xi_j^\bullet)}\\
&\quad + \frac{1}{m}\sum_{i\in M}\xi_j^{*2} e^{C(\xi_i+\xi_i^\bullet)} + \frac{|M|}{m^2}\sum_{i\neq M\cup\{j\}}\xi_j^{*2} e^{C(\xi_i+\xi_i^\bullet)}\Big).
\end{aligned} \tag{5.15}$$

Using the inequality $1 \le \frac{1}{m}\sum_{i=1}^n \xi_i \le \frac{1}{m}\sum_{i=1}^n e^{\xi_i}$ and $\xi_i^* \le \xi_i + \xi_i^\bullet + \chi$, it is not difficult to further coarsen the bound in (5.15) to

$$\mathbb{E}\Big\{\sum_{i\in N^*\cup\{J\}} e^{C(\xi_i+\xi_i^\bullet)}\mathbf{1}_A\Big|\mathcal{G}\Big\} \le \frac{C}{nm^2}\sum_{i=1}^{n}\sum_{k=1}^{n}\sum_{j\in M} e^{C(\xi_i+\xi_j+\xi_k+\xi_i^\bullet+\xi_j^\bullet+\xi_k^\bullet+\chi)}.$$

Combining this with (5.11), we have

$$\mathbb{E}\{\Delta^{*4}\mathbf{1}_A\} \le \frac{C}{\sigma^4 nm^2}\sum_{k=1}^{n}\sum_{l=1}^{n}\mathbb{E}\sum_{u\in M} e^{C(\xi_k+\xi_l+\xi_u+\xi_k^\bullet+\xi_l^\bullet+\xi_u^\bullet+\chi)}.$$

Now, let $\mathcal{G} = \sigma(I, (\iota_u)_{u\ge 1}, (\xi_u)_{1\le u\le n}, (\xi_u^\bullet)_{1\le u\le n})$; we have

$$\begin{aligned}
&\mathbb{E}\Big\{\sum_{u\in N^\bullet} e^{C(\xi_u+\xi_u^\bullet)}\Big|\mathcal{G}\Big\}\\
&\le \mathbb{E}\Big\{\sum_{u\notin N\cup\{I\}} e^{C(\xi_u+\xi_u^\bullet)}\sum_{w=1}^{|\chi-\chi^\bullet|}\mathbf{1}[\iota_w^* = u]\Big|\mathcal{G}\Big\}\\
&\le \mathbb{E}\Big\{\sum_{u\notin N\cup\{I\}} e^{C(\xi_u+\xi_u^\bullet)}\sum_{w=1}^{|\chi-\chi^\bullet|}\Big(\frac{p_u}{1-\sum_{x\in N\cup\{I\}}p_x} + \frac{\xi_u}{m-\chi-w+1}\Big)\Big|\mathcal{G}\Big\}\\
&\le \mathbb{E}\Big\{\sum_{u\notin N\cup\{I\}} e^{C(\xi_u+\xi_u^\bullet)}\Big[|\chi-\chi^\bullet|\Big(\frac{2K_2}{m} + \mathbf{1}\Big[\xi_I + 1 > \frac{m}{2K_2}\Big]\Big)\\
&\quad + \frac{\xi_u|\chi-\chi^\bullet|}{m-m/2} + 2\log(m+1)\mathbf{1}[2(\chi+\chi^\bullet) > m/2]\Big]\Big|\mathcal{G}\Big\}\\
&\le \frac{C}{m}\mathbb{E}\Big\{\sum_{u\notin N\cup\{I\}} e^{C(\xi_u+\xi_u^\bullet)}\big((\chi+\chi^\bullet)(\xi_I+1) + \xi_u(\chi+\chi^\bullet) + (\chi+\chi^\bullet)^2\big)\Big|\mathcal{G}\Big\}\\
&\le \frac{C}{m}\sum_{u\notin N\cup\{I\}} e^{C(\xi_u+\xi_u^\bullet+\chi+\chi^\bullet)}.
\end{aligned}$$

Thus, since M is the disjoint union of $N\cup\{I\}$ and $N^\bullet$, we obtain

$$\begin{aligned}
\mathbb{E}\Big\{\sum_{u\in M} e^{C(\xi_k+\xi_l+\xi_u+\xi_k^\bullet+\xi_l^\bullet+\xi_u^\bullet+\chi)}\Big\} &\le \frac{C}{m}\mathbb{E}\Big\{\sum_{u\notin N\cup\{I\}} e^{C(\xi_k+\xi_l+\xi_u+\xi_k^\bullet+\xi_l^\bullet+\xi_u^\bullet+\chi+\chi^\bullet)}\Big\}\\
&\quad + \mathbb{E}\Big\{\sum_{u\in N\cup\{I\}} e^{C(\xi_k+\xi_l+\xi_u+\xi_k^\bullet+\xi_l^\bullet+\xi_u^\bullet+\chi)}\Big\}.
\end{aligned}$$

Now, conditioning on I, ξ_I and $(\iota_i)_{i\geq 1}$, we can apply Lemma 5.8 and then Lemma 5.7 so that, for example,

$$\mathbb{E}\{e^{C(\xi_k+\xi_l+\xi_u+\xi_k^\bullet+\xi_l^\bullet+\xi_u^\bullet+\chi)}|I,\xi_I,(\iota_i)_{i\geq 1}\}\leq C\cdot C^{\xi_I}\cdot e^{C\xi_I}\leq Ce^{C\xi_I}.$$

Hence,

$$\begin{aligned}\mathbb{E}\Big\{\sum_{u\in M}e^{C(\xi_k+\xi_l+\xi_u+\xi_k^\bullet+\xi_l^\bullet+\xi_u^\bullet+\chi)}\Big\}&\leq\frac{C}{m}\mathbb{E}\Big\{\sum_{u\notin N\cup\{I\}}e^{C\xi_I}\Big\}+C\mathbb{E}\Big\{\sum_{u\in N\cup\{I\}}e^{C\xi_I}\Big\}\\&\leq C\Big(\frac{n}{m}+1\Big)\mathbb{E}e^{C\xi_I}\leq C,\end{aligned}$$

which finally leads to

$$\mathbb{E}\{\Delta^{*4}\mathbf{1}_A\}\leq\frac{C}{\sigma^4nm^2}\sum_{k=1}^n\sum_{l=1}^n C\leq\frac{C}{\sigma^4n}.$$

Again, in very much the same way we can prove that $\mathbb{E}\{G^{*4}\mathbf{1}_A\}\leq Cn^3/\sigma^4$.

Combining the bounds. Collecting the bounds we need in order to employ Lemma 5.6 below, we have

$$g_0\leq\frac{Cn}{\sigma},\qquad d_0\leq\frac{C}{\sigma},\qquad\mathbb{P}[A]\leq\frac{C}{n},\qquad g_1,g_2,g_3\leq\frac{Cn^{3/4}}{\sigma},\qquad d_1,d_2,d_3\leq\frac{C}{\sigma n^{1/4}};$$

this leads to

$$s_1\leq\frac{Cn^{1/2}}{\sigma^2},\qquad s_2,s_3,s_4\leq\frac{Cn}{\sigma^3}.$$

Applying these bounds to Theorem 5.1 proves the claim. □

5.2.1. *Technical lemmas.* The following lemma is straightforward to prove.

LEMMA 5.6. *Consider the setting of Theorem 5.1. Let A be an event such that $\{G'\neq G^*\}\cup\{\Delta'\neq\Delta^*\}\subset A$. With*

$$\begin{aligned}g_0^3&=\mathbb{E}|G|^3,\qquad g_1^4=\mathbb{E}\{G^4\mathbf{1}_A\},\qquad g_2^4=\mathbb{E}\{G'^4\mathbf{1}_A\},\qquad g_3^4=\mathbb{E}\{G^{*4}\mathbf{1}_A\},\\d_0^3&=\mathbb{E}|\Delta|^3,\qquad d_1^4=\mathbb{E}\{\Delta^4\mathbf{1}_A\},\qquad d_2^4=\mathbb{E}\{\Delta'^4\mathbf{1}_A\},\qquad d_3^4=\mathbb{E}\{\Delta^{*4}\mathbf{1}_A\},\end{aligned}$$

we have

$$\begin{aligned}s_1&\leq[g_1g_2d_1d_2+2g_1g_2d_1d_3+g_1g_3d_1d_3]^{1/2},\qquad s_2,s_3\leq g_0d_0^2,\\s_4&\leq\big(g_1g_2d_2+g_1g_2d_3+g_1g_2(d_2\wedge d_3)+g_1g_3(d_2\wedge d_3)\big)\mathbb{P}[A]^{1/4}.\end{aligned}$$

LEMMA 5.7. *Let $Z\sim\mathrm{Bi}(r,q)$ with $r\leq m$ and $q\leq c/m$. Then for any $a,b\geq 0$,*

$$\mathbb{E}\big(Z^ae^{bZ}\big)\leq C(a,b,c).$$

PROOF. This easily follows from

$$\mathbb{E}\big(Z^ae^{bZ}\big)\leq\mathbb{E}e^{(a+b)Z}=\exp\big(r\log\big(qe^{a+b}+1-q\big)\big)\leq\exp\big(rqe^{a+b}\big)\leq\exp\big(ce^{a+b}\big).\quad\square$$

LEMMA 5.8. *The family of random variables $(\xi_1,\dots,\xi_n)$ is negatively associated; that is, for nondecreasing functions f and g and disjoint sets $A,B\subset\{1,\dots,n\}$,*

$$\mathbb{E}\big\{f\big((\xi_i)_{i\in A}\big)g\big((\xi_i)_{i\in B}\big)\big\}\leq\mathbb{E}f\big((\xi_i)_{i\in A}\big)\mathbb{E}g\big((\xi_i)_{i\in B}\big).$$

LEMMA 5.9. *Let a', a^*, b, b' and b^* be real numbers. With $\bar{x} = (x \wedge 1) \vee (-1)$,*

$$\begin{aligned}
&(i) \quad |a'\bar{b}' - a^*\bar{b}^*| \\
&\qquad \le |a'|(|b' - b^*| \wedge 2) + |a' - a^*|(|b^*| \wedge 1), \\
&(ii) \quad |a'(|b| \wedge |b'| \wedge 1)\mathbf{1}[bb' > 0] - a^*(|b| \wedge |b^*| \wedge 1)\mathbf{1}[bb^* > 0]| \\
&\qquad \le |a'|(|b' - b^*| \wedge 1) + |a' - a^*|(|b| \wedge |b^*| \wedge 1), \\
&(iii) \quad |a'(|b| \wedge |b'| \wedge 1)^2\mathbf{1}[bb' > 0] - a^*(|b| \wedge |b^*| \wedge 1)^2\mathbf{1}[bb^* > 0]| \\
&\qquad \le 2|a'|(|b| \wedge 1)(|b' - b^*| \wedge 1) + |a' - a^*|(|b| \wedge |b'| \wedge 1)^2.
\end{aligned}$$

PROOF. We have

$$\begin{aligned}
|a'\bar{b}' - a^*\bar{b}^*| &\le |a'\bar{b}' - a'\bar{b}^*| + |a'\bar{b}^* - a^*\bar{b}^*| \le |a'||\bar{b}' - \bar{b}^*| + |a' - a^*||\bar{b}^*| \\
&\le |a'|(|b' - b^*| \wedge 2) + |a' - a^*|(|b^*| \wedge 1),
\end{aligned}$$

which proves (i). Moreover, by considering eight possible combinations of the signs of b, b' and b^*, we can conclude that

$$|(|b| \wedge |b'| \wedge 1)\mathbf{1}[bb' > 0] - (|b| \wedge |b^*| \wedge 1)\mathbf{1}[bb^* > 0]| \le |b' - b^*| \wedge 1. \tag{5.16}$$

This in turn ensures that

$$\begin{aligned}
&|a'(|b| \wedge |b'| \wedge 1)\mathbf{1}[bb' > 0] - a^*(|b| \wedge |b^*| \wedge 1)\mathbf{1}[bb^* > 0]| \\
&\quad \le |a'| \cdot |(|b| \wedge |b'| \wedge 1)\mathbf{1}[bb' > 0] - (|b| \wedge |b^*| \wedge 1)\mathbf{1}[bb^* > 0]| \\
&\qquad + |a' - a^*| \cdot (|b| \wedge |b^*| \wedge 1)\mathbf{1}[bb^* > 0] \\
&\quad \le |a'|(|b' - b^*| \wedge 1) + |a' - a^*|(|b| \wedge |b^*| \wedge 1),
\end{aligned}$$

as claimed in (ii). Last, using $|x^2 - y^2| = |x + y| \cdot |x - y|$ and (5.16) in the second inequality below, we obtain

$$\begin{aligned}
&|a'(|b| \wedge |b'| \wedge 1)^2\mathbf{1}[bb' > 0] - a^*(|b| \wedge |b^*| \wedge 1)^2\mathbf{1}[bb^* > 0]| \\
&\quad \le |a'| \cdot |(|b| \wedge |b'| \wedge 1)^2\mathbf{1}[bb' > 0] - (|b| \wedge |b^*| \wedge 1)^2\mathbf{1}[bb^* > 0]| \\
&\qquad + |a' - a^*| \cdot |(|b| \wedge |b^*| \wedge 1)^2\mathbf{1}[bb^* > 0]| \\
&\quad \le |a'|[(|b| \wedge |b'| \wedge 1) + (|b| \wedge |b^*| \wedge 1)](|b' - b^*| \wedge 1) \\
&\qquad + |a' - a^*|(|b| \wedge |b^*| \wedge 1)^2 \\
&\quad \le 2|a'|(|b| \wedge 1)(|b' - b^*| \wedge 1) + |a' - a^*|(|b| \wedge |b^*| \wedge 1)^2,
\end{aligned}$$

which proves (ii). □

Acknowledgments. We thank the Institute of Mathematical Sciences, NUS, for supporting the workshop *Workshop on New Directions in Stein's Method* in March 2015, during which part of this research was conducted.

Funding. This research was supported by the ARC Discovery Grants DP150101459 and DP190100613, the Singapore Ministry of Education Academic Research Fund Tier 2 Grant MOE2018-T2-2-076, and Singapore Ministry of Education Academic Research Fund Tier 1 Grants R-146-000-182-112 and R-146-000-230-114.

REFERENCES

AVRAM, F. and BERTSIMAS, D. (1993). On central limit theorems in geometrical probability. *Ann. Appl. Probab.* **3** 1033–1046. MR1241033

BARBOUR, A. D., HOLST, L. and JANSON, S. (1992). *Poisson Approximation*. *Oxford Studies in Probability* **2**. The Clarendon Press, Oxford University Press, New York. MR1163825

BARBOUR, A. D. and XIA, A. (2006). Normal approximation for random sums. *Adv. in Appl. Probab.* **38** 693–728. MR2256874 https://doi.org/10.1239/aap/1158684998

BARYSHNIKOV, Y. and YUKICH, J. E. (2005). Gaussian limits for random measures in geometric probability. *Ann. Appl. Probab.* **15** 213–253. MR2115042 https://doi.org/10.1214/105051604000000594

BŁASZCZYSZYN, B., YOGESHWARAN, D. and YUKICH, J. E. (2019). Limit theory for geometric statistics of point processes having fast decay of correlations. *Ann. Probab.* **47** 835–895. MR3916936 https://doi.org/10.1214/18-AOP1273

CHEN, L. H. Y., GOLDSTEIN, L. and SHAO, Q.-M. (2011). *Normal Approximation by Stein's Method*. *Probability and Its Applications* (*New York*). Springer, Heidelberg. MR2732624 https://doi.org/10.1007/978-3-642-15007-4

CHEN, L. H. Y. and RÖLLIN, A. (2010). Stein couplings for normal approximation. Available at arXiv:1003.6039v2.

CHEN, L. H. Y. and SHAO, Q.-M. (2004). Normal approximation under local dependence. *Ann. Probab.* **32** 1985–2028. MR2073183 https://doi.org/10.1214/009117904000000450

CHEN, L. H. Y. and XIA, A. (2004). Stein's method, Palm theory and Poisson process approximation. *Ann. Probab.* **32** 2545–2569. MR2078550 https://doi.org/10.1214/009117904000000027

DALEY, D. J. and VERE-JONES, D. (2003). *An Introduction to the Theory of Point Processes*. *Vol. I*: *Elementary Theory and Methods*, 2nd ed. *Probability and Its Applications* (*New York*). Springer, New York. MR1950431

DALEY, D. J. and VERE-JONES, D. (2008). *An Introduction to the Theory of Point Processes*. *Vol. II*: *General Theory and Structure*, 2nd ed. *Probability and Its Applications* (*New York*). Springer, New York. MR2371524 https://doi.org/10.1007/978-0-387-49835-5

DEMBO, A. and RINOTT, Y. (1996). Some examples of normal approximations by Stein's method. In *Random Discrete Structures* (*Minneapolis*, *MN*, 1993). *IMA Vol. Math. Appl.* **76** 25–44. Springer, New York. MR1395606 https://doi.org/10.1007/978-1-4612-0719-1_3

ENGLUND, G. (1981). A remainder term estimate for the normal approximation in classical occupancy. *Ann. Probab.* **9** 684–692. MR0624696

GINIBRE, J. (1965). Statistical ensembles of complex, quaternion, and real matrices. *J. Math. Phys.* **6** 440–449. MR0173726 https://doi.org/10.1063/1.1704292

GOLDMAN, A. (2010). The Palm measure and the Voronoi tessellation for the Ginibre process. *Ann. Appl. Probab.* **20** 90–128. MR2582643 https://doi.org/10.1214/09-AAP620

GOLDSTEIN, L. and RINOTT, Y. (1996). Multivariate normal approximations by Stein's method and size bias couplings. *J. Appl. Probab.* **33** 1–17. MR1371949 https://doi.org/10.1017/s0021900200103675

GOLDSTEIN, L. and XIA, A. (2006). Zero biasing and a discrete central limit theorem. *Ann. Probab.* **34** 1782–1806. MR2271482 https://doi.org/10.1214/009117906000000250

HSING, T. and LEADBETTER, M. R. (1998). On the excursion random measure of stationary processes. *Ann. Probab.* **26** 710–742. MR1626519 https://doi.org/10.1214/aop/1022855648

KALLENBERG, O. (1983). *Random Measures*, 3rd ed. Academic Press, London. MR0818219

KEELER, H. P., ROSS, N. and XIA, A. (2018). When do wireless network signals appear Poisson? *Bernoulli* **24** 1973–1994. MR3757520 https://doi.org/10.3150/16-BEJ917

KINGMAN, J. F. C. (1967). Completely random measures. *Pacific J. Math.* **21** 59–78. MR0210185

LACHIÈZE-REY, R., SCHULTE, M. and YUKICH, J. E. (2019). Normal approximation for stabilizing functionals. *Ann. Appl. Probab.* **29** 931–993. MR3910021 https://doi.org/10.1214/18-AAP1405

LE CAËR, G. and HO, J. S. (1990). The Voronoi tesselation generated from eigen-values of complex random matrices. *J. Phys. A*: *Math. Gen.* **23** 3279–3295.

MCGIVNEY, K. and YUKICH, J. E. (1999). Asymptotics for Voronoi tessellations on random samples. *Stochastic Process. Appl.* **83** 273–288. MR1708209 https://doi.org/10.1016/S0304-4149(99)00035-6

MEHTA, M. L. (1991). *Random Matrices*, 2nd ed. Academic Press, Boston, MA. MR1083764

MIYOSHI, N. and SHIRAI, T. (2014a). A cellular network model with Ginibre configured base stations. *Adv. in Appl. Probab.* **46** 832–845. MR3254344 https://doi.org/10.1239/aap/1409319562

MIYOSHI, N. and SHIRAI, T. (2014b). Cellular networks with α-Ginibre configurated base stations. In *The Impact of Applications on Mathematics*. *Math. Ind.* (*Tokyo*) **1** 211–226. Springer, Tokyo. MR3445552 https://doi.org/10.1007/978-4-431-54907-9_16

MØLLER, J. and WAAGEPETERSEN, R. P. (2004). *Statistical Inference and Simulation for Spatial Point Processes*. *Monographs on Statistics and Applied Probability* **100**. CRC Press/CRC, Boca Raton, FL. MR2004226

OSADA, H. and SHIRAI, T. (2016). Absolute continuity and singularity of Palm measures of the Ginibre point process. *Probab. Theory Related Fields* **165** 725–770. MR3520017 https://doi.org/10.1007/s00440-015-0644-6

PENROSE, M. D. (2007). Gaussian limits for random geometric measures. *Electron. J. Probab.* **12** 989–1035. MR2336596 https://doi.org/10.1214/EJP.v12-429

PENROSE, M. D. and YUKICH, J. E. (2001). Central limit theorems for some graphs in computational geometry. *Ann. Appl. Probab.* **11** 1005–1041. MR1878288 https://doi.org/10.1214/aoap/1015345393

RAFLER, M. (2011). The Pólya sum process: A Cox representation. *J. Contemp. Math. Anal.* **46** 338–345.

RAIČ, M. (2003). Normal approximation with Stein's method. In *Proceedings of the Seventh Young Statisticians Meeting*.

RIDER, B. (2004). Order statistics and Ginibre's ensembles. *J. Stat. Phys.* **114** 1139–1148. MR2035641 https://doi.org/10.1023/B:JOSS.0000012520.37908.07

SCHREIBER, T. and YUKICH, J. E. (2013). Limit theorems for geometric functionals of Gibbs point processes. *Ann. Inst. Henri Poincaré Probab. Stat.* **49** 1158–1182. MR3127918 https://doi.org/10.1214/12-AIHP500

SOSHNIKOV, A. (2000). Determinantal random point fields. *Russian Math. Surveys* **55** 923–975.

STEIN, C. (1972). A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability* (*Univ. California, Berkeley, Calif.*, 1970/1971), *Vol. II: Probability Theory* 583–602. MR0402873

TORRISI, G. L. and LEONARDI, E. (2014). Large deviations of the interference in the Ginibre network model. *Stoch. Syst.* **4** 173–205. MR3353217 https://doi.org/10.1214/13-SSY109

XIA, A. and YUKICH, J. E. (2015). Normal approximation for statistics of Gibbsian input in geometric probability. *Adv. in Appl. Probab.* **47** 934–972. MR3433291 https://doi.org/10.1239/aap/1449859795

ZESSIN, H. (2009). Der Papangelou Prozess. *J. Contemp. Math. Anal.* **44** 36–44.