University Library



A gateway to Melbourne's research publications

Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:
Barbour, AD;Braunsteins, P;Ross, N

Title:
Local limit theorems for occupancy models

Date:
2020-09-27

Citation:
Barbour, A. D., Braunsteins, P. & Ross, N. (2020). Local limit theorems for occupancy models. RANDOM STRUCTURES & ALGORITHMS, 58 (1), pp.3-33. https://doi.org/10.1002/rsa.20967.

Persistent Link:
https://hdl.handle.net/11343/276361

Check for updates

Author Manuscript

LOCAL LIMIT THEOREMS FOR OCCUPANCY MODELS

A. D. Barbour*, Peter Braunsteins†, Nathan Ross‡

June 19, 2020

Abstract

We present a rather general method for proving local limit theorems, with a good rate of convergence, for sums of dependent random variables. The method is applicable when a Stein coupling can be exhibited. Our approach involves both Stein's method for distributional approximation and Stein's method for concentration. As applications, we prove local central limit theorems with rate of convergence for the number of germs with d neighbours in a germ–grain model, and the number of degree-d vertices in an Erdős–Rényi random graph. In both cases, the error rate is optimal, up to logarithmic factors.

1 INTRODUCTION

Local central limit theorems (LCLTs) for sums of independent random variables have been well studied, largely using characteristic function techniques; see Petrov (1975, Chapter VII.1). For the standard example, if $X_1, X_2, \ldots$ are i.i.d. aperiodic integer valued random variables with finite third moment, and $W := \sum_{i=1}^n X_i$, with $\mathbb{E}W = \mu$ and $\mathrm{Var}(W) = \sigma^2$, then

$$\sup_k \left| \mathbb{P}(W = k) - \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{ -\frac{1}{2\sigma^2}(k-\mu)^2 \right\} \right| \;=\; \mathrm{O}(1/\sigma^2). \tag{1.1}$$

When the summands X_i are dependent, there are few general methods available for proving LCLTs with error bounds. In this paper, we combine results from Röllin and Ross (2015), Barbour, Ross, and Wen (2018) and Barbour, Röllin, and Ross (2019) to present an approach that is quite widely applicable. We illustrate its power in the context of certain *occupancy models*.

Our random variables of interest take the form

$$\widehat{W}_d \;:=\; \sum_{i=1}^n \mathrm{I}\{M_i = d\}, \quad d \geq 0, \tag{1.2}$$

where n is the number of *locations* and M_i is the number of *occupants* at location $i \in [n] := \{1, \ldots, n\}$ in an occupancy model. Examples of occupancy models include

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/rsa.20967

*Universität Zürich; a.d.barbour@math.uzh.ch
†University of Melbourne; braunsteins.p@unimelb.edu.au
‡University of Melbourne; nathan.ross@unimelb.edu.au

This article is protected by copyright. All rights reserved.

Author Manuscript

1. *multinomial occupancy models*, where m balls (the occupants) are placed in n urns (the locations), independently at random, and M_i is the number of balls in urn i;

2. *Erdős–Rényi random graphs*, where edges (the occupants) between distinct pairs of n vertices (the locations) are independently present or absent, with a common probability p, and M_i is the degree of vertex i;

3. *germ–grain models*, where n points (the occupants), which form the centres of balls of fixed radius (the locations), are placed uniformly at random in a bounded subset of $\mathbb{R}^2$, and M_i is the number of points that fall in ball i.

Here, we consider only sequences of *sparse* occupancy models, in which, as the number of locations n increases, the expected number of occupants at each location remains bounded.

For multinomial occupancy models, Hwang and Janson (2008) prove an LCLT with optimal error rate $\mathrm{O}(1/\sigma^2)$. Their argument relies on a simple observation: if the number of balls m were Poisson distributed rather than fixed, then the occupancy counts $\{M_i\}_{i\in[n]}$ would be independent. They prove the result by applying an LCLT for i.i.d. random variables to the Poissonized version of the problem and then use a de-Poissonization argument to transfer back to the original. However, in more complex occupancy models such as Erdős–Rényi random graphs and germ–grain models, Poissonization does not lead to a similar simplification, and it is therefore unclear how to adapt these arguments to such models. See also McDonald (2005) and references therein for other methods that have potential to prove LCLTs for sums of dependent variables, but which are not adapted to our applications.

A general method used to prove LCLTs, introduced in McDonald (1980), is to combine a (global) central limit theorem with some condition implying *smoothness* of the distribution being approximated (here, $\mathcal{L}(\widehat{W}_d)$). A common way of quantifying the smoothness, used in McDonald (1980) and Penrose and Peres (2011), is to find an embedded sum of independent random variables which are themselves smooth, in the sense that they satisfy an LCLT. Here we use a different notion of the smoothness of a distribution, given in (2.3) below, which is closely related to that given in Davis and McDonald (1995), and elaborated on in Röllin and Ross (2015). The latter paper demonstrates that bounds in a global metric (such as the Kolmogorov metric) to a target distribution, combined with appropriate bounds on the smoothness term (2.3), imply an LCLT with error. For Erdős–Rényi random graphs, for example, if the (optimal) Kolmogorov bounds in Goldstein (2013) and the smoothness bounds implicit in Lemma 4.4 of this paper are combined with the Landau–Kolmogorov inequality in Röllin and Ross (2015, Theorem 2.2), then an LCLT is obtained, but with an error bound of order $\mathrm{O}(\sigma^{-3/2})$, which is substantially worse than our target order of $\mathrm{O}(1/\sigma^2)$.

Barbour et al. (2019) (see also Röllin (2005)) combine the smoothness approach with Stein's method for distributional approximation to establish a method for proving LCLTs with error, in settings where the dependence between summands can be described in terms of a *Stein coupling* (see Chen and Röllin (2010), as well as (2.2) below). The Stein coupling most commonly applied to the occupancy counts $\widehat{W}_d$ is a *size biased coupling*; see Section 2.2. The standard method of constructing outcomes of a size biased version $\widehat{W}_d^s$ of $\widehat{W}_d$ is to take the configuration of occupants, labelled $\mathcal{X}$, and to modify it so that a single location, chosen uniformly at random, now has d occupants, and the remaining configuration has the conditional distribution, given this event. The number of locations $\widehat{W}_d^s$ with d occupants

This article is protected by copyright. All rights reserved.

Author Manuscript

in this modified configuration then has the size biased distribution of $\widehat{W}_d$. Given such a construction, Barbour et al. (2019) demonstrate that to obtain an LCLT with error for $\widehat{W}_d$ we must essentially do two things: **(i)** establish a concentration inequality for

$$\Psi \;:=\; |\mathbb{E}\{\widehat{W}_d^s - \widehat{W}_d \,|\, \mathcal{X}\} - \mathbb{E}\{\widehat{W}_d^s - \widehat{W}_d\}|,$$

and **(ii)** prove that the distribution of $\widehat{W}_d$ is smooth.

Establishing a concentration inequality for Ψ is often the hardest task. In most cases, Ψ is a complicated expression (for example, see Lemma 4.6), to which it is unclear how to apply standard concentration results, including those related to Stein's method. This significantly limits the scope of the bound in Barbour et al. (2019). Indeed, when the authors use this method to prove an LCLT with error rate $\mathrm{O}(\sigma^{-2}(\log\sigma)^{1/2})$ for the number of *isolated* ($d = 0$) vertices in an Erdős–Rényi random graph, they do so by demonstrating that, when $d = 0$, Ψ has a particularly simple expression, to which established concentration inequalities can be applied; however, when $d > 0$, Ψ is significantly more complex, and a new approach is required. In this paper we demonstrate that the recent results in Barbour et al. (2018), which provide a widely applicable method for deriving central moment inequalities, can often be used to establish the necessary concentration inequalities for these more complicated expressions. To highlight the connection to Barbour et al. (2019), in Theorem 2.2 we rewrite their general bound in terms of central moment inequalities. By applying these two bounds in tandem we obtain a relatively general method for proving LCLTs with near optimal error rate.

When we consider germ–grain models, there is an additional complication: if we apply a standard size biased coupling, as described above, then it is difficult even to find an expression for Ψ. We overcome this issue by using the *bounded size biased couplings* proposed in Bartroff, Goldstein, and Işlak (2018) for

$$W_d \;:=\; n - \widehat{W}_d \;=\; \sum_{i=1}^{n} \mathrm{I}\{M_i \neq d\},$$

the number of locations that do *not* have d occupants. There the authors demonstrate how to construct size biased versions of W_d in occupancy models, by moving at most a single occupant from its original location. Such couplings allow us not only to find an expression for Ψ in germ–grain models, but also to improve the bound in the Erdős–Rényi random graph application.

Thus, in this paper, we piece together the general LCLT bounds in Barbour et al. (2019), the central moment inequalities in Barbour et al. (2018), the bounded size biased couplings in Bartroff et al. (2018), and the smoothness terms and bounds of Röllin and Ross (2015), to establish a robust method for proving LCLTs for the number of locations with d occupants in occupancy models, together with error bounds that are of the same order as would be expected for sums of independent indicators, up to logarithmic factors. The logarithmic factors arise from the concentration inequalities used to handle Ψ, and can only be avoided using our method by modification in special examples; see Barbour et al. (2019, Remark 2.9). We emphasise that the contribution of the paper is twofold: obtaining LCLTs with good error rate in two non-trivial examples — already an interesting and difficult undertaking — and also providing a straightforward method for obtaining LCLTs with error that is applicable in a wide variety of other settings.

This article is protected by copyright. All rights reserved.

Author Manuscript

The paper is organised as follows. In Section 2, we provide the key result, Theorem 2.2, for establishing LCLTs with rate. In Section 3, we state LCLTs with rate for occupancy models obtained by applying Theorem 2.2. In Section 4, we apply the framework established in Section 2 to prove the results in Section 3. Section 5 gives a derivation of Theorem 2.2 from the results of Barbour et al. (2019), and Section 6 contains two auxiliary results used in the proofs.

2 STEIN'S METHOD

Stein's method is a powerful tool for bounding the error in the approximation of a distribution of interest by another well–understood target distribution. It was first developed for approximation by the normal distribution in Stein (1972, 1986), and by the Poisson distribution in Chen (1975); see Barbour, Holst, and Janson (1992), Barbour and Chen (2005), Chen, Goldstein, and Shao (2011) and Ross (2011) for various introductions to the method.

2.1 Stein's method for LCLTs

We use Stein's method to bound the distance between W and an integer valued target distribution Z, in the *total variation metric*,

$$d_{\mathrm{TV}}\big(\mathcal{L}(W),\mathcal{L}(Z)\big) \;:=\; \sup_{A\subseteq\mathbb{Z}} |\mathbb{P}[W\in A]-\mathbb{P}[Z\in A]|,$$

as well as in a metric to capture the local differences,

$$d_{\mathrm{loc}}\big(\mathcal{L}(W),\mathcal{L}(Z)\big) \;:=\; \sup_{a\in\mathbb{Z}} |\mathbb{P}[W=a]-\mathbb{P}[Z=a]|.$$

Following Röllin (2007), we use *translated Poisson* distributions as the approximating family, instead of the discretized normal. We say that a random variable Z has the translated Poisson distribution, and write $Z\sim \mathrm{TP}(\mu,\sigma^2)$, if $Z-s\sim \mathrm{Po}(\sigma^2+\gamma)$, where

$$s \;:=\; \lfloor \mu-\sigma^2\rfloor, \qquad \gamma \;:=\; \mu-\sigma^2-\lfloor \mu-\sigma^2\rfloor, \tag{2.1}$$

and where $\mathrm{Po}(\lambda)$ denotes the Poisson distribution with mean λ. Note that $\mathbb{E}Z=\mu$ and $\sigma^2\le \mathrm{Var}\,Z\le \sigma^2+1$. Thus, the translated Poisson distribution is a Poisson distribution translated by an integer so that both its mean and variance closely approximate μ and σ^2. The next result shows that the translated Poisson and the discretised normal are sufficiently close for the purposes of LCLTs. In particular, it implies that in the LCLTs given in Theorems 3.1 and 3.2 below, the translated Poisson distribution and the discretized normal distribution are interchangeable. The lemma follows by applying (1.1) to $(Z-s)\sim \mathrm{Po}(\sigma^2+\gamma)$ and using basic properties of the Poisson distribution and normal density.

Lemma 2.1. *If $Z\sim \mathrm{TP}(\mu,\sigma^2)$, then there exists a constant $C>0$ such that, for all $\mu\in\mathbb{R}$ and $\sigma^2\ge 1$,*

$$\sup_{n\in\mathbb{Z}}\left|\mathbb{P}(Z=n)-\frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{(n-\mu)^2}{2\sigma^2}\right)\right| \;\le\; \frac{C}{\sigma^2}.$$

This article is protected by copyright. All rights reserved.

Author Manuscript

The general LCLT theorem that we use requires that the variable of interest W is part of a *Stein coupling*. Following Chen and Röllin (2010), we say that the random variables (W, W', G, R) form an *approximate Stein coupling* if

$$\mathbb{E}[(W-\mu)f(W)] \;=\; \mathbb{E}[G(f(W')-f(W))] + \mathbb{E}[Rf(W)], \tag{2.2}$$

for all f such that the expectations exist. If $R = 0$ almost surely, we call (W, W', G) a *Stein coupling*. The final ingredients needed for the general LCLT are the following standard probabilistic measures of smoothness for integer valued distributions:

$$S_l(\mathcal{L}(W)) \;:=\; \sup_{h:\ \|h\|_\infty \le 1} |\mathbb{E}\Delta^l h(W)|, \qquad l \ge 1, \tag{2.3}$$

where $\Delta^l h(i) := \Delta^{l-1}(h(i+1)-h(i))$. Note that $S_1(\mathcal{L}(W)) = 2\, d_{\mathrm{TV}}(\mathcal{L}(W), \mathcal{L}(W+1))$.

The following is a concise and more easily applicable modification of Corollary 2.3 and Lemma 2.6 of Barbour et al. (2019), for proving LCLTs using Stein's method. As the statement here is not easily read from their results, we provide a proof below in Section 5. Here and below, for a random variable X and $q > 0$, we denote

$$\|X\|_q := \big(\mathbb{E}\big[|X|^q\big]\big)^{1/q}.$$

Theorem 2.2. *Let (W, W', G, R) be an approximate Stein coupling with W and W' integer valued, $\mathbb{E}W = \mu$ and $Var(W) = \sigma^2$. Set $D := W' - W$, and let $\mathcal{F}_1$ and $\mathcal{F}_2$ be sigma algebras such that W is $\mathcal{F}_1$-measurable and such that (G, D) is $\mathcal{F}_2$-measurable. Define*

$$\Upsilon \;:=\; \mathbb{E}\big[|GD(D-1)|\, S_2\big(\mathcal{L}(W|\mathcal{F}_2)\big)\big], \tag{2.4}$$

and

$$T := \big|\mathbb{E}[GD\,|\,\mathcal{F}_1] - \mathbb{E}[GD]\big|.$$

If c_1 is such that

$$\max\Big\{\Upsilon + 1, \|R\|_2, \sigma^{-1}\|T\|_2\Big\} \;\le\; c_1, \tag{2.5}$$

then

$$d_{\mathrm{TV}}\big(\mathcal{L}(W), TP(\mu,\sigma^2)\big) \;\le\; 5c_1\sigma^{-1}. \tag{2.6}$$

If, in addition, c_2 is such that, for $q = \lceil \log(\sigma) \rceil$,

$$\sigma^{-1}\|T\|_q \;\le\; c_2, \tag{2.7}$$

and if

$$c_1 + ec_2 \;<\; \sigma/2, \tag{2.8}$$

then

$$d_{\mathrm{loc}}\big(\mathcal{L}(W), TP(\mu,\sigma^2)\big) \;\le\; 4(4c_1 + ec_2)\sigma^{-2}. \tag{2.9}$$

Remark 2.3. In our applications, c_1 is fixed as σ increases, whereas $c_2 = c_2(\sigma)$ grows like $(\log\sigma)^\alpha$, for some fixed $\alpha > 0$; hence the total variation bound is of order $O(\sigma^{-1})$, whereas the local bound is of order $O(\sigma^{-2}(\log\sigma)^\alpha)$. For such choices, (2.8) is satisfied for

This article is protected by copyright. All rights reserved.

Author Manuscript

all σ large enough. If it were not satisfied, then the bound in (2.9) would be of order comparable to typical point probabilities, and would thus be of little use. We typically write $T = \sum_{l=1}^{k} T_l$, where the T_l, $1 \le l \le k$, have a structure enabling their norms to be bounded using a main result of Barbour et al. (2018), given as Theorem 2.4 below, for convenience and completeness. Then bounds related to T of (2.5) and (2.7) can be verified using the triangle (or Minkowski's) inequality for $\|\cdot\|_q$ for $q \ge 1$. The quantity R is zero almost surely. It remains to bound $\Upsilon + 1$ by a constant c_1. To do so, note that, in the typical regime, D is of constant order in σ and G is of order σ^2. Thus, establishing $\Upsilon + 1 \le c_1$ boils down to showing that $S_2(\mathcal{L}(W) \,|\, \mathcal{F}_2) = \mathrm{O}(\sigma^{-2})$. Conditioning on $\mathcal{F}_2$ is handled by "freezing" an asymptotically negligible part of the randomness, and then the term behaves similarly to $S_2(\mathcal{L}(W))$. We can then apply established methods for bounding $S_2(\cdot)$; see, for instance, Lindvall (2002, Chapter II.14), Mattner and Roos (2007) and Röllin and Ross (2015).

Theorem 2.4 (Theorem 2.2 of Barbour et al. (2018))**.** *Suppose that (Y, Y', G) is a Stein coupling with $\mathbb{E}\{|Y' - \mu|^r\} \le \mathbb{E}\{|Y - \mu|^r\}$, for some $r \in \mathbb{N}$. Then*

$$\|Y - \mu\|_r \;\le\; \sqrt{2(r-1)\|G\|_r \|Y' - Y\|_r}.$$

2.2 Bounded size biased couplings

The Stein couplings that we use here, when we apply Theorem 2.2, are *size biased couplings*. For a non-negative random variable W, we say that W^s has the W-size biased distribution if

$$\mathbb{E}[W f(W)] \;=\; \mu \mathbb{E}[f(W^s)] \tag{2.10}$$

for all functions f for which these expectations exist; moreover, if W^s and W are defined on the same probability space, then we say that (W, W^s) form a *size-biased coupling*. It is not difficult to verify that $(W, W', G) = (W, W^s, \mu)$ satisfies (2.2) with $R = 0$, and is therefore a Stein coupling.

The standard method used to construct sized biased couplings for occupancy counts $\widehat{W}_d$, defined at (1.2), such that $\mathbb{P}(M_i = d)$ is independent of i, relies on the identity:

$$\mathbb{E}[\widehat{W}_d f(\widehat{W}_d)] \;=\; \sum_{i=1}^{n} \mathbb{P}[M_i = d] \mathbb{E}\big(f(\widehat{W}_d) \,|\, M_i = d\big) \;=\; \mathbb{E}[\widehat{W}_d] \sum_{i=1}^{n} \frac{1}{n} \mathbb{E}\big(f(\widehat{W}_d) \,|\, M_i = d\big). \tag{2.11}$$

This identity gives the following standard method for coupling $\widehat{W}_d^s$ to $\widehat{W}_d$, by conditioning on the occupants and locations: select a location $I \in [n]$ uniformly at random; then, if $M_I > d$, select $M_I - d$ occupants uniformly at random from those at location I, and re-position them uniformly at random in $[n] \setminus I$; if $M_I < d$, select $d - M_I$ occupants uniformly at random from those at locations different from I, and re-position them at location I. The resulting number of locations with d occupants is then the corresponding outcome of $\widehat{W}_d^s$; for more details, see Chen et al. (2011, Section 2.3.4).

Bartroff et al. (2018) show that it can be advantageous to instead construct a sized biased coupling for the number of locations that do *not* contain precisely d occupants:

$$W_d \;:=\; n - \widehat{W}_d \;=\; \sum_{i=1}^{n} \mathrm{I}\{M_i \ne d\}.$$

This article is protected by copyright. All rights reserved.

Author Manuscript

As for (2.11), we have

$$\mathbb{E}[W_d f(W_d)] \;=\; \mathbb{E}[W_d] \sum_{i=1}^{n} \frac{1}{n} \mathbb{E}\big(f(W_d) \,|\, M_i \neq d\big). \tag{2.12}$$

The strategy, as before, is to choose I uniformly at random. Given $I = i$, modify M_i, by adding or removing occupants of i, in such a way that the distribution of the number of occupants of i becomes $\mathcal{L}(M_i \,|\, M_i \neq d)$. By symmetry, if occupants added are chosen from the remainder uniformly at random, and if those removed are re-distributed uniformly at random among the other locations, then the resulting distribution of the number of locations with other than d occupants becomes $\mathcal{L}(W_d \,|\, M_i \neq d)$. Of course, proving LCLTs for W_d is equivalent to doing so for $\widehat{W}_d$, but the advantage of working with W_d is that *at most one* occupant has to be moved to realize the coupling above. This is the substance of the following lemma, which is Bartroff et al. (2018, Lemma 2.1).

Lemma 2.5 (Lemma 2.1 of Bartroff et al. (2018))**.** *If M is an integer valued random variable such that $\mathcal{L}(M)$ is log–concave, and if*

$$\pi_x^{(d)} = \begin{cases} \frac{\mathbb{P}(M \geq x+1)\mathbb{P}(M=d)}{\mathbb{P}(M \geq d+1)\mathbb{P}(M=x)}, & \text{if } x \geq d \\ 0, & \text{otherwise,} \end{cases} \qquad \gamma_x^{(d)} = \begin{cases} \frac{\mathbb{P}(M \leq x-1)\mathbb{P}(M=d)}{\mathbb{P}(M \leq d-1)\mathbb{P}(M=x)}, & if x \leq d \\ 0, & \text{otherwise,} \end{cases} \tag{2.13}$$

then $\pi_x^{(d)}, \gamma_x^{(d)} \in [0,1]$ for all x. Moreover, if Z_+ and Z_- are conditionally independent given M, with $\mathcal{L}(Z_+ \,|\, M) = \mathrm{Be}(\pi_M^{(k)})$ and $\mathcal{L}(Z_- \,|\, M) = \mathrm{Be}(\gamma_M^{(k)})$; if Z is independent of Z_+ and Z_- with $\mathcal{L}(Z) = \mathrm{Be}(q)$, and

$$q \;:=\; \frac{\mathbb{P}(M \geq d+1)}{\mathbb{P}(M \neq d)}; \tag{2.14}$$

and if

$$X \;:=\; ZZ_+ - (1-Z)Z_-;$$

then $\mathcal{L}(M + X) = \mathcal{L}(M \,|\, M \neq d)$.

A distribution $\mathcal{L}(M)$ is log–concave when

$$\mathbb{P}(M = s-1)\mathbb{P}(M = s+1) \;\leq\; \mathbb{P}(M = s)^2, \qquad \text{for all integers } s; \tag{2.15}$$

binomial distributions, in particular, are log concave. The theorem shows that X occupants are to be moved, and that $X \in \{-1, 0, 1\}$.

To understand the first assertion in Lemma 2.5, observe that

$$\gamma_x^{(d)} \leq 1 \quad \text{ if } \quad \frac{\mathbb{P}(M_i \leq y-1)}{\mathbb{P}(M_i = y)} \;\leq\; \frac{\mathbb{P}(M_i \leq d-1)}{\mathbb{P}(M_i = d)}$$

for all $y \leq d-1$, which equivalent to

$$\mathbb{P}(M_i \leq y-1)\mathbb{P}(M_i = y+1) \;\leq\; \mathbb{P}(M_i \leq y)\mathbb{P}(M_i = y).$$

This article is protected by copyright. All rights reserved.

Author Manuscript

This, in turn, can be verified through repeated application of (2.15). A similar argument holds for $\pi_x^{(d)}$. To understand the second assertion of Lemma 2.5, observe that

$$\begin{aligned}\mathbb{P}(M_i + X = j) &= \mathbb{P}(M_i = j)(1 - q\pi_j^{(d)} - (1-q)\gamma_j^{(d)}) + \mathbb{P}(M_i = j-1)q\pi_{j-1}^{(d)} \\ &\quad + \mathbb{P}(M_i = j+1)(1-q)\gamma_{j-1}^{(d)} \\ &= \mathbb{P}(M_i = j \mid M_i \neq d),\end{aligned}$$

as can be verified directly. We refer the reader to the proof of Bartroff et al. (2018, Theorem 2.1) for more details.

As pointed out in Bartroff et al. (2018), constructing W_d^s in this manner leads to bounded size biased couplings, that is, couplings such that $|W_d^s - W_d| < C$ almost surely (here, for $C = 2$). The advantage of this coupling is that we need only move at most a single occupant, and this makes it easier to express the quantity T arising in Theorem 2.2, particularly in Section 4.2. Moreover, using this coupling instead of the standard one reduces the additional logarithmic factors that occur when using Theorem 2.2 to prove local limit theorems.

3 APPLICATIONS

We present LCLTs with error in two applications: the number of degree-d vertices in an Erdős–Rényi random graph, and the number of germs with d-neighbours in a germ–grain model.

3.1 Erdős–Rényi random graphs

Let $\mathbb{G}_n$ be the set of simple and undirected graphs with the vertex set $\mathcal{V} = \{v_1, \ldots, v_n\}$. We construct an Erdős–Rényi random graph $\mathcal{G}_n$ on $\mathbb{G}_n$ by letting the indicators E_{ij}, which determine the presence of an edge between v_i and v_j, be independent Bernoulli random variables with a common success probability p when $i \neq j$, and be zero when $i = j$. In this paper, we consider *sparse* Erdős–Rényi random graphs, that is, we let $p = \lambda/n$ for some constant $\lambda > 0$.

We let $M_i = \sum_{j=1}^n E_{ij}$ be the degree of vertex V_i in $\mathcal{G}_n$, and study the total number of vertices whose degree is not precisely d:

$$W_d := \sum_{i=1}^n \mathrm{I}\{M_i \neq d\}. \tag{3.1}$$

Using elementary arguments, we obtain

$$\mu_d := \mathbb{E}W_d = n(1-b_d) \qquad \text{and} \qquad \sigma^2 := \mathrm{Var}(W_d) = nb_d^2\left[\frac{(d-(n-1)p)^2}{(n-1)p(1-p)} - 1\right] + nb_d,$$

where $b_d := \binom{n-1}{d} p^d (1-p)^{n-d-1}$. Observe that μ_d and σ^2 are both of strict order n.

We are now in a position to state an LCLT for W_d. This complements the work of a number of authors, who establish central limit theorems for W_d that apply when $p = \lambda/n$: Barbour, Karoński, and Ruciński (1989) prove that W_d is asymptotically normal and, in addition, obtain bounds in the Wasserstein metric of optimal order; Kordecki (1990) obtains

This article is protected by copyright. All rights reserved.

Author Manuscript

bounds of optimal order $\mathrm{O}(\sigma^{-1})$ in the Kolmogorov metric when $d = 0$, and Goldstein (2013) obtains bounds of optimal order in the Kolmogorov metric when $d \geq 0$; Fang (2014) obtains bounds between the distribution of W_d and an appropriately discretized normal distribution in the total variation metric that are of optimal order $\mathrm{O}(\sigma^{-1})$; finally, Barbour et al. (2019) prove an LCLT with bounds of order $\mathrm{O}(\sigma^{-2}\log(\sigma)^{1/2})$, but only for the case $d = 0$.

Theorem 3.1. *For an Erdős–Rényi random graph with $p = \lambda/n$, if W_d is given by* (3.1) *and $\sigma^2 := \mathrm{Var}(W_d)$, then as $n \to \infty$, for any $d \geq 0$,*

$$\begin{aligned} d_{\mathrm{TV}}\big(\mathcal{L}(W_d), \mathrm{TP}(\mu_d, \sigma_d^2)\big) &= \mathrm{O}\big(1/\sigma\big); \\ d_{\mathrm{loc}}\big(\mathcal{L}(W_d), \mathrm{TP}(\mu_d, \sigma_d^2)\big) &= \mathrm{O}\bigg(\frac{(\log\sigma)^{5/2}}{\sigma^2}\bigg). \end{aligned}$$

3.2 Germ–grain models

To define the germ–grain models that we study, let $C_n := [0, n^{1/2})^2$ be a torus; for $x, y \in C_n$, let $D(x, y)$ denote the distance between x and y under the Euclidean toroidal metric on C_n; and, for $x \in C_n$ and $s > 0$, let $B_s(x)$ denote the ball $\{y \in C_n : D(x, y) \leq s\}$. Let $V_1, \ldots, V_n$ be independent points scattered uniformly in C_n. We refer to points in the set $\mathcal{V} := \{V_1, \ldots, V_n\}$ as *germs*. For a fixed value $r > 0$, let $B_{i,r} := B_r(V_i)$ be the r-ball that surrounds germ i. We refer to these balls as *grains*. To avoid small-n boundary effects, we assume that $\pi r^2 < n$. Let

$$\mathcal{I}_{i,r} \;:=\; \{j \neq i : V_j \in B_{i,r}\} \tag{3.2}$$

be the set of germs that fall in grain i, and $M_i := \mathrm{card}\{\mathcal{I}_{i,r}\}$ be the number of germs that fall in grain i. We study the total number of germs whose grain does not contain precisely d germs, that is,

$$W_d \;:=\; \sum_{i=1}^{n} \mathrm{I}\{M_i \neq d\}. \tag{3.3}$$

Using (3.3) and the fact that $M_i \sim \mathrm{Bi}(n-1, \pi r^2/n)$, we obtain $\mu_d := \mathbb{E}(W_d) = n(1 - b_d)$, where

$$b_d \;:=\; \binom{n-1}{d}\bigg(\frac{\pi r^2}{n}\bigg)^d\bigg(1 - \frac{\pi r^2}{n}\bigg)^{n-1-d}.$$

An expression for $\sigma^2 := \mathrm{Var}(W_d)$ is straightforward to derive, but more difficult to analyse asymptotically. For our purposes, it is enough to apply Penrose and Yukich (2001, Theorem 2.1), which implies that σ^2 is also of strict order n.

We are now in a position to state an LCLT for W_d. Our result complements the work of a number of authors: Penrose and Yukich (2001) (see also Penrose (2003)) establish general CLTs for geometric random graphs that apply to W_d, which in that setting corresponds to the number of vertices not having degree d; Chatterjee (2008) gives optimal bounds in the Wasserstein metric; Goldstein and Penrose (2010) obtain Berry–Essen bounds of optimal order $\mathrm{O}(\sigma^{-1})$ when $d = 0$ and, when combined with the bounded size-biased couplings described in Section 2.2, their method extends naturally to $d \geq 0$. Lachièze-Rey and Peccati (2017) establish Berry–Essen bounds for functionals of binomial point processes; Penrose and Peres (2011, Section 6.1) give an LCLT without rate for W_0 on the scale of its span, though it does

This article is protected by copyright. All rights reserved.

Author Manuscript

not appear that it is established that the span is 1. To the best of our knowledge, neither a bound in total variation nor an LCLT has previously been established for W_d, when $d \geq 1$.

Theorem 3.2. *In the germ–grain model described above, for any fixed $r > 0$, if W_d is given by (3.3) and $\sigma^2 := \mathrm{Var}(W_d)$, then as $n \to \infty$, for any $d \geq 0$,*

$$d_{\mathrm{TV}}\big(\mathcal{L}(W_d), TP(\mu, \sigma^2)\big) = O(1/\sigma);$$
$$d_{\mathrm{loc}}\big(\mathcal{L}(W_d), TP(\mu, \sigma^2)\big) = O\bigg(\frac{(\log \sigma)^{3/2}}{\sigma^2}\bigg).$$

4 PROOFS OF THE APPLICATIONS

We now prove Theorem 3.1 (Section 4.1) and Theorem 3.2 (Section 4.2). In each section, we split the proof into lemmas; we note that, with the exception of Lemma 4.5 (where the analogue in for Erdős–Rényi random graphs is trivial) the Lemmas in Sections 4.1 and 4.2 are in one-to-one correspondence.

4.1 Erdős–Rényi random graphs

To prove Theorem 3.1, we first construct a size biased coupling. To do so, we define a new random graph $\mathcal{G}_n^s$ with vertex set $\mathcal{V} = \{v_i\}_{i\in[n]}$ and edge indicators $\{E_{ij}^s\}_{i,j\in[n],i\neq j}$, in the following way. We let I be distributed uniformly on $[n]$, independently of everything else. Given $I = i$ and M_i, we let Z, Z_+, and Z_- be independent Bernoulli random variables with means q, $\pi_{M_i}^{(d)}$ and $\gamma_{M_i}^{(d)}$, where the expressions for q, $\pi_{M_i}^{(d)}$ and $\gamma_{M_i}^{(d)}$ are as in Lemma 2.5 ; we then let $X := ZZ_+ - (1-Z)Z_-$. If $I = i$ and $X = 0$, we set $E_{lj}^s = E_{lj}$ for all $l, j \in [n]$; if $I = i$ and $X = 1$, we let J be uniformly distributed on $\{j \in [n] : E_{ij} = 0\}$, conditionally independent of everything else, given $I = i$ and $(E_{ij},\, j \in [n] \setminus \{i\})$, and we set $E_{iJ}^s = 1$ and $E_{lj}^s = E_{lj}$ for all other pairs l, j; finally, if $I = i$ and $X = -1$, we let J be uniformly distributed on $\{j \in [n] : E_{ij} = 1\}$, conditionally independent of everything else, given $I = i$ and $(E_{ij},\, j \in [n] \setminus \{i\})$, and we set $E_{iJ}^s = 0$ and $E_{lj}^s = E_{lj}$ for all other pairs l, j. If we let W_d^s denote the number of vertices in $\mathcal{G}_n^s$ with degree different from d, then W_d^s has the size biased distribution of W_d, and (W_d, W_d^s, μ_d) is a Stein coupling.

The first lemma provides a useful expression for the quantity T that arises in Theorem 2.2, when we apply it to the Stein coupling (W_d, W_d^s, μ), in terms of *local* statistics. Fix $r \in \mathbb{N}$, and, for $G \in \mathbb{G}_n$ and each $i = 1, \ldots, n$, let $\mathcal{N}_r(i, G)$ be the 'r-neighbourhood' consisting of the vertex–labelled subgraph induced by vertices at graph distance at most r from vertex i. Observe that M_i is a function of $\mathcal{N}_1(i, \mathcal{G}_n)$, and that

$$\widehat{W}_t(v_i) \;:=\; \sum_{j=1}^{n} E_{ij}\, \mathrm{I}\{M_j = t\},$$

the number of degree t vertices connected to v_i, is a function of $\mathcal{N}_2(i, \mathcal{G}_n)$; we refer to such statistics as local. Lemma 4.1 shows that the quantity T that arises in Theorem 2.2 can be bounded in terms of a sum of centred sums of bounded local statistics.

This article is protected by copyright. All rights reserved.

Author Manuscript

Lemma 4.1. *We have*

$$|\mathbb{E}[GD\,|\,\mathcal{G}_n] - \sigma^2| \;\le\; (1-b_d)\sum_{l=1}^{6} T_l,$$

where $T_l = |T_l' - \mathbb{E}T_l'|$, $1 \le l \le 6$, *and*

$$\begin{aligned}
T_1' &= \frac{q}{n}UW_{d-1}; \qquad T_2' = \frac{q}{n}UW_d; \qquad T_6' = W_d;\\
T_3' &= q\sum_{i=1}^{n} \mathrm{I}[M_i < n/2]\,\frac{(\widehat{W}_{d-1}(v_i) - \widehat{W}_d(v_i) + \mathrm{I}[M_i = d-1] - \mathrm{I}[M_i = d])\pi_{M_i}^{(d)}}{n - M_i - 1};\\
T_4' &= q\sum_{i=1}^{n} \mathrm{I}[M_i \ge n/2]\,\frac{\pi_{M_i}^{(d)}\{W_{d-1} + \widehat{W}_{d-1}(v_i) + \mathrm{I}[M_i = d-1] - W_d - \widehat{W}_d(v_i) - \mathrm{I}[M_i = d]\}}{n - M_i - 1};\\
T_5' &= (1-q)\sum_{i=1}^{n} \frac{(\widehat{W}_{d+1}(v_i) - \widehat{W}_d(v_i))\gamma_{M_i}^{(d)}}{M_i},
\end{aligned}$$

where

$$U \;:=\; \sum_{i=1}^{n} \mathrm{I}[M_i < n/2]\,\frac{n\pi_{M_i}^{(d)}}{n - M_i - 1}.$$

Proof. By considering the degree of the vertex I chosen and which of its neighbours gain or lose an edge, we obtain

$$\begin{aligned}
\mathbb{E}[GD\,|\,\mathcal{G}_n] \;=\; (1-b_d)\sum_{i=1}^{n}\Bigg(&\mathrm{I}\{M_i = d\} + \frac{(1-q)\gamma_{M_i}^{(d)}}{M_i}\,[\widehat{W}_d(v_i) - \widehat{W}_{d+1}(v_i)]\\
&+ \frac{q\pi_{M_i}^{(d)}}{n - M_i - 1}\Big[-(n - W_{d-1}) + \widehat{W}_{d-1}(v_i) + \mathrm{I}[M_i = d-1]\\
&\qquad + (n - W_d) - \widehat{W}_d(v_i) - \mathrm{I}[M_i = d]\Big]\Bigg);
\end{aligned}$$

note that, if $M_i = 0$ or $M_i = n-1$, so that the denominator in one of the fractions is zero, the numerator is zero also, and the corresponding term is to be taken as zero. The lemma then follows by observing that $\sigma^2 = \mathbb{E}[GD]$, by rearranging the terms, and by applying the triangle inequality. □

Note that T_3', T_5', W_d and W_{d-1} are sums of local statistics that are bounded by 1, that U is a sum of local statistics bounded by 4, if $n \ge 2$, and that T_4' is at most n times the number of i such that $M_i > n/2$, whose expectation is very small. These observations help in controlling the norms of $\sigma^{-1}T_i$, $1 \le i \le 6$, as required when applying Theorem 2.2, because Theorem 2.4 can be invoked. For this example, the result of doing so has already been established as Barbour et al. (2018, Theorem 4.2), and is therefore stated here without proof.

Lemma 4.2. *Let* U *be a real valued function on all vertex–labelled graphs with one distinguished vertex* v*, and at most* $n-1$ *other vertices. Suppose that there exists constants* c *and* $\beta \ge 0$ *such that*

$$|\mathrm{U}(G, v)| \;\le\; c\,\mathrm{card}\{V(G)\}^{\beta},$$

This article is protected by copyright. All rights reserved.

for every G in the domain of U. *Fix $\lambda > 0$ and let $\mathcal{G}_n$ be an Erdős–Rényi random graph on $\mathbb{G}_n$, with edge probability $p := \lambda/n$; define*

$$X_i \;:=\; \mathrm{U}\big(\mathcal{N}_r(i, \mathcal{G}_n), i\big)$$

and $W = \sum_{i=1}^n X_i$. Then, for any $q \geq 1$,

$$n^{-1/2}\|W - \mathbb{E}W\|_q \;\leq\; cK(\beta)[C_A \max\{\lambda, q(1+\beta)\}]^{1/2+r+2r\beta},$$

where $K(\beta) := \sqrt{2}(10^{1+\beta} + 2^{1+\beta})$ and $C_A := \pi e^{e-2}/\log(e-1)$.

The implications of Lemma 4.2 for the moments of $\sigma^{-1}T_l$, $1 \leq l \leq 6$, are as follows.

Corollary 4.3. *For each $1 \leq l \leq 6$ and each $1 \leq q \leq \lceil \log \sigma \rceil$, we have $\sigma^{-1}\|T_l\|_q \leq bq^{5/2}$, for a suitable fixed choice of b.*

Proof. Taking $c = 1$, $\beta = 0$ and $r = 2$, it is immediate from Lemma 4.2 that

$$n^{-1/2}\|T_i\|_q \;\leq\; K(0)[C_A \max\{\lambda, q\}]^{5/2}, \qquad i = 3, 5,$$

and, since σ^2 is strictly of order n, it follows that, for fixed $b^{(3)}$ and $b^{(5)}$,

$$\sigma^{-1}\|T_i\|_q \;\leq\; b^{(i)}q^{5/2}, \quad i = 3, 5, \quad \text{for all } q \geq 1. \tag{4.1}$$

For W_d, W_{d-1} and U, now taking $r = 1$, it follows similarly that

$$n^{-1/2} \max\{\|W_d\|_q, \|W_{d-1}\|_q, \|U/4\|_q\} \;\leq\; b^{(0)}q^{3/2},$$

for all $q \geq 1$, for some fixed $b^{(0)}$. To convert these bounds into bounds on the moments of T_1 and T_2, we use the following inequality. Let X and Y be integrable random variables with means μ_X and μ_Y, and write $X' := X - \mu_X$ and $Y' := Y - \mu_Y$; then

$$XY - \mathbb{E}\{XY\} \;=\; X'Y' + \mu_X Y' + \mu_Y X' - \mathbb{E}\{X'Y'\}.$$

Hence, using the triangle inequality and Cauchy–Schwarz,

$$\begin{aligned} &\|XY - \mathbb{E}\{XY\}\|_q \\ &\qquad \leq\; \|X'Y'\|_q + |\mu_X|\|Y'\|_q + |\mu_Y|\|X'\|_q + \mathbb{E}|X'Y'| \\ &\qquad \leq\; 2\|X - \mathbb{E}X\|_{2q}\|Y - \mathbb{E}Y\|_{2q} + |\mathbb{E}X|\|Y - \mathbb{E}Y\|_q + |\mathbb{E}Y|\|X - \mathbb{E}X\|_q. \end{aligned} \tag{4.2}$$

Taking $X = n^{-1/2}W_{d-1}$ and $Y = n^{-1/2}U$ thus gives

$$\|T_1\|_q \;\leq\; 8\{b^{(0)}\}^2(2q)^3 + 8n^{1/2}b^{(0)}q^{3/2};$$

hence, for $q \leq \lceil \log \sigma \rceil$,

$$\sigma^{-1}\|T_1\|_q \;\leq\; b^{(1)}q^{3/2}, \tag{4.3}$$

for some fixed $b^{(1)}$, since $\sigma^{-1}\lceil \log \sigma \rceil^{3/2}$ is bounded in $\sigma \geq 1$. The same argument works also for T_2.

This article is protected by copyright. All rights reserved.

Author Manuscript

Author Manuscript

Finally, for T_4, note that $\sum_{i=1}^n I[M_i > n/2]$ has maximum value n, and that the probability that it does not take the value 0 is bounded by $\varepsilon_n := nCe^{-n/6}$, for a suitable constant $C = C(\lambda)$ and for $n > 4\lambda$, in view of a simplified Chernoff inequality read from, for example, McDiarmid (1998, Theorem 2.3(b)): for a sum S of independent Bernoulli random variables with mean μ,

$$\mathbb{P}[S > (1+\delta)\mu] \;\le\; e^{-\delta\mu/3}, \qquad \text{if } \delta \ge 1. \tag{4.4}$$

Hence it follows that, for a suitably chosen $b^{(4)}$,

$$\sigma^{-1}\|T_4\|_q \;\le\; \sigma^{-1}n^2\|\operatorname{Be}(\varepsilon_n)\|_q \;=\; \sigma^{-1}n^2\varepsilon_n^{1/q} \;\le\; b^{(4)},$$

for all $1 \le q \le \lceil \log \sigma \rceil$. Combining this with (4.1) and (4.3) completes the proof of the corollary. □

Hence, in particular, when applying Theorem 2.2, we can take $c_2 = 6b\lceil \log \sigma \rceil^{5/2}$ in (2.7), and $\sigma^{-1}\|T\|_2 \le 6 \times 2^{5/2} b$ in (2.5).

All that is now needed, in order to apply Theorem 2.2, is a bound of order O(1) for the smoothness term Υ, to be used in (2.5). We derive it by applying Röllin and Ross (2015, Theorem 3.7), using arguments that are based on those in Fang (2014, Section 2.3.1). In what follows, we define the index sets $A_I := \{I\} \cup \{j\colon E_{Ij} = 1\} \cup \{J\}$ and $B_I = \{j \notin A_I\colon E_{jk} = 1 \text{ for some } k \in A_I\}$. We then set

$$\mathcal{F}_2 \;:=\; \sigma(I, A_I, B_I, J, X),$$

observing that D (and G) are $\mathcal{F}_2$ measurable.

Lemma 4.4. *For $\Upsilon = \mathbb{E}[|GD(D-1)|S_2(\mathcal{L}(W_d \,|\, \mathcal{F}_2))]$, we have $\Upsilon = \mathrm{O}(1)$.*

Proof. Because the size biased configuration is formed by altering at most a single edge of $\mathcal{G}_n$, we have $|D| \le 2$, and hence $|D(D-1)| \le 6$. Thus

$$\begin{aligned}
\Upsilon \;&=\; \mathbb{E}[|GD(D-1)|S_2(\mathcal{L}(W_d \,|\, \mathcal{F}_2))] \;\le\; 6n(1-b_d)\mathbb{E}[S_2(\mathcal{L}(W_d \,|\, \mathcal{F}_2))] \\
&\le\; 6n(1-b_d)\mathbb{E}\big(S_2(\mathcal{L}(W_d \,|\, \mathcal{F}_2))\,\mathrm{I}[\max\{|A_I|, |B_I|\} \le \sqrt{n}]\big) \\
&\qquad\qquad + 6n(1-b_d)\mathbb{P}[\max\{|A_I|, |B_I|\} > \sqrt{n}].
\end{aligned} \tag{4.5}$$

To bound the second term in (4.5), first observe that the distribution of $|A_I|$ is stochastically dominated by $\mathrm{Bi}(n-1, p) + 2$. Thus, by (4.4), there exists $C = C(\lambda)$ such that

$$\mathbb{P}[|A_I| > \sqrt{n}] \;\le\; Ce^{-\sqrt{n}/3} \tag{4.6}$$

whenever $\sqrt{n} \ge 2(\lambda+1)$. Next observe that if $Y_1, Y_2, \ldots$ are a sequence of i.i.d. $\mathrm{Bi}(n-1, p)$–distributed random variables, then, using the standard exploration process coupling, $|B_I|$ is stochastically dominated by $\sum_{i=1}^{|A_I|} Y_i$. Using (4.4) again in (4.7), we then have

$$\begin{aligned}
\mathbb{P}\big(|B_I| > \sqrt{n}\big) \;&\le\; \mathbb{P}(|A_I| > n^{1/4}) + \mathbb{P}(|B_I| > \sqrt{n} \,|\, |A_I| \le n^{1/4}) \\
&\le\; \mathbb{P}(|A_I| > n^{1/4}) + \mathbb{P}(\exists\, i \in \{1, \ldots, \lfloor n^{1/4} \rfloor\}\colon\, Y_i > n^{1/4}) \\
&\le\; (1 + n^{1/4})C'e^{-n^{1/4}/3}
\end{aligned} \tag{4.7}$$

This article is protected by copyright. All rights reserved.

$$= \mathrm{O}(n^{-1}). \tag{4.8}$$

Combining (4.6) and (4.8), we see that the second term in (4.5) is of order O(1).

To bound the first term in (4.5), given $\mathcal{F}_2$, we define a random graph $\mathcal{G}_n^{\mathcal{F}_2}$ with vertex set $\mathcal{V}$ and edge indicators $E_{uv}^{\mathcal{F}_2}$ by letting $E_{ij}^{\mathcal{F}_2} = E_{ij}$ for $i \in A_I$ and $j \in \{1, \ldots, n\}$, and letting $E_{ij}^{\mathcal{F}_2}$ be independent Be(p) random variables for $i, j \in (A_I)^c$. If we let $W_d^{\mathcal{F}_2}$ denote the number of vertices in $\mathcal{G}_n^{\mathcal{F}_2}$ with degree different from d, then $\mathcal{L}(W_d^{\mathcal{F}_2}) = \mathcal{L}(W_d \,|\, \mathcal{F}_2)$. We now show that, for any fixed $\mathcal{F}_2$ with $\max\{|A_I|, |B_I|\} \leq \sqrt{n}$, we have

$$S_2(\mathcal{L}(W_d \,|\, \mathcal{F}_2)) \;=\; \mathrm{O}(n^{-1}), \tag{4.9}$$

by applying Röllin and Ross (2015, Theorem 3.7). For ease of notation, in the remainder of the proof, we suppress the superscript $\mathcal{F}_2$, tacitly assuming that every random quantity has distribution conditional on $\mathcal{F}_2$.

Let $\mathcal{G}_n$ be as above. Let $\mathcal{G}_n'$ be the graph obtained by choosing a pair of distinct vertices v_i and v_j with

$$\{i, j\} \;\subset\; C_I \;:=\; (A_I \cup B_I)^c,$$

uniformly at random, and resampling the edge indicator between v_i and v_j. Let $\mathcal{G}_n''$ be the graph obtained by applying the same operation to $\mathcal{G}_n'$. If we let $\widetilde{W}_d$, $\widetilde{W}_d'$ and $\widetilde{W}_d''$ be the numbers of vertices with degree different from d in $\mathcal{G}_n$, $\mathcal{G}_n'$ and $\mathcal{G}_n''$, respectively, then $(\widetilde{W}_d, \widetilde{W}_d', \widetilde{W}_d'')$ are three successive states of a reversible Markov chain. Thus, if

$$Q_{\pm 1}(G) \;:=\; \mathbb{P}\big[\widetilde{W}_d' = \widetilde{W}_d \pm 1 \,|\, \mathcal{G}_n = G\big]$$

and

$$Q_{\pm 1, \pm 1}(G) \;:=\; \mathbb{P}\big[\widetilde{W}_d' = \widetilde{W}_d \pm 1, \widetilde{W}_d'' = \widetilde{W}_d' \pm 1 \,|\, \mathcal{G}_n = G\big],$$

then, by Röllin and Ross (2015, Theorem 3.7), we have

$$\begin{aligned} S_2(\mathcal{L}(\tilde{W}_d \,|\, \mathcal{F}_2)) \;\leq\; & \frac{1}{(\mathbb{E}(Q_1(\mathcal{G}_n)))^2}\Big[2\mathrm{Var}Q_1(\mathcal{G}_n) + \mathbb{E}|Q_{1,1}(\mathcal{G}_n) - Q_1(\mathcal{G}_n)^2| \\ & + 2\mathrm{Var}Q_{-1}(\mathcal{G}_n) + \mathbb{E}|Q_{-1,-1}(\mathcal{G}_n) - Q_{-1}(\mathcal{G}_n)^2|\Big]. \end{aligned} \tag{4.10}$$

The remaining argument shows that this quantity is of order $\mathrm{O}(n^{-1})$.

We first need expressions for $Q_{\pm 1}(G)$. For $s, t \geq 0$, we let $H_{s,t}^{C_I}(G)$ ($R_{s,t}^{C_I}(G)$) denote the numbers of *connected* (*disconnected*) vertex pairs $\{v_i, v_j\}$ in G that have degrees s and t, and are such that $i, j \in C_I$. We also let $H_s^{C_I}(G) = \sum_{t \geq 0} H_{s,t}^{C_I}(G)$ ($R_s^{C_I}(G) = \sum_{t \geq 0} R_{s,t}^{C_I}(G)$) be the numbers of connected (disconnected) vertex pairs $\{v_i, v_j\}$ in G such that at least one of v_i, v_j has degree s, and such that $i, j \in C_I$. We now have

$$\begin{aligned} Q_1(G) \;&=\; p\frac{R_d^{C_I}(G) - R_{d,d}^{C_I}(G) - R_{d-1,d}^{C_I}(G)}{\binom{|C_I|}{2}} + (1-p)\frac{H_d^{C_I}(G) - H_{d,d}^{C_I}(G) - H_{d+1,d}^{C_I}(G)}{\binom{|C_I|}{2}}; \\ Q_{-1}(G) \;&=\; p\frac{R_{d-1}^{C_I}(G) - R_{d-1,d}^{C_I}(G) - R_{d-1,d-1}^{C_I}(G)}{\binom{|C_I|}{2}} \end{aligned} \tag{4.11}$$

This article is protected by copyright. All rights reserved.

Author Manuscript

$$+ (1-p)\frac{H^{C_I}_{d+1}(G) - H^{C_I}_{d,d+1}(G) - H^{C_I}_{d+1,d+1}(G)}{\binom{|C_I|}{2}}. \tag{4.12}$$

These equalities are obtained through elementary considerations. For example, the first term in (4.11) is obtained by observing that the number of degree d vertices decreases by exactly 1 when the chosen vertices are connected after resampling, and were previously disconnected, with one having had degree d and the other having had neither degree d nor $d-1$; if both had had degree d, $\widetilde{W}'_d$ would have exceeded $\widetilde{W}_d$ by 2, and if one had degree d and the other degree $d-1$ there would be no change.

To prove that $\mathbb{E}(Q_{\pm 1}(\mathcal{G}_n))$ is of strict order n^{-1}, so that the denominator in (4.10) results in a factor of order $\mathrm{O}(n^2)$, we observe that, conditional on $\mathcal{F}_2$, the degree of each vertex v_i with $i \in C_I$ is distributed as $\mathrm{Bi}(|B_I| + |C_I| - 1, p)$. Thus, if we let

$$b^{\mathcal{F}_2}_s := \binom{|B_I| + |C_I| - 2}{s} p^s (1-p)^{|B_I|+|C_I|-2-s},$$

then we have

$$\begin{aligned}
\mathbb{E}(R^{C_I}_{s,t}(\mathcal{G}_n)) &= \binom{|C_I|}{2}(1-p)[2b^{\mathcal{F}_2}_s b^{\mathcal{F}_2}_t \,\mathrm{I}\{s \neq t\} + (b^{\mathcal{F}_2}_s)^2 \,\mathrm{I}\{s = t\}], \\
\mathbb{E}(R^{C_I}_{s}(\mathcal{G}_n)) &= \binom{|C_I|}{2}(1-p)[2b^{\mathcal{F}_2}_s - (b^{\mathcal{F}_2}_s)^2], \\
\mathbb{E}(H^{C_I}_{s,t}(\mathcal{G}_n)) &= \binom{|C_I|}{2}p[2b^{\mathcal{F}_2}_{s-1} b^{\mathcal{F}_2}_{t-1} \,\mathrm{I}\{s \neq t\} + (b^{\mathcal{F}_2}_{s-1})^2 \,\mathrm{I}\{s = t\}], \\
\mathbb{E}(H^{C_I}_{s}(\mathcal{G}_n)) &= \binom{|C_I|}{2}p[2b^{\mathcal{F}_2}_{s-1} - (b^{\mathcal{F}_2}_{s-1})^2].
\end{aligned} \tag{4.13}$$

Combining (4.13) with (4.11) and (4.12), we obtain

$$\begin{aligned}
\mathbb{E}\{Q_1(\mathcal{G}_n)\} &= 2p(1-p)[b^{\mathcal{F}_2}_d + b^{\mathcal{F}_2}_{d-1} - (b^{\mathcal{F}_2}_d)^2 - (b^{\mathcal{F}_2}_{d-1})^2 - 2b^{\mathcal{F}_2}_d b^{\mathcal{F}_2}_{d-1}] \\
&= 2p(1-p)(b^{\mathcal{F}_2}_d + b^{\mathcal{F}_2}_{d-1})(1 - b^{\mathcal{F}_2}_d - b^{\mathcal{F}_2}_{d-1}) \\
&= \mathbb{E}(Q_{-1}(\mathcal{G}_n)),
\end{aligned} \tag{4.14}$$

which, for fixed d and with $p = \lambda/n$ for fixed λ, are both of strict order n^{-1}.

To bound $\mathbb{E}|Q_{1,1}(\mathcal{G}_n) - Q_1(\mathcal{G}_n)^2|$, for $G \in \mathbb{G}_n$ and $C_I \subset \mathcal{V}$, we let $\mathbb{G}^{(1)}_n(G)$ be the set of graphs in $\mathbb{G}_n$ that can be obtained by modifying at most a single edge in G whose ends are both in C_I. By observing that modifying a single edge can alter the degrees of at most two vertices, we obtain, for $G \in \mathbb{G}_n$, $G' \in \mathbb{G}^{(1)}_n(G)$, and $s, t \geq 0$,

$$\begin{aligned}
|R^{C_I}_s(G) - R^{C_I}_s(G')| &\leq 2|C_I|, \qquad |H^{C_I}_s(G) - H^{C_I}_s(G')| \leq 2s, \\
|R^{C_I}_{s,t}(G) - R^{C_I}_{s,t}(G')| &\leq 2|C_I|, \qquad |H^{C_I}_{s,t}(G) - H^{C_I}_{s,t}(G')| \leq 2\max\{s,t\}.
\end{aligned} \tag{4.15}$$

Combining (4.15) with (4.11) and (4.12) we have

$$|Q_{\pm 1}(G) - Q_{\pm 1}(G')| \leq \frac{6p|C_I| + 6(1-p)(d+1)}{\binom{|C_I|}{2}}, \tag{4.16}$$

This article is protected by copyright. All rights reserved.

Author Manuscript

for any $G \in \mathbb{G}_n$, $G' \in \mathbb{G}_n^{(1)}(G)$, and $C_I \subseteq \mathcal{V}$. Using (4.14) and (4.16) in (4.17) below, and the prescription that $p = \lambda/n$ and $\max\{|A_I|, |B_I|\} < \sqrt{n}$ in (4.18), we obtain

$$\begin{aligned}
&\mathbb{E}|Q_{1,1}(\mathcal{G}_n) - Q_1(\mathcal{G}_n)^2| \\
&\quad = \sum_{G \in \mathbb{G}_n} \mathbb{P}(\mathcal{G}_n = G)|Q_{1,1}(G) - Q_1(G)^2| \\
&\quad = \sum_{G \in \mathbb{G}_n} Q_1(G)\mathbb{P}(\mathcal{G}_n = G)\left|\left(\sum_{G' \in \mathbb{G}_n^{(1)}(G)} Q_1(G')\mathbb{P}(\mathcal{G}_n' = G'|\mathcal{G}_n = G)\right) - Q_1(G)\right| \\
&\quad \leq \sup_{G \in \mathbb{G}_n,\, G' \in \mathbb{G}_n^{(1)}(G)} \Big\{|Q_1(G') - Q_1(G)|\Big\}\mathbb{E}(Q_1(\mathcal{G}_n)) \\
&\quad \leq 12\frac{p|C_I| + (1-p)(d+1)}{\binom{|C_I|}{2}}\, p(1-p)(b_d^{\mathcal{F}_2} + b_{d-1}^{\mathcal{F}_2})(1 - b_d^{\mathcal{F}_2} - b_{d-1}^{\mathcal{F}_2}) \qquad (4.17) \\
&\quad = \mathrm{O}(n^{-3}); \qquad (4.18)
\end{aligned}$$

and the same arguments apply to $\mathbb{E}|Q_{-1,-1}(\mathcal{G}_n) - Q_{-1}(\mathcal{G}_n)^2|$. Since also

$$\mathrm{Var}(Q_{\pm 1}(\mathcal{G}_n)) \;=\; \mathrm{O}(n^{-3}), \qquad (4.19)$$

in view of Fang (2014, Pages 1416–1418), combining (4.14), (4.18) and (4.19) with (4.10) gives (4.9), and hence the result. □

Theorem 3.1 now follows directly from Theorem 2.2, in view of the bounds on the quantities appearing in (2.7) and (2.5) established in Corollary 4.3 and Lemma 4.4.

4.2 Germ–grain models

To prove Theorem 3.2, we construct a size biased coupling, based on a new configuration $\mathcal{V}^s = \{V_1^s, \ldots, V_n^s\}$. Let $\pi_{M_I}^{(d)}$, $\gamma_{M_I}^{(d)}$ and q be as given in (2.13) and (2.14). Let $Z \sim \mathrm{Be}(q)$ and I, distributed uniformly on $[n]$, be independent of everything else. Given M_I, let $Z_+ \sim \mathrm{Be}(\pi_{M_I}^{(d)})$ and $Z_- \sim \mathrm{Be}(\gamma_{M_I}^{(d)})$ be conditionally independent. Set

$$X \;:=\; ZZ_+ - (1-Z)Z_-. \qquad (4.20)$$

If $X = 0$, let $\mathcal{V}^s = \mathcal{V}$. If $X = -1$, let J be distributed uniformly on $\mathcal{I}_{I,r}$, defined in (3.2), independently of everything else; then let V_J^s be distributed uniformly on $C_n \setminus B_{I,r}$, independently of everything else, and set $V_i^s = V_i$ for all $i \neq J$. If $X = 1$, let J be distributed uniformly on (noting the modified set definition)

$$\widehat{\mathcal{I}}_{I,r}^c \;:=\; [n] \setminus (\mathcal{I}_{I,r} \cup \{I\}),$$

independently of everything else; then let V_J^s be distributed uniformly on $B_{I,r}$, and set $V_i^s = V_i$ for all $i \neq J$. By considering (2.12), if we let W_d^s be the number of germs in $\mathcal{V}^s$ whose r-grain does not contain exactly d germs, then W_d^s has the size biased distribution of W_d; see also Bartroff et al. (2018, Section 3.2.2).

This article is protected by copyright. All rights reserved.

Author Manuscript

Author Manuscript

Given this coupling, the proof of Theorem 3.2 is split into several lemmas. To state them, we first require some definitions. For $s > 0$ and $x \in C_n$, let

$$\mathcal{I}_s(x) \;:=\; \{j \in [n]\colon\, V_j \in \mathcal{V} \cap B_s(x)\} \tag{4.21}$$

be the set of germs contained in $B_s(x)$, and write

$$N_s(x) \;:=\; \mathrm{card}\{\mathcal{I}_s(x)\} \tag{4.22}$$

be the number of germs contained in $B_s(x)$. Given $\mathcal{V}$, for $i \in [n]$ and $j \in \widehat{\mathcal{I}}_{i,r}^{\,c}$, let A_{ij} be the expected increment in W_d when V_j is moved to a uniformly selected location in $B_{i,r}$ and, for $i \in [n]$ and $j \in \mathcal{I}_{i,r}$, let R_{ij} be the expected increment in W_d when V_j is moved to a uniformly selected location in $C_n \setminus B_{i,r}$. To help express A_{ij} and R_{ij}, let

$$S_j - 1 \;:=\; -\,\mathrm{I}\{M_j \neq d\} + \sum_{\ell \in \mathcal{I}_{j,r}} [\mathrm{I}\{M_\ell = d\} - \mathrm{I}\{M_\ell = d+1\}] \tag{4.23}$$

be the increment in W_d when V_j is removed from $\mathcal{V}$, and

$$H_i + 1 \;:=\; \frac{1}{\pi r^2} \int_{B_{i,r}} dx \Big\{ \mathrm{I}\{N_r(x) \neq d\} + \sum_{\ell \in \mathcal{I}_r(x)} [\mathrm{I}\{M_\ell = d\} - \mathrm{I}\{M_\ell = d-1\}] \Big\} \tag{4.24}$$

be the expected increment in W_d when an additional germ is inserted uniformly at random in $B_{i,r}$. At first, it may appear that $A_{ij} = S_j + H_i$; however, if removing V_j from $\mathcal{V}$ causes the value of H_i to change, then this is not the case. We let

$$\begin{aligned} Q_{ij} \;&:=\; \frac{1}{\pi r^2} \int_{B_{i,r}} dx \Big\{ \mathrm{I}\{x \in B_{j,r}\} [-\,\mathrm{I}\{N_r(x) \neq d\} + \mathrm{I}\{N_r(x) \neq d+1\}] \\ &\qquad + \sum_{\ell \in \mathcal{I}_r(x) \cap \mathcal{I}_{j,r}} [\mathrm{I}\{M_\ell = d+1\} - 2\,\mathrm{I}\{M_\ell = d\} + \mathrm{I}\{M_\ell = d-1\}] \Big\} \\ &=\; \frac{1}{\pi r^2} \int_{B_{i,r}} dx \Big\{ \mathrm{I}\{x \in B_{j,r}\} [\mathrm{I}\{N_r(x) = d\} - \mathrm{I}\{N_r(x) = d+1\}] \\ &\qquad + \sum_{\ell \in \mathcal{I}_r(x) \cap \mathcal{I}_{j,r}} [\mathrm{I}\{M_\ell = d+1\} - 2\,\mathrm{I}\{M_\ell = d\} + \mathrm{I}\{M_\ell = d-1\}] \Big\} \end{aligned} \tag{4.25}$$

be the increment in H_i caused by removing V_j from $\mathcal{V}$. Observe that

$$Q_{ij} \;=\; 0 \quad \text{if} \quad D(V_i, V_j) \;>\; 3r. \tag{4.26}$$

With these definitions, we now have

$$A_{ij} \;=\; H_i + S_j + Q_{ij}. \tag{4.27}$$

To express R_{ij}, first note that, when a germ is inserted uniformly into $B_{i,r}^c$, then the expected change in W_d is given by

$$K_i + 1 \;:=\; \frac{1}{n - \pi r^2} \int_{B_{i,r}^c} dx \Big\{ \mathrm{I}\{N_r(x) \neq d\} + \sum_{\ell \in \mathcal{I}_r(x)}^{n} [\mathrm{I}\{M_\ell = d\} - \mathrm{I}\{M_\ell = d-1\}] \Big\}$$

This article is protected by copyright. All rights reserved.

Author Manuscript

$$
\begin{aligned}
&= \frac{1}{n-\pi r^2}\Bigg[\int_{C_n} dx\Big\{\mathrm{I}\{N_r(x)\neq d\} + \sum_{\ell\in\mathcal{I}_r(x)}^{n}[\mathrm{I}\{M_\ell \neq d-1\} - \mathrm{I}\{M_\ell \neq d\}]\Big\} \\
&\quad - \int_{B_{i,r}} dx\Big\{\mathrm{I}\{N_r(x)\neq d\} + \sum_{\ell\in\mathcal{I}_r(x)}^{n}[\mathrm{I}\{M_\ell = d\} - \mathrm{I}\{M_\ell = d-1\}]\Big\}\Bigg] \\
&= \frac{\pi r^2}{n-\pi r^2}\left(W_{d-1} - W_d - H_i\right) - \frac{1}{n-\pi r^2}Y_d + 1, \qquad (4.28)
\end{aligned}
$$

where

$$
Y_d \;:=\; \int_{C_n} \mathrm{I}\{N_r(x) = d\}\, dx.
$$

As before, if removing V_j from $\mathcal{V}$ does not cause the value of K_i to change, then $R_{ij} = S_j + K_i$. To account for instances where K_i does change, for $i \in [n]$ and $j \in \mathcal{I}_{i,r}$, we let

$$
\begin{aligned}
E_{ij} \;:=\; &\frac{1}{\pi r^2}\int_{B_{j,r}\setminus B_{i,r}} dx\Big\{\mathrm{I}\{N_r(x)\neq d+1\} - \mathrm{I}\{N_r(x)\neq d\} \\
&\quad + \sum_{\ell\in\mathcal{I}_r(x)}[\mathrm{I}\{M_\ell = d+1\} - 2\,\mathrm{I}\{M_\ell = d\} + \mathrm{I}\{M_\ell = d-1\}]\Big\}, \\
= \;&\frac{1}{\pi r^2}\int_{B_{j,r}\setminus B_{i,r}} dx\Big\{\mathrm{I}\{N_r(x) = d\} - \mathrm{I}\{N_r(x) = d+1\} \\
&\quad + \sum_{\ell\in\mathcal{I}_r(x)}[\mathrm{I}\{M_\ell = d+1\} - 2\,\mathrm{I}\{M_\ell = d\} + \mathrm{I}\{M_\ell = d-1\}]\Big\}, \qquad (4.29)
\end{aligned}
$$

be the change in K_i caused by removing V_j from $\mathcal{V}$. Using the expression for K_i in (4.28), we now have

$$
R_{ij} \;=\; S_j + \frac{\pi r^2}{n-\pi r^2}\left(W_d - W_{d-1} - H_i\right) - \frac{1}{n-\pi r^2}\,Y_d + E_{ij}. \qquad (4.30)
$$

The following lemma is similar to Bartroff et al. (2018, Lemma 3.3), and shows that the values of S_j, H_i, Q_{ij} and E_{ij}, as well as of $\sum_{j\in\mathcal{I}_{i,r}} S_j$, are uniformly bounded. The bounds are expressed in terms of κ_s, where κ_s is the maximum number of disjoint unit balls that can be packed inside a ball of radius s. A crude bound on κ_s, which is sufficient for our purposes, is

$$
\kappa_s \;\le\; \frac{\mathrm{Leb}\{B_s(\cdot)\}}{\mathrm{Leb}\{B_1(\cdot)\}} \;=\; s^2,
$$

where Leb denotes Lebesgue measure.

Lemma 4.5. *For every i, j and configuration $\mathcal{V}$,*

$$
|S_j| \le \kappa_3(d+2), \quad |H_i| \le \kappa_3(d+1), \quad |Q_{ij}| \le 2\kappa_3(d+2), \quad |E_{ij}| \le 2\kappa_3(d+2),
$$

and

$$
\Big|\sum_{j\in\mathcal{I}_r(V_i)} S_j\Big| \;\le\; \kappa_5(d+2)^2.
$$

This article is protected by copyright. All rights reserved.

Author Manuscript

Proof. For $s > 0$, let

$$\Gamma_{s,u}(x) \;:=\; \mathcal{I}_{sr}(x) \cap \{j \in [n]\colon\; M_j = u\}$$

be the set of indices of all degree u vertices within distance sr of x, so that

$$\sum_{j \in \mathcal{I}_{sr}(x)} \mathrm{I}\{M_i = u\} \;=\; \mathrm{card}\{\Gamma_{s,u}(x)\}.$$

To bound $\mathrm{card}\{\Gamma_{s,u}(x)\}$, first observe that

$$M_i \;=\; \sum_{j \neq i} \mathrm{I}\{B_{r/2}(V_i) \cap B_{r/2}(V_j) \neq \emptyset\}; \tag{4.31}$$

that is, M_i is given by the number of $r/2$-*grains* that intersect the $r/2$-grain of germ i. Now let $\mathcal{R}_{s,u}(x)$ be a subset of $\Gamma_{s,u}(x)$ with maximal size such that its corresponding $r/2$-grains are pairwise disjoint. Because the $r/2$-grains of elements in $\mathcal{R}_{s,u}(x)$ are contained within $B_{sr+r/2}(x)$, we have $\mathrm{card}\{\mathcal{R}_{s,u}(x)\} \leq \kappa_{2s+1}$. By the maximality of $\mathcal{R}_{s,u}(x)$, the $r/2$-grain of each germ in $\Gamma_{s,u}(x) \backslash \mathcal{R}_{s,u}(x)$ must intersect that of a germ in $\mathcal{R}_{s,u}(x)$. Because the $r/2$-grain of each germ in $\mathcal{R}_{s,u}(x)$ intersects u other $r/2$-grains, we then have $\mathrm{card}\{\Gamma_{s,u}(x)\} \leq \kappa_{2s+1}(u+1)$, which implies that

$$\sum_{j \in \mathcal{I}_{sr}(x)} \mathrm{I}\{M_j = u\} \;\leq\; \kappa_{2s+1}(u+1). \tag{4.32}$$

Applying (4.32) to expressions (4.23), (4.24), (4.25) and (4.29) yields the bounds on S_j, H_i, Q_{ij} and E_{ij}, respectively.

To bound $\sum_{j \in \mathcal{I}_{i,r}} S_j$, observe that

$$\begin{aligned}\sum_{j \in \mathcal{I}_{i,r}} S_j \;&\leq\; \sum_{j \in \mathcal{I}_r(V_i)} \sum_{\ell \in \mathcal{I}_r(V_j)} \mathrm{I}\{M_\ell = d\} \;\leq\; \sum_{\ell \in \mathcal{I}_{2r}(V_i)} \mathrm{I}\{M_\ell = d\} \sum_{j \in \mathcal{I}_r(V_\ell)} 1 \\ &=\; \sum_{\ell \in \mathcal{I}_{2r}(V_i)} \mathrm{I}\{M_\ell = d\}(M_\ell + 1) \;\leq\; \kappa_5(d+1)^2.\end{aligned}$$

A corresponding lower bound can be obtained by applying the same arguments, but with d replaced by $d+1$. □

We say that the *radius* of a random variable indexed by i is the smallest value of ρ such that X_i is determined by the positions of $B_\rho(V_i) \cap \mathcal{V}$ relative to V_i. Observe that radii of $\mathrm{I}\{M_i = d\}$, $\pi^{(d)}_{M_i}$ and $\gamma^{(d)}_{M_i}$ are all equal to r, that the radius of S_i is $2r$, that the radii of H_i and Q_{ij} (by (4.26)) are both $3r$, and that the radius of $\sum_{j \in \mathcal{I}_{i,r}} E_{ij}$ is $4r$. Because each of these random variables has a finite radius, we refer to them as *local statistics*. In light of Lemma 4.5, the next lemma demonstrates that the T arising in Theorem 2.2, when we use the size biased coupling described above, can be bounded by absolute values of centred variables of the following forms: sums of uniformly bounded local statistics $(T_1, T_2, T_5, T_6, T_9, T_{11})$, $1/n$ times the products of sums of bounded local statistics (T_3, T_7, T_8, T_{10}), and sums of terms which are only non-zero on the rare events $\mathrm{I}\{M_i > n/2\}$ (T_4).

This article is protected by copyright. All rights reserved.

Author Manuscript

Lemma 4.6. *Using the notation defined above, we have*

$$\left|\mathbb{E}[GD|\mathcal{V}] - \sigma^2\right| \le (1-b_d)\sum_{i=1}^{11} T_i,$$

where $T_i := |T_i' - \mathbb{E}T_i'|$, $1 \le i \le 11$, *and*

$$\begin{aligned}
T_1' &= q\sum_{i=1}^{n}\pi_{M_i}^{(d)}H_i; & T_2' &= q\sum_{i=1}^{n}\frac{\pi_{M_i}^{(d)}}{n-M_i-1}\sum_{j\in\widehat{\mathcal{I}}_{i,r}^c}Q_{ij};\\
T_3' &= \frac{2q}{n}\sum_{i=1}^{n}\frac{n\pi_{M_i}^{(d)}\,\mathrm{I}\{M_i\le n/2\}}{2(n-M_i-1)}\sum_{j=1}^{n}S_j; & T_4' &= q\sum_{i=1}^{n}\frac{\pi_{M_i}^{(d)}\,\mathrm{I}\{M_i> n/2\}}{n-M_i-1}\sum_{j=1}^{n}S_j;\\
T_5' &= q\sum_{i=1}^{n}\frac{\pi_{M_i}^{(d)}}{n-M_i-1}\sum_{j\in\mathcal{I}_r(V_i)}S_j; & T_6' &= (1-q)\sum_{i=1}^{n}\frac{\gamma_{M_i}^{(d)}}{M_i}\sum_{j\in\mathcal{I}_{i,r}}S_j;\\
T_7' &= \frac{(1-q)\pi r^2}{n-\pi r^2}W_d\sum_{i=1}^{n}\gamma_{M_i}^{(d)}; & T_8' &= \frac{(1-q)\pi r^2}{n-\pi r^2}W_{d-1}\sum_{i=1}^{n}\gamma_{M_i}^{(d)};\\
T_9' &= \frac{(1-q)\pi r^2}{n-\pi r^2}\sum_{i=1}^{n}\gamma_{M_i}^{(d)}H_i; & T_{10}' &= \frac{(1-q)}{n-\pi r^2}Y_d\sum_{i=1}^{n}\gamma_{M_i}^{(d)};\\
T_{11}' &= (1-q)\sum_{i=1}^{n}\frac{\gamma_{M_i}^{(d)}}{M_i}\sum_{j\in\mathcal{I}_{i,r}}E_{ij},
\end{aligned}$$

adopting the convention that summands with zero numerator are zero, for example if $M_i = 0$ *in* T_{11}.

Proof. Applying (4.27) and (4.30) in (4.33) and (4.34), respectively, we obtain

$$\begin{aligned}
\mathbb{E}(GD\,|\,\mathcal{V}) &= \mu\mathbb{E}(W_d^s - W_d\,|\,\mathcal{V})\\
&= n(1-b_d)\sum_{i=1}^{n}\frac{1}{n}\left[\frac{q\pi_{M_i}^{(d)}}{n-M_i-1}\sum_{j\in\widehat{\mathcal{I}}_{i,r}^c}A_{ij} + \frac{(1-q)\gamma_{M_i}^{(d)}}{M_i}\sum_{j\in\mathcal{I}_{i,r}}R_{ij}\right]\\
&= (1-b_d)q\sum_{i=1}^{n}\frac{\pi_{M_i}^{(d)}}{n-M_i-1}\sum_{j\in\widehat{\mathcal{I}}_{i,r}^c}(H_i+S_j+Q_{ij}) \qquad (4.33)\\
&\quad + (1-b_d)(1-q)\sum_{i=1}^{n}\frac{\gamma_{M_i}^{(d)}}{M_i}\sum_{j\in\mathcal{I}_{i,r}}\Big(S_j + \frac{\pi r^2}{n-\pi r^2}(W_d - W_{d-1} - H_i)\\
&\qquad\qquad - \frac{1}{n-\pi r^2}Y_d + E_{ij}\Big) \qquad (4.34)\\
&= (1-b_d)q\left[\sum_{i=1}^{n}\pi_{M_i}^{(d)}H_i + \sum_{i=1}^{n}\frac{\pi_{M_i}^{(d)}}{n-M_i-1}\sum_{j\in\widehat{\mathcal{I}}_{i,r}^c}S_j + \sum_{i=1}^{n}\frac{\pi_{M_i}^{(d)}}{n-M_i-1}\sum_{j\in\widehat{\mathcal{I}}_{i,r}^c}Q_{ij}\right] \qquad (4.35)
\end{aligned}$$

This article is protected by copyright. All rights reserved.

$$+ (1-b_d)(1-q)\Bigg[\sum_{i=1}^n \frac{\gamma_{M_i}^{(d)}}{M_i} \sum_{j\in\mathcal{I}_{i,r}} S_j + \frac{\pi r^2(W_d - W_{d-1})}{n-\pi r^2}\sum_{i=1}^n \gamma_{M_i}^{(d)}$$
$$- \frac{\pi r^2}{n-\pi r^2}\sum_{i=1}^n \gamma_{M_i}^{(d)} H_i - \frac{Y_d}{n-\pi r^2}\sum_{i=1}^n \gamma_{M_i}^{(d)} + \sum_{i=1}^n \frac{\gamma_{M_i}^{(d)}}{M_i}\sum_{j\in\mathcal{I}_{i,r}} E_{ij}\Bigg].$$

We rearrange the second term in (4.35) to obtain

$$\begin{aligned}\sum_{i=1}^n \frac{\pi_{M_i}^{(d)}}{n-M_i-1}\sum_{j\in\widehat{\mathcal{I}}_{i,r}^c} S_j &= \sum_{i=1}^n \frac{\pi_{M_i}^{(d)}}{n-M_i-1}\left(\sum_{j=1}^n S_j - \sum_{j\in\mathcal{I}_r(V_i)} S_j\right)\\ &= \sum_{i=1}^n \frac{\pi_{M_i}^{(d)}\,\mathrm{I}\{M_i\le n/2\}}{n-M_i-1}\sum_{j=1}^n S_j + \sum_{i=1}^n \frac{\pi_{M_i}^{(d)}\,\mathrm{I}\{M_i > n/2\}}{n-M_i-1}\sum_{j=1}^n S_j\\ &\quad - \sum_{i=1}^n \frac{\pi_{M_i}^{(d)}}{n-M_i-1}\sum_{j\in\mathcal{I}_r(V_i)} S_j.\end{aligned}$$

After observing that $\sigma^2 = \mathbb{E}[GD]$, the result follows by applying the triangle inequality. □

The next lemma is used to show that the quantities T_i' are weakly concentrated about their means. Note that, in view of Lemma 4.5, we only apply this result when $\beta = 0$; however, proving the result in the more general form stated below requires little additional effort.

Lemma 4.7. *Let* U *be a real-valued function on finite subsets of C_n with a distinguished point v, whose value is determined by the positions of points relative to v. Suppose also that there exist constants $c, \beta \ge 0$ such that, for any set $A \subset C_n$ with distinguished point, we have*

$$|\mathrm{U}(A, v)| \le c\,\mathrm{card}\{A\}^\beta. \tag{4.36}$$

Fixing $s > 0$, define

$$X_i := \mathrm{U}(\mathcal{V}\cap B_{i,s}, V_i) \qquad \text{and} \qquad W := \sum_{i=1}^n X_i.$$

Then, for any $q \in \mathbb{N}$,

$$n^{-1/2}\|W - \mathbb{E}W\|_q \le 2\sqrt{6}c[C_A \max\{9\pi s^2, q(1+\beta)\} + 1]^{\beta+3/2},$$

where $C_A := \pi e^{e-2}/\log(e-1)$.

Proof. To obtain moment bounds for W, we use Theorem 2.4, together with a suitable Stein coupling that makes use of the dependence structure. For each $j = 1, \dots, n$ we generate a new configuration $\widetilde{\mathcal{V}}^{(j)} = \{\widetilde{V}_1^{(j)}, \dots, \widetilde{V}_n^{(j)}\}$ from $\mathcal{V}$. To generate $\widetilde{\mathcal{V}}^{(j)}$, for each $i \in \mathcal{I}_{j,s}$ we let $\widetilde{V}_i^{(j)}$ be uniformly distributed on C_n, independently of everything else; for each $i \in \widehat{\mathcal{I}}_{j,s}^c = \{1, \dots, n\} \setminus (\mathcal{I}_{j,s} \cup \{j\})$, we let $\widetilde{V}_i^{(j)} = V_i$ with probability $1 - \pi s^2/n$, and otherwise let $\widetilde{V}_i^{(j)}$ be

This article is protected by copyright. All rights reserved.

Author Manuscript

uniformly distributed on $B_{j,s}$, where both randomizations occur independently of everything else; finally, we let $\widetilde{V}_j^{(j)} = V_j$. Let J be uniform on the set $\{1,\dots,n\}$, independently of the random objects above, and define $\widetilde{W} := \widetilde{W}^{(J)}$ and $\widetilde{G} := -n(X_J - \mathbb{E}X_J)$. To show that $(W,\widetilde{W},\widetilde{G})$ is an exact Stein coupling, first observe that for each $j \in [n]$, $\{V_i : i \in \mathcal{I}_{j,s}\}$ and $\widetilde{\mathcal{V}}^{(j)}$ are conditionally independent given V_j. Therefore, X_j and $\widetilde{W}^{(j)}$ are also conditionally independent given V_j. Since X_j depends on the positions of the points of $\mathcal{V}$ only *relative* to V_j, it follows that $\mathcal{L}(X_j|V_j) = \mathcal{L}(X_j)$, and thus combining the arguments above, that X_j is independent of $\widetilde{W}^{(j)}$. Consequently,

$$\begin{aligned}\mathbb{E}[\widetilde{G}(f(\widetilde{W}) - f(W))] &= \mathbb{E}[-n(X_J - \mathbb{E}X_J)(f(\widetilde{W}^{(J)}) - f(W))] \\ &= \sum_{j=1}^{n} -\mathbb{E}[(X_j - \mathbb{E}X_j)(f(\widetilde{W}^{(j)}) - f(W))] = \mathbb{E}[(W - \mathbb{E}W)f(W)];\end{aligned}$$

that is, $(W,\widetilde{W},\widetilde{G})$ is an exact Stein coupling. Moreover, because $\mathcal{L}(\widetilde{W}) = \mathcal{L}(W)$, the central moments of $\widetilde{W}$ and W are equal, and hence Theorem 2.4 applies.

To apply Theorem 2.4 to bound the q-th central moment, we need to bound

$$\|\widetilde{G}\|_q^q \;=\; n^{q-1}\sum_{i=1}^{n}\mathbb{E}|X_i - \mathbb{E}X_i|^q, \tag{4.37}$$

and, for $\widetilde{D} := \widetilde{W} - W$,

$$\begin{aligned}\|\widetilde{D}\|_q^q \;&=\; \frac{1}{n}\sum_{j=1}^{n}\mathbb{E}|\widetilde{W}^{(j)} - W|^q \\ &=\; \frac{1}{n}\sum_{j=1}^{n}\mathbb{E}\Big|\sum_{i=1}^{n}(\widetilde{X}_i^{(j)} - X_i)\Big|^q \;=\; \mathbb{E}\Big|\sum_{i=1}^{n}(\widetilde{X}_i^{(1)} - X_i)\Big|^q.\end{aligned} \tag{4.38}$$

To bound (4.37), note that, in view of (4.36) and Lemma 6.2,

$$\|X_i\|_q \;\le\; c\|N_s(V_i)^\beta\|_q \;=\; c\|N_s(V_i)\|_{\beta q}^\beta \;\le\; c[C_A\max\{\pi s^2, q\beta\}+1]^\beta. \tag{4.39}$$

Using (4.39) and Minkowski's inequality, this gives

$$\mathbb{E}|X_i - \mathbb{E}X_i|^q \;=\; \|X_i - \mathbb{E}X_i\|_q^q \le (\|X_i\|_q + \|X_i\|_1)^q \;\le\; (2c)^q[C_A\max\{\pi s^2, q\beta\}+1]^{\beta q}.$$

Thus, from (4.37), we can bound

$$\|\widetilde{G}\|_q \;\le\; 2nc(C_A\max\{\pi s^2, q\beta\}+1)^\beta. \tag{4.40}$$

To bound (4.38), for $\ell > 0$, let

$$\widetilde{\mathcal{I}}_{j,\ell}^{(1)} \;:=\; \Big\{i \neq j : \widetilde{V}_i^{(1)} \in B_\ell(\widetilde{V}_j^{(1)})\Big\}$$

contain the indices of the germs located in $B_\ell(\widetilde{V}_j^{(1)})$ in configuration $\widetilde{\mathcal{V}}^{(1)}$. Observe that $|X_i - \widetilde{X}_i^{(1)}| = 0$ if germ i is located (i) at least $2s$ units from V_1, (ii) at least s units from

This article is protected by copyright. All rights reserved.

Author Manuscript

the previous location of a germ that has been moved inside $B_s(V_1)$, and (iii) at least s units from any germ which has been moved outside $B_s(V_1)$.

We define $\mathcal{M}$ as follows, so that $\{1,\ldots,n\}\setminus\mathcal{M}$ contains the indices of such germs:

$$\begin{aligned}\mathcal{M} &= (\{1\}\cup\mathcal{I}_{1,2s})\bigcup\Big(\cup_{i\in\widetilde{\mathcal{I}}^{(1)}_{1,s}}\mathcal{I}_s(V_i)\Big)\bigcup\Big(\cup_{i\in\mathcal{I}_{1,s}}\mathcal{I}_s(\widetilde{V}^{(1)}_i)\Big)\\ &= (\{1\}\cup\mathcal{I}_{1,2s})\bigcup\Big(\cup_{i\in\widetilde{\mathcal{I}}^{(1)}_{1,s}}(\mathcal{I}_{i,s}\cup\{i\})\Big)\bigcup\Big(\cup_{i\in\mathcal{I}_{1,s}}\widetilde{\mathcal{I}}^{(1)}_{i,s}\Big) \qquad (4.41)\\ &= \Big(\{1\}\cup\widetilde{\mathcal{I}}^{(1)}_{1,2s}\Big)\bigcup\Big(\cup_{i\in\mathcal{I}_{1,s}}(\widetilde{\mathcal{I}}^{(1)}_{i,s}\cup\{i\})\Big)\bigcup\Big(\cup_{i\in\widetilde{\mathcal{I}}^{(1)}_{1,s}}\mathcal{I}_{i,s}\Big), \qquad (4.42)\end{aligned}$$

where the last equality is by considering the the groups $(i),(ii),(iii)$ relative to the new configuration $\widetilde{\mathcal{V}}^{(1)}$. Even though these sets are equivalent, we think of (4.41) as $\mathcal{M}$ and (4.42) as $\widetilde{\mathcal{M}}^{(1)}$. We now have

$$\begin{aligned}\Big|\sum_{i=1}^n(\widetilde{X}^{(1)}_i-X_i)\Big| &= \Big|\sum_{i\in\mathcal{M}}(X_i-\widetilde{X}^{(1)}_i)\Big| \le \sum_{i\in\mathcal{M}}|X_i|+\sum_{i\in\widetilde{\mathcal{M}}^{(1)}}|\widetilde{X}^{(1)}_i|\\ &\le \sum_{i\in\mathcal{I}_{1,2s}\cup\{1\}}|X_j|+\sum_{i\in\widetilde{\mathcal{I}}^{(1)}_{1,2s}\cup\{1\}}|\widetilde{X}^{(1)}_j| \qquad (4.43)\\ &\quad+\sum_{i\in\widetilde{\mathcal{I}}^{(1)}_{1,s}}\sum_{j\in\mathcal{I}_{i,s}\cup\{i\}}|X_j|+\sum_{i\in\mathcal{I}_{1,s}}\sum_{j\in\widetilde{\mathcal{I}}^{(1)}_{i,s}\cup\{i\}}|\widetilde{X}^{(1)}_j| \qquad (4.44)\\ &\quad+\sum_{i\in\mathcal{I}_{1,s}}\sum_{j\in\widetilde{\mathcal{I}}^{(1)}_{i,s}\setminus(\mathcal{I}_{1,2s}\cup\{1\})}|X_j|+\sum_{i\in\widetilde{\mathcal{I}}^{(1)}_{1,s}}\sum_{j\in\mathcal{I}_{i,s}\setminus(\widetilde{\mathcal{I}}^{(1)}_{1,2s}\cup\{1\})}|\widetilde{X}^{(1)}_j|, \qquad (4.45)\end{aligned}$$

where we remove some double counting in (4.45). By considering the process of generating $\widetilde{\mathcal{V}}^{(1)}$ from $\mathcal{V}$ in reverse, observe that the left and right hand terms of (4.43)–(4.45) have the same distributions, and thus we only need to bound the qth-moment of one term of each.

To bound the first term of (4.43), observe that, for each $i\in\mathcal{I}_{1,2s}$,

$$|X_i| \le c(N_s(V_i))^\beta \le c(N_{3s}(V_1))^\beta,$$

and that there are $N_{2s}(V_1)$ such indices; hence

$$\sum_{i\in\mathcal{I}_{1,2s}\cup\{1\}}|X_j| \le cN_{2s}(V_1)(N_{3s}(V_1))^\beta \le c(N_{3s}(V_1))^{\beta+1},$$

implying that

$$\Big\|\sum_{i\in\mathcal{I}_{1,2s}\cup\{1\}}|X_j|\Big\|_q \le c\|(N_{3s}(V_1))^{\beta+1}\|_q = c\|N_{3s}(V_1)\|^{\beta+1}_{(\beta+1)q}.$$

Because $\mathcal{L}(N_{3s}(V_1)) = Bi(n-1,9\pi s^2/n)+1$, we may then apply Lemma 6.2 to the right hand side of this last display to obtain

$$\Big\|\sum_{i\in\mathcal{I}_{1,2s}\cup\{1\}}|X_j|\Big\|_q \le c(C_A\max\{9\pi s^2,(\beta+1)q\}+1)^{\beta+1}. \qquad (4.46)$$

This article is protected by copyright. All rights reserved.

Author Manuscript

Author Manuscript

As previously mentioned, this also bounds the second term of (4.43).

To bound the second term of (4.44), for $x \in C_n$, let $\widetilde{N}_s^{(1)}(x)$ be the number of points in configuration $\widetilde{\mathcal{V}}^{(1)}$ that fall in $B_s(x)$. Observe that, by the same arguments as for (4.43),

$$\sum_{i\in\mathcal{I}_{1,s}} \sum_{j\in\widetilde{\mathcal{I}}_{i,s}^{(1)}\cup\{i\}} |\widetilde{X}_j^{(1)}| \;\le\; \sum_{i\in\mathcal{I}_{1,s}} c(\widetilde{N}_{2s}^{(1)}(\widetilde{V}_i^{(1)}))^{\beta+1}.$$

By construction, $(\widetilde{N}_{2s}^{(1)}(\widetilde{V}_i^{(1)}))_{i\in\mathcal{I}_{1,s}}$ has the same distribution as the counts of the first (say) $|\mathcal{I}_{1,s}|$ $2s$-neighbourhoods in an independent uniform n-configuration, and so we can apply Lemma 6.1. Applying Lemmas 6.1 and 6.2 thus gives

$$\begin{aligned}\Big\|\sum_{i\in\mathcal{I}_{1,s}} c(\widetilde{N}_{2s}^{(1)}(\widetilde{V}_i^{(1)}))^{\beta+1}\Big\|_q \;&\le\; c\|N_s(V_1)-1\|_q\|(\widetilde{N}_{2s}^{(1)}(\widetilde{V}_i^{(1)}))^{\beta+1}\|_q\\ &\le\; c(C_A\max\{4\pi s^2,(\beta+1)q\}+1)^{\beta+2}. \qquad (4.47)\end{aligned}$$

To bound the first term of (4.45), first note that, for each $i \in \mathcal{I}_{1,s}$ and $j \in \widetilde{\mathcal{I}}_{i,s}^{(1)} \setminus \mathcal{I}_{1,2s}$, we have $|X_j| \le c(N_{2s}(\widetilde{V}_i^{(1)}))^\beta$ (recall that $N_{2s}(\widetilde{V}_i^{(1)})$ gives the number of germs that fall in $B_{2s}(\widetilde{V}_i^{(1)})$ *in configuration* $\mathcal{V}$, *not* $\widetilde{\mathcal{V}}^{(1)}$). However, because $N_{2s}(\widetilde{V}_i^{(1)})$ and $\mathcal{I}_{1,s}$ are dependent (to see why, consider the distribution of $N_{2s}(\widetilde{V}_i^{(1)})$ given $\mathrm{card}\{\mathcal{I}_{1,s}\} = n$), we are unable to apply Lemma 6.1 directly. To construct a bound on $\sum_{j\in\widetilde{\mathcal{I}}_{i,s}^{(1)}\setminus(\mathcal{I}_{1,2s}\cup\{1\})}|X_j|$ that is independent of $\mathcal{I}_{1,s}$, first note that, for each $i \in \mathcal{I}_{1,s}$ and $j \in \widetilde{\mathcal{I}}_{i,s}^{(1)} \setminus (\mathcal{I}_{1,2s}\cup\{1\})$, we have

$$|X_j| \;\le\; c(N_s(V_j))^\beta \;=\; c\,\mathrm{card}\{\mathcal{I}_s(V_j)\setminus\mathcal{I}_s(V_1)\}^\beta,$$

since $\mathcal{I}_s(V_j)\cap\mathcal{I}_s(V_1)=\emptyset$ when $J\notin\mathcal{I}_{1,2s}\cup\{1\}$. Then, because $j\in\widetilde{\mathcal{I}}_{i,s}^{(1)}$, $V_j\in B_s(\widetilde{V}_i^{(1)})$, and so $\mathcal{I}_s(V_j)\subset\mathcal{I}_{2s}(\widetilde{V}_i^{(1)})$; hence

$$|X_j| \;\le\; c\,\mathrm{card}\{\mathcal{I}_{2s}(\widetilde{V}_i^{(1)})\setminus\mathcal{I}_s(V_1)\}^\beta.$$

We now construct a new configuration, $\widehat{\mathcal{V}}^{(1)} = \{\widehat{V}_1^{(1)},\dots,\widehat{V}_n^{(1)}\}$, from $\mathcal{V}$, such that, for $i\in\mathcal{I}_{1,s}^c$, $\widehat{V}_i^{(1)} = V_j$ and, for $i\in\mathcal{I}_{1,s}\cup\{1\}$, $\widehat{V}_i^{(1)}$ is distributed uniformly on $C_n\setminus B_{i,s}$, independently of everything else. Extending our notation naturally, we now have

$$|X_j| \;\le\; c\,\mathrm{card}\{\mathcal{I}_{2s}(\widetilde{V}_i^{(1)})\setminus\mathcal{I}_s(V_1)\}^\beta \;\le\; c(\widehat{N}_{2s}(\widetilde{V}_i^{(1)}))^\beta.$$

By construction, $\widehat{N}_{2s}(\widetilde{V}_i^{(1)})$ is independent of $\mathcal{I}_{1,s}$ and is stochastically dominated by the distribution $\mathrm{Bi}(n, 4\pi s^2/(n-4\pi s^2))$. We may now apply Lemmas 6.1 and 6.2 to obtain

$$\begin{aligned}\Big\|\sum_{i\in\mathcal{I}_{1,s}}\sum_{j\in\mathcal{I}_{i,s}^{(1)}\setminus\mathcal{I}_{1,2s}}|X_j|\Big\|_q \;&\le\; c\|N_s(V_1)-1\|_q\,\|\widehat{N}_{2s}(V_\ell^{(1)})\|_{q(\beta+1)}^{\beta+1}\\ &\le\; c(C_A\max\{4\pi s^2,(\beta+1)q\}+1)^{\beta+2}. \qquad (4.48)\end{aligned}$$

Combining (4.38) with (4.43)–(4.48) we obtain

$$\|\widetilde{D}\|_q \;\le\; 6c(C_A\max\{9\pi s^2,(\beta+1)q\}+1)^{\beta+2}. \qquad (4.49)$$

Thus, using (4.40) and (4.49), we can apply Theorem 2.4 to obtain the required result. □

This article is protected by copyright. All rights reserved.

Corollary 4.8. *For each $1 \le l \le 11$ and for $q \in \{2, \lceil \log \sigma \rceil\}$, we have $\sigma^{-1}\|T_l\|_q \le bq^{3/2}$, for a suitable fixed choice of b.*

Proof. . For $i = 1, 2, 5, 6, 9, 11$, Lemma 4.5 implies that the conditions of Lemma 4.7 hold with $s = 4r$, $c = \kappa_5(d+2)^2$ and $\beta = 0$, so that there are constants b_i, $i = 1, 2, 5, 6, 9, 11$, not depending on σ, such that, for any $q \in \mathbb{N}$,

$$\sigma^{-1}\|T_i\|_q \;\le\; b_i q^{3/2}. \tag{4.50}$$

For $i = 3, 7, 8, 10$, we use (4.2), taking $X = \sum_{i=1}^n \frac{n\pi_{M_i}^{(d)} \mathrm{I}\{M_i \le n/2\}}{2(n-M_i-1)}$, W_d, W_{d-1}, Y_d and $Y = \sum_{j=1}^n S_j$, $\sum_{i=1}^n \gamma_{M_i}^{(d)}$, $\sum_{i=1}^n \gamma_{M_i}^{(d)}$, $\sum_{i=1}^n \gamma_{M_i}^{(d)}$, respectively; with the exception of Y_d, we can apply Lemma 4.7 with $s = 2r$, $c = \kappa_3(d+2)$ and $\beta = 0$ to prove that there is a constant $\hat{c} > 0$ such that, for all $k \in \mathbb{N}$,

$$\sigma^{-1}\|X - \mathbb{E}X\|_k \;\le\; \hat{c}k^{3/2}, \quad \text{and} \quad \sigma^{-1}\|Y - \mathbb{E}Y\|_k \;\le\; \hat{c}k^{3/2}. \tag{4.51}$$

For Y_d, rather than defining a Stein coupling and using Theorem 2.2, we quote Bartroff et al. (2018, Theorem 3.2), which implies that, for some $C > 0$ and all $t > 0$,

$$\mathbb{P}(|Y_d - \mathbb{E}Y_d| > t) \;\le\; \exp\left(-\frac{t^2}{2C\mathbb{E}Y_d}\right). \tag{4.52}$$

Since $\mathbb{E}Y_d \le n$, a standard calculation implies that (4.51) also holds for $Y = Y_d$. Now applying (4.2), and using the fact that, for $i = 3, 7, 8, 10$, $n^{-1}\mathbb{E}X \le 1$ and $n^{-1}\mathbb{E}Y \le \kappa_{3/2}(d+1)$, we deduce that there are constants $b_i, i = 3, 7, 8, 10$ such that

$$\sigma^{-1}\|T_i\|_2 \le b_i, \tag{4.53}$$

and, for $q = \lceil \log(\sigma) \rceil$, and for all σ large enough,

$$\sigma^{-1}\|T_i\|_q \;\le\; b_i \max\{\sigma^{-1}(\log\sigma)^3, (\log\sigma)^{3/2}\} \;\le\; b_i \log(\sigma)^{3/2}. \tag{4.54}$$

To bound T_4, we use (4.4), which implies that, for some constant $C = C_{r,d}$,

$$\mathbb{P}\left(\sum_{i=1}^n \mathrm{I}\{M_i > n/2\} \ne 0\right) \;\le\; Cn\exp(-n/6). \tag{4.55}$$

In addition, applying Lemma 4.5 to bound $S_{(\cdot)}$, we have

$$\left|\sum_{i=1}^n \frac{\pi_{M_i}^{(d)} \mathrm{I}\{M_i > n/2\}}{n - M_i - 1} \sum_{j=1}^n S_j\right| \;\le\; n^2\kappa_3(d+1) \quad \text{a.s.}, \tag{4.56}$$

since the interpretation of any summand with $M_i = n-1$ is zero, as remarked in Lemma 4.7. Thus $\sigma^{-1}\|T_4\|_q \le b_4$ for all $q \in \mathbb{N}$, and the corollary is proved. □

This article is protected by copyright. All rights reserved.

Author Manuscript

Hence, in particular, when applying Theorem 2.2, we can take $c_2 = 11b\lceil \log \sigma \rceil^{3/2}$ in (2.7), and $\sigma^{-1}\|T\|_2 \leq 11 \times 2^{3/2} b$ in (2.5).

Since $R = 0$ a.s., all that remains to be proved, in order to apply Theorem 2.2, is a bound of order O(1) for the smoothness term Υ, to be used in (2.5). Using the notation I, J, X and V_j^s from around (4.20), we define

$$\mathcal{F}_2 \;:=\; \sigma(I, X, J, V_I, V_J, V_J^s, \mathcal{V} \cap B_{2r}(V_I), \mathcal{V} \cap B_r(V_J), \mathcal{V} \cap B_r(V_J^s));$$

observe that D (and G) are $\mathcal{F}_2$-measurable. We adopt the convention that $J = I$ when $X = 0$.

Lemma 4.9. *For Υ defined in Theorem 2.2, we have $\Upsilon = O(1)$.*

Proof. From Lemma 4.5, we see that there is a constant K such that $|D| \leq K$, uniformly in n (see also the proof of Bartroff et al. (2018, Theorem 3.3)). Letting $N^{\mathcal{F}_2} := \mathrm{card}\{\mathcal{V} \cap (B_{2r}(V_I) \cup B_r(V_J) \cup B_r(V_J^s))\}$, we have

$$\begin{aligned}
\Upsilon &= \mathbb{E}[|GD(D-1)| S_2(\mathcal{L}(W_d \,|\, \mathcal{F}_2)] \\
&\leq \; \mu K(K+1)\big\{\mathbb{E}\big[S_2(\mathcal{L}(W_d \,|\, \mathcal{F}_2)\, \mathrm{I}\{N^{\mathcal{F}_2} \leq \sqrt{n}\}\big] + \mathbb{E}\big[S_2(\mathcal{L}(W_d \,|\, \mathcal{F}_2)\, \mathrm{I}\{N^{\mathcal{F}_2} > \sqrt{n}\}\big]\big\} \\
&\leq \; \mu K(K+1)\mathbb{E}\big[S_2(\mathcal{L}(W_d \,|\, \mathcal{F}_2)\, \mathrm{I}\{N^{\mathcal{F}_2} \leq \sqrt{n}\}\big] \qquad (4.57) \\
&\quad + \mu K(K+1)\mathbb{P}(N^{\mathcal{F}_2} > \sqrt{n}). \qquad (4.58)
\end{aligned}$$

To bound (4.58), observe that $\mathcal{L}(|\mathcal{V} \cap B_{2r}(V_I)|) = Bi(n-1, 4\pi r^2) + 1$. If $X = 1$, then $\mathcal{V} \cap B_r(V_J^s) \subseteq \mathcal{V} \cap B_{2r}(V_I)$ and $\mathrm{card}\{\mathcal{V} \cap B_r(V_J) \setminus B_{2r}(V_I)\}$ is stochastically smaller than a random variable with the distribution $\mathrm{Bi}(n-1, \pi r^2/(n - \pi r^2)) + 1$; and if $X = -1$, then $\mathcal{V} \cap B_r(V_J) \subseteq \mathcal{V} \cap B_{2r}(V_I)$ and $\mathrm{card}\{\mathcal{V} \cap B_r(V_J^s) \setminus B_{2r}(V_I)\}$ is stochastically smaller than a random variable with the distribution $\mathrm{Bi}(n-1, \pi r^2/(n - \pi r^2)) + 1$. Letting Y denote a random variable with distribution $\mathrm{Bi}(n-1, \pi r^2/(n - \pi r^2))$, and applying (4.4), we obtain

$$\mathbb{P}(N^{\mathcal{F}_2} > \sqrt{n}) \;\leq\; \mathbb{P}(N_{2r}(V_I) \geq \sqrt{n}/2) + \mathbb{P}(Y^{(J)} \geq \sqrt{n}/2) \;\leq\; C_r e^{-\sqrt{n}/6}, \qquad (4.59)$$

for some $C_r > 0$, uniformly in n. This implies that (4.58) is of order O(1).

To bound (4.57), given $\mathcal{F}_2$, we define a new configuration $\mathcal{V}^{\mathcal{F}_2} = \{V_1^{\mathcal{F}_2}, \ldots, V_n^{\mathcal{F}_2}\}$ by letting $V_i^{\mathcal{F}_2} = V_i$ if $i \in \mathcal{I}_{2r}(V_I) \cup \mathcal{I}_r(V_J) \cup \mathcal{I}_r(V_J^s)$, and otherwise letting $V_i^{\mathcal{F}_2}$ be uniformly distributed on $C_n \setminus (B_{2r}(V_I) \cup B_r(V_J) \cup B_r(V_J^s))$, independently of everything else. Observe that, if we let $W_d^{\mathcal{F}_2}$ be the number of germs in $\mathcal{V}^{\mathcal{F}_2}$ whose grain does not contain precisely d germs, then $\mathcal{L}(W_d^{\mathcal{F}_2}) = \mathcal{L}(W_d \,|\, \mathcal{F}_2)$. To establish the bound on Υ, we prove that, for any fixed event in $\mathcal{F}_2$ with $N^{\mathcal{F}_2} \leq \sqrt{n}$, we have

$$S_2(\mathcal{L}(\tilde{W}_d \,|\, \mathcal{F}_2)) \;=\; \mathrm{O}(n^{-1}). \qquad (4.60)$$

For ease of notation, during the remainder of the proof, we tacitly assume that every random quantity has distribution conditional on $\mathcal{F}_2$.

We first establish (4.60) under the assumption that $d \geq 1$. Divide the space C_n into disjoint rectangles R_i with height $7r/3$ and width $13r/3$; if we ignore any left over space, then there are $\left\lfloor \frac{\sqrt{n}}{7r/3} \right\rfloor \left\lfloor \frac{\sqrt{n}}{13r/3} \right\rfloor$ such rectangles. Let $\mathcal{R} := \{i\colon\ R_i \cap (B_{2r}(V_I) \cup B_r(V_J) \cup B_r(V_J^s)) = \emptyset\}$ be

This article is protected by copyright. All rights reserved.

Author Manuscript

the set of rectangles that do not intersect $B_{2r}(V_I) \cup B_r(V_J) \cup B_r(V_J^s)$. Letting $n_r := \mathrm{card}\{\mathcal{R}\}$, we see that

$$\left\lfloor \frac{\sqrt{n}}{7r/3} \right\rfloor \left\lfloor \frac{\sqrt{n}}{13r/3} \right\rfloor - 10 \leq n_r \;\leq\; \left\lfloor \frac{\sqrt{n}}{7r/3} \right\rfloor \left\lfloor \frac{\sqrt{n}}{13r/3} \right\rfloor, \tag{4.61}$$

since, for instance, a ball of radius $2r$ can intersect at most 6 of the rectangles; this means that n_r is of strict order n. Let $(x_i, y_i) \in C_n$ denote the coordinates of the bottom lefthand corner of rectangle i, and, for $j = 1, \ldots, 4$, let $c_{i,j} = (x_i + r/2 + 2rj/3,\, y_i + 7r/6)$. We say that the *good event* occurs in R_i, denoted GE_i, if *(i)* $N_{r/6}(c_{i,1}) = 2$, *(ii)* $N_{r/6}(c_{i,2}) = d$, *(iii)* $N_{r/6}(c_{i,3}) + N_{r/6}(c_{i,4}) = 1$, and *(iv)* $\mathcal{V} \cap \big(R_i \setminus \cup_{j=1}^4 B_{r/6}(c_{i,j})\big) = \emptyset$. Observe that, if the good event occurs in R_i, then the total number of germs in R_i whose grain contains exactly d germs is 1 if $N_{r/6}(c_{i,3}) = 1$, and is 0 if $N_{r/6}(c_{i,4}) = 1$, regardless of the configuration of germs outside R_i.

Let $\mathrm{I}\{GE_i\}$ denote the indicator of GE_i, and let $\mathcal{R}_{GE} := \{i \in \mathcal{R}\colon\ \mathrm{I}\{GE_i\} = 1\}$, $N_{GE} := \mathrm{card}\{\mathcal{R}_{GE}\}$; define $U_{I,J} = \mathrm{Leb}\{B_{2r}(V_I) \cup B_r(V_J) \cup B_r(V_J^s)\} \leq 5\pi r^2$, noting that it is $\mathcal{F}_2$-measurable. For $i \in \mathcal{R}$, through elementary calculations, we obtain

$$\begin{aligned} \xi \;:=\; \mathbb{E}\,\mathrm{I}\{GE_i\} \;&=\; \binom{n - N^{\mathcal{F}_2}}{d+3} \left(\frac{|R_i|}{n - U_{I,J}}\right)^{d+3} \\ &\quad \times \left(\frac{n - U_{I,J} - |R_i|}{n - U_{I,J}}\right)^{n - N^{\mathcal{F}_2} - (d+3)} \frac{(d+3)!}{d!} \left(\frac{\pi(r/6)^2}{|R_i|}\right)^{d+3}. \end{aligned}$$

Under the assumption that $N^{\mathcal{F}_2} \leq \sqrt{n}$, we conclude that ξ converges to a positive constant, for d fixed, as $n \to \infty$. When combined with (4.61), this gives $\mathbb{E}(N_{GE}) = n_r \xi \asymp n$. We can then (suppressing the conditioning on $\mathcal{F}_2$ in the expectations) write

$$\begin{aligned} S_2(W_d \,|\, \mathcal{F}_2) \;&=\; \sup_{h:\|h\| \leq 1} \mathbb{E}[\Delta^2 h(W_d)] \\ &=\; \sup_{h:\|h\| \leq 1} \mathbb{E}[\Delta^2 h(W_d)\ \mathrm{I}\{|N_{GE} - \mathbb{E}(N_{GE})| \geq \mathbb{E}(N_{GE})/2\}] \\ &\quad + \sup_{h:\|h\| \leq 1} \mathbb{E}[\Delta^2 h(W_d)\ \mathrm{I}\{|N_{GE} - \mathbb{E}(N_{GE})| < \mathbb{E}(N_{GE})/2\}] \\ &\leq\; 4\mathbb{P}[|N_{GE} - n_r\xi| \geq n_r\xi/2] \qquad (4.62) \\ &\quad + \sup_{h:\|h\| \leq 1} \mathbb{E}\left[\Delta^2 h(W_d) \,\middle|\, |N_{GE} - n_r\xi| \leq \frac{n_r\xi}{2}\right]. \qquad (4.63) \end{aligned}$$

We establish (4.60) by separately demonstrating that both (4.62) and (4.63) are of order $\mathrm{O}(n^{-1})$.

To bound (4.62), observe that, for $i \neq j \in \mathcal{R}$,

$$\begin{aligned} \mathbb{E}(\mathrm{I}\{GE_i\}\,\mathrm{I}\{GE_j\}) \;&=\; \xi \mathbb{E}(\mathrm{I}\{GE_j\} |\, \mathrm{I}\{GE_i\} = 1) \\ &=\; \xi \binom{n - N^{\mathcal{F}_2} - (d+3)}{d+3} \left(\frac{|R_i|}{n - U_{I,J} - |R_i|}\right)^{d+3} \\ &\quad \times \left(\frac{n - U_{I,J} - 2|R_i|}{n - U_{I,J} - |R_i|}\right)^{n - N^{\mathcal{F}_2} - 2(d+3)} \frac{(d+3)!}{d!} \left(\frac{\pi(r/6)^2}{|R_i|}\right)^{d+3}. \end{aligned}$$

This article is protected by copyright. All rights reserved.

Author Manuscript

Author Manuscript

Under the assumption that $N^{\mathcal{F}_2} \leq \sqrt{n}$, we easily deduce that

$$\mathbb{E}(\mathrm{I}\{GE_i\}\,\mathrm{I}\{GE_j\}) \;=\; \xi^2(1+O(1/n)),$$

which leads to

$$\begin{aligned}\mathrm{Var}(N_{GE}) \;&=\; \sum_{i,j}\mathbb{E}(\mathrm{I}\{GE_i\}\,\mathrm{I}\{GE_j\}) - n_r^2\xi^2\\ &=\; n_r(n_r-1)\xi^2(1+O(1/n)) + n_r\xi - n_r^2\xi^2\\ &=\; \mathrm{O}(n).\end{aligned}$$

By Chebyshev's inequality, it follows that

$$\mathbb{P}(|N_{GE}-n_r\xi| > n_r\xi/2) \;\leq\; \frac{\mathrm{Var}(N_{GE})}{(n_r\xi/2)^2} \;=\; \mathrm{O}(1/n).$$

To bound (4.63), we let $X_i = \sum_{j=1}^n \mathrm{I}\{V_j \in R_i\}\,\mathrm{I}\{M_j = d\}$ be the number of germs in R_i whose grain contains d other germs, and define

$$Z_d \;:=\; W_d - \sum_{i\in\mathcal{R}_{GE}} X_i.$$

For the reasons described above, $\mathcal{L}(X_i \,|\, i \in \mathcal{R}_{GE}) = \mathrm{Be}(1/2)$ and, given $\mathcal{R}_{GE}$, $(X_i)_{i\in\mathcal{R}_{GE}}$ are conditionally i.i.d. and independent of Z_d.

By a standard argument, for $B \sim \mathrm{Bi}(m,1/2)$, $\sup_{h:\,\|h\|\leq 1}\mathbb{E}[\Delta^2 h(B)] \leq Cm^{-1}$, for a universal constant C. Since, on $|N_{GE}-n_r\xi| \leq n_r\xi/2$, we have $N_{GE} \geq n_r\xi/2 \asymp n$, it follows that, for all h with $\|h\| \leq 1$, we have

$$\mathbb{E}\left[\Delta^2 h(W_d)\,\middle|\,\mathcal{R}_{GE}, |N_{GE}-n_r\xi| \leq \frac{n_r\xi}{2}, Z_d\right] \;=\; \mathrm{O}(n^{-1}),$$

where the constant implied in the order term can be taken to be uniform in n and in the realizations of the conditioning random variables. Taking expectations establishes (4.60), and thus the lemma in proved in the case $d \geq 1$.

The proof in case where $d = 0$ follows the same lines, once the definition of the good event in R_i is modified to *(i)* $N_{r/6}(c_{i,1}) = 2$, *(ii)* $N_{r/6}(c_{i,2}) + N_{r/6}(c_{i,3}) = 1$, *(iii)* $N_{r/6}(c_{i,4}) = 0$, and *(iv)* $\mathcal{V} \cap \big(R_i \setminus \cup_{j=1}^4 B_{r/6}(c_{i,j})\big) = \emptyset$. □

In view of Corollary 4.8 and Lemma 4.9, and because $R = 0$ a.s., Theorem 3.2 is proved.

5 PROOF OF THEOREM 2.2

We begin by giving a high level overview of Stein's method for LCLTs. Stein's method is used to bound distances that can be expressed in the form

$$d_{(*)}(\mathcal{L}(W),\mathcal{L}(Z)) \;=\; \sup_{h\in\mathcal{H}_{(*)}}\{\mathbb{E}h(W) - \mathbb{E}h(Z)\};$$

This article is protected by copyright. All rights reserved.

Author Manuscript

here, we consider $\mathcal{H}_{\mathrm{TV}} = \{\mathbf{1}_A \colon A \subset \mathbb{Z}\}$ and $\mathcal{H}_{\mathrm{loc}} = \{\pm\mathbf{1}_{\{a\}} \colon a \in \mathbb{Z}\}$. To apply Stein's method, we first find an operator $\mathcal{A}$ such that $\mathbb{E}\mathcal{A}f(Z) = 0$ for all functions f for which this expectation exists; we then solve the *Stein equation*

$$\mathcal{A}f_h \;=\; h - \mathbb{E}h(Z),$$

for all $h \in \mathcal{H}_{(*)}$, yielding the set of solutions $\mathcal{F}_{(*)}$; finally, we bound

$$d_{(*)}(\mathcal{L}(W), \mathcal{L}(Z)) \;=\; \sup_{f_h \in \mathcal{F}_{(*)}} \mathbb{E}\mathcal{A}f_h(W) \tag{5.1}$$

by deriving properties of the solutions $f_h \in \mathcal{F}_{(*)}$ and by exploiting probabilistic properties of W. In the case of approximation by translated Poisson distributions, deriving properties of $f_h \in \mathcal{F}_{(*)}$ reduces to studying the solutions to the Poisson Stein equation

$$\lambda f_h(i+1) - i f_h(i) \;=\; h(i) - \mathbb{E}h(Y), \quad i \geq 0; \quad Y \sim \mathrm{Po}(\lambda). \tag{5.2}$$

Indeed, recalling the definition of the translated Poisson distribution in Section 2.1 and the notation s and γ of (2.1), we need only to take $\lambda := \sigma^2 + \gamma$ and to replace $f_h(x)$ by $g_h(x) = f_h(x-s)$ for $x \in \mathbb{Z}$. If $h = \mathbf{1}_{\{a\}}$ and f_a is the corresponding solution to (5.2), then the primary property of f_a used to establish the local bound in Theorem 2.2 (see (5.7)–(5.9)) is

$$|\Delta f_a(x)| \;\leq\; \frac{1}{\lambda^{3/2}\sqrt{2e}} + \frac{|\lambda - x|}{\lambda^2} + \frac{1}{\lambda}\mathbf{1}_{\{a\}}(x), \tag{5.3}$$

where Δ denotes the first difference operator $\Delta g(k) \;:=\; g(k+1) - g(k)$. We refer the reader to Barbour et al. (2019, Lemma 3.3) for the derivation of (5.3).

After deriving the necessary properties of $\mathcal{F}_{(*)}$, it still remains to bound (5.1) using probabilistic properties of W, for which we use a Stein coupling. If (W, W', G) is a Stein coupling and $D := W' - W$, then

$$\begin{aligned} &|\mathbb{E}[\lambda g_h(W+1) - W g_h(W)]| &(5.4)\\ &\quad= \;|\mathbb{E}[\lambda \Delta g_h(W) + G(g_h(W') - g_h(W))]| \\ &\quad\leq\; |\mathbb{E}[(\lambda - GD)\Delta g_h(W)]| + |\mathbb{E}[GD\Delta g_h(W) - G(g_h(W') - g_h(W))]|. &(5.5)\end{aligned}$$

If we can find a Stein coupling, it allows us to bound (5.5), which is generally easier to bound than (5.4). Roughly speaking, if we continue this derivation and apply (5.3) to each g_a, then we obtain the following result, which is relatively straightforward to put together from Barbour et al. (2018).

Theorem 5.1 (Corollary 2.3 and Lemma 2.6 of Barbour et al. (2018))**.** *Let (W, W', G, R) be an approximate Stein coupling with W and W' integer valued, $\mathbb{E}W = \mu$ and $Var(W) = \sigma^2$. Set $D := W' - W$, and let $\mathcal{F}_1$ and $\mathcal{F}_2$ be sigma algebras such that W is $\mathcal{F}_1$-measurable and such that (G, D) is $\mathcal{F}_2$-measurable. Define*

$$\Upsilon \;:=\; \mathbb{E}[|GD(D-1)|\, S_2(\mathcal{L}(W \,|\, \mathcal{F}_2))],$$

and

$$T := |\mathbb{E}[GD \,|\, \mathcal{F}_1] - \mathbb{E}[GD]|.$$

This article is protected by copyright. All rights reserved.

Then

$$d_{\mathrm{TV}}\big(\mathcal{L}(W), TP(\mu,\sigma^2)\big) \;\le\; \frac{1}{\sigma}\big(\sigma^{-1}\|T\|_1 + 2\|R\|_2 + 2(\Upsilon+1)\big). \tag{5.6}$$

Moreover, for any any positive t,

$$\begin{aligned}
\delta_{\mathrm{loc}} \;&:=\; d_{\mathrm{loc}}\big(\mathcal{L}(W), TP(\mu,\sigma^2)\big)\\
&\le\; \frac{1}{\sigma^2}\bigg(2\sigma^{-1}\|T\|_2 + t\bigg(t^{-1}\mathbb{E}\big[T\,\mathrm{I}[\sigma^{-1}T \ge t]\big] + \sigma\sup_{a\in\mathbb{Z}}\mathbb{P}(W=a)\bigg)\bigg) &(5.7)\\
&\quad + \frac{\|R\|_2}{\sigma^2}\bigg(3 + \sigma\sup_{a\in\mathbb{Z}}\mathbb{P}(W=a)\bigg) &(5.8)\\
&\quad + \frac{2(\Upsilon+1)}{\sigma^2}. &(5.9)
\end{aligned}$$

Proof of Theorem 2.2. The condition (2.5) of Theorem 2.2 and the Cauchy–Schwarz inequality imply that (5.6) is bounded by $5c_1\sigma^{-1}$, which is the total variation bound (2.6). For the local bound, it is immediate that (5.9) is bounded by $2c_1\sigma^{-2}$. Next, note that

$$\sup_{a\in\mathbb{Z}}\mathbb{P}(W=a) \le \delta_{\mathrm{loc}} + \sup_{a\in\mathbb{Z}}\mathrm{TP}(\mu,\sigma^2)\big(\{a\}\big) \le \delta_{\mathrm{loc}} + \sigma^{-1}, \tag{5.10}$$

and so (5.8) is bounded by $4c_1\sigma^{-2} + c_1\sigma^{-1}\delta_{\mathrm{loc}}$. For (5.7), the first term is easily seen to be bounded by $2c_1\sigma^{-2}$. To bound the second term, note that Markov's inequality implies that, with $q := \lceil \log\sigma \rceil$,

$$t^{-1}\mathbb{E}\big[T\,\mathrm{I}[\sigma^{-1}T \ge t]\big] \;\le\; \mathbb{E}\big[(\sigma^{-1}T)^q\big]\,\frac{\sigma}{t^q} \;\le\; \sigma(c_2/t)^q. \tag{5.11}$$

Choosing $t = ec_2$ implies that (5.11) is bounded by 1. Hence, invoking (5.10), we have

$$t\bigg(t^{-1}\mathbb{E}\big[T\,\mathrm{I}[\sigma^{-1}T \ge t]\big] + \sigma\sup_{a\in\mathbb{Z}}\mathbb{P}(W=a)\bigg) \;\le\; (2+\sigma\delta_{\mathrm{loc}})ec_2,$$

and hence the second term in (5.7) is bounded by $2ec_2\sigma^{-2} + ec_2\sigma^{-1}\delta_{\mathrm{loc}}$. Since, by Assumption (2.8), $(c_1 + ec_2)\sigma^{-1} \le 1/2$, it follows from the bounds on (5.7)–(5.9) that

$$(1/2)\delta_{\mathrm{loc}} \;\le\; \{8c_1 + 2ec_2\}\sigma^{-2},$$

completing the proof of the local bound (2.9). □

6 RANDOM SUM AND BINOMIAL MOMENT BOUNDS

The following two results are Barbour et al. (2018, Lemmas 5.1 and 5.2), and are stated without proof.

Lemma 6.1. *Let $\mathcal{I}$ be a finite index set, $\mathcal{E}$ be a random (possibly empty) subset of $\mathcal{I}$, and define $E := |\mathcal{E}|$. Let $\{Y_i\}_{i\in\mathcal{I}}$ be a collection of random variables independent of $\mathcal{E}$ and, for $\ell \in \mathbb{N}$, let $\max_{i\in\mathcal{I}}\|Y_i\|_\ell \le y$. Then*

$$\|\sum_{i\in\mathcal{E}} Y_j\|_\ell \le y\|E\|_\ell.$$

This article is protected by copyright. All rights reserved.

Author Manuscript

Author Manuscript

Lemma 6.2. *Let $n \in \mathbb{N}$, $0 \leq p \leq 1$, $Y \sim Bi(n,p)$, and $\ell \in \mathbb{N}$. Then $\|Y\|_\ell \leq A(np,\ell)$, where*

$$A(x,\ell) := \pi e^{e-2} \times \begin{cases} \ell/\log((e-1)), & \ell > x, \\ x, & \ell \leq x. \end{cases}$$

In particular, $A(x,l) \leq C_A(x \vee l) \leq C_A(x+l)$, where $C_A := \pi e^{e-2}/\log(e-1)$.

7 ACKNOWLEDGMENTS

The authors were partially supported by the Australian Research Council (ARC) Grant DP150101459 and the ARC Centre of Excellence for Mathematical and Statistical Frontiers, CE140100049. PB was partially supported by ARC Grant FL130100039. We thank the referees for their comments.

REFERENCES

A. D. Barbour and L. H. Y. Chen, editors, 2005. *An introduction to Stein's method*, volume 4 of *Lecture Notes Series. Institute for Mathematical Sciences. National University of Singapore*. Singapore University Press, Singapore.

A. D. Barbour, M. Karoński, and A. Ruciński, 1989. A central limit theorem for decomposable random variables with applications to random graphs. *Journal of Combinatorial Theory, Series B*, 47(2):125–145.

A. D. Barbour, L. Holst, and S. Janson, 1992. *Poisson approximation*, volume 2 of *Oxford Studies in Probability*. The Clarendon Press, Oxford.

A. D. Barbour, N. Ross, and Y. Wen, 2018. Central moment inequalities using Stein's method. *arXiv preprint arXiv:1802.10225*.

A. D. Barbour, A. Röllin, and N. Ross, 2019. Error bounds in local limit theorems using Stein's method. *Bernoulli*, 25(2):1076–1104.

J. Bartroff, L. Goldstein, and U. Işlak, 2018. Bounded size biased couplings, log concave distributions and concentration of measure for occupancy models. *Bernoulli*, 24(4B):3283–3317.

S. Chatterjee, 2008. A new method of normal approximation. *The Annals of Probability*, 36 (4):1584–1610.

L. H. Chen and A. Röllin, 2010. Stein couplings for normal approximation. *arXiv preprint arXiv:1003.6039*.

L. H. Y. Chen, 1975. Poisson approximation for dependent trials. *The Annals of Probability*, 3(3):534–545.

L. H. Y. Chen, L. Goldstein, and Q.-M. Shao, 2011. *Normal approximation by Stein's method*. Probability and its Applications (New York). Springer, Heidelberg.

This article is protected by copyright. All rights reserved.

Author Manuscript

B. Davis and D. McDonald, 1995. An elementary proof of the local central limit theorem. *Journal of Theoretical Probability*, 8(3):693–701.

X. Fang, 2014. Discretized normal approximation by Stein's method. *Bernoulli*, 20(3):1404–1431.

L. Goldstein, 2013. A Berry–Esseen bound with applications to vertex degree counts in the Erdős–Rényi random graph. *The Annals of Applied Probability*, 23(2):617–636.

L. Goldstein and M. D. Penrose, 2010. Normal approximation for coverage models over binomial point processes. *The Annals of Applied Probability*, 20(2):696–721.

H.-K. Hwang and S. Janson, 2008. Local limit theorems for finite and infinite urn models. *The Annals of Probability*, 36(3):992–1022.

W. Kordecki, 1990. Normal approximation and isolated vertices in random graphs. In *Random graphs*, volume 87, pages 131–139.

R. Lachièze-Rey and G. Peccati, 2017. New Berry–Esseen bounds for functionals of binomial point processes. *The Annals of Applied Probability*, 27(4):1992–2031.

T. Lindvall, 2002. *Lectures on the coupling method*. Courier Corporation.

L. Mattner and B. Roos, 2007. A shorter proof of Kanter's Bessel function concentration bound. *Probability Theory and Related Fields*, 139(1-2):191–205.

C. McDiarmid, 1998. Concentration. In *Probabilistic methods for algorithmic discrete mathematics*, pages 195–248. Springer.

D. R. McDonald, 1980. On local limit theorem for integer-valued random variables. *Theory of Probability & Its Applications*, 24(3):613–619.

D. R. McDonald, 2005. The local limit theorem: A historical perspective. *JIRSS*, 4(2).

M. Penrose, 2003. *Random geometric graphs*, volume 5 of *Oxford Studies in Probability*. Oxford University Press, Oxford.

M. D. Penrose and Y. Peres, 2011. Local central limit theorems in stochastic geometry. *Electronic Journal of Probability*, 16:no. 91, 2509–2544.

M. D. Penrose and J. E. Yukich, 2001. Central limit theorems for some graphs in computational geometry. *The Annals of Applied Probability*, 11(4):1005–1041.

V. V. Petrov, 1975. *Sums of independent random variables*. Springer-Verlag, New York-Heidelberg. Translated from the Russian by A. A. Brown, Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 82.

A. Röllin, 2005. Approximation of sums of conditionally independent variables by the translated Poisson distribution. *Bernoulli*, 11(6):1115–1128.

This article is protected by copyright. All rights reserved.

Author Manuscript

A. Röllin, 2007. Translated Poisson approximation using exchangeable pair couplings. *The Annals of Applied Probability*, 17(5/6):1596–1614.

A. Röllin and N. Ross, 2015. Local limit theorems via Landau–Kolmogorov inequalities. *Bernoulli*, 21(2):851–880.

N. Ross, 2011. Fundamentals of Stein's method. *Probabability Surveys*, 8:210–293.

C. Stein, 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*, pages 583–602. Univ. California Press, Berkeley, Calif.

C. Stein, 1986. *Approximate computation of expectations.* Institute of Mathematical Statistics Lecture Notes—Monograph Series, 7. Institute of Mathematical Statistics, Hayward, CA.

This article is protected by copyright. All rights reserved.