The Annals of Applied Probability
2019, Vol. 29, No. 2, 1046–1069
https://doi.org/10.1214/18-AAP1425

CENTRAL LIMIT THEOREMS IN THE CONFIGURATION MODEL

BY A. D. BARBOUR[1] AND ADRIAN RÖLLIN[2]

Universität Zürich and National University of Singapore

We prove a general normal approximation theorem for local graph statistics in the configuration model, together with an explicit bound on the error in the approximation with respect to the Wasserstein metric. Such statistics take the form $T := \sum_{v\in V} H_v$, where V is the vertex set, and H_v depends on a neighbourhood in the graph around v of size at most ℓ. The error bound is expressed in terms of ℓ, $|V|$, an almost sure bound on H_v, the maximum vertex degree $d_{\max}$ and the variance of T. Under suitable assumptions on the convergence of the empirical degree distributions to a limiting distribution, we deduce that the size of the giant component in the configuration model has asymptotically Gaussian fluctuations.

1. Introduction. Random graphs with a prescribed degree sequence have been intensively studied in recent years. This is largely because the binomial degree distributions that are automatic in Erdős–Rényi random graphs do not correspond well with those of many networks observed in applications, making more plausible null models essential for assessing statistical significance. One of the first results was obtained by Bender and Canfield (1978), who investigated the total number of graphs with a prescribed degree sequence. Their result was later generalised by Bollobás (1980), who made use of the *configuration model*; this is a random multigraph (i.e., a graph possibly containing loops and multiple edges), obtained from a randomly chosen perfect matching of the elements of the set of half-edges attached to each vertex. The importance of the configuration model lies in the fact that, conditionally on there being no loops or multiple edges, the resulting graph is distributed as a uniformly chosen simple graph with the given degree sequence.

Under certain conditions on the degree sequence, Molloy and Reed (1995) showed that the configuration model has a giant component which spans a fixed fraction of the vertices, in the limit when the number of vertices tends to infinity,

Received October 2017; revised July 2018.
[1]Supported in part by Australian Research Council Grants Nos. DP120102728, DP120102398, DP150101459 and DP150103588, and by their Centre of Excellence for Mathematical and Statistical Frontiers.
[2]Supported in part by NUS Research Grant R-155-000-167-112.
MSC2010 subject classifications. Primary 60F05; secondary 05C80.
Key words and phrases. Configuration random graph model, giant component, central limit theorem, Stein's method.

with the degree distribution converging to a fixed probability distribution. However, relatively little progress has been made towards understanding the *second-order fluctuations* of the size of the giant component. Only recently, Ball and Neal (2017) were able to take a first step in that direction, by providing an asymptotic expression for the variance of the size of the giant component, and—by means of numerical simulations—conjecturing a central limit theorem (CLT).

For the Erdős–Rényi random graph $G(n, p)$, CLTs are readily available for many graph statistics; see, for example, Ruciński (1988) and Barbour, Karoński and Ruciński (1989) for subgraph counts, Janson and Luczak (2008a) for the so-called *susceptibility* and Pittel and Wormald (2005) and Bollobás and Riordan (2012) for the size of the giant component, to cite just a few examples. In contrast, for the configuration model, the literature on CLTs is rather sparse; we are only aware of the following CLTs: Janson and Luczak (2008b) for the k-core, Angel, van der Hofstad and Holmgren (2016) for the number of loops and multiple edges in the infinite-variance case, KhudaBukhsh et al. (2017) for certain statistics arising from the SI-epidemic on the configuration graph and Riordan (2012) for the size of the giant component in the barely supercritical case. Riordan's proof is quite different from ours. It is based on a careful analysis of the exploration process, and makes use of the graph being almost critical. Finally, Athreya and Yogeshwaran (2018) have proved a CLT for additive summary statistics in the subcritical case. Their proof is based on the martingale CLT and, therefore, also quite different from ours.

The purpose of this article is to prove a general CLT for statistics which can be expressed as sums of "local" vertex statistics. By "local", we mean statistics that are determined by a limited part of the neighbourhood of a vertex that is close to the vertex itself—for instance, by the ℓ closest neighbours with respect to graph distance, where a suitable rule is used if necessary to choose among neighbours that are at the same distance from the vertex. Then, as an application of the general result, we prove a CLT for the size of the giant component in the configuration model. Following the strategy of Ball and Neal (2017), instead of approximating the size of the giant component directly, we approximate the number of vertices in components of size at most n^β, for some small $\beta > 0$, which is asymptotically equivalent to the number of vertices *not* in the giant component.

2. Main results. Let $d = (d_1, \dots, d_n)$ be such that $m := \sum_{v=1}^n d_v$ is even, a "vector of degrees", and write $d_{\max} := \max_{1 \le v \le n} d_v$. Let $\mathcal{G} \sim \mathrm{CM}(d)$ be a realisation of the configuration multigraph on n vertices, labelled 1 through n, where vertex v has degree d_v. For each vertex $v \in [n] := \{1, 2, \dots, n\}$, let $\mathcal{T}(v)$ denote the rooted component in $\mathcal{G}$ containing v, with v being assigned the root label, making $\mathcal{T}(v)$ a finite, connected, vertex-labelled and rooted multigraph.

Let $h(\mathcal{T})$ be a real-valued function on finite, connected, vertex-labelled and rooted multigraphs $\mathcal{T}$. We are interested in the random quantity

$$U_{d,h} = \sum_{v \in [n]} h\big(\mathcal{T}(v)\big), \tag{2.1}$$

and the corresponding centred and normalised version $\hat{U}_{d,h} = (U_{d,h} - \mu_{d,h})/\sigma_{d,h}$, where $\mu_{d,h} = \mathbb{E}U_{d,h}$ and $\sigma^2_{d,h} = \operatorname{Var} U_{d,h}$.

2.1. *A central limit theorem for local graph statistics.* For two distributions F and G on $\mathbb{R}$, denote by $d_{\mathrm{W}}(F, G)$ the Wasserstein distance between F and G; that is, $d_{\mathrm{W}}(F, G) := \sup_{f\in\mathcal{F}_{\mathrm{W}}} |\int f dF - \int f dG|$, where $\mathcal{F}_{\mathrm{W}}$ denotes the set of real valued functions with Lipschitz constant at most 1. Write $\mathrm{N}(0, 1)$ for the standard normal distribution. For any multigraph $\mathcal{T}$, we use $|\mathcal{T}|$ to denote the number of vertices in $\mathcal{T}$.

Let h be a nonnegative function on finite, connected, vertex-labelled and rooted multigraphs, and let $\ell \geq 1$. We say that h *only depends on the first* ℓ *vertices away from the root*, if the following holds. Let $\mathcal{T}$ and $\mathcal{T}'$ be two finite, connected, vertex-labelled and rooted multigraphs. Let $\mathcal{T}_\ell$ and $\mathcal{T}'_\ell$ be the respective subgraphs induced by exploring the respective multigraphs, starting with the root, breadth first and, for each wave of exploration, smallest label first, until at most ℓ vertices (including the root) have been explored. If $\mathcal{T}_\ell = \mathcal{T}'_\ell$, then $h(\mathcal{T}_\ell) = h(\mathcal{T}'_\ell)$.

THEOREM 2.1. *Assume that h only depends on the first ℓ vertices away from the root. Then, if $2 \leq d_{\max} \leq n^{1/4}$, $12 \leq \ell \leq n^{1/4}$ and $m \geq \max\{n, 7d^2_{\max}\ell\}$, it follows that*

$$d_{\mathrm{W}}\big(\mathscr{L}(\hat{U}_{d,h}), \mathrm{N}(0, 1)\big) \leq \frac{\|h\|^3 d^2_{\max}\ell^{10} n}{4536\sigma^3_{d,h}} + \frac{\|h\|^2 d^2_{\max}\ell^8 n^{1/2}}{78\sigma^2_{d,h}},$$

where $\|h\|$ denotes supremum norm.

Theorem 2.1 can potentially be applied to many graph statistics: sub-graph counts, the number of small components and the susceptibility, to name but a few. The upper bounds on $d_{\max}$ and ℓ are largely unimportant, since the bound given in Theorem 2.1 would typically be larger than 1 if they were violated. Evaluating the order of the approximation error depends on obtaining lower bounds on the variance $\sigma^2_{d,h}$. For the size of the giant component, it is of strict order n; see also the discussion in Section 4.3.

2.2. *Size of the giant component.* We now use Theorem 2.1 to prove our second main result. For each $n \geq 1$, let $d^{(n)} = (d^{(n)}_1, \ldots, d^{(n)}_n)$ be a vector of degrees on n vertices; let $\pi = (\pi_j)_{j\geq 0}$ be a probability distribution on the nonnegative integers. We denote by $\pi^{(n)} := \{n^{-1} \sum_{i=1}^n \mathrm{I}[d^{(n)}_i = j], 0 \leq j \leq n\}$ the empirical degree distribution of $d^{(n)}$. Let the random variable D_n have distribution $\pi^{(n)}$, and let D have distribution π. Denote by D^* a random variable having the size-bias distribution of D, that is, $\mathbb{P}[D^* = j] = j\pi_j/\mathbb{E}D$, $j \geq 1$, and note that $\mathbb{E}D^* = \mathbb{E}D^*/\mathbb{E}D$. The following is the main result of Ball and Neal (2017); all limits are to be understood as $n \to \infty$.

THEOREM 2.2 (Ball and Neal (2017), Theorem 2.1). *Assume that*:

(1) $\mathbb{E}D^* > 2$,
(2) $\pi_1 > 0$,
(3) $\mathbb{E}D^3 < \infty$,
(4) $\mathscr{L}(D_n) \to \mathscr{L}(D)$,
(5) $\mathbb{E}D_n^3 \to \mathbb{E}D^3$;
(6) *there exists* $\beta > 0$ *such that* $d_{\mathrm{TV}}(\mathscr{L}(D_n^*), \mathscr{L}(D^*)) = \mathrm{O}(n^{-\beta})$ *and such that* $|\mathbb{E}D_n - \mathbb{E}D| = \mathrm{O}(n^{-\beta})$;
(7) *there exists* $\delta > 0$ *such that* $d_{\max}^{(n)} = \mathrm{O}(n^{1/4-\delta})$.

Let R_n *denote the size of the largest component of* $\mathrm{CM}(d^{(n)})$. *Then there exists a positive constant* b^2 *such that, as* $n \to \infty$,

$$\operatorname{Var} R_n \sim nb^2.$$

REMARK 2.3. Condition (1) is equivalent to the usual threshold condition of Molloy and Reed (1995), which guarantees the existence of a giant component with high probability. Condition (2) ensures that, under Conditions (3), (4) and (5), with high probability not all vertices are either isolated or part of the same giant component; if $\pi_1 = 0$, $\operatorname{Var} R_n$ may be of smaller asymptotic order than n. Condition (3) is the moment condition of Ball and Neal (2017); Condition (4) is the condition that the empirical degree distribution converges; Conditions (5) and (6), respectively, are equivalent to Conditions (a)(ii) and (b) of Ball and Neal (2017); Condition (7) is Condition (c) of Ball and Neal (2017). Moreover, it is straightforward to check that Conditions (3), (4) and (5) imply Condition (a)(i) of Ball and Neal (2017), which is, as is easily verified, redundant there. Conditions (3), (5), (6) and (7) are required in their current forms in the proof given by Ball and Neal (2017), and we use their variance asymptotics in proving Theorem 2.4. We also need something close to Condition (7) to ensure that the bound in Theorem 2.1 is small. However, the conditions could possibly be relaxed somewhat.

THEOREM 2.4. *Assume that the conditions of Theorem 2.2 hold. Let* $\hat{R}_n$ *denote the centred and normalized version of* R_n. *Then*

$$\mathscr{L}(\hat{R}_n) \to N(0,1) \qquad \text{as } n \to \infty.$$

PROOF. As mentioned in the Introduction, instead of counting the number of vertices in the giant component, we proceed as Ball and Neal (2017) and count the number of vertices that are not in the giant component, which is just $n - R_n =: S_n$; a CLT for S_n obviously implies a CLT for R_n. Let $h_\ell(\mathcal{C})$ be the function that equals 1 if $|\mathcal{C}| \le \ell$ and equals 0 otherwise; hence, $U_n := U_{d^{(n)},h_\ell}$ from (2.1) is the number of vertices in components of size less or equal to ℓ in $\mathcal{G} \sim \mathrm{CM}(d^{(n)})$. Under the conditions of Theorem 2.4, it follows from the proof of Molloy and Reed

(1995), Lemma 11 (see also the discussion after Theorem 2.1 of Ball and Neal (2017)) that, for any fixed $\delta > 0$, the probability that a vertex lies in a component larger than $n^{\delta/10} =: \ell$, but not in the largest component, is bounded by Cn^{-2}, for some constant $C > 0$ independent of n; hence,

$$\mathbb{P}[S_n \neq U_n] \leq \frac{C}{n}. \tag{2.2}$$

In what follows, C will denote a generic constant that may differ from line to line, but is always independent of n. Let $\lambda_n := \mathbb{E}S_n$ and let $\tau_n^2 := \operatorname{Var} S_n$, and with $\mu_n = \mathbb{E}U_n$ and $\sigma_n^2 = \operatorname{Var} U_n$, let

$$a_n = \frac{\mu_n - \lambda_n}{\tau_n}, \qquad b_n^2 = \frac{\sigma_n^2}{\tau_n^2}.$$

Since, under the conditions of Theorem 2.4, $\sigma_n^2 \sim \tau_n^2$ (see the discussion after Theorem 2.1 of Ball and Neal (2017)) and since (2.2) implies $|\lambda_n - \mu_n| = \mathrm{O}(1)$, we have

$$|a_n| \leq \frac{C}{\sqrt{n}}, \qquad \left|1 - b_n^2\right| = \mathrm{o}(1).$$

It is a standard exercise in Stein's method for normal approximation to show that $d_{\mathrm{W}}(\mathrm{N}(0,1), \mathrm{N}(a, b^2)) \leq |a| + |1 - b^2|$. Now

$$\frac{U_n - \lambda_n}{\tau_n} = b_n\left(\frac{U_n - \mu_n}{\sigma_n}\right) + a_n,$$

so that, from the definition of d_{W},

$$d_{\mathrm{W}}\big(\mathscr{L}(\tau_n^{-1}(U_n - \lambda_n)), \mathrm{N}(a_n, b_n^2)\big) = b_n d_{\mathrm{W}}\big(\mathscr{L}(\sigma_n^{-1}(U_n - \mu_n)), \mathrm{N}(0,1)\big).$$

Now, let h be a real valued function bounded in modulus by 1 and with Lipschitz constant at most 1, and let Z have a standard normal distribution. We have

$$\begin{aligned}
&\left|\mathbb{E}h(\hat{S}_n) - \mathbb{E}h(Z)\right| \\
&\quad \leq 2\mathbb{P}[S_n \neq U_n] + d_{\mathrm{W}}\big(\mathscr{L}(\tau_n^{-1}(U_n - \lambda_n)), \mathrm{N}(a_n, b_n^2)\big) \\
&\qquad + d_{\mathrm{W}}\big(\mathrm{N}(a_n, b_n^2), \mathrm{N}(0,1)\big) \\
&\quad \leq C\left(\frac{1}{n} + \left(\frac{d_{\max}^2 \ell^{10}}{n^{1/2}} + \frac{d_{\max}^2 \ell^8}{n^{1/2}}\right)\right) + \mathrm{o}(1) = \mathrm{o}(1).
\end{aligned}$$

Since the set of bounded and Lipschitz continuous functions characterises convergence in distribution, the claim follows. □

REMARK 2.5. By conditioning on the multigraph to be simple, results obtained for the configuration model can often be transferred to simple graphs (see, e.g., Janson (2010)). However, as pointed out by Janson [(2010), Remark 1.4), distributional limit results cannot be transferred in general, so the question about the fluctuations of the size of the giant component in simple graphs with a prescribed degree sequence remains open.

3. Technical preliminaries. We begin by discussing some technical material as a preparation for the proof of Theorem 2.1, which is presented in the next section. Our main tool is Stein's method for normal approximation (see Stein (1972) and Chen, Goldstein and Shao (2011)). In particular, we make use of *Stein couplings*, for which we refer to Chen and Röllin (2010) for a detailed discussion.

3.1. *An abstract normal approximation theorem.* We say that a triple of random variables (W, W', G) is a *Stein coupling* if

$$\mathbb{E}\{Gf(W') - Gf(W)\} = \mathbb{E}\{Wf(W)\} \tag{3.1}$$

for all functions f for which the expectations exists. The proof of the following result is standard; see also Chen and Röllin (2010), Corollary 2.2.

THEOREM 3.1. *Let (W, W', G) be a Stein coupling with* $\operatorname{Var} W = 1$. *Then*

$$d_{\mathrm{W}}(\mathscr{L}(W), \mathrm{N}(0,1)) \le 0.8\sqrt{\operatorname{Var}\mathbb{E}(G\Delta|W)} + \mathbb{E}|G\Delta^2|,$$

where $\Delta = W' - W$.

Since Stein couplings are rather abstract, we now present a general construction that leads to the Stein coupling that is used in the proof of Theorem 2.1. It is not clear how other couplings that have appeared in the literature could have been applied. For instance, Stein's exchangeable pair coupling (see Stein (1986)) requires a certain linearity condition to be satisfied, which would be difficult to verify here. The flexibility offered by Stein couplings is thus important. Moreover, the following construction, which is a variant of Construction 2A of Chen and Röllin (2010), reduces the effort needed in deriving the bounds, because the X_i appearing in the construction need not be centred.

LEMMA 3.2. *Let $X_1, \dots, X_n$ be random variables and let $U = \sum_{i=1}^n X_i$. Let $W = \sigma^{-1}(U - \mathbb{E}U)$, where $\sigma^2 := \operatorname{Var} U < \infty$. Assume that, for each i, we can construct a random variable W_i' such that X_i and W_i' are independent and such that $\mathscr{L}(W_i') = \mathscr{L}(W)$. Let I be a random variable uniformly distributed on $[n]$ and independent of all else. Let*

$$G_i = -\frac{n}{\sigma}X_i, \qquad G = G_I, \qquad W' = W_I'.$$

Then (W, W', G) is a Stein coupling.

PROOF. Indeed,

$$\mathbb{E}\{Gf(W)\} = -\frac{1}{\sigma}\sum_{i=1}^n \mathbb{E}\{X_i f(W)\}$$

and

$$\mathbb{E}\{Gf(W')\} = -\frac{1}{\sigma}\sum_{i=1}^{n}\mathbb{E}\{X_i f(W_i')\} = -\frac{1}{\sigma}\sum_{i=1}^{n}\mathbb{E}X_i\mathbb{E}f(W),$$

so that

$$\mathbb{E}\{Gf(W') - Gf(W)\} = \frac{1}{\sigma}\sum_{i=1}^{n}\mathbb{E}\{(X_i - \mathbb{E}X_i)f(W)\} = \mathbb{E}\{Wf(W)\},$$

establishing (3.1). □

Broadly speaking, the difference $W - W_i'$ measures how much the sum W needs to be changed in order to become independent of its summand X_i. If this difference is small, the influence of X_i on W is weak. Randomizing over the summands, we can similarly argue that if $W - W_I'$ is small, then the influence of most X_i on W is weak, indicating that W is a sum of weakly dependent random variables and, therefore, making a CLT for W plausible.

3.2. *Configurations*. Instead of working with multigraphs, we follow the standard procedure and work instead with matchings of coloured balls, where the colours represent the individual vertices, and the balls represent the half-edges coming out of each vertex. For simplicity, we represent the colours by the numbers in $[n]$, and we represent the labels of the balls by the numbers in $[m]$. Let $d = (d_i)_{1\le i\le n}$ be a vector of degrees, so that $m = \sum_i d_i$ is even. One may think of balls 1 to d_1 having colour 1, balls $d_1 + 1$ to $d_1 + d_2$ having colour 2, and so forth. A configuration $\mathcal{G}$ on this set of balls is simply a perfect matching of the balls (note that a perfect matching consists of an unordered set of $m/2$ nonoverlapping pairs); denote by $\mathcal{M}_0(d)$ the set of all such perfect matchings, and denote by $\mathcal{U}_0(d)$ the uniform distribution on $\mathcal{M}_0(d)$ (the reason for the "0" in the notation becomes clear in the next paragraph). Now, in this notation, saying that $\mathcal{G}$ is a configuration random graph CM(d) is equivalent to saying that $\mathcal{G} \sim \mathcal{U}_0(d)$.

We also need to consider sub-configurations of $\mathcal{G}$, and these may contain *unpaired* balls. To make this precise, let $C \subset [n]$ be a subset of colours. We denote by $\mathcal{G}|_C$ the sub-configuration of $\mathcal{G}$ restricted to the set of balls which have any of the colours in C. This comprises the pairs for which the colours of *both* balls are in C, together with unpaired balls, with colours in C, whose partners in $\mathcal{G}$ have colour in C^c. Let $s(C)$ be the number of unpaired balls in $\mathcal{G}|_C$. Since in the base configuration $\mathcal{G}$ all balls are matched, we have $s(C) = s(C^c)$. Moreover, denote by $d(C) = (d_1(C), \dots, d_n(C))$ the degree sequence of length n, now restricted to the colours in C, so that $d_v(C) = d_v$ if $v \in C$ and $d_v(C) = 0$ if $v \notin C$. Finally, for any fixed s, let $\mathcal{M}_s(d(C))$ denote the set of matchings of the balls having colours in C, with s balls left unpaired, and denote by $\mathcal{U}_s(d(C))$ the uniform distribution on $\mathcal{M}_s(d(C))$.

The following lemma is a consequence of the uniform distribution on matchings.

LEMMA 3.3. *Let $C \subset [n]$ be a fixed subset of colours, and let $\mathcal{G} \sim \mathcal{U}_0(d)$. Then, conditionally on $s(C)$, we have that $\mathcal{G}|_C$ and $\mathcal{G}|_{C^c}$ are independent with*

$$\mathcal{G}|_C \sim \mathcal{U}_{s(C)}\big(d(C)\big) \quad \text{and} \quad \mathcal{G}|_{C^c} \sim \mathcal{U}_{s(C)}\big(d\big(C^c\big)\big). \tag{3.2}$$

In other words, the lemma says that, given $s(C)$, the two sub-configurations $\mathcal{G}|_C$ and $\mathcal{G}|_{C^c}$ are independent and themselves uniformly distributed, but on different vectors of degrees, and with unpaired balls if $s(C) > 0$.

3.3. *Truncated components*. The key assumption in Theorem 2.1 is that the function h only depends on the first ℓ vertices away from the root. It is therefore enough to explore the component for each vertex breadth first, up to the point where ℓ vertices have been explored—we denote this truncated component by $\mathcal{T}_\ell(v)$.

ALGORITHM A.

1. Let $j \leftarrow 1$, let $C_1 \leftarrow \{v\}$, and let $\mathcal{S}_1 \leftarrow \mathcal{G}|_{\{v\}}$.
2. If $j = \ell$ or if the sub-configuration $\mathcal{S}_j$ has no unpaired balls, then proceed to Step 7.
3. Among all the unpaired balls in $\mathcal{S}_j$, take the one with the *smallest ball label* from the colour with the *smallest colour label* among the colours closest to v, reveal its partner, and denote by w the colour of that partner.
4. Let $j \leftarrow j + 1$.
5. Let $C_j \leftarrow C_{j-1} \cup \{w\}$, and let $\mathcal{S}_j \leftarrow \mathcal{G}|_{C_j}$.
6. Return to Step 2.
7. Let $\mathcal{T}_\ell(v) \leftarrow \mathcal{S}_j$, and stop.

REMARK 3.4. A few comments are in order.

1. With $|\mathcal{T}_\ell(v)|$ denoting the number of colours in $\mathcal{T}_\ell(v)$, we have $|\mathcal{T}_\ell(v)| \leq \ell$.
2. At any stage j in the algorithm, each of the unpaired balls in $\mathcal{S}_j$ is paired (from the viewpoint of $\mathcal{G}$) with a colour not yet in $\mathcal{S}_j$, since by the definition of sub-configurations, once a new colour is added, all loops and all pairings with colours already in the configuration are automatically revealed.
3. It is not difficult to see that $\mathcal{T}_\ell(v)$ contains unpaired balls if and only if $\mathcal{T}(v)$ contains strictly more than ℓ vertices.
4. Under the assumptions on h, it is clear that $h(\mathcal{T}(v)) = h(\mathcal{T}_\ell(v))$; hence

$$U_{d,h} = \sum_{v \in [n]} h\big(\mathcal{T}_\ell(v)\big).$$

REMARK 3.5. We need to be able to condition the configuration $\mathcal{G}$ on the realization of $\mathcal{T}_\ell(v)$, or rather on the set of colours in $\mathcal{T}_\ell(v)$ along with the number of unpaired balls in $\mathcal{T}_\ell(v)$. To shorten notation, we let ξ_v denote the set of colours contained in $\mathcal{T}_\ell(v)$. Then, in order to construct the couplings of Lemma 3.2, we shall use statements of the form

$$\mathscr{L}(\mathcal{G}|_{\xi_v^c} | \xi_v, s(\xi_v)) = \mathcal{U}_{s(\xi_v)}(d(\xi_v^c)), \tag{3.3}$$

which is the analogue of the strong Markov property, but which does not follow immediately from Lemma 3.3. However, (3.3) can be rigorously established, and has in fact been used implicitly in the literature many times. We refer to Rozanov (1982) for the general theory of *stopping sets*; to establish (3.3), we refer in particular to Rozanov (1982), Lemma 1, page 75.

3.4. *Stein coupling.* We proceed to construct the Stein coupling necessary for the proof of the main theorem. In order to shorten notation, we consider n, ℓ and the degree sequence d to be fixed, and throughout the remainder of the article, we let ξ_v denote the set of colours contained in $\mathcal{T}_\ell(v)$ as before; in particular, note that $|\xi_v| \le \ell$.

Let $\mathcal{G} \sim \mathcal{U}_0(d)$. Let $X_v = h(\mathcal{T}_\ell(v))$ for every $v \in [n]$, and let $W = \sum_{v\in[n]} X_v$. Consider v as being fixed for now. Starting from $\mathcal{G}$, we construct a new configuration $\mathcal{G}'_v$ in such a way that $\mathcal{G}'_v$ is independent of $\mathcal{T}_\ell(v)$ and, therefore, independent of X_v. We establish independence by showing that the conditional distribution of the graph $\mathcal{G}'_v$, given $\mathcal{T}_\ell(v)$, does not depend on $\mathcal{T}_\ell(v)$; specifically, we now show that we have $\mathscr{L}(\mathcal{G}'_v|\mathcal{T}_\ell(v)) = \mathscr{L}(\mathcal{G})$.

We start with the sub-configuration $\mathcal{G}|_{\xi_v^c}$. By (3.3), $\mathcal{G}|_{\xi_v^c}$ contains $s(\xi_v)$ unpaired balls, which, conditionally on $\mathcal{T}_\ell(v)$, are uniformly distributed among the balls in $\mathcal{G}|_{\xi_v^c}$. We can thus add all unpaired balls from $\mathcal{G}|_{\xi_v}$ into $\mathcal{G}|_{\xi_v^c}$, and pair them with the partners that they were already paired with in $\mathcal{G}$; denote the resulting sub-configuration by $\mathcal{G}'_{v,0}$.

In order to add the remaining balls to $\mathcal{G}'_{v,0}$, we proceed step-wise, one pair at a time. In what follows, we will make random choices; so denote by B_v a suitable source of random numbers (e.g., B_v could just be a sequence of independent uniform random variables), in such a way that the choices made are a deterministic function of B_v.

Let K_v denote the number of pairings among the remaining balls, and, if $K_v > 0$, repeat the following procedure K_v times (if $K_v = 0$, there is nothing to be done). The result is a sequence $\mathcal{G}'_{v,0}, \mathcal{G}'_{v,1}, \dots, \mathcal{G}'_{v,K_v}$ with the property that the matchings in $\mathcal{G}'_{v,k}$ are uniformly distributed among all matchings between the respective balls; in particular, $\mathcal{G}'_{v,K_v} \sim \mathcal{U}_0(d)$, irrespective of $\mathcal{T}_\ell(v)$.

Constructing $\mathcal{G}'_{v,k}$ from $\mathcal{G}'_{v,k-1}$ for $1 \le k \le K_v$. Among the balls not in $\mathcal{G}'_{v,k-1}$, pick a pair of matched balls. Independently of all else, toss a coin that shows heads with probability $1/(m - 2(K_v - k) - 1)$ and that shows tails with probability $(m - 2(K_v - k) - 2)/(m - 2(K_v - k) - 1)$. If the coin shows heads, add the two balls to $\mathcal{G}'_{v,k-1}$, match them with each other and denote the resulting configuration by $\mathcal{G}'_{v,k}$. If the coin shows tails, pick one ball from $\mathcal{G}'_{v,k-1}$ uniformly at random, call it I and call its partner J, and break up the bond between I and J. Add one of the balls to be added to $\mathcal{G}'_{v,k-1}$ and match it with I, and then add the remaining ball to $\mathcal{G}'_{v,k-1}$ and match it with J; denote the resulting configuration by $\mathcal{G}'_{v,k}$.

It is not difficult to convince oneself that the matchings in each $\mathcal{G}'_{v,k}$ are uniformly distributed. Assuming that they are for $\mathcal{G}'_{v,k-1}$, we add a pair of balls, and leave them paired with probability $1/(m-2(K_v-k)-1)$; this is exactly the probability that two randomly chosen balls are matched in $\mathcal{G}'_{v,k}$. Otherwise, we match each of the two balls with a ball randomly chosen from $\mathcal{G}'_{v,k-1}$. The way to do this is to pick a random pair from $\mathcal{G}'_{v,k-1}$, break it up, and pair the two balls individually with the balls that have just been added.

We need to keep track of the colours involved in the construction. Let H_v denote the union of the colours of the balls I and J picked in the steps $1 \le k \le K_v$; that is, H_v consists of the colours of all the balls used in applying the above construction. Then define $\eta_v := \xi_v \cup H_v$. Note that ξ_v is a deterministic function of $\mathcal{G}$, whereas η_v is a deterministic function of both $\mathcal{G}$ and B_v; moreover, (3.3) also holds if ξ_v is replaced by η_v throughout.

Finally, let $\mathcal{G}'_v := \mathcal{G}'_{v,K_v}$, which is the configuration obtained after all balls have been put back. As mentioned before, conditionally on $\mathcal{T}_\ell(v)$, $\mathcal{G}'_v$ is distributed as $\mathcal{U}_0(d)$; hence, $\mathcal{G}'_v$ is independent of $\mathcal{T}_\ell(v)$, and hence, independent of X_v, since the latter is a function of $\mathcal{T}_\ell(v)$.

Now, for each $w \in [n]$, let $\mathcal{T}^v_\ell(w)$ denote the truncated component of w in $\mathcal{G}'_v$, that is, the truncated component obtained by applying Algorithm A in $\mathcal{G}'_v$, starting with w; let

$$W'_v = \frac{1}{\sigma_{d,h}}\left(\sum_{w=1}^{n} h\big(\mathcal{T}^v_\ell(w)\big) - \mu_{d,h}\right).$$

Since W'_v is independent of X_v and has the same distribution as W, we apply Lemma 3.2 to obtain a Stein coupling (W, W', G).

4. Proof of Theorem 2.1. As a first step, we show that $K_v - \ell$ is bounded with high probability, uniformly in $v \in [n]$, as long as $d_{\max}\ell$ grows no faster than a small power of m.

LEMMA 4.1. *For $k \ge 1$ and $m \ge 8(k \vee \ell)$,*

$$\mathbb{P}[K_v \ge \ell + k - 1] \le \frac{d_{\max}^{2k}\ell^{2k}}{k!m^k}.$$

In particular, if $m \ge n$,

$$\mathbb{P}\big[K_v \ge \ell + k - 1 \text{ for some } v \in [n]\big] \le \frac{d_{\max}^{2k}\ell^{2k}}{k!m^{k-1}}.$$

PROOF. At Step 3, Algorithm A reveals the partner of an unmatched ball; there are at most $\ell - 1$ of pairings revealed in this manner. In addition, at Step 5, other pairs may be revealed; if w denotes the colour of the partner revealed at Step 3, then some of its remaining $d_w - 1$ balls may be paired with unpaired balls having the previously chosen colours. At any stage, there are no more than $d_{\max}\ell$ unpaired balls. The chance of any two given balls being paired is at most $1/(m - 2\ell + 1)$, and the chance that any k given sets of two balls are each paired is at most $\{\prod_{j=1}^{k}(m - 2\ell - 2j + 3)\}^{-1}$. Hence the expected number of k-tuples of matched pairs in $\mathcal{T}_\ell(v)$, other than those revealed at Step 3, is at most

$$\begin{aligned}
&\binom{\binom{d_{\max}\ell}{2}}{k}\frac{1}{\prod_{j=1}^{k}(m-2\ell-2j+3)}\\
&\quad\le \binom{d_{\max}\ell}{2}^{k}\frac{1}{k!\prod_{j=1}^{k}(m-2\ell-2j+3)}\\
&\quad\le \frac{d_{\max}^{2k}\ell^{2k}}{k!2^k\prod_{j=1}^{k}(m-2\ell-2j+3)}\\
&\quad\le \frac{d_{\max}^{2k}\ell^{2k}}{k!m^k2^k\prod_{j=1}^{k}(1-2\ell/m-2j/m+3/m)} \le \frac{d_{\max}^{2k}\ell^{2k}}{k!m^k};
\end{aligned}$$

the fact that $2^k\prod_{j=1}^{k}(1 - 2\ell/m - 2j/m + 3/m) \ge 1$ follows from the fact that, under the given assumptions, $(1 - 2\ell/m - 2j/m + 3/m) \ge 1/2$. The two claims now easily follow. □

In what follows, we define the event A by

$$A := \{K_v \le \ell + 6 \text{ for all } v \in [n]\}, \tag{4.1}$$

and write

$$\gamma := \mathbb{P}[A^c] \le \frac{d_{\max}^{16}\ell^{16}}{8!m^7}, \tag{4.2}$$

where the upper bound in (4.2) follows directly from Lemma 4.1 with $k = 8$.

4.1. *Bounds on intersection probabilities.*

LEMMA 4.2. *Let $\mathcal{G} \sim \mathcal{U}_0(d)$, and assume that a random set of colours α has been obtained, perhaps using $\mathcal{G}$, but in such a way that*

$$\mathscr{L}(\mathcal{G}|_{\alpha^c}|\alpha, s(\alpha)) = \mathcal{U}_{s(\alpha)}(d(\alpha^c)).$$

Then, for every $v \in [n]$,

$$\mathbb{P}[\xi_v \cap \alpha \ne \varnothing|\alpha, s(\alpha)] \le I_{v\in\alpha} + \frac{2d_{\max}|\alpha|(\ell-1)}{m} \tag{4.3}$$

and

$$\mathbb{E}\{|\xi_v \cap \alpha| \big| \alpha, s(\alpha)\} \le \ell I_{v\in\alpha} + \frac{2d_{\max}|\alpha|(\ell-1)}{m}, \tag{4.4}$$

whenever $m \ge 2d_{\max}(|\alpha| + \ell)$. *Similarly, we have*

$$\mathbb{P}[\eta_v \cap \alpha \ne \varnothing, A \big| \alpha, s(\alpha)] \le I_{v\in\alpha} + \frac{2d_{\max}|\alpha|(3\ell+11)}{m} \tag{4.5}$$

and

$$\mathbb{E}\{|\eta_v \cap \alpha| I_A \big| \alpha, s(\alpha)\} \le (3\ell+12) I_{v\in\alpha} + \frac{2d_{\max}|\alpha|(3\ell+11)}{m}. \tag{4.6}$$

PROOF. Note that the event $\{\xi_v \cap \alpha \ne \varnothing\}$ happens either if $v \in \alpha$ or if $v \notin \alpha$ and, during the exploration of $\mathcal{T}_\ell(v)$, a ball whose partner is to be revealed in Step 3 of Algorithm A is an unpaired ball in $\mathcal{G}|_{\alpha^c}$, since then the partner is of a colour in α. If $v \in \alpha$, the bound is trivial; so, assume $v \notin \alpha$. Before the exploration starts, colour v is already considered explored, and since there could be loops, the pairings of up to $d_{\max}$ balls could have been revealed. If there are still unpaired balls at this phase, the exploration process starts. At this point, the probability that a specific ball is unpaired in $\mathcal{G}|_{\alpha^c}$ is at most $s(\alpha)/(m - d_{\max}|\alpha| - d_{\max})$, so the first ball whose partner is to be revealed has at most this probability of being an unpaired ball in $\mathcal{G}|_{\alpha^c}$. If the process continues, the next ball whose partner is to be revealed has a probability of at most $s(\alpha)/(m - d_{\max}|\alpha| - 2d_{\max})$ of being unpaired in $\mathcal{G}|_{\alpha^c}$, and so forth. The process continues for at most $\ell - 1$ steps, so that the probability of a ball being unpaired in $\mathcal{G}|_{\alpha^c}$ never exceeds

$$\frac{s(\alpha)}{m - |\alpha| d_{\max} - d_{\max}(\ell-1)} \le \frac{2s(\alpha)}{m} \le \frac{2d_{\max}|\alpha|}{m} \tag{4.7}$$

if $m > 2d_{\max}(|\alpha| + \ell - 1)$. Hence, whenever $v \notin \alpha$,

$$\mathbb{E}\{|\xi_v \cap \alpha| \big| \alpha, s(\alpha)\} \le \frac{2d_{\max}|\alpha|(\ell-1)}{m},$$

and (4.4) follows; bounding probabilities by expectations, (4.3) also follows.

In order to bound (4.6), we proceed in a similar manner. The bound is immediate if $v \in \alpha$, since $|\eta_v \cap \alpha| \le |\xi_v| + 2K_v \le \ell + 2(\ell+6)$ on the event A; so we now suppose that $v \notin \alpha$. We have $|\eta_v \cap \alpha| = |\xi_v \cap \alpha| + |(\eta_v \setminus \xi_v) \cap \alpha|$. The expectation of $|\xi_v \cap \alpha|$ can be bounded by (4.4), so we only need to consider the expectation of the latter summand on the event A. So, assume that $\mathcal{T}_\ell(v)$ has been explored. We now repeatedly perform the swapping of paired balls K_v times. Each time that a swapping is performed, we could pick a colour from α in one of two ways: either we pick a ball from $\mathcal{G}|_{\alpha^c}$ which is unpaired in $\mathcal{G}|_{\alpha^c}$, or we pick a ball directly from $\mathcal{G}|_\alpha$. The probability of the former is no greater than

$$\frac{s(\alpha)}{m - d_{\max}(|\alpha| + \ell)} \le \frac{2s(\alpha)}{m} \le \frac{2d_{\max}|\alpha|}{m},$$

whenever $m > 2d_{\max}(|\alpha| + \ell)$ (note that more and more pairs are being put back now, so that the denominator is in fact increasing), and the probability of the latter is no greater than

$$\frac{d_{\max}|\alpha|}{m - d_{\max}(|\alpha| + \ell)} \le \frac{2d_{\max}|\alpha|}{m}.$$

Since, on the event A, we have $K_v \le \ell + 6$, we deduce that the expected number of balls from $\mathcal{G}|_\alpha$ being reached is no greater than

$$\frac{4d_{\max}|\alpha|K_v}{m} \le \frac{4d_{\max}|\alpha|(\ell+6)}{m},$$

from which, when adding the last term of (4.4), (4.6) follows. Using once again expectations to bound probabilities, (4.5) also follows. □

COROLLARY 4.3. *Under the conditions of Lemma 4.2, and assuming that* $d_{\max} \ge 2$, *that* $\ell \ge 12$ *and that* $m \ge \max\{n, 2d_{\max}(|\alpha| + \ell)\}$, *we have*

$$\sum_{v=1}^{n} \mathbb{P}[\xi_v \cap \alpha \ne \varnothing | \alpha, s(\alpha)] \le |\alpha| + \frac{2d_{\max}|\alpha|(\ell-1)n}{m} \le 2|\alpha|d_{\max}\ell; \tag{4.8}$$

$$\sum_{v=1}^{n} \mathbb{E}\{|\xi_v \cap \alpha| | \alpha, s(\alpha)\} \le \ell|\alpha| + \frac{2d_{\max}|\alpha|(\ell-1)n}{m} \le 3|\alpha|d_{\max}\ell; \tag{4.9}$$

$$\sum_{v=1}^{n} \mathbb{P}[\eta_v \cap \alpha \ne \varnothing, A | \alpha, s(\alpha)] \le |\alpha| + \frac{2d_{\max}|\alpha|(3\ell+11)n}{m} \le 8|\alpha|d_{\max}\ell; \tag{4.10}$$

$$\sum_{v=1}^{n} \mathbb{E}\{|\eta_v \cap \alpha| I_A | \alpha, s(\alpha)\} \le (3\ell+12)|\alpha| + \frac{2d_{\max}|\alpha|(3\ell+11)n}{m} \le 10|\alpha|d_{\max}\ell. \tag{4.11}$$

We also need the following Efron–Stein type variance bound; see Chen, Goldstein and Röllin (in preparation).

LEMMA 4.4. *Let π be a uniform permutation on $[m]$, let $B = (B_1, \dots, B_n)$ be a sequence of i.i.d. random elements taking values in some suitable space, and let $f(\pi, B)$ be a real-valued function. Let $\tau_1, \dots, \tau_{m-1}$ be independent transpositions, also independent of all else, where τ_j transposes j and a randomly chosen integer in the set $\{j, \dots, m\}$, let B' be an independent copy of B, and for $1 \le i \le n$, let $B^i = (B_1, \dots, B_{i-1}, B'_i, B_i, \dots, B_n)$. Then*

$$\begin{aligned}\operatorname{Var} f(\pi, B) \le \frac{1}{2} \sum_{j=1}^{m-1} \mathbb{E}\bigl(f(\pi, B) - f(\pi\tau_j, B)\bigr)^2 \\ + \frac{1}{2} \sum_{i=1}^{n} \mathbb{E}\bigl(f(\pi, B) - f\bigl(\pi, B^i\bigr)\bigr)^2.\end{aligned}$$

The appearance of the particular transpositions in the lemma comes from using a well-known construction of uniform random permutations; see Knuth (1969).

4.2. *Completing the proof of Theorem* 2.1. Using these preliminary results, we can now bound the two terms appearing in Theorem 3.1.

Bounding $\mathbb{E}|G\Delta^2|$. The bound

$$|G_v| = \frac{n}{\sigma_{d,h}} \bigl|h\bigl(\mathcal{T}_\ell(v)\bigr)\bigr| \le \frac{n\|h\|}{\sigma_{d,h}} \tag{4.12}$$

is straightforward. Now, it is easy to see that we can write

$$\Delta_v = \frac{1}{\sigma_{d,h}} \sum_{w \in Q_v} \bigl(h\bigl(\mathcal{T}_\ell^v(w)\bigr) - h\bigl(\mathcal{T}_\ell(w)\bigr)\bigr), \tag{4.13}$$

where

$$Q_v := \{w : \xi_w \cap \eta_v \neq \varnothing\}.$$

Put in words, Q_v is the set of those colours whose truncated components are *potentially* affected when changing the underlying graph from $\mathcal{G}$ to $\mathcal{G}'_v$. We emphasize "potentially" here since, even if we have $\mathcal{T}_\ell(w) \neq \mathcal{T}_\ell^v(w)$, it is still possible that $h(\mathcal{T}_\ell(w)) = h(\mathcal{T}_\ell^v(w))$.

Note first that we have the crude bound $|\Delta_v| \le 2n\sigma_{d,h}^{-1}\|h\|$, so that, using (4.2),

$$\mathbb{E}\{|G|\Delta^2 I_{A^c}\} \le \frac{4n^3\|h\|^3\gamma}{\sigma_{d,h}^3} \le \frac{4\|h\|^3 d_{\max}^{16}\ell^{16}}{8!n^4\sigma_{d,h}^3}. \tag{4.14}$$

For the main part, we observe that

$$
\begin{aligned}
\mathbb{E}\{\Delta_v^2 I_A\} &\le \frac{4\|h\|^2}{\sigma_{d,h}^2}\mathbb{E}\{|Q_v|^2 I_A\} \\
&\le \frac{4\|h\|^2}{\sigma_{d,h}^2}\sum_{w=1}^n\sum_{w'=1}^n \mathbb{P}[\xi_{w'}\cap\eta_v\ne\varnothing, \xi_w\cap\eta_v\ne\varnothing, A].
\end{aligned} \tag{4.15}
$$

Enlarging the set with which $\xi_{w'}$ is allowed to overlap, we have

$$
\begin{aligned}
&\mathbb{P}[\xi_{w'}\cap\eta_v\ne\varnothing, \xi_w\cap\eta_v\ne\varnothing, A] \\
&\qquad \le \mathbb{E}\{\mathrm{I}[\xi_{w'}\cap(\xi_w\cup\eta_v)\ne\varnothing]\mathrm{I}[\xi_w\cap\eta_v\ne\varnothing]I_A\}.
\end{aligned} \tag{4.16}
$$

Noting that $|\xi_w|\le\ell$ and $|\eta_v|\le\ell+2(\ell+6)\le 4\ell$, the latter on the event A and for $\ell\ge 12$, we now apply (4.8) to the right-hand side of (4.16) twice, once for $\alpha=\xi_w\cup\eta_v$ and once for $\alpha=\eta_v$. Hence, the double sum on the right-hand side of (4.15) becomes

$$
\begin{aligned}
&\sum_{w=1}^n\sum_{w'=1}^n \mathbb{P}[\xi_w\cap\eta_v\ne\varnothing, \xi_{w'}\cap\eta_v\ne\varnothing, A] \\
&\qquad \le 2d_{\max}\ell\sum_{w=1}^n \mathbb{E}\{|\xi_w\cup\eta_v|\mathrm{I}[\xi_w\cap\eta_v\ne\varnothing]I_A\} \\
&\qquad \le 10d_{\max}\ell^2\sum_{w=1}^n \mathbb{E}\{\mathbb{P}[\xi_w\cap\eta_v\ne\varnothing, A|\eta_v, s(\eta_v)]\}\le 80d_{\max}^2\ell^4
\end{aligned} \tag{4.17}
$$

(this line of argument is frequently repeated in the calculations that follow). Thus,

$$
\mathbb{E}\{|Q_v|^2 I_A\}\le 80d_{\max}^2\ell^4, \tag{4.18}
$$

which yields

$$
\mathbb{E}\{|G|\Delta^2 I_A\} = \frac{1}{n}\sum_{v=1}^n \mathbb{E}\{|G_v|\Delta_v^2 I_A\}\le\frac{320\|h\|^3 d_{\max}^2\ell^4 n}{\sigma_{d,h}^3};
$$

hence

$$
\mathbb{E}\{|G|\Delta^2\}\le\frac{\|h\|^3 d_{\max}^2\ell^{10}n}{\sigma_{d,h}^3}\left(\frac{320}{12^6}+\frac{4d_{\max}^{14}\ell^6}{8!n^5}\right)\le\frac{\|h\|^3 d_{\max}^2\ell^{10}n}{4536\sigma_{d,h}^3},
$$

if $\max\{d_{\max},\ell\}\le n^{1/4}$.

Bounding $\operatorname{Var}\mathbb{E}(G\Delta|W)$. In order to bound $\operatorname{Var}\mathbb{E}(G\Delta|W)$, note that we can generate the random configuration $\mathcal{G}$ by means of a uniformly chosen random permutation π of $[m]$ by pairing balls $\pi(1)$ and $\pi(2)$, pairing balls $\pi(3)$ and $\pi(4)$ and so forth. But also note that the same configuration can be generated by more than one permutation: For example, both $(\pi(2), \pi(1), \pi(3), \dots, \pi(m))$ and $(\pi(3), \pi(4), \pi(1), \pi(2), \pi(5), \dots, \pi(m))$ represent the same graph. Moreover, recall that in the construction of $\mathcal{G}'_v$, we encode the choices made in Algorithm A by $B_1, \dots, B_n$. With this in mind, we define

$$X_v^\pi = h\big(\mathcal{T}_\ell(v)\big), \qquad X_w^{\pi,B_v} = h\big(\mathcal{T}_\ell^v(w)\big),$$

and we write Q_v^{π,B_v} instead of Q_v (and we do the same with other quantities) to make the dependence on π and B_v explicit. Now, with $B = (B_1, \dots, B_n)$, write

$$\mathbb{E}(G\Delta|\pi, B) = \frac{1}{\sigma_{d,h}^2} \sum_{v=1}^n X_v^\pi \sum_{w \in Q_v^{\pi,B_v}} \big(X_w^\pi - X_w^{\pi,B_v}\big) =: \frac{1}{\sigma_{d,h}^2} f(\pi, B),$$

so that we can apply Lemma 4.4. For $1 \le i < j \le m$, let $\pi_{ij} = \pi\tau_{ij}$, where τ_{ij} is the transposition switching i and j. Now, it is clear that

$$\mathbb{E}\big(f(\pi, B) - f(\pi_{ij}, B)\big)^2 = \mathbb{E}\big(f(\pi, B) - f(\pi_{13}, B)\big)^2 \tag{4.19}$$

(unless $i = 2k+1$ and $j = 2k+2$ for some $0 \le k < m$, in which case the expectation on the left-hand side vanishes, since the transposition does not affect the underlying graph). Hence, it is enough to bound the right-hand side of (4.19). Adding and subtracting corresponding terms, we have

$$\begin{aligned}
&f(\pi, B) - f(\pi_{13}, B) \\
&\quad = \sum_{v=1}^n X_v^\pi \sum_{w \in Q_v^{\pi,B_v}} \big(X_w^\pi - X_w^{\pi,B_v}\big) - \sum_{v=1}^n X_v^{\pi_{13}} \sum_{w \in Q_v^{\pi_{13},B_v}} \big(X_w^{\pi_{13}} - X_w^{\pi_{13},B_v}\big) \\
&\quad = \sum_{v=1}^n \big(X_v^\pi - X_v^{\pi_{13}}\big) \sum_{w \in Q_v^{\pi,B_v}} \big(X_w^\pi - X_w^{\pi,B_v}\big) \\
&\qquad + \sum_{v=1}^n X_v^{\pi_{13}} \bigg(\sum_{w \in Q_v^{\pi,B_v}} \big(X_w^\pi - X_w^{\pi,B_v}\big) - \sum_{w \in Q_v^{\pi_{13},B_v}} \big(X_w^\pi - X_w^{\pi,B_v}\big) \bigg) \\
&\qquad + \sum_{v=1}^n X_v^{\pi_{13}} \sum_{w \in Q_v^{\pi_{13},B_v}} \big(X_w^\pi - X_w^{\pi_{13}}\big) \\
&\qquad - \sum_{v=1}^n X_v^{\pi_{13}} \sum_{w \in Q_v^{\pi_{13},B_v}} \big(X_w^{\pi,B_v} - X_w^{\pi_{13},B_v}\big).
\end{aligned}$$

Letting χ be the set of colours of the balls $\pi(1)$, $\pi(2)$, $\pi(3)$ and $\pi(4)$, we obtain

$$\mathbb{E}\big\{\big(f(\pi, B) - f(\pi_{13}, B)\big)^2 I_A\big\} \le 4\|h\|^4 \mathbb{E}\big\{\big(R_1^2 + R_2^2 + R_3^2 + R_4^2\big) I_A\big\}, \tag{4.20}$$

where the event A is as in (4.1) and where

$$R_1 = \sum_{v=1}^{n} \mathrm{I}[\chi \cap \xi_v^\pi \neq \varnothing] |Q_v^{\pi,B_v}|,$$

$$R_2 = \sum_{v=1}^{n} |Q_v^{\pi,B_v} \triangle Q_v^{\pi_{13},B_v}|,$$

$$R_3 = \sum_{v=1}^{n} \sum_{w \in Q_v^{\pi,B_v}} \mathrm{I}[\chi \cap \xi_w^\pi \neq \varnothing],$$

$$R_4 = \sum_{v=1}^{n} \sum_{w \in Q_v^{\pi,B_v}} \mathrm{I}[\mathcal{T}_\ell^{v,\pi,B_v}(w) \neq \mathcal{T}_\ell^{v,\pi_{13},B_v}(w)],$$

(note that for R_3 and R_4, we have replaced π_{13} by π and vice versa, since the two random permutations are exchangeable). We now proceed to bound the four error terms individually. In order to keep the formulae short, we abbreviate multiple sums such as $\sum_{v=1}^n \sum_{w=1}^n$ to $\sum_{v,w}$, where it is understood that summation always ranges from 1 to n.

We begin with some preliminary calculations involving χ. First, let $\mathcal{G}$ be the configuration generated by π as before, and let the set of colours α be a function of $\mathcal{G}$ (but not of π directly). Let $E_1(\alpha)$ denote the set of edges incident to α, and let $\mathcal{E} := \{\{\pi(1), \pi(2)\}, \{\pi(3), \pi(4)\}\}$. Then $\chi \cap \alpha \neq \varnothing$ if $\mathcal{E} \cap E_1(\alpha) \neq \varnothing$. Hence, since a given edge has probability $1/(m/2)$ of being represented by $\{\pi(2i-1), \pi(2i)\}$ for any $1 \le i \le m/2$, and since $|E_1(\alpha)| \le d_{\max}|\alpha|$, we have

$$\mathbb{P}[\chi \cap \alpha \neq \varnothing | \mathcal{G}] \le 2\frac{|\alpha| d_{\max}}{m/2} = \frac{4|\alpha| d_{\max}}{m}. \tag{4.21}$$

We also need to bound probabilities of the form

$$\mathbb{P}[\chi \cap \alpha_1 \neq \varnothing, \chi \cap \alpha_2 \neq \varnothing | \mathcal{G}],$$

for pairs of colour sets α_1, α_2 as above. To this end, let $E_2(\alpha_1, \alpha_2)$ denote the set of edges in $\mathcal{G}$ that join a vertex in α_1 to one in α_2, and define $E(\alpha_1, \alpha_2) := E_1(\alpha_1 \cap \alpha_2) \cup E_2(\alpha_1, \alpha_2)$. Then note that

$$\begin{aligned} &\mathrm{I}[\chi \cap \alpha_1 \neq \varnothing, \chi \cap \alpha_2 \neq \varnothing] \\ &\quad \le \mathrm{I}\bigl[\mathcal{E} \subset E_1(\alpha_1 \cup \alpha_2)\bigr] + \mathrm{I}\bigl[\mathcal{E} \cap E(\alpha_1, \alpha_2) \neq \varnothing\bigr]. \end{aligned} \tag{4.22}$$

Now it is easy to see that

$$\mathbb{P}\bigl[\mathcal{E} \subset E_1(\alpha_1 \cup \alpha_2) \big| \mathcal{G}\bigr] \le \frac{(d_{\max}|\alpha_1 \cup \alpha_2|)^2}{(m/2)(m/2-1)} \le \frac{5 d_{\max}^2 |\alpha_1 \cup \alpha_2|^2}{m^2}, \tag{4.23}$$

and that

$$\mathbb{P}[\mathcal{E} \cap E(\alpha_1, \alpha_2) \neq \varnothing | \mathcal{G}] \leq \mathbb{E}\{|\mathcal{E} \cap E(\alpha_1, \alpha_2)||\mathcal{G}\} \\ \leq \frac{4}{m}(|E_1(\alpha_1 \cap \alpha_2)| + |E_2(\alpha_1, \alpha_2)|). \tag{4.24}$$

Note, in particular, that (4.9) implies that, for any $1 \leq v \leq n$,

$$\sum_{w=1}^{n} \mathbb{E}|E_1(\xi_w \cap \xi_v)| \leq d_{\max} \sum_{w=1}^{n} \mathbb{E}|\xi_w \cap \xi_v| \leq 3d_{\max}^2 \ell^2; \tag{4.25}$$

then, taking α to be the set of all colours joined to ξ_v in $\mathcal{G}$, so that $|\alpha| \leq d_{\max}\ell$, (4.9) also implies that

$$\sum_{w=1}^{n} \mathbb{E}|E_2(\xi_w, \xi_v)| \leq 3d_{\max}^2 \ell^2. \tag{4.26}$$

In similar fashion, using (4.11), we also have

$$\sum_{w=1}^{n} \mathbb{E}|E_1(\eta_w \cap \eta_v)| \leq 40d_{\max}^2 \ell^2, \qquad \sum_{w=1}^{n} \mathbb{E}|E_2(\eta_w, \eta_v)| \leq 40d_{\max}^2 \ell^2. \tag{4.27}$$

Bound on $\mathbb{E}\{R_1^2 I_A\}$. First, write

$$R_1 = \sum_{v,w} \mathrm{I}[\chi \cap \xi_v^\pi \neq \varnothing]\mathrm{I}[\xi_w^\pi \cap \eta_v^{\pi,B_v} \neq \varnothing].$$

Then, using (4.22)–(4.24), we have

$$\begin{aligned} &\mathbb{E}\{R_1^2 I_A\} \\ &\quad = \mathbb{E} \sum_{v,v',w,w'} \mathrm{I}[\chi \cap \xi_{v'}^\pi \neq \varnothing]\mathrm{I}[\chi \cap \xi_v^\pi \neq \varnothing]\mathrm{I}[\xi_{w'}^\pi \cap \eta_{v'}^{\pi,B_{v'}} \neq \varnothing] \\ &\qquad\qquad \times \mathrm{I}[\xi_w^\pi \cap \eta_v^{\pi,B_v} \neq \varnothing] I_A \\ &\quad \leq \frac{20\ell^2 d_{\max}^2}{m^2} \mathbb{E} \sum_{v,v',w,w'} \mathrm{I}[\xi_{w'} \cap \eta_{v'} \neq \varnothing]\mathrm{I}[\xi_w \cap \eta_v \neq \varnothing] I_A \\ &\qquad + \frac{4}{m} \mathbb{E} \sum_{v,v',w,w'} (|E_1(\xi_{v'}^\pi \cap \xi_v^\pi)| + |E_2(\xi_{v'}^\pi, \xi_v^\pi)|)\mathrm{I}[\xi_{w'} \cap \eta_{v'} \neq \varnothing] \\ &\qquad\qquad \times \mathrm{I}[\xi_w \cap \eta_v \neq \varnothing] I_A. \end{aligned}$$

Much as for (4.17), we use inequality analogous to (4.16) to show that the first term yields at most

$$\frac{20d_{\max}^2 \ell^2}{m^2} \sum_{v,v'} 144 d_{\max}^2 \ell^4 \leq \frac{2880 n d_{\max}^4 \ell^6}{m},$$

where we also used that $n \le m$. For the second term, again using inequalities in the spirit of (4.16), we have

$$\begin{aligned}
&\frac{4}{m}\mathbb{E}\sum_{v,v',w,w'}\big(\big|E_1(\xi_{v'}^{\pi}\cap\xi_v^{\pi})\big|+\big|E_2(\xi_{v'}^{\pi},\xi_v^{\pi})\big|\big)\mathrm{I}[\xi_{w'}\cap\eta_{v'}\neq\varnothing]\mathrm{I}[\xi_w\cap\eta_v\neq\varnothing]I_A\\
&\qquad\le\frac{4}{m}\sum_{v,v'}144d_{\max}^2\ell^4\mathbb{E}\big\{\big|E_1(\xi_{v'}^{\pi}\cap\xi_v^{\pi})\big|+\big|E_2(\xi_{v'}^{\pi},\xi_v^{\pi})\big|\big\}\le\frac{3456nd_{\max}^4\ell^6}{m},
\end{aligned}$$

where the last line follows from (4.25) and (4.26), and the two bounds combine to give

$$\mathbb{E}\{R_1^2I_A\}\le\frac{6336nd_{\max}^4\ell^6}{m}. \tag{4.28}$$

Bound on $\mathbb{E}\{R_2^2I_A\}$. First, note that

$$\begin{aligned}
&\mathrm{I}\big[w\in Q_v^{\pi,B_v}\triangle Q_v^{\pi_{13},B_v}\big]\\
&\qquad\le\big(\mathrm{I}\big[\chi\cap\xi_w^{\pi}\neq\varnothing\big]+\mathrm{I}\big[\chi\cap\eta_v^{\pi,B_v}\neq\varnothing\big]\big)\mathrm{I}\big[\xi_w^{\pi}\cap\eta_v^{\pi,B_v}\neq\varnothing\big]\\
&\qquad\quad+\big(\mathrm{I}\big[\chi\cap\xi_w^{\pi_{13}}\neq\varnothing\big]+\mathrm{I}\big[\chi\cap\eta_v^{\pi_{13},B_v}\neq\varnothing\big]\big)\mathrm{I}\big[\xi_w^{\pi_{13}}\cap\eta_v^{\pi_{13},B_v}\neq\varnothing\big].
\end{aligned}$$

Hence, using the inequality $(a_1+\cdots+a_k)^2\le k(a_1^2+\cdots+a_k^2)$ and the exchangeability of π and π_{13},

$$\begin{aligned}
&\mathbb{E}\{R_2^2I_A\}\\
&\quad\le\mathbb{E}\bigg(\sum_{v,w}I_A\Big[\big(\mathrm{I}\big[\chi\cap\xi_w^{\pi}\neq\varnothing\big]+\mathrm{I}\big[\chi\cap\eta_v^{\pi,B_v}\neq\varnothing\big]\big)\\
&\qquad\qquad\qquad\qquad\times\mathrm{I}\big[\xi_w^{\pi}\cap\eta_v^{\pi,B_v}\neq\varnothing\big]\\
&\qquad\qquad+\big(\mathrm{I}\big[\chi\cap\xi_w^{\pi_{13}}\neq\varnothing\big]+\mathrm{I}\big[\chi\cap\eta_v^{\pi_{13},B_v}\neq\varnothing\big]\big)\\
&\qquad\qquad\qquad\qquad\times\mathrm{I}\big[\xi_w^{\pi_{13}}\cap\eta_v^{\pi_{13},B_v}\neq\varnothing\big]\Big]\bigg)^2\\
&\quad\le8\mathbb{E}\bigg(\sum_{v,w}\mathrm{I}\big[\chi\cap\xi_w^{\pi}\neq\varnothing\big]\mathrm{I}\big[\xi_w^{\pi}\cap\eta_v^{\pi,B_v}\neq\varnothing\big]I_A\bigg)^2\\
&\qquad+8\mathbb{E}\bigg(\sum_{v,w}\mathrm{I}\big[\chi\cap\eta_v^{\pi,B_v}\neq\varnothing\big]\mathrm{I}\big[\xi_w^{\pi}\cap\eta_v^{\pi,B_v}\neq\varnothing\big]I_A\bigg)^2\\
&\quad=8\mathbb{E}\sum_{v,v',w,w'}\mathrm{I}\big[\chi\cap\xi_{w'}^{\pi}\neq\varnothing\big]\mathrm{I}\big[\chi\cap\xi_w^{\pi}\neq\varnothing\big]\\
&\qquad\qquad\times\mathrm{I}\big[\xi_w^{\pi}\cap\eta_v^{\pi,B_v}\neq\varnothing\big]\mathrm{I}\big[\xi_{w'}^{\pi}\cap\eta_{v'}^{\pi,B_{v'}}\neq\varnothing\big]I_A
\end{aligned} \tag{4.29}$$

$$
\begin{aligned}
&+8\mathbb{E}\sum_{v,v',w,w'} \mathrm{I}[\chi \cap \eta_{v'}^{\pi,B_{v'}} \neq \varnothing]\mathrm{I}[\chi \cap \eta_v^{\pi,B_v} \neq \varnothing] \\
&\qquad \times \mathrm{I}[\xi_w^\pi \cap \eta_v^{\pi,B_v} \neq \varnothing]\mathrm{I}[\xi_{w'}^\pi \cap \eta_{v'}^{\pi,B_{v'}} \neq \varnothing] I_A.
\end{aligned}
\tag{4.30}
$$

We now use (4.22)–(4.24), together with inequalities such as in (4.16), to give

$$
\begin{aligned}
&\sum_{v,v'} \mathbb{E}\{\mathrm{I}[\chi \cap \xi_{w'}^\pi \neq \varnothing]\mathrm{I}[\chi \cap \xi_w^\pi \neq \varnothing]\mathrm{I}[\xi_w^\pi \cap \eta_v^{\pi,B_v} \neq \varnothing]\mathrm{I}[\xi_{w'}^\pi \cap \eta_{v'}^{\pi,B_{v'}} \neq \varnothing] I_A\} \\
&\quad \le \sum_{v,v'} \mathbb{E}\biggl\{\biggl(\frac{20d_{\max}^2\ell^2}{m^2} + \frac{4}{m}\bigl(|E_1(\xi_w \cap \xi_{w'})| + |E_2(\xi_w, \xi_{w'})|\bigr)\biggr) \\
&\qquad\qquad \times \mathrm{I}[\eta_v \cap (\xi_w \cup \xi_{w'} \cup \eta_{v'}) \neq \varnothing]\mathrm{I}[\eta_{v'} \cap (\xi_w \cup \xi_{w'}) \neq \varnothing] I_A\biggr\} \\
&\quad \le \mathbb{E}\biggl\{\frac{20d_{\max}^2\ell^2}{m^2} + \frac{4}{m}\bigl(|E_1(\xi_w \cap \xi_{w'})| + |E_2(\xi_w, \xi_{w'})|\bigr)\biggr\} 768 d_{\max}^2 \ell^4,
\end{aligned}
$$

this last using (4.10) twice. Now sum over w and w', using (4.25) and (4.26), to give a contribution to $\mathbb{E}\{R_2^2 I_A\}$ from (4.29) of at most

$$
\frac{8n}{m}(20d_{\max}^2\ell^2 + 12d_{\max}^2\ell^2 + 12d_{\max}^2\ell^2)768d_{\max}^2\ell^4 = \frac{270{,}336 n d_{\max}^4 \ell^6}{m}.
$$

The term in (4.30) is treated analogously, using (4.27) in place of (4.25) and (4.26), giving a further

$$
\frac{8n}{m}(320 + 160 + 160)144 d_{\max}^4 \ell^6 = \frac{737{,}280 n d_{\max}^4 \ell^6}{m},
$$

so that

$$
\mathbb{E}\{R_2^2 I_A\} \le \frac{1{,}007{,}616 n d_{\max}^4 \ell^6}{m}. \tag{4.31}
$$

Bound on $\mathbb{E}\{R_3^2 I_A\}$. For this term, we have

$$
\begin{aligned}
&\mathbb{E}\{R_3^2 I_A\} \\
&\quad = \mathbb{E}\biggl(\sum_{v=1}^n \sum_{w=1}^n \mathrm{I}[\chi \cap \xi_w^\pi \neq \varnothing]\mathrm{I}[\xi_w^\pi \cap \eta_v^{\pi,B_v} \neq \varnothing] I_A\biggr)^2 \\
&\quad \le \mathbb{E}\sum_{v,v',w,w'} \mathrm{I}[\chi \cap \xi_w \neq \varnothing]\mathrm{I}[\chi \cap \xi_{w'} \neq \varnothing]\mathrm{I}[\xi_w \cap \eta_v \neq \varnothing]\mathrm{I}[\xi_{w'} \cap \eta_{v'} \neq \varnothing] I_A.
\end{aligned}
$$

This is the term in (4.29), but without the factor of 8, giving

$$
\mathbb{E}\{R_3^2 I_A\} \le \frac{33{,}792 n d_{\max}^4 \ell^6}{m}. \tag{4.32}
$$

Bound on $\mathbb{E}\{R_4^2 I_A\}$. In order to bound $\mathbb{E}\{R_4^2 I_A\}$, note that

$$\begin{aligned}
&\mathrm{I}[w \in Q_v^{\pi,B_v}]\mathrm{I}[\xi'^{\pi,B_v}_{v,w} \neq \xi'^{\pi_{13},B_v}_{v,w}] \\
&\qquad \leq \mathrm{I}[\xi_w^\pi \cap \eta_v^{\pi,B_v} \neq \varnothing](\mathrm{I}[\chi \cap \xi_w^\pi \neq \varnothing] + \mathrm{I}[\chi \cap \eta_v^{\pi,B_v} \neq \varnothing]),
\end{aligned}$$

so that

$$\begin{aligned}
\mathbb{E}\{R_4^2 I_A\} &\leq \mathbb{E}\biggl(\sum_{v,w} \mathrm{I}[\xi_w^\pi \cap \eta_v^{\pi,B_v} \neq \varnothing] I_A(\mathrm{I}[\chi \cap \xi_w^\pi \neq \varnothing] + \mathrm{I}[\chi \cap \eta_v^{\pi,B_v} \neq \varnothing])\biggr)^2 \\
&\leq 4\mathbb{E}\biggl(\sum_{v,w} \mathrm{I}[\xi_w^\pi \cap \eta_v^{\pi,B_v} \neq \varnothing]\mathrm{I}[\chi \cap \xi_w^\pi \neq \varnothing] I_A\biggr)^2 \\
&\quad + 4\mathbb{E}\biggl(\sum_{v,w} \mathrm{I}[\xi_w^\pi \cap \eta_v^{\pi,B_v} \neq \varnothing]\mathrm{I}[\chi \cap \eta_v^{\pi,B_v} \neq \varnothing] I_A\biggr)^2.
\end{aligned}$$

This is half the sum of the quantities given in (4.29) and (4.30), and hence yields

$$\mathbb{E}\{R_4^2 I_A\} \leq \frac{503{,}808 n d_{\max}^4 \ell^6}{m}. \tag{4.33}$$

Substituting (4.28), (4.31), (4.32) and (4.33) into (4.20) gives

$$\mathbb{E}\{(f(\pi, B) - f(\pi_{13}, B))^2 I_A\} \leq \frac{6{,}206{,}208 \|h\|^4 n d_{\max}^4 \ell^6}{m} \leq \frac{\|h\|^4 n d_{\max}^4 \ell^{16}}{9976m}.$$

The calculation on A^c is based on the crude bound

$$|f(\pi, B)| \leq 2n^2 \|h\|^2,$$

together with (4.2), giving

$$\begin{aligned}
\mathbb{E}\{(f(\pi, B) - f(\pi_{13}, B))^2 \mathrm{I}_{A^c}\} &\leq 16n^4 \|h\|^4 \gamma \\
&\leq \frac{16n^4 \|h\|^4 d_{\max}^{16} \ell^{16}}{8! m^7} \\
&\leq \frac{\|h\|^4 n d_{\max}^4 \ell^{16}}{2520m} \times \frac{d_{\max}^{12}}{m^2 n}.
\end{aligned}$$

Thus, for $d_{\max} \leq n^{1/4}$ and $m \geq n$, we have

$$\mathbb{E}(f(\pi, B) - f(\pi_{13}, B))^2 \leq \frac{\|h\|^4 n d_{\max}^4 \ell^{16}}{2011m}. \tag{4.34}$$

For the second sum in Lemma 4.4, we have

$$f(\pi, B) - f(\pi, B^v) = X_v^\pi \biggl(\sum_{w \in Q_v^{\pi,B_v}} (X_w^\pi - X_w^{\pi,B_v}) - \sum_{w \in Q_v^{\pi,B'_v}} (X_w^\pi - X_w^{\pi,B'_v}) \biggr),$$

so that

$$\big|f(\pi, B) - f\big(\pi, B^v\big)\big| \le 2\|h\|^2\big(\big|Q_v^{\pi,B_v}\big| + \big|Q_v^{\pi,B'_v}\big|\big).$$

Hence, using exchangeability to replace B'_v by B_v, we deduce that

$$\mathbb{E}\big\{\big(f(\pi, B) - f\big(\pi, B^v\big)\big)^2 I_A\big\} \le 16\|h\|^4\mathbb{E}\big\{|Q_v|^2 I_A\big\},$$

and, using (4.18), this gives

$$\mathbb{E}\big\{\big(f(\pi, B) - f\big(\pi, B^v\big)\big)^2 I_A\big\} \le 16\|h\|^4 80 d_{\max}^2 \ell^4 \le \frac{320}{12^{12}}\|h\|^4 d_{\max}^4 \ell^{16},$$

where we used that $d_{\max} \ge 2$ and $\ell \ge 12$. Then, since $|f(\pi, B) - f(\pi, B^v)| \le 4n\|h\|^2$, it is immediate that

$$\begin{aligned}\mathbb{E}\big\{\big(f(\pi, B) - f\big(\pi, B^v\big)\big)^2 I_{A^c}\big\} &\le 16n^2\|h\|^4\gamma \le \frac{16n^2\|h\|^4 d_{\max}^{16}\ell^{16}}{8!m^7} \\ &\le \frac{16}{8!}\|h\|^4 d_{\max}^4 \ell^8 \times \frac{d_{\max}^{12}\ell^8}{m^5} \le \frac{16}{8!12^8}\|h\|^4 d_{\max}^4 \ell^{16},\end{aligned}$$

where we used that $\max\{d_{\max}, \ell\} \le n^{1/4}$. Hence,

$$\mathbb{E}\big(f(\pi, B) - f\big(\pi, B^v\big)\big)^2 \le \frac{1}{10^{10}}\|h\|^4 d_{\max}^4 \ell^{16}. \tag{4.35}$$

Substituting these bounds into Lemma 4.4, we obtain that

$$\operatorname{Var} f(\pi, B) \le \frac{1}{4021} n\|h\|^4 d_{\max}^4 \ell^{16},$$

and hence that

$$\sqrt{\operatorname{Var}\mathbb{E}(G\Delta|W)} \le \frac{\sqrt{n}\|h\|^2 d_{\max}^2 \ell^8}{63\sigma_{d,h}^2},$$

completing the proof of Theorem 2.1.

4.3. *The variance* $\sigma_{d,h}^2$. By substituting $f(w) = 1$ and then $f(w) = w$ into the Stein coupling (3.1), it follows that $\operatorname{Var} W = \mathbb{E}\{G\Delta\}$. Recalling the definitions (4.12) and (4.13) of G_v and Δ_v, it then follows that

$$\sigma_{d,h}^2 = -n\mathbb{E}\bigg\{h\big(\mathcal{T}_\ell(I)\big)\sum_{w\in Q_I}\big(h\big(\mathcal{T}_\ell^I(w)\big) - h\big(\mathcal{T}_\ell(w)\big)\big)\bigg\}, \tag{4.36}$$

where I denotes a randomly chosen vertex in $[n]$. Under asymptotic circumstances in which the expectation in (4.36) remains of order $\mathrm{O}(1)$ as $n \to \infty$, this yields a variance $\sigma_{d,h}^2$ of order $\mathrm{O}(n)$. Broadly speaking, such circumstances are those in which the value of $h(\mathcal{T}_\ell(v))$ is not much influenced by vertices far from v. As far as the accuracy in Theorem 2.1 is concerned, it is advantageous to have $n^{-1}\sigma_{d,h}^2$ bounded below as $n \to \infty$. This is equivalent to requiring that the expectation in (4.36) does not tend to zero as $n \to \infty$, which might usually be supposed to be the case. If, however, $h(\mathcal{T}_\ell(v)) := \mathrm{I}[d_v = k]$ for some k, then $h(\mathcal{T}_\ell^v(w)) = h(\mathcal{T}_\ell(w))$ for all v, w, and the expectation would be exactly zero—as it has to be, since the number of vertices of any given degree k is fixed in the model. So, in practice, this condition has to be checked.

Acknowledgements. We thank the referees for their careful reading, helpful comments and additional references. We also thank Siva Athreya and D. Yogeshwaran for sharing an early draft of their manuscript with us. This work was started while A. D. Barbour was Saw Swee Hock Professor of Statistics at the National University of Singapore. A. D. Barbour thanks the Department of Statistics and Applied Probability at the National University of Singapore and the mathematics departments of the University of Melbourne and Monash University for their kind hospitality.

REFERENCES

ANGEL, O., VAN DER HOFSTAD, R. and HOLMGREN, C. (2016). Limit laws for self-loops and multiple edges in the configuration model. Available at arXiv:1603.07172.

ATHREYA, S. and YOGESHWARAN, D. (2018). Central limit theorem for statistics of subcritical configuration models. Available at arXiv:1808.06778.

BALL, F. and NEAL, P. (2017). The asymptotic variance of the giant component of configuration model random graphs. *Ann. Appl. Probab.* **27** 1057–1092. MR3655861

BARBOUR, A. D., KAROŃSKI, M. and RUCIŃSKI, A. (1989). A central limit theorem for decomposable random variables with applications to random graphs. *J. Combin. Theory Ser. B* **47** 125–145. MR1047781

BENDER, E. A. and CANFIELD, E. R. (1978). The asymptotic number of labeled graphs with given degree sequences. *J. Combin. Theory Ser. A* **24** 296–307. MR0505796

BOLLOBÁS, B. (1980). A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European J. Combin.* **1** 311–316. MR0595929

BOLLOBÁS, B. and RIORDAN, O. (2012). Asymptotic normality of the size of the giant component via a random walk. *J. Combin. Theory Ser. B* **102** 53–61. MR2871766

CHEN, L. H. Y., GOLDSTEIN, L. and RÖLLIN, A. (2017). Stein couplings and the inductive method. In preparation.

CHEN, L. H. Y., GOLDSTEIN, L. and SHAO, Q.-M. (2011). *Normal Approximation by Stein's Method*. Springer, Heidelberg. MR2732624

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

JANSON, S. (2010). Asymptotic equivalence and contiguity of some random graphs. *Random Structures Algorithms* **36** 26–45. MR2591045

JANSON, S. and LUCZAK, M. J. (2008a). Susceptibility in subcritical random graphs. *J. Math. Phys.* **49** 125207. MR2484338

JANSON, S. and LUCZAK, M. J. (2008b). Asymptotic normality of the k-core in random graphs. *Ann. Appl. Probab.* **18** 1085–1137. MR2418239

KHUDABUKHSH, W. R., WOROSZYLO, C., REMPAŁA, G. A. and KOEPPL, H. (2017). Functional central limit theorem for susceptible–infected process on configuration model graphs. Available at arXiv:1703.06328.

KNUTH, D. E. (1969). *The Art of Computer Programming*. *Vol*. 2: *Seminumerical Algorithms*. Addison-Wesley Co., Reading, MA. MR0286318

MOLLOY, M. and REED, B. (1995). A critical point for random graphs with a given degree sequence. In *Random Structures Algorithms* **6** 161–179.

PITTEL, B. and WORMALD, N. C. (2005). Counting connected graphs inside-out. *J. Combin. Theory Ser. B* **93** 127–172. MR2117934

RIORDAN, O. (2012). The phase transition in the configuration model. *Combin. Probab. Comput.* **21** 265–299. MR2900063

ROZANOV, Y. A. (1982). *Markov Random Fields*. Springer, New York. MR0676644

RUCIŃSKI, A. (1988). When are small subgraphs of a random graph normally distributed? *Probab. Theory Related Fields* **78** 1–10. MR0940863

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

STEIN, C. (1986). *Approximate Computation of Expectations*. *Institute of Mathematical Statistics Lecture Notes—Monograph Series* **7**. IMS, Hayward, CA. MR0882007

INSTITUT FÜR MATHEMATIK
UNIVERSITÄT ZÜRICH
WINTERTHURERSTRASSE 190
8057 ZÜRICH
SWITZERLAND
E-MAIL: a.d.barbour@math.uzh.ch

DEPARTMENT OF STATISTICS
AND APPLIED PROBABILITY
NATIONAL UNIVERSITY OF SINGAPORE
6 SCIENCE DRIVE 2
117546 SINGAPORE
E-MAIL: adrian.roellin@nus.edu.sg