Stein's method via induction

Louis H. Y. Chen[*]   Larry Goldstein[†]   Adrian Röllin[‡]

Abstract

Applying an inductive technique for Stein and zero bias couplings yields Berry-Esseen theorems for normal approximation for two new examples. The conditions of the main results do not require that the couplings be bounded. Our two applications, one to the Erdős-Rényi random graph with a fixed number of edges, and one to Jack measure on tableaux, demonstrate that the method can handle non-bounded variables with non-trivial global dependence, and can produce bounds in the Kolmogorov metric with the optimal rate.

1 Introduction

We present new Berry-Esseen theorems for sums Y of possibly dependent variables by combining both the Stein and zero bias couplings of Stein's method with the inductive technique of Bolthausen (1984) originally developed for the combinatorial central limit theorem. We apply these results to obtain normal approximations in the Kolmogorov metric for two new examples.

Stein's method (Stein, 1972, 1986) typically proceeds by coupling a random variable Y of interest to a related variable Y'; for an overview see Chen, Goldstein and Shao (2011) and Ross (2011). Here we develop results that can be applied to the Stein couplings of Chen and Röllin (2010) and to the zero bias couplings of Goldstein and Reinert (1997), thus encompassing most of the known couplings that have appeared in the literature, including settings not typically framed in terms of couplings, such as local dependence. The innovation here is the widened scope of the couplings that can be handled that permit applications when the difference $|Y - Y'|$ between Y and the coupled Y' is not almost surely bounded by a constant, or where the bound on this difference increases

---

[*]National University of Singapore. E-mail: matchyl@nus.edu.sg
[†]University of Southern California. E-mail: larry@usc.edu
[‡]National University of Singapore. E-mail: adrian.roellin@nus.edu.sg

in the problem size. This work is a broad extension and continuation of Ghosh (2009), applying induction and the zero bias coupling for the combinatorial central limit theorem where the random permutations are involutions, and of Goldstein (2013) using the size bias coupling to study degree counts in the Erdős-Rényi random graph; the inductive method considered here is inspired by Bolthausen (1984), but goes ultimately back to Bergström (1944).

At the center of Stein's method is the characterization that Z is a standard normal random variable if and only if

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

for all locally absolutely continuous functions f for which the above expectations exist. Given a standardized variable W whose distribution is to be compared to Z, and a test function h on which to evaluate the difference $\mathbb{E}h(W) - \mathbb{E}h(Z)$, one solves the Stein equation

$$f'(w) - wf(w) = h(w) - \mathbb{E}h(Z)$$

for f. The difference $\mathbb{E}h(W) - \mathbb{E}h(Z)$ may then be evaluated by substituting W for w and taking expectation on the left hand side of (1.1), rather than the right. One explanation of why the expectation of the left hand side may simpler to compute, or bound, than that of the right is that it depends only on the distribution of W, whereas the right also depends on that of Z. In particular, on the left hand side one may apply couplings of W to auxiliary random variables having properties that allow for convenient manipulations.

In Theorem 1.1 we present results for situations in which one can form a Stein coupling as defined by Chen and Röllin (2010). Following the treatment there, we say that the triple (W, W', G) of random variables is a Stein coupling when

$$\mathbb{E}\{Gf(W') - Gf(W)\} = \mathbb{E}\{Wf(W)\} \tag{1.2}$$

for all functions f for which the expectations above exist. It is not difficult to see that the canonical exchangeable pair coupling of Stein (1986), and the size bias coupling of Goldstein and Rinott (1996) are both special cases of Stein couplings. Indeed, recall that for $\lambda \in (0, 1]$ we say (W, W') is a λ-*Stein pair* if (W, W') is exchangeable and

$$\mathbb{E}\{W'|W\} = (1 - \lambda)W. \tag{1.3}$$

In this case, it is easily verified that (1.2) is satisfied with

$$G = \frac{1}{2\lambda}(W' - W).$$

Likewise, for a non-negative random variable Y with finite mean μ, we say that (Y, Y') is a size bias coupling of Y when Y' has the Y-size bias distribution, that is, when

$$\mathbb{E}\{Yf(Y)\} = \mu\,\mathbb{E}\{f(Y')\}$$

for all functions f for which these expectations exist. Again, it is easy to verify that for such couplings (1.2) is satisfied with

$$W = Y - \mu, \quad W' = Y' - \mu \quad \text{and} \quad G = \mu.$$

In particular, Theorem 1.1 extend results in Goldstein (2013) for the size bias coupling.

Theorem 1.2 provides a parallel result for the zero bias coupling (W, W^*) of Goldstein and Reinert (1997). Recall that for a non-trivial mean zero, variance σ^2 random variable W, we say that W^* has the W-zero biased distribution if

$$\mathbb{E}\{Wf(W)\} = \sigma^2\,\mathbb{E}\{f'(W^*)\} \tag{1.4}$$

for all functions f for which the quantities above exist.

In Stein's method in general, simplification occurs when a coupling of W to an appropriate W' can be achieved in such a way that the difference is almost surely bounded, or bounded uniformly in the size of the problem. However, in many situations appropriately bounded couplings may be difficult to construct, whereas unbounded couplings seem to appear naturally. Hence Theorems 1.1 and 1.2, which do not impose restrictive boundedness conditions, may be applied to produce new results in a variety of examples.

General Framework. Let $(\Theta, \mathcal{T})$ and $(\Omega, \mathcal{F})$ be two measurable spaces, the *parameter space* and the *sample space*, respectively. All random variables are understood to be real valued measurable functions from the product space $(\Theta \times \Omega, \mathcal{T} \otimes \mathcal{F})$. The distribution of a random variable X is determined by a parameter $\theta \in \Theta$ through a given transition kernel $\mathbb{P}_\theta$ from Θ to Ω. That is, for each $\theta \in \Theta$, $\mathbb{P}_\theta[\cdot]$ is a probability measure on $(\Omega, \mathcal{F})$, and for each $A \in \mathcal{F}$, the map $\mathbb{P}_\cdot[A]$ is $\mathcal{T}$-measurable. Depending on context and emphasis, we may also write X as $X(\theta, \omega)$ or $X_\theta(\omega)$, so that, for instance, $\mathbb{E}_\theta X = \int_\Omega X(\theta, \omega)\mathbb{P}_\theta[d\omega]$.

These measurability conditions are needed to assure the measurability of mappings that appear later, such as of the mean μ_θ, the variance σ_θ^2 of Y, and of $Y_{\Psi(\theta, \omega)}(\omega)$, which represents the value of Y at the parameter used in the inductive step. These conditions will not always be invoked explicitly below; we illustrate their use by showing in the Appendix, Section 4, that this latter variable in particular is measurable.

Our goal is to obtain bounds on the Kolmogorov distance between the standardized version W of a random variable Y and the normal distribution in terms of the parameter θ. Theorems 1.1 and 1.2 below yield a bound of the form C/r_θ for r_θ a positive 'rate' function of θ and C a constant not depending on θ.

As noted, one main step our method requires is to couple W to a random variable W', which satisfies either the Stein coupling relation (1.2) or the zero bias coupling relation (1.4). In order to apply induction, we identify a subset $\odot \subset \Theta$ in Condition (G1), consisting of the 'nicely behaved' parameters; its complement plays the role of the base case, on which the bound C/r_θ may be trivial. For our bound to be informative, it is necessary that the rate function r_θ be unbounded on $\odot$.

For the induction step, we also introduce a sub σ-algebra $\mathcal{F}_\theta$ that, roughly speaking, captures the information about the changes that were necessary to construct W' from W (or equivalently, Y' from Y); the coarser $\mathcal{F}_\theta$ is, the better the normal approximation will be. A certain tension is created here, as $\mathcal{F}_\theta$ must be large enough to contain the variables describing the changes from Y to Y', but small enough so that the conditional distribution of Y on $\mathcal{F}_\theta$, is sufficiently close to its original one.

Conditional on $\mathcal{F}_\theta$, the variable Y may no longer have its original distribution, but induction is viable when one can identify within Y another variable V that has a distribution similar to the original Y; when the parameter space Θ is ordered, V typically has a smaller parameter. For a successful induction, the parameter of the smaller problem should not stray too far from that of Y. There is some leeway here, as it suffices to have control over an event $F_{\theta,1}$, as specified in Condition (G4). Intuitively, the event $F_{\theta,1}$ should contain the bulk of the support of the variables that generate $\mathcal{F}_\theta$, and not their extremes. For instance, for the Erdős-Rényi graph problem considered, $\mathcal{F}_\theta$ contains the label and degree of a chosen vertex on which the coupling is based, and $F_{\theta,1}$ is an even on which its degree is 'not too large'.

Relaxing the condition that the difference $D = W' - W$ be bounded, we control the magnitude of this difference by its moments. Moreover, we upper bound D by $\overline{D}$, and in the case of a Stein coupling, also G by $\overline{G}$, where these majorizing variables are required to be $\mathcal{F}_\theta$ measurable; we are able to handle exceptional or boundary cases as these

upper bounds are only required to hold on $F_{\theta,1}$. We will also require the existence of a random variable B that bounds the absolute difference $|Y - V|$, and which is not 'too large.' See Conditions (G3), (G4) and (G6) for the case of Stein couplings.

There is also some leeway in that the distribution of V, conditionally on $\mathcal{F}_\theta$, only needs to be close to that of Y on an event $F_{\theta,2} \in \mathcal{F}_\theta$. Precisely, for the Stein coupling case, with similar remarks also applying to zero bias couplings, we impose in Condition (G5) that

$$\mathscr{L}_\theta(V|\mathcal{F}_\theta) = \mathscr{L}_{\Psi_\theta}(Y) \quad \text{on } F_{\theta,2}, \tag{1.5}$$

where Ψ_θ is the (typically random) parameter capturing the conditional distribution of the embedded variable V. For clarification, by (1.5) we mean

$$\mathbb{P}_\theta[V \in \cdot \,|\mathcal{F}_\theta](\omega) = \mathbb{P}_{\Psi_\theta(\omega)}[Y \in \cdot] \quad \text{for all } \omega \in F_{\theta,2}.$$

With the help of V, a recursive inequality for a bound on the distance between W and the normal can be produced.

Before attempting to apply the methods presented in this article, it is advisable that a user first 'test the waters' by constructing a Stein or zero-bias coupling and proving a normal approximation for a smooth metric such as the Wasserstein distance; see Chen and Röllin (2010), or Goldstein (2007), respectively. Once this goal has been achieved, the sigma-algebra $\mathcal{F}_\theta$ will typically arise naturally from the coupling construction, and one may then proceed to identify a suitable variable V whose conditional distribution given $\mathcal{F}_\theta$ is within the same class of distributions determined by Θ and close to that of Y. For instance, in occupancy problems, a Stein coupling or zero-bias coupling typically involves moving around a small number of balls among a small number of urns, and V will typically again represent an occupancy problem, but on fewer balls and fewer urns.

1.1 Abstract approximation theorems

We now state the conditions required for our main results. The inverse rate function r_θ is assumed to be a positive function, measurable in θ, a condition satisfied for all natural examples, including the ones considered here. The mean $\mu_\theta = \mathbb{E}_\theta Y$ and variance $\sigma_\theta^2 = \text{Var}_\theta(Y)$ are measurable by the conditions in our General Framework. To avoid repetition, the distribution of random variables indicated after $\theta \in \Theta$ has been fixed is with respect to $\mathscr{L}_\theta(\cdot)$. The random variable Z will always denote the standard normal.

The variable Y denotes the unstandardized random variable of interest. Theorem 1.1 shows that the following set of conditions are sufficient for the Kolmogorov distance between the standardized version W of Y and the normal to be bounded by C/r_θ for some universal constant C.

(G1) Let r_θ be a positive measurable function, let $\bar{r}$ be a positive number, and let

$$\text{☺} = \{\theta \in \Theta : r_\theta > \bar{r}\}. \tag{1.6}$$

Assume that $\bar{r}$ is chosen such that $\text{Var}_\theta Y > 0$ for all $\theta \in \text{☺}$.

(G2) For all $\theta \in \Theta$, let $\mu_\theta = \mathbb{E}_\theta Y$ and $\sigma_\theta^2 = \text{Var}_\theta Y$, and define

$$W = \frac{Y - \mu_\theta}{\sigma_\theta}$$

whenever $\sigma_\theta > 0$, and set $W = 0$ otherwise. Let W' and G be two random variables such that, for each $\theta \in \text{☺}$, (W, W', G) is a Stein coupling, in the sense of (1.2), with respect to $\mathbb{P}_\theta$.

(G3) With $D = W' - W$ assume that

$$\sup_{\theta \in \smiley} r_\theta \, \mathbb{E}_\theta \big| \mathbb{E}_\theta(1 - GD|W) \big| < \infty \quad \text{and} \quad \sup_{\theta \in \smiley} r_\theta \, \mathbb{E}_\theta \big\{ (1 + |W|)|G|D^2 \big\} < \infty. \quad (1.7)$$

(G4) For each $\theta \in \smiley$, let $\mathcal{F}_\theta \subset \mathcal{F}$ be a sub-σ-algebra. Let $\overline{G}$ and $\overline{D}$ be random variables such that, for each $\theta \in \smiley$, the mappings $\overline{G}(\theta, \cdot)$ and $\overline{D}(\theta, \cdot)$ are $\mathcal{F}_\theta$-measurable and such that, on some event $F_{\theta,1}$ which need not be in $\mathcal{F}_\theta$, we have $|G| \leq \overline{G}$, $|D| \leq \overline{D}$, and

$$\sup_{\theta \in \smiley} r_\theta^2 \, \mathbb{E}_\theta \big\{ |G|D^2 (1 - I_{F_{\theta,1}}) \big\} < \infty \quad \text{and} \quad \sup_{\theta \in \smiley} r_\theta \, \mathbb{E}_\theta \big\{ \overline{G}\,\overline{D}^2 \big\} < \infty. \quad (1.8)$$

(G5) Let Ψ be a Θ-valued random element such that, for each $\theta \in \smiley$, $\Psi(\theta, \cdot)$ is $\mathcal{F}_\theta$-measurable. Let V be a random variable, and for each $\theta \in \smiley$, let $F_{\theta,2} \in \mathcal{F}_\theta$ be such that

$$\mathscr{L}_\theta(V|\mathcal{F}_\theta) = \mathscr{L}_\Psi(Y) \qquad \text{on } F_{\theta,2}, \quad (1.9)$$

and

$$\sup_{\theta \in \smiley} r_\theta^2 \, \mathbb{E}_\theta \big\{ |G|D^2(1 - I_{F_{\theta,2}}) \big\} < \infty. \quad (1.10)$$

(G6) Let $\overline{B}$ be a random variable such that, for each $\theta \in \smiley$, $\overline{B}(\theta, \cdot)$ is $\mathcal{F}_\theta$-measurable,

$$\sigma_\theta^{-1}|Y - V| \leq \overline{B} \quad \text{on } F_{\theta,1}, \quad \text{and} \quad \sup_{\theta \in \smiley} r_\theta^2 \, \mathbb{E}_\theta \big\{ \overline{G}\,\overline{D}^2 \overline{B} I_{F_{\theta,2}} \big\} < \infty. \quad (1.11)$$

(G7) Assume

$$\sup_{\theta \in \smiley} \, \operatorname*{ess\,sup}_{\omega \in F_{\theta,2} \cap \{\Psi \in \smiley\}} \frac{\sigma_\theta^2}{\sigma_{\Psi(\theta,\omega)}^2} < \infty, \quad (1.12)$$

$$\sup_{\theta \in \smiley} \, \operatorname*{ess\,sup}_{\omega \in F_{\theta,2}} \frac{r_\theta}{r_{\Psi(\theta,\omega)}} < \infty, \qquad \sup_{\theta \in \smiley} \, \operatorname*{ess\,sup}_{\omega \in F_{\theta,2} \cap \{\Psi \in \smiley\}} \frac{r_{\Psi(\theta,\omega)}}{r_\theta} < \infty, \quad (1.13)$$

where the essential suprema are taken with respect to $\mathbb{P}_\theta$.

Theorem 1.1. *If Conditions* (G1)–(G7) *are satisfied, then there exists a constant C, independent of θ, such that*

$$\sup_{z \in \mathbb{R}} \big| \mathbb{P}_\theta[W \leq z] - \mathbb{P}[Z \leq z] \big| \leq \frac{C}{r_\theta} \qquad \text{for all } \theta \in \Theta. \quad (1.14)$$

Theorem 1.1 extends Theorem 1.1 in Goldstein (2013), which produces a Kolmogorov bound equivalent up to constants to the bound in Chen and Röllin (2010) for the Wasserstein distance to the normal for bounded size bias couplings. In addition, the bound produced by Bartroff and Goldstein (2013) by an application of Theorem 1.1 of Goldstein (2013) to counts in a multinomial occupancy model was shown there to be of optimal order by the lower bound (1.6) of Englund (1981), see also (1.7) of Bartroff and Goldstein (2013); the bound of Theorem 1.2 of Goldstein (2013), using also Theorem 1.1 of that same work, for degree counts in the Erdős-Rényi random graph can also be shown to be optimal up to constant factors in the same manner.

When higher moments exist a number of the conditions of the theorem may be verified using simpler expressions, obtained via standard inequalities. For instance, using $f(w) = w$ and that $\mathrm{Var}_\theta(W) = 1$ in (1.2) shows that $\mathbb{E}_\theta(GD) = 1$, hence applying

the Cauchy-Schwarz inequality to the first expression in (1.7) in Condition (G3) above, followed by a consequence of the conditional variance formula, we obtain

$$\mathbb{E}_\theta |\mathbb{E}_\theta(1 - GD|W)| \le \sqrt{\operatorname{Var}_\theta(\mathbb{E}_\theta(GD|W))} \le \sqrt{\operatorname{Var}_\theta(\mathbb{E}_\theta(GD|\mathcal{H}))}, \qquad (1.15)$$

where $\mathcal{H}$ is any σ-algebra with respect to which W is measurable.

We now state a parallel result for zero bias couplings.

(Z1) Let r_θ be a positive measurable function, let $\overline{r}$ a positive number, and let

$$\smiley = \{\theta \in \Theta : r_\theta > \overline{r}\}.$$

Assume that $\overline{r}$ is chosen such that $\operatorname{Var}_\theta Y > 0$ for all $\theta \in \smiley$.

(Z2) Let $\mu_\theta = \mathbb{E}_\theta Y$ and $\sigma_\theta^2 = \operatorname{Var}_\theta Y$, and define

$$W = \frac{Y - \mu_\theta}{\sigma_\theta}$$

whenever $\sigma_\theta > 0$ and $W = 0$ otherwise. Let W^* be defined on Ω, such that for each $\theta \in \smiley$ the variable W^* has the W-zero bias distribution as in (1.4) with respect to $\mathbb{P}_\theta$.

(Z3) For each $\theta \in \smiley$ let $\mathcal{F}_\theta$ be a sub-sigma algebra of $\mathcal{F}$, let $D = W^* - W$, and let $\overline{D}$ be a random variable such that $\overline{D}(\theta, \cdot)$ is $\mathcal{F}_\theta$-measurable, and let $F_{\theta,1}$ be an event, which need not be in $\mathcal{F}_\theta$, on which $|D| \le \overline{D}$ and such that

$$\sup_{\theta \in \smiley} r_\theta^2 \, \mathbb{E}_\theta\big\{|D|(1 - I_{F_{\theta,1}})\big\} < \infty \quad \text{and} \quad \sup_{\theta \in \smiley} r_\theta E_\theta\big\{|DW| + \overline{D}\big\} < \infty. \qquad (1.16)$$

(Z4) Let V be a random variable, and let Ψ be a Θ-valued random element such that, for each $\theta \in \smiley$, $\Psi(\theta, \cdot)$ is $\mathcal{F}_\theta$-measurable. For each $\theta \in \smiley$, let $F_{\theta,2}$ be an event in $\mathcal{F}_\theta$ such that

$$\mathscr{L}_\theta(V|\mathcal{F}_\theta) = \mathscr{L}_\Psi(Y) \qquad \text{on } F_{\theta,2}, \qquad (1.17)$$

and

$$\sup_{\theta \in \smiley} r_\theta^2 \, \mathbb{E}_\theta\big\{|D|(1 - I_{F_{\theta,2}})\big\} < \infty. \qquad (1.18)$$

(Z5) Let $\overline{B}$ be a random variable such that, for each $\theta \in \smiley$, $\overline{B}(\theta, \cdot)$ is $\mathcal{F}_\theta$-measurable, and

$$\sigma_\theta^{-1}|Y - V| \le \overline{B} \quad \text{on } F_{\theta,1}, \quad \text{and} \quad \sup_{\theta \in \smiley} r_\theta^2 \, \mathbb{E}_\theta\big\{\overline{D}\big(\overline{B} + \overline{D}\big) I_{F_{\theta,2}}\big\} < \infty. \qquad (1.19)$$

Theorem 1.2. *If Conditions* (Z1)–(Z5) *and* (G7) *are satisfied, then there exists a constant C, independent of θ, such that*

$$\sup_{z \in \mathbb{R}} \big|\mathbb{P}_\theta[W \le z] - \mathbb{P}[Z \le z]\big| \le \frac{C}{r_\theta}, \qquad \text{for all } \theta \in \Theta.$$

Many of the conditions of Theorem 1.2, as for Theorem 1.1, can be shown to be satisfied using inequalities on moments. The proofs of Theorems 1.1 and 1.2 appear in Section 4.

1.2 Applications

We apply Theorems 1.1 and 1.2 to obtain new results in two examples; the proofs are deferred to Sections 2 and 3.

The first examples invokes Theorem 1.1 for Stein couplings for the normal approximation of the number Y of isolated vertices in the Erdős-Rényi graph $\mathcal{G} \sim \mathrm{ER}(n, m)$ on n vertices, having exactly m edges, distributed uniformly at random. This model is related to the one where edges between each pair of vertices are chosen independently with some fixed probability, but in the model we consider the indicators that vertices are isolated exhibit a non-trivial global dependence since the total number of edges is fixed. In fact, while in the model with independent edges these indicators are positively correlated, the effect of the global dependence in $\mathrm{ER}(n, m)$ is stronger, resulting in a negative correlation; see proof of Lemma 2.5.

Related work was done by Kordecki (1987) on the number of isolated vertices in the Erdős-Rényi graph model, although his general framework is not applicable here.The boundedness of the second derivative of the solution to the Stein equation on page 132 is shown only for the points where the second derivative exists, whereas, in order to perform the Taylor expansion on page 135, it is needed to hold everywhere; we were thus not able to reproduce his final results. In addition, the fixed number of edges model does not appear to satisfy the condition on page 134 of his work. We also mention the work by Goldstein (2013), who considered vertex degrees in general, though it only addressed the independent edge model.

Theorem 1.3 provides the following bound on the Kolmogorov distance between the standardized variable Y and the normal.

Theorem 1.3. *Let Y count the number of isolated vertices in the Erdős-Rényi graph $\mathcal{G} \sim \mathrm{ER}(n, m)$ on n vertices, having exactly m edges, distributed uniformly at random. Then, with $\mu_{n,m}$ and $\sigma_{n,m}^2$ the mean and variance of Y, letting $W = (Y - \mu_{n,m})/\sigma_{n,m}$ when $\sigma_{n,m} > 0$ and zero otherwise, with*

$$\Theta = \left\{ (n, m) \, : \, n \geq 3, \, 0 < m < \binom{n}{2} \right\}, \tag{1.20}$$

there exists a universal constant $C > 0$ such that, for all $(n, m) \in \Theta$,

$$\sup_{z \in \mathbb{R}} \left| \mathbb{P}_{n,m}[W \leq z] - \Phi(z) \right| \leq \frac{C}{r_{n,m}}$$

where

$$r_{n,m} = \frac{\sigma_{n,m}^3}{\mu_{n,m}\left(1 + \frac{m^2}{n^2}\right)}. \tag{1.21}$$

Remark 1.4. In order to better understand the bounds obtained in Theorem 1.3, we now discuss in more detail the different regimes at which m and n can tend to infinity. To this end, denote by $a(n) \sim b(n)$ that $\lim a(n)/b(n) = 1$, and by $a(n) \asymp b(n)$ that $\liminf a(n)/b(n) > 0$ and $\limsup a(n)/b(n) < \infty$. By Lemma 2.7, if n and m tend to infinity so that $\max\{m/n^2, m^2/n^3\} \to 0$, then

$$\mu_{n,m} \sim n e^{-2m/n} \quad \text{and} \quad \sigma_{n,m}^2 \sim n\varphi(2m/n) \quad \text{for} \quad \varphi(x) = e^{-x}(1 - e^{-x}(1 + x)).$$

Hence, we have

$$\frac{\sigma_{n,m}^2}{\mu_{n,m}} \sim 1 - e^{-2m/n}\left(1 + \frac{2m}{n}\right),$$

so that

$$r_{n,m} \sim \sigma_{n,m} \frac{1 - e^{-2m/n}(1 + \frac{2m}{n})}{1 + (m/n)^2}.$$

For $n \asymp m$, the central domain, it follows that $r_{n,m} \asymp \sigma_{n,m}$, and moreover, in the special case where $m \sim cn$,

$$\mu_{n,m} \sim ne^{-2c} \quad \text{and} \quad \sigma_{n,m}^2 \sim ne^{-2c}\big(1 - e^{-2c}(1+2c)\big).$$

Regarding lower bounds, Englund (1981, Section 6) shows that for the standardized number of occupied cells in a uniform occupancy model with n balls and m boxes,

$$\sup_{z \in \mathbb{R}} |\mathbb{P}[W_{n,m} \le z] - \mathbb{P}[Z \le z]| \ge 0.087 / \max(3, \sigma_{n,m}).$$

Englund's argument holds without changes for any random variable with finite variance supported on the integers, and so also for the number of isolated vertices in our model. Hence, since in the central domain $r_{n,m} \asymp \sigma_{n,m}$, the rate function is of optimal order.

If $m \to \infty$ and $m/n \to 0$, the left domain, say, then

$$r_{n,m} \asymp \frac{\sigma_{n,m}m^2}{n^2} \asymp \frac{m^3}{n^{5/2}}$$

since $1 - e^{-x}(1+x) \sim x^2/2$ as $x \to 0$ for the first relation, and $\sigma_{n,m}^2 \asymp m^2/n$ for the second. In this case, Englund's lower bound is not achieved since $r_{n,m} = o(\sigma_{n,m})$. Nonetheless, the bound is informative as long as $r_{n,m} \to \infty$, which is the case as long as $m/n^{5/6} \to \infty$, such as when $m = cn^\alpha$ for $c > 0$ and $5/6 < \alpha < 1$.

If $m/n \to \infty$, the right domain, using $\sigma_{n,m}^2 \asymp ne^{-2m/n}$ for the second relation we have

$$r_{n,m} \sim \frac{\sigma_{n,m}n^2}{m^2} \asymp \frac{e^{-m/n}n^{5/2}}{m^2},$$

so Englund's lower bound is not attained. However, $r_{n,m}$ goes to infinity when $m \le \alpha n \log n$ for $0 < \alpha < 1/2$.

In the second example, we use the zero bias coupling constructed in Fulman and Goldstein (2011, Theorem 3.1) in Theorem 1.2 to give a bound on the normal approximation of the content Y of a Young tableux under Jack_α measure over a range of large α. In more detail, we recall that a partition of a positive integer n can be represented as a vector $\Lambda = (\lambda_1, \dots, \lambda_p)$ of non-increasing, positive integers summing to n, where p is the number of parts of the partition. For instance, $\Lambda = (4, 2, 1)$ corresponds to a partition of $n = 7$ with $p = 3$. In turn, the partition Λ can be represented by a tableaux with p rows of equal sized boxes, whose j^{th} row is of length λ_j, such as in (1.23).

The Jack_α measure on tableaux, defined for $\alpha > 0$, recovers the Plancherel measure when specializing to the case $\alpha = 1$. Under Jack_α, see Fulman (2004) for instance, the probability of a partition Λ of n is given by

$$\text{Jack}_\alpha(\Lambda) = \frac{\alpha^n n!}{\prod_{x \in \Lambda}(\alpha a(x) + l(x) + 1)(\alpha a(x) + l(x) + \alpha)}, \tag{1.22}$$

where the product is over all boxes x in the partition, $a(x)$ denotes the number of boxes in the same row of x and to the right of x (the "arm" of x), and $l(x)$ denotes the number of boxes in the same column of x and below x (the "leg" of x). For each tableaux representing a partition of n we may define the α-content of any individual box by

$$c_\alpha(x) = \alpha(\text{column number of } x - 1) - (\text{row number of } x - 1),$$

as depicted in the following tableaux for the partition $(4, 2, 1)$ of 7:

0	α	2α	3α
-1	$\alpha-1$		
-2			

$$\tag{1.23}$$

Here we study the distribution of the standardized sum of the α-contents over all boxes in the tableaux, that is,

$$W = \frac{Y}{\sqrt{\alpha\binom{n}{2}}}, \quad \text{where} \quad Y = \sum_{x \in \Lambda_n} c_\alpha(x) \tag{1.24}$$

and where the partition Λ_n of n is sampled from the Jack$_\alpha$ measure in (1.22).

Fulman (2004) proved an $O(n^{-1/4})$ bound for the error in the Kolmogorov metric for the normal approximation of W, improved by Fulman (2006) using martingales to $O(n^{-1/2+\varepsilon})$ for any $\varepsilon > 0$, and by Fulman (2006) to $O(n^{-1/2})$ using Bolthausen's inductive approach and Stein's method, but without an explicit constant. Hora and Obata (2007) prove a central limit theorem, with no error bound, for $W_{n,\alpha}$ using quantum probability.

Fulman and Goldstein (2011) prove the bound

$$d_1(W, Z) \leq \sqrt{\frac{2}{n}}\left(2 + \sqrt{2 + \frac{\max(\alpha, 1/\alpha)}{n-1}}\right) \quad \text{for all } n \geq 2, \alpha > 0, \tag{1.25}$$

in the Wasserstein metric d_1, where Z is a standard normal variable. In addition to providing explicit constants, this bound also highlights the role of α. A natural question it brings is whether a bound in the Kolmogorov metric can be shown that has this same dependence on α. A few weeks before the current work was posted, Chen, Raič and Thành (2020+, Theorem 1.1) proved the bound

$$\sup_{x \in \mathbb{R}} |\mathbb{P}_{n,\alpha}[W \leq x] - \mathbb{P}[Z \leq x]| \leq 9\left(\frac{1}{\sqrt{n}} \vee \frac{(\sqrt{\alpha} \vee 1/\sqrt{\alpha})\log n}{n}\right),$$

which achieves this goal with an explicit constant to within a logarithmic factor.

Here, given any $\varepsilon \in (0, 1)$, we show that, in the 'large α' region $\alpha \geq n^{1+\varepsilon}$, this log factor may be removed, resulting in the bound having the same α dependence as (1.25). That is, as $\alpha \geq n$ over the region we consider, the ratio between the right hand sides of (1.25) and (1.26) is bounded away from zero and infinity. This same result, with an explicit constant, was also achieved by Chen, Raič and Thành (2020+, Proposition 4.1) by applying a different approach. We do not consider $\varepsilon > 1$, as Theorem 3.1 below shows that this case is degenerate.

Theorem 1.5. *For W as given in (1.24) with Λ_n sampled according to Jack$_\alpha$ measure for some $n \geq 2$, for every $\varepsilon \in (0, 1)$ there exists a constant C depending only on ε such that*

$$\sup_{z \in \mathbb{R}} |\mathbb{P}_{n,\alpha}[W \leq z] - \mathbb{P}[Z \leq z]| \leq \frac{C\sqrt{\alpha}}{n} \quad \text{for all } n \geq 2 \text{ and } \alpha \geq n^{1+\varepsilon}. \tag{1.26}$$

We remark that by applying the reasoning at the end of the proof of Theorem 4.1 of Fulman and Goldstein (2011) the result holds also for $\alpha \leq n^{-1-\varepsilon}$ when replacing the α on the right hand side by $1/\alpha$. In the computations that follow, C without subscript will denote a universal constant whose value may change from line to line, and for n a non-negative integer, $[n]$ will denote the set $\{1, \ldots, n\}$.

2 Isolated vertices in the Erdős-Rényi random graph

In this section we prove Theorem 1.3. We begin by reviewing Construction 2A of Chen and Röllin (2010) for Stein couplings. Let $\mathbf{X} = (X_1, \ldots, X_n)$ be a collection of mean zero random variables, and let I be a random index uniformly distributed over $[n]$,

independent of $\mathbf{X}$. Let $W = \sum_{i \in [n]} X_i$ and suppose that for each $i = 1, \ldots, n$ there exists W_i' such that

$$\mathbb{E}\{X_i | W_i'\} = 0. \tag{2.1}$$

Then, with $G = -nX_I$, the triple (W, W_I', G) is a Stein coupling. To verify the claim, first note that

$$\mathbb{E}\{Gf(W_I')\} = -\mathbb{E}\{nX_I f(W_I')\} = -\mathbb{E}\sum_{i \in [n]} X_i f(W_i')$$

$$= -\mathbb{E}\sum_{i \in [n]} \mathbb{E}\{X_i | W_i'\} f(W_i') = 0.$$

On the other hand,

$$-\mathbb{E}\{Gf(W)\} = \mathbb{E}\{nX_I f(W)\} = \mathbb{E}\sum_{i \in [n]} X_i f(W) = \mathbb{E}\{Wf(W)\};$$

so (1.2) holds.

2.1 Isolated vertices in $\mathrm{ER}(n, m)$

Consider the Erdős and Rényi (1960) random graph $\mathcal{G} \sim \mathrm{ER}(n, m)$ on n vertices, having exactly m edges, distributed uniformly at random. Let d_v be the degree of vertex $v \in [n]$, and consider the number of isolated vertices

$$Y = \sum_{v=1}^{n} \mathrm{I}[d_v = 0].$$

With $N = \binom{n}{2}$, the mean and variance of Y are given by, respectively,

$$\mu_{n,m} = n \frac{\binom{N-(n-1)}{m}}{\binom{N}{m}} \quad \text{and} \quad \sigma_{n,m}^2 = \mu_{n,m} + n(n-1)\frac{\binom{N-(2n-3)}{m}}{\binom{N}{m}} - \mu_{n,m}^2.$$

We remark that though there may be a choice of couplings for a given situation, the coupling we have chosen will work for the more general problem where Y is a sum

$$Y = \sum_{v=1}^{n} h_v(d_v)$$

of functions h_v of the degree d_v of vertex v. For instance, the size bias coupling will work, as in Goldstein (2013), for counting the number of vertices having specified degrees, but not in this greater generality.

Proof of Theorem 1.3. The proof consists of the setting up the framework, and then checking that Conditions (G1)–(G7) hold, with Condition (G2) requiring the construction of a Stein coupling. First, let $\mathcal{E}_n$ be the enumeration of all N unordered pairs $\{v, w\} \subset [n]$ with $v \neq w$, given by

$$\mathcal{E}_n = \big(\{1, 2\}, \ldots, \{1, n\}, \{2, 3\}, \ldots, \{2, n\}, \ldots, \{n-1, n\}\big). \tag{2.2}$$

Let π be a uniformly chosen random permutation of $[N]$. We will describe the construction of a graph $\mathcal{G}(m, \pi)$, determined by m and π, that has distribution $\mathrm{ER}(n, m)$. As n is determined by N, and hence by π, n may be omitted in the notation for the graph; the same principle will be applied without comment for like quantities that appear later.

We construct $\mathcal{G}(m,\pi)$ as follows. For each $\{v,w\} \subset [n]$ with $v < w$, connect vertices v and w with an edge if and only if

$$\pi^{-1}(i) \leq m, \tag{2.3}$$

where i is the index in the enumeration (2.2) corresponding to the pair $\{v,w\}$. Clearly this construction results in a graph with m edges — those with labels $\{\pi(1),\ldots,\pi(m)\}$. Since π is uniform it is immediate that $\mathcal{G}(m,\pi) \sim \mathrm{ER}(n,m)$. Let $d_v(m,\pi)$ be the degree of vertex $v \in [n]$ in $\mathcal{G}(m,\pi)$, let

$$I_v(m,\pi) = \mathrm{I}[d_v(m,\pi) = 0] \quad \text{and} \quad Y(m,\pi) = \sum_{v=1}^{n} I_v(m,\pi). \tag{2.4}$$

We now verify the conditions of Theorem 1.1 with Θ and $r_{n,m}$ as given in (1.20) and (1.21), respectively.

Condition (G1). Let n_0, m_0, c_0 and C_0 be as in Lemma 2.7. Now obtain $\bar{r}$ in the definition (1.6) of ☺ through Lemma 2.8 and the choices

$$\bar{n} = (2n_0) \vee 344, \qquad \overline{m} = (2m_0) \vee (8C_0) \vee 28, \quad \text{and} \quad \bar{c} = 1 \wedge \frac{c_0}{2} \wedge \frac{1}{3C_0^{1/2}}. \tag{2.5}$$

Since our definition of $r_{n,m}$ in (1.21) implies that $r_{n,m} = 0$ whenever $\sigma_{n,m}^2 = 0$, the condition that $\sigma_{n,m}^2 > 0$ on ☺ is satisfied. Note that by Lemma 2.8

$$n \geq \bar{n} \quad \text{and} \quad \overline{m} \leq m \leq \bar{c}n^{3/2} \quad \text{whenever } (n,m) \in \text{☺}. \tag{2.6}$$

Condition (G2). For $(n,m) \in$ ☺, let

$$W = \frac{Y(m,\pi) - \mu_{n,m}}{\sigma_{n,m}}, \tag{2.7}$$

and set $W = 0$ otherwise. Assume $(n,m) \in$ ☺. Let $\Sigma = (\sigma_1,\ldots,\sigma_n)$ be a collection of uniform random permutations of $[N]$, with $\pi,\sigma_1,\ldots,\sigma_n$ mutually independent. The purpose of the following algorithm is to take the graph $\mathcal{G}(m,\pi)$ as input and to construct, for each vertex $v \in [n]$, a graph $\mathcal{G}^v(m,\pi,\sigma_v)$ on the $n-1$ vertices $[n] \setminus \{v\}$, having distribution $\mathrm{ER}(n-1,m)$, independent of $d_v(m,\pi)$, and which can be closely coupled to $\mathcal{G}(m,\pi)$.

We first describe the algorithm in words: Initialise counters k and i that respectively record the number of edges successfully relocated, and the index of a candidate edge for possible addition to the new graph; for each given vertex $v \in [n]$, begin with $\mathcal{G}(m,\pi)$ and relocate the $d_v(m,\pi)$ edges incident to v uniformly by, incrementing i when needed, adding $\mathcal{E}_n(\sigma_v(i))$ as a new edge when it connects two vertices, neither of which are incident to v (Step 6), and which are not already connected (Step 7). The counter k records the number of edges successfully relocated, and the set $L^v(m,\pi,\sigma_v)$ holds their locations (that is, indices) in $\mathcal{E}_n$. At termination, the set $L^v(m,\pi,\sigma_v)$ will have size $d_v(m,\pi)$.

Algorithm 1. Fix $v \in [n]$.

  1. Let $L^v(m,\pi,\sigma_v) \leftarrow \emptyset$

  2. Let $\mathcal{G}'$ be equal to $\mathcal{G}(m,\pi)$, but with vertex v and all $d_v(m,\pi)$ edges incident to v removed.

  3. Let $k \leftarrow 0$ and $i \leftarrow 0$.

  4. If $k = d_v(m,\pi)$, then denote the resulting graph by $\mathcal{G}^v(m,\pi,\sigma_v)$, and stop.

5. Let $i \leftarrow i + 1$.

6. If $v \in \mathcal{E}_n(\sigma_v(i))$, then return to Step 5.

7. If $\pi^{-1}(\sigma_v(i)) \leq m$, that is, if $\mathcal{E}_n(\sigma_v(i))$ is an edge in $\mathcal{G}(m, \pi)$, then return to Step 5.

8. In $\mathcal{G}'$ connect the vertices in $\mathcal{E}_n(\sigma_v(i))$ by an edge, and moreover, let $L^v(m, \pi, \sigma_v) \leftarrow L^v(m, \pi, \sigma_v) \cup \{\sigma_v(i)\}$.

9. Let $k \leftarrow k + 1$.

10. Return to Step 4.

It is not difficult to see that the algorithm will succeed in redistributing the edges incident on v if and only if $m \leq \binom{n-1}{2}$, which is guaranteed by our choice of ☺. Note that, given m, π and σ_v, the construction of $\mathcal{G}^v(m, \pi, \sigma_v)$ from $\mathcal{G}(m, \pi)$ is *deterministic* and hence, for given m, π and σ_v, will always result in the same graph $\mathcal{G}^v(m, \pi, \sigma_v)$. Note also that, although $\mathcal{G}^v(m, \pi, \sigma_v)$ has only $n - 1$ vertices, we keep the labeling from the original graph $\mathcal{G}(m, \pi)$. Since the order at which potential locations where the $d_v(m, \pi)$ edges are added are sampled uniformly at random without replacement (via σ_v), it is clear that $\mathcal{G}^v(m, \pi, \sigma_v) \sim \text{ER}(n - 1, m)$, up to vertex labeling.

Now, let $W = W(m, \pi)$ as in (2.7). With $\mathcal{V}$ a uniformly chosen vertex from $[n]$, independent of $\pi, \sigma_1, \ldots, \sigma_n$, and recalling the notation in (2.4), let

$$G = -\frac{n}{\sigma_{n,m}}(I_{\mathcal{V}}(m, \pi) - \mu_{n,m}/n). \tag{2.8}$$

For $w \neq v$, let $d_w^v(m, \pi, \sigma_v)$ be the degree of vertex w in the graph $\mathcal{G}^v(m, \pi, \sigma_v)$, let

$$I_w^v(m, \pi, \sigma_v) = I[d_w^v(m, \pi, \sigma_v) = 0], \qquad Y^v(m, \pi, \sigma_v) = \sum_{w \in [n] \setminus v} I_w^v(m, \pi, \sigma_v),$$

and

$$W' = \frac{Y^{\mathcal{V}}(m, \pi, \sigma_{\mathcal{V}}) - \mu_{n,m}}{\sigma_{n,m}}, \quad \text{and hence,} \quad D = \frac{Y^{\mathcal{V}}(m, \pi, \sigma_{\mathcal{V}}) - Y(m, \pi)}{\sigma_{n,m}}. \tag{2.9}$$

Since the distribution of $\mathcal{G}^v(m, \pi, \sigma_v)$ is the same regardless of the value of $d_v(m, \pi)$, we conclude that $I_v(m, \pi) - \mu_{n,m}/n$ and $Y^v(m, \pi, \sigma_v)$ are independent, so (2.1) holds, implying (W, W', G) is a Stein coupling.

Condition (G3). In what follows, consider a fixed $(n, m) \in$ ☺, and drop the subscript θ in the expectations that follow. As W is a function of (π, Σ), using (1.15) we have

$$\mathbb{E}\big|\mathbb{E}(1 - GD|W)\big| \leq \big(\text{Var}\,\mathbb{E}(GD|\pi, \Sigma)\big)^{1/2}. \tag{2.10}$$

Now, from (2.8) and (2.9), we have

$$\text{Var}\,\mathbb{E}(GD|\pi, \Sigma) = \frac{1}{\sigma_{n,m}^4}\,\text{Var}\sum_{v \in [n]}(I_v(m, \pi) - \mu_{n,m}/n)(Y(m, \pi) - Y^v(m, \pi, \sigma_v)).$$

Splitting the sum into two and using $\text{Var}(X + Y) \leq 2\,\text{Var}\,X + 2\,\text{Var}\,Y$, we have

$$\text{Var}\,\mathbb{E}(GD|\pi, \Sigma) \leq \frac{2}{\sigma_{n,m}^4}\,\text{Var}(f_m(\pi, \Sigma)) + \frac{2\mu_{n,m}^2}{n^2\sigma_{n,m}^4}\,\text{Var}(g_m(\pi, \Sigma)),$$

where

$$f_m(\pi, \Sigma) = \sum_{v \in [n]} I_v(m, \pi)B_v(m, \pi, \sigma_v) \quad \text{and} \quad g_m(\pi, \Sigma) = \sum_{v \in [n]} B_v(m, \pi, \sigma_v) \tag{2.11}$$

with $B_v(m, \pi, \sigma_v) = Y(m, \pi) - Y^v(m, \pi, \sigma_v)$. Note that $f_m(\pi, \Sigma)$ and $g_m(\pi, \Sigma)$ are deterministic functions of m, π and Σ. Applying Lemma 2.1 and using the notation as there, we obtain

$$\operatorname{Var}\mathbb{E}(GD|\pi, \Sigma) \leq \frac{1}{\sigma_{n,m}^4}\big(R_{g,1} + R_{f,1} + R_{g,2} + R_{f,2}\big) \tag{2.12}$$

where

$$R_{g,1} = \frac{\mu_{n,m}^2}{n^2} \sum_{i=1}^{n} \mathbb{E}\big(g_m(\pi, \Sigma) - g_m(\pi, \Sigma_i')\big)^2,$$

$$R_{f,1} = \sum_{i=1}^{n} \mathbb{E}\big(f_m(\pi, \Sigma) - f_m(\pi, \Sigma_i')\big)^2,$$

$$R_{g,2} = \frac{\mu_{n,m}^2}{n^2} \sum_{j=1}^{N-1} \mathbb{E}\big(g_m(\pi, \Sigma) - g_m(\pi\tau_j, \Sigma)\big)^2,$$

$$R_{f,2} = \sum_{j=1}^{N-1} \mathbb{E}\big(f_m(\pi, \Sigma) - f_m(\pi\tau_j, \Sigma)\big)^2.$$

Bounding $R_{g,1}$ Note that

$$g_m(\pi, \Sigma) - g_m(\pi, \Sigma_i') = B_i(m, \pi, \sigma_i) - B_i(m, \pi, \sigma_i'), \tag{2.13}$$

since all differences arising from the first sum in (2.11) cancel except the one with index $v = i$. Applying the simple bound

$$|B_v(m, \pi, \sigma_v)| = |Y(m, \pi) - Y^v(m, \pi, \sigma_v)| \leq 1 + 2d_i(m, \pi)$$

we obtain

$$|B_i(m, \pi, \sigma_i) - B_i(m, \pi, \sigma_i')| \leq 2 + 4d_i(m, \pi). \tag{2.14}$$

Let $\operatorname{Hyp}(N, m, n)$ count the number of white balls among m draws from an urn with N balls, n of which are white and $N - n$ black. Note that the marginal distribution of the degree of any vertex in $\mathcal{G}(m, \pi)$ is $\operatorname{Hyp}(N, m, n-1)$, and hence has mean $2m/n$, since the graph's m edges are uniformly sampled among all N possibilities, and exactly $n-1$ of them are associated with a specific vertex. Hence, applying Lemma 2.2, (2.13) and (2.14), we obtain

$$\mathbb{E}\big(g_m(\pi, \Sigma) - g_m(\pi, \Sigma_i')\big)^2 \leq C\Big(1 + \frac{m^2}{n^2}\Big),$$

where we recall C denotes a universal constant, whose value may change from line to line. Thus, as $\mu_{n,m} \leq n$,

$$R_{g,1} \leq \frac{C\mu_{n,m}^2}{n}\Big(1 + \frac{m^2}{n^2}\Big) \leq C\mu_{n,m}\Big(1 + \frac{m^2}{n^2}\Big). \tag{2.15}$$

Bounding $R_{f,1}$ As for g_m, we likewise have

$$f_m(\pi, \Sigma) - f_m(\pi, \Sigma_i') = I_i(m, \pi)\big(B_i(m, \pi, \sigma_i) - B_i(m, \pi, \sigma_i')\big).$$

Noting that, if $I_i(m, \pi) = 1$, we have $d_i(m, \pi) = 0$ and hence $B_i(m, \pi, \sigma_i) = B_i(m, \pi, \sigma_i') = 1$, it is immediate that

$$R_{f,1} = 0. \tag{2.16}$$

Bounding $R_{g,2}$ In order to bound $R_{g,2}$, with τ_{ij} the transposition of i and j, note first that

$$g_m(\pi\tau_{ij}, \Sigma) = g_m(\pi, \Sigma), \qquad \text{if } i,j \le m \text{ or } i,j > m, \tag{2.17}$$

since g_m is a function of the graph $\mathcal{G}(m, \pi)$ and Σ, and by (2.3), the graph $\mathcal{G}(m, \pi)$ obtained from π does not change when swapping edge with edge or non-edge with non-edge. Hence, averaging over τ_j, a transposition of j and a uniformly chosen index in $\{j, \dots, N\}$, yields

$$R_{g,2} = \frac{\mu_{n,m}^2}{n^2} \sum_{j=1}^m \frac{1}{N-j+1} \sum_{i=m+1}^N \mathbb{E}\big(g_m(\pi, \Sigma) - g_m(\pi\tau_{ij}, \Sigma)\big)^2.$$

By exchangeability the expectation on the right hand side is constant for $j \le m$ and $i \ge m+1$; hence, for such i and j,

$$\mathbb{E}\big(g_m(\pi, \Sigma) - g_m(\pi\tau_{ij}, \Sigma)\big)^2 = \mathbb{E}\big(g_m(\pi, \Sigma) - g_m(\pi\tau_{1,m+1}, \Sigma)\big)^2,$$

so that

$$R_{g,2} = \frac{\mu_{n,m}^2}{n^2} \sum_{j=1}^m \frac{N-m}{N-j+1} \mathbb{E}\big(g_m(\pi, \Sigma) - g_m(\pi\tau_{1,m+1}, \Sigma)\big)^2$$

$$\le \frac{\mu_{n,m}^2 m}{n^2} \mathbb{E}\big(g_m(\pi, \Sigma) - g_m(\pi\tau_{1,m+1}, \Sigma)\big)^2.$$

Now,

$$\mathbb{E}\big(g_m(\pi, \Sigma) - g_m(\pi\tau_{1,m+1}, \Sigma)\big)^2$$
$$\le 2\,\mathbb{E}\big(g_m(\pi, \Sigma) - g_{m+1}(\pi\tau_{1,m+1}, \Sigma)\big)^2$$
$$\quad + 2\,\mathbb{E}\big(g_{m+1}(\pi\tau_{1,m+1}, \Sigma) - g_m(\pi\tau_{1,m+1}, \Sigma)\big)^2$$
$$= 2\,\mathbb{E}\big(g_m(\pi, \Sigma) - g_{m+1}(\pi, \Sigma)\big)^2 + 2\,\mathbb{E}\big(g_{m+1}(\pi, \Sigma) - g_m(\pi, \Sigma)\big)^2$$
$$= 4\,\mathbb{E}\big(g_m(\pi, \Sigma) - g_{m+1}(\pi, \Sigma)\big)^2;$$

here, we have first applied the inequality $(x + y)^2 \le 2x^2 + 2y^2$, followed by (2.17) with m replaced by $m+1$ to the first expectation in the expression that results to yield that $g_{m+1}(\pi\tau_{1,m+1}, \Sigma) = g_{m+1}(\pi, \Sigma)$, and $\pi\tau_{1,m+1} =_d \pi$ to the second expectation, where $=_d$ denotes equality in distribution. Hence,

$$R_{g,2} \le \frac{4\mu_{n,m}^2 m}{n^2} \mathbb{E}\big(g_m(\pi, \Sigma) - g_{m+1}(\pi, \Sigma)\big)^2$$
$$= \frac{4\mu_{n,m}^2 m}{n^2} \mathbb{E}\bigg(\sum_{v \in [n]} \big(B_v(m, \pi, \sigma_v) - B_v(m+1, \pi, \sigma_v)\big) \bigg)^2. \tag{2.18}$$

Now, recalling that $L^v(m, \pi, \sigma_v)$ is the set of indices of edges to which those edges adjacent to vertex v were relocated, let

$$N^v(m, \pi, \sigma_v) = \bigcup_{i \in L^v(m, \pi, \sigma_v)} \mathcal{E}_n(i)$$
$$= \{w \in [n] \colon \exists i \in L^v(m, \pi, \sigma_v) \text{ such that } w \in \mathcal{E}_n(i)\}, \tag{2.19}$$

the set of vertices that received at least one additional edge when redistributing those edges. Also, let

$$M^v(m, \pi, \sigma_v) = \{w \in [n] \setminus v \,:\, \{w, v\} \in \mathcal{G}(m, \pi), w \notin N^v(m, \pi, \sigma_v)\}, \tag{2.20}$$

the neighbours of v that did not receive a new edge when redistributing the edges incident on v.

Note that the chosen vertex v will increase the difference $Y(m,\pi) - Y^v(m,\pi,\sigma_v)$ by one if it is isolated in $\mathcal{G}(m,\pi)$. A vertex $w \neq v$, will have this same effect if w is isolated in $\mathcal{G}(m,\pi)$ but then has an edge attached to it in the redistribution of the removed edges of v. On the other hand, a vertex $w \neq v$ will decrease this difference by one when w is connected to v, and has degree 1 in $\mathcal{G}(m,\pi)$, and does not have such an edge reattached. Hence, this difference is given by

$$B_v(m,\pi,\sigma_v) = I_v(m,\pi) + \sum_{\substack{w \in \\ N^v(m,\pi,\sigma_v)}} I_w(m,\pi) - \sum_{\substack{w \in \\ M^v(m,\pi,\sigma_v)}} I_{w,1}(m,\pi), \qquad (2.21)$$

where $I_{w,1}(m,\pi) = \mathrm{I}[d_w(m,\pi) = 1]$. Letting $\triangle$ denote set difference, we obtain

$$|B_v(m,\pi,\sigma_v) - B_v(m+1,\pi,\sigma_v)|$$

$$\leq |I_v(m,\pi) - I_v(m+1,\pi)|$$

$$+ \left| \sum_{\substack{w \in \\ N^v(m,\pi,\sigma_v)}} I_w(m,\pi) - \sum_{\substack{w \in \\ N^v(m+1,\pi,\sigma_v)}} I_w(m+1,\pi) \right|$$

$$+ \left| \sum_{\substack{w \in \\ M^v(m,\pi,\sigma_v)}} I_{w,1}(m,\pi) - \sum_{\substack{w \in \\ M^v(m+1,\pi,\sigma_v)}} I_{w,1}(m+1,\pi) \right|$$

$$\leq \mathrm{I}[v \in \mathcal{E}_n(\pi(m+1))] \qquad (2.22)$$

$$+ |N^v(m,\pi,\sigma_v) \cap \mathcal{E}_n(\pi(m+1))|$$

$$+ \sum_{\substack{w \in \\ N^v(m,\pi,\sigma_v) \triangle N^v(m+1,\pi,\sigma_v)}} I_w(m+1,\pi)$$

$$+ \sum_{\substack{w \in \\ M^v(m,\pi,\sigma_v)}} |I_{w,1}(m,\pi) - I_{w,1}(m+1,\pi)|$$

$$+ \sum_{\substack{w \in \\ M^v(m,\pi,\sigma_v) \triangle M^v(m+1,\pi,\sigma_v)}} I_{w,1}(m+1,\pi).$$

For the first term in (2.22), we have used that for any vertex $w \in [n]$ we can only have $I_w(m,\pi) \neq I_w(m+1,\pi)$ when w is an endpoint of the additional edge determined by $\pi(m+1)$, that is, when $w \in \mathcal{E}_n(\pi(m+1))$. For the second term in (2.22) we have used similarly that

$$\sum_{\substack{w \in \\ N^v(m,\pi,\sigma_v)}} |I_w(m,\pi) - I_w(m+1,\pi)| \leq |N^v(m,\pi,\sigma_v) \cap \mathcal{E}_n(\pi(m+1))|.$$

Moving now to the third term in (2.22), if $v \notin \mathcal{E}_n(\pi(m+1))$ and $\pi(m+1) \notin L^v(m,\pi,\sigma_v)$, then $L^v(m+1,\pi,\sigma_v) = L^v(m,\pi,\sigma_v)$; indeed, if $v \notin \mathcal{E}_n(\pi(m+1))$, vertex v has the same degree in both $\mathcal{G}(m,\pi)$ and $\mathcal{G}(m+1,\pi)$, and if also $\pi(m+1) \notin L^v(m,\pi,\sigma_v)$, then Algorithm 1 will redistribute the edges adjacent to v to the same available pairs of vertices when v has degree m or $m+1$; indeed, note that between the two cases m and $m+1$, Step 7 changes only if $\sigma_v(i) = \pi(m+1)$ for any of the i tested there, which is equivalent to $\pi(m+1) \in L^v(m,\pi,\sigma_v)$). Therefore, if $L^v(m,\pi,\sigma_v) \neq L^v(m+1,\pi,\sigma_v)$, we must either have $v \in \mathcal{E}_n(\pi(m+1))$ or $\pi(m+1) \in L^v(m,\pi,\sigma_v)$. Now, if $v \in \mathcal{E}_n(\pi(m+1))$,

then the degree of v in $\mathcal{G}(m+1,\pi)$ is one more than its degree in $\mathcal{G}(m,\pi)$, so $L^v(m+1,\pi,\sigma_v)$ will contain one more edge than $L^v(m,\pi,\sigma_v)$. And if $\pi(m+1) \in L^v(m,\pi,\sigma_v)$, then $|L^v(m,\pi,\sigma_v)\triangle L^v(m+1,\pi,\sigma_v)| = 2$ since $\pi(m+1)$ will be found blocked when forming $\mathcal{G}^v(m+1,\pi,\sigma_v)$ and a new non-edge has to be found. Hence,

$$\sum_{\substack{w \in \\ N^v(m+1,\pi,\sigma_v)\triangle N^v(m,\pi,\sigma_v)}} I_w(m+1,\pi)$$
$$\leq |N^v(m+1,\pi,\sigma_v)\triangle N^v(m,\pi,\sigma_v)|$$
$$\leq 2\,\mathrm{I}[v \in \mathcal{E}_n(\pi(m+1))] + 4\,\mathrm{I}[\pi(m+1) \in L^v(m,\pi,\sigma_v)].$$

For the fourth term in (2.22) we apply the bound

$$\sum_{\substack{w \in \\ M^v(m,\pi,\sigma_v)}} |I_{w,1}(m,\pi) - I_{w,1}(m+1,\pi)|$$
$$\leq \sum_{\substack{w: \\ \{w,v\}\in\mathcal{G}(m,\pi)}} |I_{w,1}(m,\pi) - I_{w,1}(m+1,\pi)|.$$

Finally, for the last term, similarly as for the third, if both $v \notin \mathcal{E}_n(\pi(m+1))$ and $\pi(m+1) \notin L^v(m,\pi,\sigma_v)$, it is easy to see that $M^v(m+1,\pi,\sigma_v) = M^v(m,\pi,\sigma_v)$; indeed, under these conditions, the set of vertices adjacent to v does not change with the addition of edge $m+1$, and moreover, $L^v(m+1,\pi,\sigma_v) = L^v(m,\pi,\sigma_v)$, which implies $N^v(m+1,\pi,\sigma_v) = N^v(m,\pi,\sigma_v)$, so that $M^v(m+1,\pi,\sigma_v) = M^v(m,\pi,\sigma_v)$. Hence, if $M^v(m,\pi,\sigma_v) \neq M^v(m+1,\pi,\sigma_v)$, we must either have $v \in \mathcal{E}_n(\pi(m+1))$ or $\pi(m+1) \in L^v(m,\pi,\sigma_v)$.

If $v \in \mathcal{E}_n(\pi(m+1))$, then v has one more neighbour in $\mathcal{G}(m+1,\pi)$ than in $\mathcal{G}(m,\pi)$, and so $L^v(m+1,\pi,\sigma_v)$ will contain one more edge than $L^v(m,\pi,\sigma_v)$. In this case, $M^v(m,\pi,\sigma_v)$ and $M^v(m+1,\pi,\sigma_v)$ can differ by at most three elements. Indeed, they may only differ by the additional neighbour in $\mathcal{G}(m+1,\pi)$, and by at most two existing neighbours of v in $\mathcal{G}(m,\pi)$ which were not assigned an edge in $L^v(m,\pi,\sigma_v)$, but were so assigned in $L^v(m+1,\pi,\sigma_v)$.

If $\pi(m+1) \in L^v(m,\pi,\sigma_v)$, then $|L^v(m,\pi,\sigma_v)\triangle L^v(m+1,\pi,\sigma_v)| = 2$, which means that $M^v(m,\pi,\sigma_v)$ and $M^v(m+1,\pi,\sigma_v)$ can differ by at most four elements; hence

$$\sum_{\substack{w \in \\ M^v(m,\pi,\sigma_v)\triangle M^v(m+1,\pi,\sigma_v)}} I_{w,1}(m+1,\pi)$$
$$\leq |M^v(m,\pi,\sigma_v)\triangle M^v(m+1,\pi,\sigma_v)|$$
$$\leq 3\,\mathrm{I}[v \in \mathcal{E}_n(\pi(m+1))] + 4\,\mathrm{I}[\pi(m+1) \in L^v(m,\pi,\sigma_v)].$$

Now recalling (2.18), summing (2.22) over $v \in [n]$ and noting that

$$\sum_{v\in[n]} \mathrm{I}[v \in \mathcal{E}_n(\pi(m+1))] \leq 2,$$

we obtain

$$R_{g,2} \leq \frac{C\mu_{n,m}^2 m}{n^2}\, \mathbb{E}\{1 + R_{g,2,1}^2 + R_{g,2,2}^2 + R_{g,2,3}^2\}, \tag{2.23}$$

where

$$R_{g,2,1} = \sum_{v \in [n]} |N^v(m, \pi, \sigma_v) \cap \mathcal{E}_n(\pi(m+1))|$$

$$R_{g,2,2} = \sum_{v \in [n]} \mathrm{I}[\pi(m+1) \in L^v(m, \pi, \sigma_v)]$$

$$R_{g,2,3} = \sum_{v \in [n]} \sum_{w:\{w,v\} \in \mathcal{G}(m,\pi)} |I_{w,1}(m, \pi) - I_{w,1}(m+1, \pi)|.$$

For the first term,

$$\begin{aligned}
\mathbb{E}R_{g,2,1}^2 = {} & n\, \mathbb{E}|N^1(m, \pi, \sigma_1) \cap \mathcal{E}_n(\pi(m+1))|^2 \\
& + n(n-1)\, \mathbb{E}\{|N^1(m, \pi, \sigma_1) \cap \mathcal{E}_n(\pi(m+1))| \\
& \qquad\qquad \times |N^2(m, \pi, \sigma_2) \cap \mathcal{E}_n(\pi(m+1))|\}.
\end{aligned} \tag{2.24}$$

Note that each vertex has at most $n-1$ potential edges available where the new edge $\pi(m+1)$ can be placed. Hence, since $N^1(m, \pi, \sigma_1) \leq 2d_1(m, \pi)$, there are at most $2d_1(m, \pi)(n-1)$ potential edges with one end in $N^1(m, \pi, \sigma_1)$, and so

$$\mathbb{P}[N^1(m, \pi, \sigma_1) \cap \mathcal{E}_n(\pi(m+1)) \neq \emptyset | d_1(m, \pi), N^1(m, \pi, \sigma_1)] \leq \frac{2d_1(m, \pi)(n-1)}{N-m}.$$

Noting that $|N^1(m, \pi, \sigma_v) \cap \mathcal{E}_n(\pi(m+1))|$ is bounded by $2\,\mathrm{I}[N^1(m, \pi, \sigma_v) \cap \mathcal{E}_n(\pi(m+1)) \neq \emptyset]$, recalling that $d_1(m, \pi) \sim \mathrm{Hyp}(N, m, n-1)$ and using Lemma 2.2, and also (2.6) of Condition (G1), which gives that $m \leq n^{3/2}$ as $\bar{c} \leq 1$ by (2.5), we therefore have

$$n\, \mathbb{E}|N^1(m, \pi, \sigma_v) \cap \mathcal{E}_n(\pi(m+1))|^2 \leq C\Big(1 + \frac{m}{n}\Big).$$

Moreover, with $\mathbb{P}_{12}[\cdot] = \mathbb{P}[\cdot | d_1(m, \pi), d_2(m, \pi), N^1(m, \pi, \sigma_1), N^2(m, \pi, \sigma_2)]$,

$$\begin{aligned}
& \mathbb{P}_{12}[N^1(m, \pi, \sigma_1) \cap \mathcal{E}_n(\pi(m+1)) \neq \emptyset, N^2(m, \pi, \sigma_2) \cap \mathcal{E}_n(\pi(m+1)) \neq \emptyset] \\
& \quad \leq \frac{4d_1(m, \pi)d_2(m, \pi)}{N-m},
\end{aligned}$$

since there are at most $2d_1(m, \pi) \times 2d_2(m, \pi)$ potential edges with one end in $N^1(m, \pi, \sigma_1)$ and the other end in $N^2(m, \pi, \sigma_2)$. Hence, again using $m \leq n^{3/2}$ and Lemma 2.2, and also Cauchy-Schwarz, we obtain

$$\begin{aligned}
& n(n-1)\, \mathbb{E}\{|N^1(m, \pi, \sigma_1) \cap \mathcal{E}_n(\pi(m+1))||N^2(m, \pi, \sigma_2) \cap \mathcal{E}_n(\pi(m+1))|\} \\
& \quad \leq C\Big(1 + \frac{m^2}{n^2}\Big),
\end{aligned}$$

so that (2.24) results in the bound

$$\mathbb{E}R_{g,2,1}^2 \leq C\Big(1 + \frac{m^2}{n^2}\Big). \tag{2.25}$$

Next, we have

$$\begin{aligned}
\mathbb{E}R_{g,2,2}^2 = {} & n\mathbb{P}[\pi(m+1) \in L^1(m, \pi, \sigma_1)] \\
& + n(n-1)\mathbb{P}[\pi(m+1) \in L^1(m, \pi, \sigma_1) \cap L^2(m, \pi, \sigma_2)].
\end{aligned} \tag{2.26}$$

To calculate the first probability, we condition on π and average over σ_1. If $1 \in \mathcal{E}_n(\pi(m+1))$, then the conditional probability vanishes, as no edge incident on the (removed) vertex v gets redistributed. Hence, take π such that $1 \notin \mathcal{E}_n(\pi(m+1))$. To compute $\mathbb{P}[\pi(m+1) \in$

$L^1(m, \pi, \sigma_1)|\pi]$, note that there are $N - m$ non-edges of $\mathcal{G}(m, \pi)$, out of which $n - 1 - d_1(m, \pi)$ involve vertex 1 and can therefore not be used during the redistribution of the $d_1(m, \pi)$ edges incident to vertex 1, which is to be removed. This leaves $N - m - n + 1 + d_1(m, \pi)$ potential edges from which to draw our sample of $d_1(m, \pi)$ non-edges. By uniformity, the probability that $\pi(m + 1)$ is in this sample is given by

$$p(\pi) := \mathbb{P}[\pi(m+1) \in L^1(m, \pi, \sigma_1)|\pi]$$
$$= \frac{d_1(m, \pi)}{N - m - n + 1 + d_1(m, \pi)} \leq \frac{d_1(m, \pi)}{N - m - n} \leq \frac{Cd_1(m, \pi)}{N}, \qquad (2.27)$$

as we only ask for the probability that one special object is included in a simple random sample of $d_1(m, \pi)$ objects from a population of size $N - m - n + 1 + d_1(m, \pi)$, and where in the final inequality we have used (2.6) of Condition (G1). Averaging over π, for the first term in (2.26) we obtain the bound

$$n \, \mathbb{E}p(\pi) \leq \frac{Cn \, \mathbb{E}d_1(m, \pi)}{N} = \frac{Cn(n-1)m}{N^2} \leq \frac{Cm}{n^2}. \qquad (2.28)$$

Next, as the events $\pi(m + 1) \in L^1(m, \pi, \sigma_1)$ and $\pi(m + 1) \in L^2(m, \pi, \sigma_2)$ are conditionally independent given π, we may handle the second, off diagonal term of (2.26) by using Lemma 2.2 to give that

$$\mathbb{E}H^2 \leq C\left(1 + \frac{m^2}{n^2}\right) \quad \text{when} \quad H \sim \mathrm{Hyp}(N, m, n), \qquad (2.29)$$

which, recalling (2.27), results in the bound

$$n^2 \, \mathbb{E}\left\{\mathbb{P}[\pi(m+1) \in L^1(m, \pi, \sigma_1)|\pi]\mathbb{P}[\pi(m+1) \in L^1(m, \pi, \sigma_2)|\pi]\right\}$$
$$= n^2 \, \mathbb{E}p(\pi)^2 \leq \frac{Cn^2 \, \mathbb{E}d_1(m, \pi)^2}{N^2} \leq \frac{Cn^2}{N^2}\left(1 + \frac{m^2}{n^2}\right).$$

Thus, using (2.26), (2.28) and the inequality directly above, we obtain

$$\mathbb{E}R_{g,2,2}^2 \leq C\left(1 + \frac{m^2}{n^2}\right). \qquad (2.30)$$

Finally, in order to bound $R_{g,2,3}$, note that the double sum is simply twice the sum over all the vertices of edges in $\mathcal{G}(m, \pi)$. Note also that, as w must have degree at least one to be included in the sum, $I_{w,1}(m, \pi) \neq I_{w,1}(m+1, \pi)$ only if w has degree 1 in $\mathcal{G}(n, m)$ and it receives the additional edge $\pi(m + 1)$. Thus, since the additional edge has two endpoints, it is immediate that $R_{g,2,3}$ can be no more than 4, so that

$$\mathbb{E}R_{g,2,3}^2 \leq 16. \qquad (2.31)$$

Recalling (2.23) and applying (2.25), (2.30) and (2.31) yields

$$R_{g,2} \leq \frac{C\mu_{n,m}^2 m}{n^2}\left(1 + \frac{m^2}{n^2}\right).$$

Now, by Lemma 2.6, we have $\mu_{n,m}/n \leq \exp(-2m/n)$, and since $x \exp(-2x)$ remains bounded on the positive real numbers, it follows that $m\mu_{n,m}/n^2$ is bounded; hence,

$$R_{g,2} \leq C\mu_{n,m}\left(1 + \frac{m^2}{n^2}\right). \qquad (2.32)$$

Bounding $R_{f,2}$ Using the same arguments as those used for $R_{g,2}$ to reach (2.18), we can show that

$$R_{f,2} \leq Cm \, \mathbb{E}\bigg(\sum_{v\in[n]} \big(I_v(m,\pi)B_v(m,\pi,\sigma_v) - I_v(m+1,\pi)B_v(m+1,\pi,\sigma_v)\big)\bigg)^2$$

Adding and subtracting $I_v(m+1,\pi)B_v(m,\pi,\sigma_v)$, and splitting the sum, we obtain

$$R_{f,2} \leq Cm \, \mathbb{E}\big\{R_{f,2,1}^2 + R_{f,2,2}^2\big\},$$

where

$$R_{f,2,1} = \sum_{v\in[n]} \big(I_v(m,\pi) - I_v(m+1,\pi)\big)B_v(m,\pi,\sigma_v),$$

$$R_{f,2,2} = \sum_{v\in[n]} I_v(m+1,\pi)\big(B_v(m,\pi,\sigma_v) - B_v(m+1,\pi,\sigma_v)\big).$$

In order to bound $R_{f,2,1}$, note first that $I_v(m,\pi) - I_v(m+1,\pi)$ is non-zero, and in that case equals one, exactly when vertex v is isolated in $\mathcal{G}(m,\pi)$ and the $(m+1)^{\text{th}}$ added edge is incident on v; that is,

$$I_v(m,\pi) - I_v(m+1,\pi) = I_v(m,\pi)\,\mathrm{I}[v \in \mathcal{E}_n(\pi(m+1))].$$

And since $I_v(m,\pi) = 1$ implies $B_v(m,\pi,\sigma_v) = 1$, we have

$$R_{f,2,1} = \sum_{v\in[n]} I_v(m,\pi)\,\mathrm{I}[v \in \mathcal{E}_n(\pi(m+1))].$$

Squaring, taking expectation and using exchangeability, we obtain

$$\begin{aligned}
\mathbb{E}R_{f,2,1}^2 \\
= \, & n\,\mathbb{E}\big\{I_1(m,\pi)\,\mathrm{I}[1 \in \mathcal{E}_n(\pi(m+1))]\big\} \\
& + n(n-1)\,\mathbb{E}\big\{I_1(m,\pi)I_2(m,\pi)\,\mathrm{I}[1 \in \mathcal{E}_n(\pi(m+1))]\,\mathrm{I}[2 \in \mathcal{E}_n(\pi(m+1))]\big\} \\
=: \, & R_{f,2,1,1} + R_{f,2,1,2}.
\end{aligned}$$

For the first term, we have

$$R_{f,2,1,1} = \mu_{n,m}\mathbb{P}[1 \in \mathcal{E}_n(\pi(m+1))|d_1(m,\pi) = 0] = \frac{\mu_{n,m}(n-1)}{N-m} \leq \frac{C\mu_{n,m}}{n},$$

while for the second term

$$\begin{aligned}
R_{f,2,1,2} &= n(n-1)\mathbb{P}[d_1(m,\pi) = d_2(m,\pi) = 0] \\
&\quad \times \mathbb{P}[\mathcal{E}_n(\pi(m+1)) = \{1,2\}|d_1(m,\pi) = d_2(m,\pi) = 0] \\
&= \mathrm{Hyp}(N,m,2n-3)(\{0\})\frac{n(n-1)}{N-m} \leq \frac{C\mu_{n,m}}{n},
\end{aligned}$$

where we have used that $\mathrm{Hyp}(N,m,n)(\{0\})$, the probability that a hypergeometric variable with the given parameters takes the value 0, is a decreasing function of the number of special items n. Hence,

$$\mathbb{E}R_{f,2,1}^2 \leq \frac{C\mu_{n,m}}{n}$$

In order to handle $R_{f,2,2}$, note that if $I_v(m+1,\pi) = 1$, we necessarily have $I_v(m,\pi) = 1$, so that $B_v(m,\pi,\sigma_v) = B_v(m+1,\pi,\sigma_v) = 1$ whenever $I_v(m+1,\pi) = 1$; it follows that

$$R_{f,2,2} = 0.$$

Therefore,

$$R_{f,2} \leq C\mu_{n,m}\frac{m}{n} \leq C\mu_{n,m}\left(1 + \frac{m^2}{n^2}\right). \tag{2.33}$$

Combining the bounds (2.15), (2.16), (2.32) and (2.33) as in (2.12), and then recalling (2.10), we obtain

$$\mathbb{E}\big|\mathbb{E}(1 - GD|W)\big| \leq \frac{C\sqrt{\mu_{n,m}}}{\sigma_{n,m}^2}\sqrt{1 + \frac{m^2}{n^2}}.$$

Recalling (1.21) and noting that $\sigma_{n,m}^2 \leq \mu_{n,m}$ by Lemma 2.5, the first condition in (1.7) holds, as

$$\sup_{(n,m)\in\text{☺}} r_{n,m}\,\mathbb{E}\big|\mathbb{E}(1 - GD|W)\big| < \infty.$$

Next, it clearly suffices to verify the second condition in (1.7) of (G3) with D replaced by its absolute upper bound

$$\overline{D} = \frac{1 + 2d_{\mathcal{V}}}{\sigma_{n,m}} \tag{2.34}$$

obtained in (2.14), and splitting the resulting expression to be bounded into two terms, we have

$$\mathbb{E}\big\{(1 + |W|)|G|\overline{D}^2\big\} = \mathbb{E}\big\{|G|\overline{D}^2\big\} + \mathbb{E}\big\{|W||G|\overline{D}^2\big\}. \tag{2.35}$$

Now, let $a \geq 1$. Using the given form (2.8) of G, we obtain

$$\begin{aligned}
\mathbb{E}\big\{|G|\overline{D}^a\big\} &= \frac{n}{\sigma_{n,m}}\mathbb{E}\big\{|I_{\mathcal{V}} - \frac{\mu_{n,m}}{n}|\overline{D}^a\big\} \\
&\leq \frac{n}{\sigma_{n,m}}\mathbb{E}\big\{I_{\mathcal{V}}\overline{D}^a\big\} + \frac{\mu_{n,m}}{\sigma_{n,m}}\mathbb{E}\overline{D}^a \leq \frac{C\mu_{n,m}}{\sigma_{n,m}^{1+a}}\left(1 + \left(\frac{m}{n}\right)^a\right),
\end{aligned} \tag{2.36}$$

where, for the final inequality, we used that $\overline{D} = 1/\sigma_{n,n}$ when $I_{\mathcal{V}} = 1$ and that $\mathbb{E}I_{\mathcal{V}} = \mu_{n,m}/n$ on the first summand, and Lemma 2.2 on the second summand. Setting $a = 2$ we obtain the bound

$$\mathbb{E}\big\{|G|\overline{D}^2\big\} \leq \frac{C\mu_{n,m}}{\sigma_{n,m}^3}\left(1 + \left(\frac{m}{n}\right)^2\right) \tag{2.37}$$

on the first term of (2.35).

The second term in (2.35) likewise leads to two terms, corresponding to the two in the second line of (2.36), but with an additional factor of $|W|$. Now setting $a = 2$, for the first we have, by applying Cauchy-Schwarz,

$$\frac{n}{\sigma_{n,m}}\mathbb{E}\big\{|W|I_{\mathcal{V}}\overline{D}^2\big\} \leq \frac{\mu_{n,m}}{\sigma_{n,m}^3}\mathbb{E}\big\{|W|\big|I_{\mathcal{V}} = 1\big\} \leq \frac{\mu_{n,m}}{\sigma_{n,m}^3}\sqrt{\mathbb{E}\big\{W^2\big|I_{\mathcal{V}} = 1\big\}}. \tag{2.38}$$

Conditional on vertex $\mathcal{V}$ being isolated, the distribution of the number of isolated vertices in the $\text{ER}(n,m)$ model is one more than the number of isolated vertices in the $\text{ER}(n-1,m)$ model. Hence, writing

$$W = \frac{\sigma_{n-1,m}}{\sigma_{n,m}} \times \frac{Y - \mu_{n-1,m}}{\sigma_{n-1,m}} + \frac{\mu_{n-1,m} - \mu_{n,m}}{\sigma_{n,m}},$$

and using $(x+y)^2 \leq 2(x^2 + y^2)$ twice, we obtain

$$\mathbb{E}_{n,m}\big\{W^2\big|I_\mathcal{V} = 1\big\}$$

$$\leq 2\left(\frac{\sigma_{n-1,m}}{\sigma_{n,m}}\right)^2 \mathbb{E}_{n-1,m}\left(\frac{Y - \mu_{n-1,m} + 1}{\sigma_{n-1,m}}\right)^2 + 2\left(\frac{\mu_{n-1,m} - \mu_{n,m}}{\sigma_{n,m}}\right)^2$$

$$\leq \frac{4\sigma_{n-1,m}^2}{\sigma_{n,m}^2} + \frac{4 + 2(\mu_{n-1,m} - \mu_{n,m})^2}{\sigma_{n,m}^2}.$$

Lemma 2.9 yields that the first term is bounded by a constant. For the second term, by removing all edges from the nth vertex and relocating them among the remaining vertices, we have a coupling of $\mathrm{ER}(n, m)$ and $\mathrm{ER}(n-1, m)$ which yields $|Y_{n-1,m} - Y_{n,m}| \leq 1 + 2d_n$, so that

$$|\mu_{n-1,m} - \mu_{n,m}| \leq 1 + 2\,\mathbb{E}d_n \leq 1 + \frac{2mn}{N} \leq C\Big(1 + \frac{m}{n}\Big).$$

Using that $r_{n,m}$ in (1.21) is lower bounded by $\bar{r}$, which is at least 1 by Lemma 2.8, and that $\mu_{n,m} \geq \sigma_{n,m}^2$ by Lemma 2.5 yields $\sigma_{n,m} \geq (1 + (m/n)^2)$, and using also (2.38), we conclude that

$$\mathbb{E}[W^2|I_\mathcal{V} = 1] \leq C \quad \text{and hence} \quad \frac{n}{\sigma_{n,m}}\mathbb{E}\big\{|W|I_\mathcal{V}\overline{D}^2\big\} \leq \frac{\mu_{n,m}}{\sigma_{n,m}^3}. \tag{2.39}$$

For the corresponding second term of (2.36), with $a = 2$ and the additional factor of $|W|$, using Cauchy-Schwarz and $\mathbb{E}W^2 = 1$,

$$\frac{\mu_{n,m}}{\sigma_{n,m}}\mathbb{E}\big[|W|D^2\big] \leq \frac{\mu_{n,m}}{\sigma_{n,m}}\sqrt{\mathbb{E}[D^4]} \leq \frac{C\mu_{n,m}}{\sigma_{n,m}^3}\left(1 + \Big(\frac{m}{n}\Big)^2\right), \tag{2.40}$$

applying Lemma 2.2. Combining with (2.37) and (2.39) we see the sum is of the order of (2.40) and it follows that

$$\sup_{(n,m)\in\smiley} r_{n,m}\,\mathbb{E}_{n,m}\big\{(1 + |W|)|G|D^2\big\} < \infty.$$

Condition (G4). Let $(n, m) \in \smiley$, and define

$$\mathcal{F}_{n,m} = \sigma\big(\mathcal{V}, d_\mathcal{V}(m, \pi)\big), \tag{2.41}$$

the σ-algebra generated by the identity of the vertex chosen to be removed in the coupling and its degree. Letting $\overline{G} = |G|$, and $\overline{D}$ be as in (2.34), we see that both are clearly $\mathcal{F}_{n,m}$-measurable.

For the first condition in (1.8), let

$$F_{n,m,1} = \{d_\mathcal{V}(m, \pi) \leq t(n, m)\} \quad \text{where} \quad t(n, m) = \min\{n, m\}/4. \tag{2.42}$$

Recall (2.5) and (2.6); in particular, on $\smiley$, we have $n \geq 344$ and $28 \leq m \leq n^{3/2}$. It is straightforward to check that under these conditions,

$$\underline{t}(n, m) := \frac{4m}{n} + 2\log(m \wedge n) \leq t(n, m) \quad \text{for all } (n, m) \in \smiley. \tag{2.43}$$

Indeed, if for $m \leq n$, the bound follows using that $2\log m \leq (1/4 - 4/344)m$ for $m \geq 28$, while for $n \leq m$ one verifies, for $n \geq 344$, that $4\sqrt{n} + 2\log n \leq n/4$.

Now, bounding D by $\overline{D}$ as given in (2.34), writing F as short for $F_{n,m,1}$ and using that $\mathbb{P}[I_\mathcal{V} = 0] = 1 - \mu_{n,m}/n$ in the final inequality, we obtain

$$\mathbb{E}\big\{|G|D^2(1 - I_F)\big\} \leq \frac{n}{\sigma_{n,m}}\mathbb{E}\big\{|I_\mathcal{V} - \mu_{n,m}/n|\overline{D}^2(1 - I_F)\big\}$$

$$\leq \frac{n}{\sigma_{n,m}}\mathbb{E}\big\{I_\mathcal{V}\overline{D}^2(1 - I_F)\big\} + \frac{\mu_{n,m}}{\sigma_{n,m}}\mathbb{E}\big\{\overline{D}^2(1 - I_F)\big\}.$$

Since $\mathcal{V}$ cannot be both isolated and have positive degree, we have $I_\mathcal{V}(1 - I_F) = 0$ almsot surely, and so the first term is zero. Applying Cauchy-Schwarz to the second term and then invoking Lemma 2.2,

$$\mathbb{E}\{|G|D^2(1 - I_F)\} \leq \frac{\mu_{n,m}}{\sigma_{n,m}}\left(\mathbb{E}\overline{D}^4 \mathbb{E}(1 - I_F)\right)^{1/2}$$

$$\leq \frac{C\mu_{n,m}\left(1 + \left(\frac{m}{n}\right)^2\right)}{\sigma_{n,m}^3}\mathbb{P}[d_\mathcal{V}(m, \pi) > t(n, m)]^{1/2}. \qquad (2.44)$$

By Lemma 2.2 with $\gamma = 2m/n$ being the mean of $d_1(m, \pi)$, we have for any $t > \gamma$ that

$$\mathbb{P}[d_\mathcal{V}(m, \pi) > t] \leq \mathbb{P}[d_\mathcal{V}(m, \pi) > \gamma + (t - \gamma)]$$

$$\leq \exp\left(-\frac{(t - \gamma)^2}{t + \gamma}\right) \leq \exp\left(-\frac{t - 2\gamma}{2}\right);$$

trivially, the final expression upper bounds the left hand side for $t \leq \gamma$ as well and hence holds for all $t \geq 0$. Hence, with $\underline{t}(n, m)$ as in (2.43), by (2.44) and recalling $r_{n,m}$ in (1.21), we obtain

$$r_{n,m}^2 \mathbb{E}\{|G|D^2(1 - I_F)\} \leq \frac{C\sigma_{n,m}^3}{\mu_{n,m}(1 + (\frac{m}{n})^2)}\exp\left(-\frac{t(n, m) - 2\gamma}{4}\right)$$

$$\leq \frac{C(m \wedge n)^{1/2}}{1 + (\frac{m}{n})^2}\exp\left(-\frac{t(n, m) - 2\gamma}{4}\right) \qquad (2.45)$$

$$= \frac{C(m \wedge n)^{1/2}}{1 + \left(\frac{m}{n}\right)^2}\exp\left(-\frac{1}{2}\log(m \wedge n)\right) = \frac{C}{1 + \left(\frac{m}{n}\right)^2} \leq C,$$

where we have used that $\sigma_{n,m}^2 \leq \min\{\mu_{n,m}, 2m\}$ via Lemma 2.5, and trivially $\mu_{n,m} \leq n$, for the second inequality, thus showing the first condition in (1.8) is satisfied.

From (2.36) with $a = 2$ it follows that

$$\sup_{(n,m)\in\smiley} r_{n,m} \mathbb{E}_{n,m}\{\overline{GD}^2\} < \infty,$$

thus showing that the second condition in (1.8) is also satisfied.

Condition (G5). Denote by $\mathcal{G}^{\mathrm{emb},\mathcal{V}}$ the "embedded" graph obtained by removing vertex $\mathcal{V}$ and all its incident edges; we keep the original vertex labeling. As the remaining $m - d_\mathcal{V}(m, \pi)$ edges are uniformly distributed over the remaining $n - 1$ vertices, conditional on $\mathcal{F}_{n,m}$ in (2.41), the resulting graph has conditional distribution

$$\mathscr{L}(\mathcal{G}^{\mathrm{emb},\mathcal{V}}|\mathcal{F}_{n,m}) \sim \mathrm{ER}(n - 1, m - d_\mathcal{V}(m, \pi)) \qquad (2.46)$$

almost surely; this identity is again to be understood up to labeling. In particular, letting $d_w^{\mathrm{emb},\mathcal{V}}$ be the degree of vertex w in graph $\mathcal{G}^{\mathrm{emb},\mathcal{V}}$,

$$V = \sum_{w:w\neq\mathcal{V}} \mathrm{I}[d_w^{\mathrm{emb},\mathcal{V}} = 0]$$

is the number of isolated vertices of $\mathcal{G}^{\mathrm{emb},\mathcal{V}}$, and (2.46) implies

$$\mathscr{L}(V|\mathcal{F}_{n,m}) = \mathscr{L}_\Psi(Y) \quad \text{where} \quad \Psi = (n - 1, m - d_\mathcal{V}(m, \pi)).$$

Clearly Ψ is $\mathcal{F}_{n,m}$-measurable. Now set $F_{n,m,2} = F_{n,m,1}$ as in (2.42), which is also clearly $\mathcal{F}_{n,m}$ measurable. Condition (1.10) is clearly equivalent to the first condition in (1.8), which was verified in (2.45).

Condition (G6). Let

$$\overline{B} = \frac{d_{\mathcal{V}}(m, \pi) + 1}{\sigma_{n,m}},$$

which is clearly $\mathcal{F}_{n,m}$-measurable. Moreover, $\sigma_{n,m}^{-1}|Y - V| \leq \overline{B}$ since removing any edge connected to vertex $\mathcal{V}$ can make at most one vertex, other than $\mathcal{V}$, isolated; the additional term of one accounts for the case when vertex $\mathcal{V}$ is isolated. Since $\overline{B} \leq \overline{D}$, as given in (2.34), by setting $a = 3$ in (2.36) we obtain

$$r_{n,m}^2 \, \mathbb{E}\{\overline{GD}^2 \overline{B} I_{F_{n,m,2}}\} \leq r_{n,m}^2 \, \mathbb{E}\{\overline{GD}^2 \overline{B}\} \leq r_{n,m}^2 \, \mathbb{E}\{\overline{GD}^3\}$$
$$\leq r_{n,m}^2 \, \frac{C\mu_{n,m}(1 + (\frac{m}{n})^3)}{\sigma_{n,m}^4} = \frac{C\sigma_{n,m}^2}{\mu_{n,m}(1 + \frac{m}{n})}.$$

As $\sigma_{n,m}^2 \leq \mu_{n,m}$ via Lemma 2.5, the second bound in (1.11) holds.

Condition (G7). We verify the stronger conditions that (1.12) and the second bound of (1.13) hold when taking the larger supremum obtained when removing the intersection with $\{\Psi \in \text{☺}\}$. This stronger version of (1.12) is an immediate consequence of Lemma 2.9. As this same lemma shows that the ratios in (1.13) involving means and variances are bounded by a constant, it is only required to bound the ratios of the remaining factor. For $r_{n,m}/r_{n-1,m-d}$, we have

$$\frac{(1 + ((m-d)/(n-1))^2)}{(1 + (m/n)^2)} \leq \frac{1 + 2(m/n)^2}{1 + (m/n)^2} \leq 2,$$

and for the reciprocal, using that $m/n \leq 2(m-d)/(n-1)$ for $d \leq m/4$,

$$\frac{(1 + (m/n)^2)}{(1 + ((m-d)/(n-1))^2)} \leq 4.$$

Conditions (G1)–(G7) have been verified, and Theorem 1.3 now follows from Theorem 1.1. $\qquad\square$

2.2 Technical results

Lemma 2.1 (Efron-Stein-type variance bound)**.** *Let π and the components of the vector $\Sigma = (\sigma_1, \ldots, \sigma_n)$ be independent uniform random permutations of $[N]$, and let $h(\pi, \Sigma)$ be a real-valued function. Let $\tau_1, \ldots, \tau_{N-1}$ be random transpositions independent of each other and of (π, Σ), where τ_j transposes j and a uniformly chosen integer in the set $\{j, \ldots, N\}$. Let $\Sigma' = (\sigma_1', \ldots, \sigma_n')$ be an independent copy of Σ and let $\Sigma_i' = (\sigma_1, \ldots, \sigma_{i-1}, \sigma_i', \sigma_{i+1}, \ldots, \sigma_n)$. Then*

$$\text{Var}\, h(\pi, \Sigma) \leq \frac{1}{2} \sum_{i=1}^n \mathbb{E}\big(h(\pi, \Sigma) - h(\pi, \Sigma_i')\big)^2 + \frac{1}{2} \sum_{j=1}^{N-1} \mathbb{E}\big(h(\pi, \Sigma) - h(\pi\tau_j, \Sigma)\big)^2.$$

Proof. Without loss of generality assume $\mathbb{E}h(\pi, \Sigma) = 0$. Let $\pi_0 = \pi$ and $\Sigma_0 = \Sigma$, and let

$$\pi_j = \pi_0 \tau_j \cdots \tau_1, \qquad 1 \leq j \leq N-1,$$

and

$$\Sigma_i = (\sigma_1', \ldots, \sigma_i', \sigma_{i+1}, \ldots, \sigma_n), \qquad 1 \le i \le n.$$

Let B be uniform on $\{0, 1\}$, let I be uniform on $\{1, \ldots, n\}$, let J be uniform on the set $\{1, \ldots, N-1\}$, and assume B, I and J are mutually independent and independent of all else. Let $W = h(\pi_0, \Sigma_0)$, let $W_{1,i}' = h(\pi_0, \Sigma_i')$, and let $W_{2,j}' = h(\pi_0 \tau_j, \Sigma_0)$, and $W' = B W_{1,I}' + (1 - B) W_{2,J}'$. Let $G_{1,i} = n\big(h(\pi_{N-1}, \Sigma_i) - h(\pi_{N-1}, \Sigma_{i-1})\big)$, let $G_{2,j} = (N-1)\big(h(\pi_j, \Sigma_0) - h(\pi_{j-1}, \Sigma_0)\big)$, and let $G = B G_{1,I} + (1 - B) G_{2,J}$. Let $g : \mathbb{R} \to \mathbb{R}$ be any bounded measurable function. Then, on the one hand,

$$
\begin{aligned}
-\mathbb{E}\{Gg(W)\} &= -\frac{1}{2} \sum_{i=1}^{n} \mathbb{E}\big\{\big(h(\pi_{N-1}, \Sigma_i) - h(\pi_{N-1}, \Sigma_{i-1})\big)g(W)\big\} \\
&\quad - \frac{1}{2} \sum_{j=1}^{N-1} \mathbb{E}\big\{\big(h(\pi_j, \Sigma_0) - h(\pi_{j-1}, \Sigma_0)\big)g(W)\big\} \\
&= -\frac{1}{2} \mathbb{E}\big\{\big(h(\pi_{N-1}, \Sigma_n) - h(\pi_{N-1}, \Sigma_0)\big)g(W)\big\} \\
&\quad - \frac{1}{2} \mathbb{E}\big\{\big(h(\pi_{N-1}, \Sigma_0) - h(\pi_0, \Sigma_0)\big)g(W)\big\} \\
&= -\frac{1}{2} \mathbb{E}h(\pi_{N-1}, \Sigma_n)\, \mathbb{E}g(W) + \frac{1}{2} \mathbb{E}\big\{h(\pi_0, \Sigma_0)g(W)\big\} \\
&= \frac{1}{2} \mathbb{E}\{Wg(W)\},
\end{aligned}
$$

where we used that (π_{N-1}, Σ_n) is equal in distribution to and independent of (π_0, Σ_0); this follows e.g. from *Algorithm P* of Knuth (1969, p. 147) since the distribution of π_{N-1} is uniform conditionally on π_0, and therefore independent of π_0.

On the other hand, for all i we have $(\Sigma_i, \Sigma_{i-1}, \Sigma_i') =_d (\Sigma_{i-1}, \Sigma_i, \Sigma_0)$ since $\Sigma =_d \Sigma_i'$, and for all j that $(\pi_j, \pi_{j-1}, \pi_0 \tau_j) =_d (\pi_{j-1}, \pi_j, \pi_0)$, by recalling the definition of π_j and observing that π_0 and $\pi_0 \tau_j$ have the same distribution, and that both are independent of $\tau_{j-1} \cdots \tau_1$, so

$$
\begin{aligned}
\mathbb{E}\{Gg(W')\} &= \frac{1}{2} \sum_{i=1}^{n} \mathbb{E}\big\{\big(h(\pi_{N-1}, \Sigma_i) - h(\pi_{N-1}, \Sigma_{i-1})\big)g(h(\pi_0, \Sigma_i'))\big\} \\
&\quad + \frac{1}{2} \sum_{j=1}^{N-1} \mathbb{E}\big\{\big(h(\pi_j, \Sigma_0) - h(\pi_{j-1}, \Sigma_0)\big)g(h(\pi_0 \tau_j, \Sigma_0))\big\} \\
&= \frac{1}{2} \sum_{i=1}^{n} \mathbb{E}\big\{\big(h(\pi_{N-1}, \Sigma_{i-1}) - h(\pi_{N-1}, \Sigma_i)\big)g(h(\pi_0, \Sigma_0))\big\} \\
&\quad + \frac{1}{2} \sum_{j=1}^{N-1} \mathbb{E}\big\{\big(h(\pi_{j-1}, \Sigma_0) - h(\pi_j, \Sigma_0)\big)g(h(\pi_0, \Sigma_0))\big\} \\
&= \frac{1}{2} \mathbb{E}\big\{(h(\pi_{N-1}, \Sigma_0) - h(\pi_{N-1}, \Sigma_n))g(h(\pi_0, \Sigma_0))\big\} \\
&\quad + \frac{1}{2} \mathbb{E}\big\{(h(\pi_0, \Sigma_0) - h(\pi_{N-1}, \Sigma_0))g(h(\pi_0, \Sigma_0))\big\} \\
&= \frac{1}{2} \mathbb{E}\{Wg(W)\}.
\end{aligned}
$$

Therefore, (W, W', G) is a Stein coupling and, specializing (1.2) to the case $f(x) = x$ and

applying the Cauchy Schwarz inequality and noting that $(\Sigma_i, \Sigma_{i-1}) =_d (\Sigma_i', \Sigma)$, we have

$$
\begin{aligned}
\operatorname{Var} W &= \mathbb{E}\{G(W' - W)\} \\
&= \frac{1}{2} \sum_{i=1}^{n} \mathbb{E}\big\{ \big(h(\pi_{N-1}, \Sigma_i) - h(\pi_{N-1}, \Sigma_{i-1})\big)\big(h(\pi_0, \Sigma_i') - h(\pi_0, \Sigma)\big)\big\} \\
&\quad + \frac{1}{2} \sum_{j=1}^{N-1} \mathbb{E}\big\{ \big(h(\pi_j, \Sigma_0) - h(\pi_{j-1}, \Sigma_0)\big)\big(h(\pi_0 \tau_j, \Sigma_0) - h(\pi_0, \Sigma_0)\big)\big\} \\
&\le \frac{1}{2} \sum_{i=1}^{n} \Big(\mathbb{E}\big(h(\pi_{N-1}, \Sigma_i) - h(\pi_{N-1}, \Sigma_{i-1})\big)^2 \mathbb{E}\big(h(\pi_0, \Sigma_i') - h(\pi_0, \Sigma)\big)^2 \Big)^{1/2} \\
&\quad + \frac{1}{2} \sum_{j=1}^{N-1} \Big(\mathbb{E}\big(h(\pi_j, \Sigma) - h(\pi_{j-1}, \Sigma)\big)^2 \mathbb{E}\big(h(\pi_0 \tau_j, \Sigma) - h(\pi_0, \Sigma)\big)^2 \Big)^{1/2},
\end{aligned}
$$

from which the claim follows. □

Lemma 2.2 (Tail and moment bounds for the hypergeometric distribution). *Let H have the hypergeometric distribution $\mathrm{Hyp}(N, m, n)$, which equals the number of white balls among m draws from an urn with N balls, n of which are white and $N - n$ black. Let $\gamma = \mathbb{E}H = nm/N$. Then, for any $t > 0$,*

$$
\mathbb{P}[H \ge \gamma + t] \le \exp\left(\frac{-t^2}{2\gamma + t}\right) \tag{2.47}
$$

Moreover, for any $k \ge 1$, there is a constant C_k independent of γ such that

$$
\mathbb{E}H^k \le C_k(\gamma^k + 1).
$$

Proof. To construct a bounded size bias coupling, index the white balls by $[n]$, and write $H = \sum_{i=1}^{n} I_i$ where I_i is the indicator that the ith white ball is sampled. Construct H^s with the H-size biased distribution by uniformly sampling a random index J from 1 to n independently of $I_1, \ldots, I_n$; if $I_J = 1$, set $H^s = H$, otherwise independently and uniformly select a ball from the sample and swap it with the Jth white ball. It is easy to see that H^s has the size-bias distribution, see for instance, Lemma 2.1 of Goldstein and Rinott (1996). Moreover, $H^s = H + 1$ if a sampled black ball was swapped with the Jth white ball, and $H^s = H$ otherwise. Hence, $|H^s - H| \le 1$, and the tail-bound (2.47) follows readily from Theorem 1.1 of Ghosh and Goldstein (2011).

Now, it is straightforward to check that $t^2/(2\gamma + t) \ge (t - 1)/(\gamma + 1)$ whenever $t \ge 1$ and $\gamma > 0$, so that

$$
\mathbb{P}[H \ge \gamma + t] \le \exp\left(\frac{-(t - 1)}{\gamma + 1}\right) \quad \text{for all } t \ge 1.
$$

Hence, $H - \gamma - 1$ is stochastically dominated by an exponential random variable X with mean $1/(\gamma + 1)$, and in particular

$$
\mathbb{E}H^k \le \mathbb{E}(X + \gamma + 1)^k \le 3^{k-1}(\mathbb{E}X^k + \gamma^k + 1) = 3^{k-1}(k!(\gamma + 1)^k + \gamma^k + 1),
$$

from which the second claim easily follows. □

A bound similar to (2.47) can be obtained from Greene and Wellner (2017, Corollary 1) with better constants, but under additional conditions on the parameters of the hypergeometric distribution

Lemma 2.3. *If $H \sim \mathrm{Hyp}(N, m, n)$, then*

$$\frac{mn}{N} - \frac{m^2 n^2}{2N^2} \;\leq\; 1 - e^{-mn/N} \;\leq\; \mathbb{P}[H > 0] \;\leq\; \frac{mn}{N}$$

and

$$e^{-mn/(N-m-n+1)} \;\leq\; \mathbb{P}[H = 0] \;\leq\; e^{-mn/N},$$

where the lower bound on $\mathbb{P}[H = 0]$ is valid whenever $m + n - 1 < N$.

Proof. Since $\mathbb{P}[H > 0] \leq \mathbb{E}H$, the upper bound on $\mathbb{P}[H > 0]$ immediately follows. Using the usual exponential upper bound for the final inequality,

$$\left(1 - \frac{n}{N-m+1}\right)^m \leq \mathbb{P}[H = 0] = \left(1 - \frac{n}{N}\right) \cdots \left(1 - \frac{n}{N-m+1}\right)$$
$$\leq \left(1 - \frac{n}{N}\right)^m \leq e^{-mn/N}, \tag{2.48}$$

from which the upper bound on $\mathbb{P}[H = 0]$ and first lower bound on $\mathbb{P}[H > 0]$ follow. The second lower bound on $\mathbb{P}[H > 0]$ follows from the first lower bound and the inequality $e^{-x} \leq 1 - x + x^2/2$ when $x \geq 0$. The lower bound on $\mathbb{P}[H = 0]$ follows from the inequality $\log(1 + x) \geq x/(1 + x)$ for $x > -1$ and the lower bound in (2.48), which together yield

$$\mathbb{P}[H = 0] \geq \exp\left(-\frac{mn}{(N-m+1)\left(1 - \frac{n}{N-m+1}\right)}\right) = \exp\left(-\frac{mn}{N-m-n+1}\right). \qquad \square$$

Lemma 2.4. *For any $x \geq 0$*

$$\frac{\min\{x^2, 1\}}{4} \leq 1 - e^{-x}(1 + x) \leq \frac{\min\{x^2, 2\}}{2}.$$

Proof. The upper and lower bounds hold trivially at $x = 0$. With $\psi(x) = 1 - e^{-x}(1 + x)$, by Talyor's expansion around zero, for all $x > 0$ there exists $\xi_x \in (0, x)$ such that

$$\psi(x) = \psi(0) + x\psi'(0) + \frac{x^2}{2}\psi''(\xi_x) = \frac{x^2}{2}\psi''(\xi_x),$$
$$\text{where} \quad \psi'(x) = xe^{-x} \quad \text{and} \quad \psi''(x) = e^{-x}(1 - x).$$

For $y \in [0, 2]$ we have $|\psi''(y)| \leq |1 - y| \leq 1$, thus proving the upper bound $x^2/2$ over this interval. As $\psi'''(y) = e^{-y}(y - 2) \geq 0$ for all $y \geq 2$, the function $\psi''(y)$ is non-decreasing for $y \geq 2$. As $\psi''(2) = -e^{-2} \in (-1, 0)$, and $\lim_{y \to \infty} \psi''(y) = 0$, we have $\psi''(y) \in (-1, 0)$ for all $y \geq 2$, thus proving the upper bound $x^2/2$ on $(2, \infty)$. As $\psi'(x) \geq 0$ for all $x \geq 0$, the function is non-decreasing on $[0, \infty)$, and as $\psi(x) \to 1$ as $x \to \infty$, we have $\psi(x) \leq 1$ for all $x \geq 0$.

For the lower bound, for $x > 0$ letting

$$q(x) = \frac{1 - e^{-x}(1 + x)}{x^2}, \quad \text{we have} \quad q'(x) = \frac{e^{-x}(x^2 + 2x + 2) - 2}{x^3}.$$

With $p(x) = e^{-x}(x^2 + 2x + 2) - 2$ we have $p'(x) = -x^2 e^{-x} \leq 0$, so $q(x)$ is decreasing for $x > 0$. In particular, $q(x) \geq q(1) = 1 - 2e^{-1} \geq 1/4$ for $x \in [0, 1]$. As $\psi'(x) = xe^{-x}$, the function $\psi(x)$ is non-decreasing, and hence for $x \geq 1$ we have $\psi(x) \geq \psi(1) = 1 - 2e^{-1} \geq 1/4$, completing the proof of the lower bound. $\qquad \square$

Lemma 2.5. *For all $(n, m) \in \Theta$ and distinct vertices v and w, the indicators $\mathrm{I}[d_w = 0]$ and $\mathrm{I}[d_v = 0]$ that v and w are isolated are negatively correlated, that is,*

$$\mathbb{P}[d_v = 0, d_w = 0] \leq \mathbb{P}[d_v = 0]\mathbb{P}[d_w = 0], \quad and \quad \sigma_{n,m}^2 \leq \min\{\mu_{n,m}, 2m\}.$$

Proof. Vertex v is isolated if and only if none of the $n - 1$ edges that connect v to another vertex is included in the set of m edges selected. Likewise, distinct vertices v and w are both isolated if and only if none of a particular set of $(n - 2) + (n - 2) + 1$ edges is selected. Hence, the first claim is equivalent to

$$\frac{\binom{N-2n+3}{m}}{\binom{N}{m}} \leq \frac{\binom{N-n+1}{m}^2}{\binom{N}{m}^2} \quad \text{or} \quad \binom{N}{m}\binom{N-2n+3}{m} \leq \binom{N-n+1}{m}^2.$$

Expanding the binomial coefficients and canceling common factors yields the equivalent form

$$(N)_m(N - 2n + 3)_m \leq (N - n + 1)_m^2,$$

where $(n)_k = n(n-1)\cdots(n-k+1)$, and pairing up the kth factors of the falling factorials we obtain

$$\prod_{k=0}^{m-1} (N - k)(N - 2n + 3 - k) \leq \prod_{k=0}^{m-1} (N - n + 1 - k)^2.$$

It suffices to show the inequality holds termwise. Expanding both sides of the kth term of each side and simplifying yields

$$N + 2n \leq n^2 + 1 + k.$$

The case $k = 0$ implies all others, and reduces to $0 \leq n^2 - 3n + 2 = (n - 2)(n - 1)$, and so holds for all $n \geq 2$, thus proving the first claim.

Since the indicators of vertices being isolated are negatively correlated, we have

$$\sigma_{n,m}^2 \leq n\mathbb{P}[d_v = 0]\mathbb{P}[d_v > 0] \leq \min\{n\mathbb{P}[d_v = 0], n\mathbb{P}[d_v > 0]\},$$

from which $\sigma_{n,m}^2 \leq \mu_{n,m}$ is immediate. As $d_v \sim \mathrm{Hyp}(N, m, n - 1)$ for $N = n(n - 1)/2$, using Lemma 2.3 we have

$$\sigma_{n,m}^2 \leq n\mathbb{P}[d_v > 0] \leq n\frac{m(n - 1)}{N} = 2m,$$

as claimed. $\qquad\square$

Lemma 2.6. *For $n \geq 6$ and $0 \leq m \leq n^2/4 - 3n/2$, we have*

$$\exp\left(-\frac{2m}{n} - \frac{8m(m + n)}{n^3}\right) \leq \frac{\mu_{n,m}}{n} \leq \exp\left(-\frac{2m}{n}\right), \tag{2.49}$$

and

$$\mu_{n,m}\left[1 - \frac{\mu_{n,m}}{n}\left(1 + \frac{2m}{n} + \frac{78m(m + n)}{n^3}\right)\right]$$
$$\leq \sigma_{n,m}^2 \leq \mu_{n,m}\left[1 - \frac{\mu_{n,m}}{n}\left(1 + \frac{2m}{n} - \frac{48m(m + n)}{n^3}\right)\right]. \tag{2.50}$$

Proof. Since the distribution of each individual degree is $\text{Hyp}(N, m, n-1)$, and as the hypothesis of Lemma 2.3 holds due to the restriction assumed on m, it follows from that lemma that

$$\exp\Big(-\frac{m(n-1)}{N-m-n+2}\Big) \leq \frac{\mu_{n,m}}{n} \leq \exp\Big(-\frac{2m}{n}\Big),$$

yielding the upper bound in (2.49). Since under the assertions on m and n we have

$$n^2 - 2m - 3n + 4 \geq n^2/2, \tag{2.51}$$

it follows that

$$\frac{m(n-1)}{N-m-n+2} = \frac{2m}{n} + \frac{4m(m+n-2)}{n(n^2-2m-3n+4)} \leq \frac{2m}{n} + \frac{8m(m+n)}{n^3},$$

from which we obtain the lower bound in (2.49).

In order to prove the upper and lower bounds on the variance, we use the fact that $\text{Var}(W) = \mathbb{E}\{G(W'-W)\}$ when (W, W', G) is a Stein coupling for a mean zero random variable W; this identity follows immediately upon setting $f(x) = x$ in (1.2). Now recall (2.8), (2.9) and (2.21), and that $N^v(m, \pi, \sigma_v)$ in (2.19) is the set of vertices that receive at least one edge when forming $\mathcal{G}^v(m, \pi, \sigma_v)$, and that $M^v(m, \pi, \sigma_v)$ in (2.20) is the set of all vertices $w \neq v$ such that $\{v, w\}$ is an edge in $\mathcal{G}(m, \pi)$, and does not receive a redistributed edge. As when $I_v(m, \pi) = 1$ the sets $N^v(m, \pi, \sigma_v)$ and $M^v(m, \pi, \sigma_v)$ are empty, and recalling that $I_{w,1}(m, \pi) = \mathrm{I}[d_w(m, \pi) = 1]$, we have

$$
\begin{aligned}
&\sigma_{n,m}^2 \\
&= \mathbb{E}\sum_{v\in[n]} \Big(I_v(m,\pi) - \frac{\mu_{n,m}}{n}\Big) \\
&\qquad \times \Big(I_v(m,\pi) + \sum_{\substack{w\in \\ N^v(m,\pi,\sigma_v)}} I_w(m,\pi) - \sum_{\substack{w\in \\ M^v(m,\pi,\sigma_v)}} I_{w,1}(m,\pi)\Big) \\
&= n\,\mathbb{E}\Big\{\Big(I_1(m,\pi) - \frac{\mu_{n,m}}{n}\Big) \\
&\qquad \times \Big(I_1(m,\pi) + \sum_{\substack{w\in \\ N^1(m,\pi,\sigma_1)}} I_w(m,\pi) - \sum_{\substack{w\in \\ M^1(m,\pi,\sigma_1)}} I_{w,1}(m,\pi)\Big)\Big\} \\
&= n\,\mathbb{E}\Big\{I_1(m,\pi)\Big(1 - \frac{\mu_{n,m}}{n}\Big) - \frac{\mu_{n,m}}{n} \sum_{\substack{w\in \\ N^1(m,\pi,\sigma_1)}} I_w(m,\pi) \\
&\qquad\qquad\qquad\qquad + \frac{\mu_{n,m}}{n} \sum_{\substack{w\in \\ M^1(m,\pi,\sigma_1)}} I_{w,1}(m,\pi)\Big\} \\
&= \mu_{n,m}\Big(1 - \frac{\mu_{n,m}}{n} - \mathbb{E}\sum_{\substack{w\in \\ N^1(m,\pi,\sigma_1)}} I_w(m,\pi) + \mathbb{E}\sum_{\substack{w\in \\ M^1(m,\pi,\sigma_1)}} I_{w,1}(m,\pi)\Big).
\end{aligned}
\tag{2.52}
$$

Now consider the first sum in (2.52). Note that when $d_1(m, \pi) = k$, of the potential N edges, $n - 1$ have vertex 1 as an endpoint, and an additional $m - k$ edges remain in

$\mathcal{G}(m, \pi)$ and are not redistributed. Hence,

$$
\begin{aligned}
\mathbb{E} & \sum_{w \in N^1(m,\pi,\sigma_1)} I_w(m,\pi) \\
&= (n-1)\mathbb{P}[2 \in N^1(m,\pi,\sigma_1), d_2(m,\pi) = 0] \\
&= \mu_{n,m}\frac{n-1}{n}\mathbb{P}[2 \in N^1(m,\pi,\sigma_1)|d_2(m,\pi) = 0] \\
&= \mu_{n,m}\frac{n-1}{n}\sum_{k=0}^{n-2}\mathbb{P}[2 \in N^1(m,\pi,\sigma_1)|d_1(m,\pi) = k, d_2(m,\pi) = 0] \\
&\qquad\qquad\qquad\qquad\qquad \times \mathbb{P}[d_1(m,\pi) = k|d_2(m,\pi) = 0] \\
&= \mu_{n,m}\frac{n-1}{n}\sum_{k=0}^{n-2}\mathbb{P}\big[\mathrm{Hyp}\big(N - (n-1) - (m-k), k, n-2\big) > 0\big] \\
&\qquad\qquad\qquad\qquad\qquad \times \mathbb{P}[d_1(m,\pi) = k|d_2(m,\pi) = 0].
\end{aligned}
\tag{2.53}
$$

To arrive at the hypergeometric expression in the sum in the last equality from the conditional probability that vertex 2 is incident on any of the k redistributed edges that were removed from vertex 1 when making the new graph, note that the total number of edges available is reduced from N first by $n - 1$, as vertex 1 has been removed, and also due to the $m - k$ edges that were part of the original graph that are not changed. Of these remaining edges, $n - 2$ are incident on vertex 2, which is one fewer than their original number of $n - 1$, due to the removal of vertex 1.

Using Lemma 2.3,

$$
\begin{aligned}
\frac{k(n-2)}{N-n-m+k+1} &- \frac{k^2(n-2)^2}{2(N-n-m+k+1)^2} \\
&\leq \mathbb{P}\big[\mathrm{Hyp}\big(N - (n-1) - (m-k), k, n-2\big) > 0\big] \\
&\qquad\qquad\qquad \leq \frac{k(n-2)}{N-n-m+k+1},
\end{aligned}
\tag{2.54}
$$

from which we obtain the upper bound

$$
\begin{aligned}
\mathbb{E} & \sum_{w \in N^1(m,\pi,\sigma_1)} I_w(m,\pi) \\
&\leq \frac{\mu_{n,m}}{n}\frac{(n-1)(n-2)}{(N-n-m+2)}\sum_{k=1}^{n-1}k\mathbb{P}[d_1(m,\pi) = k|d_2(m,\pi) = 0] \\
&= \frac{\mu_{n,m}}{n}\frac{(n-1)(n-2)}{(N-n-m+2)}\mathbb{E}\{d_1(m,\pi)|d_2(m,\pi) = 0\}.
\end{aligned}
$$

Given $d_2(m,\pi) = 0$, we have $d_1(m,\pi) \sim \mathrm{Hyp}(N - (n-1), m, n-2)$, hence

$$
\mathbb{E}\{d_1(m,\pi)|d_2(m,\pi) = 0\} = \frac{m(n-2)}{N-n+1},
$$

and so,

$$
\begin{aligned}
\mathbb{E} \sum_{w \in N^1(m,\pi,\sigma_1)} I_w(m,\pi) &\leq \frac{\mu_{n,m}}{n}\frac{m(n-1)(n-2)^2}{(N-n-m+2)(N-n+1)} \\
&= \frac{\mu_{n,m}}{n}\left[\frac{4m}{n} + \frac{4m(2m+n-4)}{n(n^2-2m-3n+4)}\right] \\
&\leq \frac{\mu_{n,m}}{n}\left[\frac{4m}{n} + \frac{16m(m+n)}{n^3}\right].
\end{aligned}
\tag{2.55}
$$

Similarly, using the second moment expression from (2.29)

$$\mathbb{E}\big\{d_1(m,\pi)^2\big|d_2(m,\pi)=0\big\} = \frac{m(n-2)(N+mn-2n-3m+3)}{(N-n)(N-n+1)},$$

and so from (2.54) we obtain the lower bound

$$\mathbb{E}\sum_{w\in N^1(m,\pi,\sigma_1)} I_w(m,\pi)$$

$$\geq \mu_{n,m}\frac{n-1}{n}\sum_{k=1}^{n-2}\Big(\frac{k(n-2)}{N-n-m+k+1} - \frac{k^2(n-2)^2}{2(N-n-m+k+1)^2}\Big)$$
$$\times \mathbb{P}[d_1(m,\pi)=k|d_2(m,\pi)=0]$$

$$\geq \mu_{n,m}\frac{n-1}{n}\sum_{k=1}^{n-2}\Big(\frac{k(n-2)}{N-m-1} - \frac{k^2(n-2)^2}{2(N-n-m+2)^2}\Big)$$
$$\times \mathbb{P}[d_1(m,\pi)=k|d_2(m,\pi)=0]$$

$$= \frac{\mu_{n,m}}{n}\Big(\frac{(n-1)(n-2)}{N-m-1}\mathbb{E}\big\{d_1(m,\pi)\big|d_2(m,\pi)=0\big\}$$
$$- \frac{(n-1)(n-2)^2}{2(N-n-m+2)^2}\mathbb{E}\big\{d_1(m,\pi)^2\big|d_2(m,\pi)=0\big\}\Big)$$

$$= \frac{\mu_{n,m}}{n}\Big(\frac{(n-1)(n-2)}{N-m-1}\frac{m(n-2)}{N-n+1}$$
$$- \frac{(n-1)(n-2)^2}{2(N-n-m+2)^2}\frac{m(n-2)(N+mn-2n-3m+3)}{(N-n)(N-n+1)}\Big).$$

Now, for the first term in the brackets we have

$$\frac{m(n-1)(n-2)^2}{(N-m-1)(N-n+1)} = \frac{4m}{n} + \frac{4m(2m-n+2)}{n(n^2-2m-n-2)}$$
$$\geq \frac{4m}{n} - \frac{4mn}{n(n^2-2m-n-2)} \geq \frac{4m}{n} - \frac{8m}{n^2},$$

where we have used (2.51) for the last inequality. For the second term in the brackets,

$$\frac{m(n-1)(n-2)^3(N+mn-2n-3m+3)}{2(N-n)(N-n+1)(N-n-m+2)^2} = \frac{4m(n-2)^2(2m+n-2)}{n(n^2-2m-3n+4)^2}$$
$$\leq \frac{8mn^2(m+n)}{n(n^2-2m-3n+4)^2} \leq \frac{32m(m+n)}{n^3},$$

where again we have used (2.51) for the last inequality. Hence, together with the upper bound (2.55), we arrive at

$$\frac{\mu_{n,m}}{n}\Big[\frac{4m}{n} - \frac{40m(m+n)}{n^3}\Big]$$
$$\leq \mathbb{E}\sum_{w\in N^1(m,\pi,\sigma_1)} I_w(m,\pi) \leq \frac{\mu_{n,m}}{n}\Big[\frac{4m}{n} + \frac{16m(m+n)}{n^3}\Big]. \tag{2.56}$$

Now considering the second sum in (2.52), we can write

$$\sum_{\substack{w\in\\M^1(m,\pi,\sigma_1)}} I_{w,1}(m,\pi) = \sum_{\substack{w:\{w,1\}\\\in\mathcal{G}(m,\pi)}} I_{w,1}(m,\pi) - \sum_{\substack{w\in\\M^{c,1}(m,\pi,\sigma_1)}} I_{w,1}(m,\pi), \tag{2.57}$$

where $M^{c,1}(m, \pi, \sigma_1) = \{w : \{w, 1\} \in \mathcal{G}(m, \pi), w \in N^1(m, \pi, \sigma_1)\}$. Taking expectation of the first sum on the right hand side of (2.57) and noting that the distributions of the degrees in the graph are hypergeometric, we obtain that

$$
\mathbb{E} \sum_{\substack{w:\{w,1\} \\ \in \mathcal{G}(m,\pi)}} I_{w,1}(m, \pi)
$$

$$
= (n-1)\mathbb{P}[\{1, 2\} \in \mathcal{G}(m, \pi), d_2(m, \pi) = 1]
$$

$$
= (n-1)\mathbb{P}[d_2(m, \pi) = 1 | \{1, 2\} \in \mathcal{G}(m, \pi)] \, \mathbb{P}[\{1, 2\} \in \mathcal{G}(m, \pi)]
$$

$$
= (n-1)\mathbb{P}[\mathrm{Hyp}(N-1, m-1, n-2) = 0]\frac{m}{N}
$$

$$
= \frac{\mu_{n,m}}{n}\frac{N(n-1)}{N-m-n+2}\frac{m}{N}
$$

$$
= \frac{\mu_{n,m}}{n}\left[\frac{2m}{n} + \frac{4m(m+n-2)}{n(n^2 - 2m - 3n + 4)}\right].
$$

From this equality and using the assertions on m and n, we obtain

$$
\frac{\mu_{n,m}}{n}\left[\frac{2m}{n} + \frac{2m(m+n)}{n^3}\right] \ \leq \ \mathbb{E} \sum_{\substack{w:\{w,1\} \\ \in \mathcal{G}(m,\pi)}} I_{w,1}(m, \pi) \ \leq \ \frac{\mu_{n,m}}{n}\left[\frac{2m}{n} + \frac{8m(m+n)}{n^3}\right]. \quad (2.58)
$$

Now taking expectation of the second sum of (2.57),

$$
\mathbb{E} \sum_{\substack{w\in \\ M^{c,1}(m,\pi,\sigma_1)}} I_{w,1}(m, \pi)
$$

$$
= (n-1)\mathbb{P}[\{1, 2\} \in \mathcal{G}(m, \pi), 2 \in N^1(m, \pi, \sigma_1), d_2(m, \pi) = 1]
$$

$$
= (n-1)\sum_{k=1}^{n-1}\mathbb{P}[\{1, 2\} \in \mathcal{G}(m, \pi), 2 \in N^1(m, \pi, \sigma_1), d_2(m, \pi) = 1, d_1(m, \pi) = k]
$$

$$
= \frac{(n-1)m}{N}\sum_{k=1}^{n-1}\mathbb{P}[2 \in N^1(m, \pi, \sigma_1) | d_2(m, \pi) = 1, d_1(m, \pi) = k, \{1, 2\} \in \mathcal{G}(m, \pi)]
$$

$$
\times \mathbb{P}[d_1(m, \pi) = k | d_2(m, \pi) = 1, \{1, 2\} \in \mathcal{G}(m, \pi)]]
$$

$$
\times \mathbb{P}[d_2(m, \pi) = 1 | \{1, 2\} \in \mathcal{G}(m, \pi)]]
$$

$$
= \frac{(n-1)m}{N}\sum_{k=1}^{n-1}\mathbb{P}[\mathrm{Hyp}(N-(n-1)-(m-k), k, n-2) > 0]
$$

$$
\times \mathbb{P}[\mathrm{Hyp}(N-(n-1), m-1, n-2) = k-1]
$$

$$
\times \mathbb{P}[\mathrm{Hyp}(N-1, m-1, n-2) = 0].
$$

We arrive at the first Hypergeomtric expression in the sum in the last equality by the same reasoning as that given following (2.53); the remaining two expressions in the sum follow by similar, and simpler, means.

Now, for the first and last terms, using Lemma 2.3 for the upper bound, we have

$$
\mathbb{P}[\mathrm{Hyp}(N-(n-1)-(m-k), k, n-2) > 0] \leq \frac{k(n-2)}{N-n-m+k+1}
$$

$$
\mathbb{P}[\mathrm{Hyp}(N-1, m-1, n-2) = 0] = \frac{\mu_{n,m}}{n}\frac{N}{N-m-n+2},
$$

and thus, using in the final inequality that $n^2 \leq 4(N-m-n+2)$, which holds via the assumption that $m \leq n^2/4 - 3n/2$, and that $n^2 \leq 4(N-n)$, which holds as $n \geq 6$, true by

assumption, we obtain

$$
\mathbb{E} \sum_{\substack{w: \\ M^{c,1}(m,\pi,\sigma_1)}} I_{w,1}(m,\pi)
$$

$$
\leq \frac{\mu_{n,m}}{n} \frac{(n-1)^2 m}{(N-m-n+2)^2}
$$

$$
\times \sum_{k=0}^{n-2} (k+1)\mathbb{P}[\mathrm{Hyp}(N-(n-1),m-1,n-2)=k] \tag{2.59}
$$

$$
= \frac{\mu_{n,m}(n-1)^2 m}{n(N-m-n+2)^2} \left(\frac{(m-1)(n-2)}{(N-(n-1))} + 1 \right) \leq \frac{\mu_{n,m}}{n} \frac{16m}{n^2} \left(\frac{4m}{n} + 1 \right)
$$

$$
\leq \frac{\mu_{n,m}}{n} \frac{16m}{n^2} \left(\frac{4m}{n} + \frac{4n}{n} \right) \leq \frac{\mu_{n,m}}{n} \frac{64(m+n)}{n^3}.
$$

Using the estimates from (2.58) and (2.59) in the difference (2.57), and then applying that result and (2.56) in (2.52) yields the claim. $\qquad\square$

Lemma 2.7. *There exist universal integers m_0 and n_0, and positive constants C_0 and c_0 such that, whenever*

$$
n \geq n_0 \quad \text{and} \quad m_0 \leq m \leq c_0 n^{3/2}, \tag{2.60}
$$

we have

$$
\left| \frac{\mu_{n,m}}{ne^{-2m/n}} - 1 \right| \leq C_0 \left(\frac{m}{n^2} + \frac{m^2}{n^3} \right) \tag{2.61}
$$

and

$$
\left| \frac{\sigma_{n,m}^2}{n\varphi(2m/n)} - 1 \right| \leq C_0 \left(\frac{1}{m} + \frac{m^2}{n^3} \right), \tag{2.62}
$$

where

$$
\varphi(x) = e^{-x}(1 - e^{-x}(1+x)).
$$

Proof. It is easy to verify that

$$
n^{3/2} \leq \frac{n^2}{4} - \frac{3n}{2} \quad \text{for all } n \geq 27. \tag{2.63}
$$

Hence, with the first inequality in (2.60) holding with n_0 replaced by 27, and taking $c_0 \leq 1$, Lemma 2.6 can be invoked to yield

$$
\left| \frac{\mu_{n,m}}{ne^{-2m/n}} - 1 \right| \leq 1 - \exp\left(-\frac{8m(m+n)}{n^3} \right) \leq \frac{8m(m+n)}{n^3},
$$

from which (2.61) now follows for any $C_0 \geq 8$.

Turning to (2.62), we first show that the lower bound in (2.50) is positive whenever $n \geq 78$ and $m \geq 78$. Indeed, that lower bound is positive whenever

$$
\frac{\mu_{n,m}}{n} \left(1 + \frac{2m}{n} + \frac{78m(m+n)}{n^3} \right) < 1,
$$

which, recalling the upper bound (2.49), is implied whenever

$$
e^{-x}(1 + x + y) < 1 \tag{2.64}
$$

with

$$
x = \frac{2m}{n} \quad \text{and} \quad y = \frac{78m(m+n)}{n^3}. \tag{2.65}
$$

Since (2.64) is equivalent to the inequality $y < e^x - x - 1$, which in turn is satisfied if $y \le x^2/2$, since $x^2/2 < e^x - x - 1$, we arrive at the sufficient condition

$$\frac{78m(m+n)}{n^3} \le \frac{2m^2}{n^2},$$

which is equivalent to $39 \le m(2 - 39/n)$. This inequality holds whenever both $n \ge 78$ and $m \ge 78$.

We now proceed to bound the ratio between the upper and lower bounds, say $\overline{\sigma}_{n,m}^2$ and $\underline{\sigma}_{n,m}^2$, respectively, of (2.50). Using the identity $(1-a)/(1-b) = 1 + (b-a)/(1-b)$, we have

$$\frac{\overline{\sigma}_{n,m}^2}{\underline{\sigma}_{n,m}^2} = \frac{1 - \frac{\mu_{n,m}}{n}\left(1 + \frac{2m}{n} - \frac{48m(m+n)}{n^3}\right)}{1 - \frac{\mu_{n,m}}{n}\left(1 + \frac{2m}{n} + \frac{78m(m+n)}{n^3}\right)} = 1 + \frac{\frac{\mu_{n,m}}{n}\frac{126m(m+n)}{n^3}}{1 - \frac{\mu_{n,m}}{n}\left(1 + \frac{2m}{n} + \frac{78m(m+n)}{n^3}\right)}. \quad (2.66)$$

We proceed to lower bound the denominator in (2.66). Letting x and y be as in (2.65), and applying the upper bound in (2.49), we may write

$$1 - \frac{\mu_{n,m}}{n}\left(1 + \frac{2m}{n} + \frac{78m(m+n)}{n^3}\right)$$
$$\ge 1 - e^{-2m/n}\left(1 + \frac{2m}{n} + \frac{78m(m+n)}{n^3}\right)$$
$$= 1 - e^{-x}(1 + x + y) \ge 1 - e^{-x}(1 + x) - y.$$

If $2m/n \le 1$, we have $0 \le x \le 1$ and thus $1 - e^{-x}(1+x) \ge x^2/4$ from Lemma 2.4, so that

$$1 - e^{-x}(1+x) - y \ge \frac{x^2}{4} - y = \frac{m^2}{n^2}\left(1 - \frac{78}{n} - \frac{78}{m}\right) \ge \frac{1}{8}\left(\frac{2m}{n}\right)^2$$

when $\min(n,m) \ge 312$. If $2m/n > 1$ and so $x > 1$, we simply use the lower bound

$$1 - e^{-x}(1+x) \ge \frac{1}{4},$$

and for any positive c_0 we can take n_0 large enough so that

$$1 - e^{-x}(1+x) - y \ge \frac{1}{4} - y = \frac{1}{4} - \frac{78m^2}{n^3} - \frac{78m}{n^2} \ge \frac{1}{8}.$$

Hence, writing $O(\cdot)$ with the understanding that the implied bound holds with universal constants, recalling (2.66), and using Lemma 2.6 to bound $\mu_{n,m}/n$ in its numerator, we have

$$\frac{\overline{\sigma}_{n,m}^2}{\underline{\sigma}_{n,m}^2} \le 1 + \frac{8e^{-2m/n}}{\left(\frac{2m}{n} \wedge 1\right)^2}\left(\frac{126m^2}{n^3} + \frac{126m}{n^2}\right) = \begin{cases} 1 + O\left(\dfrac{1}{m}\right) & \text{if } 2m/n \le 1 \\[2mm] 1 + O\left(\dfrac{m^2}{n^3}\right) & \text{if } 2m/n > 1 \,, \end{cases} \quad (2.67)$$

where both the $O(\cdot)$ terms are non-negative.

Next, with $\varphi(x) = e^{-x}(1 - e^{-x}(1+x))$, we show that

$$\frac{\overline{\sigma}_{n,m}^2}{n\varphi(2m/n)} = \begin{cases} 1 + O\left(\dfrac{1}{m}\right) & \text{if } 2m/n \le 1 \\[2mm] 1 + O\left(\dfrac{m^2}{n^3}\right) & \text{if } 2m/n > 1. \end{cases} \quad (2.68)$$

Using (2.61) for the second equality, (2.49) for the third, then (2.61) again and the lower bound of Lemma 2.4 for the fourth, we obtain

$$\frac{\overline{\sigma}_{n,m}^2}{n\varphi(2m/n)}$$

$$= \frac{\mu_{n,m}}{ne^{-2m/n}} \times \left[\frac{1 - \frac{\mu_{n,m}}{n}\left(1 + \frac{2m}{n} + O\left(\frac{m}{n^2} + \frac{m^2}{n^3}\right)\right)}{1 - e^{-2m/n}\left(1 + \frac{2m}{n}\right)} \right]$$

$$= \left(1 + O\left(\frac{m}{n^2} + \frac{m^2}{n^3}\right)\right)$$

$$\times \left[1 - \frac{\frac{\mu_{n,m}}{n}\left(1 + \frac{2m}{n} + O\left(\frac{m}{n^2} + \frac{m^2}{n^3}\right)\right) - e^{-2m/n}\left(1 + \frac{2m}{n}\right)}{1 - e^{-2m/n}\left(1 + \frac{2m}{n}\right)} \right]$$

$$= \left(1 + O\left(\frac{m}{n^2} + \frac{m^2}{n^3}\right)\right)$$

$$\times \left[1 - \frac{\left(\frac{\mu_{n,m}}{n} - e^{-2m/n}\right)\left(1 + \frac{2m}{n}\right)}{1 - e^{-2m/n}\left(1 + \frac{2m}{n}\right)} + O\left(\frac{e^{-2m/n}\left(\frac{m}{n^2} + \frac{m^2}{n^3}\right)}{1 - e^{-2m/n}\left(1 + \frac{2m}{n}\right)}\right) \right]$$

$$= \left(1 + O\left(\frac{m}{n^2} + \frac{m^2}{n^3}\right)\right)$$

$$\times \left[1 + O\left(\frac{e^{-2m/n}\left(\frac{m}{n^2} + \frac{m^2}{n^3}\right)\left(1 + \frac{2m}{n}\right)}{\left(\frac{2m}{n} \wedge 1\right)^2}\right) + O\left(\frac{e^{-2m/n}\left(\frac{m}{n^2} + \frac{m^2}{n^3}\right)}{\left(\frac{2m}{n} \wedge 1\right)^2}\right) \right]$$

$$=: (1 + R_1)(1 + R_2 + R_3) = O((1 + R_1)(1 + R_2)),$$

as $R_3 = O(R_2)$. In the case $2m/n \leq 1$, we have

$$R_1 = O\left(\frac{m}{n^2} + \frac{m^2}{n^3}\right) = O\left(\frac{1}{m}\left(\frac{m^2}{n^2} + \frac{m^3}{n^3}\right)\right) = O\left(\frac{1}{m}\right) \quad \text{and}$$

$$R_2 = O\left(\frac{m/n^2 + m^2/n^3}{(2m/n)^2}\right) = O\left(\frac{1}{m} + \frac{1}{n}\right) = O\left(\frac{1}{m}\right),$$

showing the first bound in (2.68). In the case $2m/n > 1$,

$$R_1 = O\left(\frac{m}{n^2} + \frac{m^2}{n^3}\right) = O\left(\frac{m^2}{n^3}\right),$$

and using that $x\exp(-x)$ is bounded over $[0, \infty)$,

$$R_2 = O\left(e^{-2m/n}\left(\frac{m}{n^2} + \frac{m^2}{n^3}\right)\left(1 + \frac{2m}{n}\right)\right) = O\left(\left(\frac{m^2}{n^3}\right)e^{-2m/n}\left(\frac{2m}{n}\right)\right) = O\left(\frac{m^2}{n^3}\right).$$

Applying (2.63), the second bound in (2.68) is shown. Now, using that $\overline{\sigma}_{n,m}^2 \geq \underline{\sigma}_{n,m}^2$, and writing

$$\frac{\overline{\sigma}_{n,m}^2}{\underline{\sigma}_{n,m}^2} = 1 + a, \qquad \frac{\overline{\sigma}_{n,m}^2}{n\varphi(2m/n)} = 1 + b$$

and observing that, because the implicit constants in the bounds (2.67) and (2.68) are universal, and using that the $O(\cdot)$ terms in (2.67) are non-negative, we can choose c_0

small enough and m_0 large enough to guarantee that $0 \le a < 1$ and $-1 < b < 1$, and hence obtain the upper and lower bounds

$$(1-a)(1+b) \le \left(1 - \frac{a}{1+a}\right)(1+b) = \frac{\sigma_{n,m}^2}{\overline{\sigma}_{n,m}^2} \frac{\overline{\sigma}_{n,m}^2}{n\varphi(2m/n)} = \frac{\sigma_{n,m}^2}{n\varphi(2m/n)}$$

$$\le \frac{\sigma_{n,m}^2}{n\varphi(2m/n)} \le \frac{\overline{\sigma}_{n,m}^2}{n\varphi(2m/n)} = 1 + b,$$

from which the estimate (2.62) follows.

$\square$

Lemma 2.8. *Let $r_{n,m}$ be defined as in (1.21). For any integers $\overline{n}$ and $\overline{m}$ and any positive constant $\overline{c} > 0$, there exists $\overline{r} \ge 1$ such that $r_{n,m} > \overline{r}$ implies*

$$n \ge \overline{n} \quad \text{and} \quad \overline{m} \le m \le \overline{c}n^{3/2}. \tag{2.69}$$

Proof. We will show that $r_{n,m} \le \overline{r}$ for $\overline{r} = \max\{\overline{n}^{1/2}, (2\overline{m})^{3/2}, 1/\overline{c}^2, 1\}$ if (2.69) is violated. Indeed, if $n < \overline{n}$, we have by Lemma 2.5, and that $\mu_{n,m} \le n$, then

$$r_{n,m} = \frac{\sigma_{n,m}^3}{\mu_{n,m}(1 + (m/n)^2)} \le \sigma_{n,m} \le \min\{\overline{n}^{1/2}, (2\overline{m})^{1/2}\}.$$

Finally, if $m > \overline{c}n^{3/2}$, then similarly

$$r_{n,m} = \frac{\sigma_{n,m}^3}{\mu_{n,m}(1 + (m/n)^2)} \le \frac{\sqrt{\mu_{n,m}}}{1 + (m/n)^2}$$

$$\le \frac{\sqrt{n}}{1 + (\overline{c}n^{3/2}/n)^2} = \frac{\sqrt{n}}{1 + (\overline{c}\sqrt{n})^2} \le \frac{1}{\overline{c}^2}. \qquad \square$$

Lemma 2.9. *Letting ☺ be as in Condition (G1), it holds that*

$$\sup_{\substack{(n,m)\in☺ \\ 0 \le d \le \min\{n,m\}/4}} \left(\frac{\mu_{n,m}^2}{\mu_{n-1,m-d}^2} \vee \frac{\mu_{n-1,m-d}^2}{\mu_{n,m}^2}\right) < \infty,$$

and

$$\sup_{\substack{(n,m)\in☺ \\ 0 \le d \le \min\{n,m\}/4}} \left(\frac{\sigma_{n,m}^2}{\sigma_{n-1,m-d}^2} \vee \frac{\sigma_{n-1,m-d}^2}{\sigma_{n,m}^2}\right) < \infty. \tag{2.70}$$

Proof. First, note that if $(n,m) \in ☺$, then from (2.5) and (2.6) the conclusion of Lemma 2.7 holds. For the ratio of means, from Lemma 2.6, to upper bound $\mu_{n,m}/\mu_{n-1,m-d}$ it suffices to upper bound the ratio

$$-\frac{2m}{n} + \frac{2(m-d)}{n-1} + \frac{8(m-d)(m-d+n-1)}{(n-1)^3}$$

$$= \frac{2m - 2nd}{n(n-1)} + \frac{8(m-d)(m-d+n-1)}{(n-1)^3} \le 8\left(\frac{2m}{n^2} + \frac{8m(m+n)}{n^3}\right),$$

which is bounded by a constant via $m \le c_0 n^{3/2}$, as in (2.60). Similarly, to upper bound $\mu_{n-1,m-d}/\mu_{n,m}$ it suffices to upper bound the ratio

$$-\frac{2(m-d)}{n-1} + \frac{2m}{n} + \frac{8m(m+n)}{n^3}$$

$$= \frac{2nd - 2m}{n(n-1)} + \frac{8m(m+n)}{n^3} \le 2\left(\frac{2d}{n} + \frac{8m(m+n)}{n^3}\right),$$

which, here using that $d \leq n/4$, we see is also so bounded.

For the ratios of variances, for $0 \leq d \leq \min\{n, m\}/4$ let

$$x = \frac{2m}{n}, \qquad y = \frac{2m}{n} - \frac{2(m - d)}{n - 1} = \frac{2(nd - m)}{n(n - 1)},$$

let $\varphi(x) = e^{-x}(1 - e^{-x}(1 + x))$, and write

$$\frac{\sigma^2_{n,m}}{\sigma^2_{n-1,m-d}} = \frac{\sigma^2_{n,m}}{n\varphi(x)} \times \frac{(n-1)\varphi(x-y)}{\sigma^2_{n-1,m-d}} \times \frac{n}{n-1} \times \frac{\varphi(x)}{\varphi(x-y)} =: R_1 \times R_2 \times R_3 \times R_4.$$

We show that these four terms, and their reciprocals, can be uniformly bounded over the range of the supremum in (2.70). Since (2.60) holds for n, and m, we can apply Lemma 2.7, and also (2.5) for the first and final bounds, and obtain

$$\frac{1}{2} \leq 1 - C_0\Big(\frac{1}{m} + \frac{m^2}{n^3}\Big) \leq \frac{\sigma^2_{n,m}}{n\varphi(x)} \leq 1 + C_0\Big(\frac{1}{m} + \frac{m^2}{n^3}\Big) \leq \frac{3}{2}. \tag{2.71}$$

Next, since $m \geq 2m_0$ by (2.5) and $d \leq m/4$, we have that $m - d \geq 3m/4 \geq 3m_0/2 \geq m_0$. Since $n \geq 2n_0$, again by (2.5), we have that $n - 1 \geq n_0$, and since $m \leq (c_0/2)n^{3/2}$ by (2.5) and $(n/(n-1))^{3/2} \leq 2$ for $n \geq 3$, we have that $m - d \leq m \leq c_0(n-1)^{3/2}$. It follows that $(m-d, n-1)$ also satisfies the hypotheses of Lemma 2.7. Using the lower bound on $\overline{n}$ from (2.5), we have $1/(n-1)^3 \leq 2/n^3$, and also from (2.6) that $C_0(2/m + 2m^2/n^3) \leq 1/2$, so also using $d \leq m/4$ for the second and second to last inequality,

$$\frac{1}{2} \leq 1 - C_0\Big(\frac{2}{m} + \frac{2m^2}{n^3}\Big) \leq 1 - C_0\Big(\frac{1}{m-d} + \frac{(m-d)^2}{(n-1)^3}\Big)$$

$$\leq \frac{\sigma^2_{n-1,m-d}}{(n-1)\varphi(x-y)} \tag{2.72}$$

$$\leq 1 + C_0\Big(\frac{1}{m-d} + \frac{(m-d)^2}{(n-1)^3}\Big) \leq 1 + C_0\Big(\frac{2}{m} + \frac{2m^2}{n^3}\Big) \leq \frac{3}{2}.$$

Hence, (2.71) and (2.72) imply that

$$\frac{1}{2} \leq R_1 \leq \frac{3}{2} \quad \text{and} \quad \frac{2}{3} \leq R_2 \leq 2.$$

Clearly, $1 \leq R_3 \leq 2$ for $n \geq 2$. Lastly,

$$R_4 = \frac{e^{-x}(1 - e^{-x}(1 + x))}{e^{-x+y}(1 - e^{-x+y}(1 + x - y))} = \frac{1 - e^{-x}(1 + x)}{e^y(1 - e^{-x+y}(1 + x - y))}.$$

Note that by (2.6), and by (2.5) that gives that $\overline{c} \leq 1$, and also using $d \leq n/4$,

$$-\frac{4}{n_0^{1/2}} \leq -\frac{4}{n^{1/2}} \leq -\frac{2m}{n(n-1)} \leq y = \frac{2(nd - m)}{n(n-1)} \leq \frac{2d}{n-1} \leq 1. \tag{2.73}$$

It follows that $1/e^y$ remains bounded on ☺, and therefore, to show R_4 is bounded it suffices to show that

$$\frac{1 - e^{-x}(1 + x)}{1 - e^{-x+y}(1 + x - y)}$$

remains bounded. Using Lemma 2.4,

$$\frac{1 - e^{-x}(1 + x)}{1 - e^{-x+y}(1 + x - y)} \leq \frac{2\min\{x^2, 2\}}{\min\{(x - y)^2, 1\}} = \frac{2\min\{x^2, 2\}}{\min\{x^2(1 - y/x)^2, 1\}}.$$

But this ratio remains bounded from above, away from 1, as $d \leq m/4$ implies

$$\frac{y}{x} = \frac{nd - m}{m(n - 1)} \leq \frac{2d}{m} - \frac{1}{n - 1} \leq \frac{1}{2}.$$

The reciprocal $1/R_4$ is bounded similarly, using that (2.73) shows that e^y is bounded. □

3   Jack measure on tableaux

We now turn to the study of the distribution of the standardized sum of the α-contents over all boxes in a tableaux whose shape is determined by the partition Λ_n of n, that is, to

$$W = \frac{Y}{\sqrt{\alpha \binom{n}{2}}}, \quad \text{where} \quad Y = \sum_{x \in \Lambda_n} c_\alpha(x), \tag{3.1}$$

where

$$c_\alpha(x) = \alpha(\text{column number of } x - 1) - (\text{row number of } x - 1),$$

and where the partition Λ_n is sampled from the Jack$_\alpha$ measure in (1.22), as described in detail in the introduction; see (1.23) for an illustration of $c_\alpha(x)$, where $x \in \Lambda_7$.

Our bound is based on the zero bias construction in Fulman and Goldstein (2011), which itself depends on an exchangeable pair constructed using Kerov's growth process, a sequential procedure for growing a random partition distributed according to Jack$_\alpha$ measure.

The state of Kerov's growth process at times $n = 1, 2, \ldots$ is a partition of n, starting at time 1 with the unique partition (1) of 1. To describe its transition rule from time $n-1$ to n for $n \geq 2$, given a box x in the diagram of a partition Λ_n of n, let $a(x)$ denote the number of boxes in the same row of x and to the right of x (the "arm" of x), and let $l(x)$ denote the number of boxes in the same column of x and below x (the "leg" of x), as in (1.22). Now set

$$c_\Lambda(\alpha) = \prod_{x \in \Lambda} (\alpha a(x) + l(x) + 1), \quad c'_\Lambda(\alpha) = \prod_{x \in \Lambda} (\alpha a(x) + l(x) + \alpha)$$

and, for Λ_{n-1} a partition of $n-1$ obtained from Λ_n by removing a single corner box, let

$$\psi'_{\Lambda_n/\Lambda_{n-1}}(\alpha)$$
$$= \prod_{x \in C_{\Lambda_n/\Lambda_{n-1}} - R_{\Lambda_n/\Lambda_{n-1}}} \frac{(\alpha a_\lambda(x) + l_\lambda(x) + 1)}{(\alpha a_\lambda(x) + l_\lambda(x) + \alpha)} \frac{(\alpha a_{\Lambda_{n-1}}(x) + l_{\Lambda_{n-1}}(x) + \alpha)}{(\alpha a_{\Lambda_{n-1}}(x) + l_{\Lambda_{n-1}}(x) + 1)},$$

where $C_{\Lambda_n/\Lambda_{n-1}}$ is the union of columns of Λ_n that intersect $\Lambda_n - \Lambda_{n-1}$ and $R_{\Lambda_n/\Lambda_{n-1}}$ is the union of rows of Λ_n that intersect $\Lambda_n - \Lambda_{n-1}$. If at stage $n-1$ the state of the process is the partition Λ_{n-1}, a transition to the partition Λ_n occurs with probability

$$\frac{c_{\Lambda_{n-1}}(\alpha)}{c_{\Lambda_n}(\alpha)} \psi'_{\Lambda_n/\Lambda_{n-1}}(\alpha).$$

It is shown in Kerov (1994), see also Fulman (2006), that if Λ_{n-1} is distributed according to Jack$_\alpha$ measure on partitions of $n-1$, then the partition Λ_n obtained by this process at time n has the Jack$_\alpha$ distribution.

In the proof of Theorem 3.1 of Fulman and Goldstein (2011), a variable having the zero bias distribution of W was constructed as follows. Fix n and α and let Λ_k be the state of Kerov's growth process at time k, and set

$$V = \sum_{x \in \Lambda_{n-1}} c_\alpha(x). \tag{3.2}$$

Denoting by $c_\alpha(x_n)$ the content of the box x_n added at time n to form Λ_n, we can now write

$$W = \frac{V}{\sqrt{\alpha \binom{n}{2}}} + T, \quad \text{where} \quad T = \frac{c_\alpha(x_n)}{\sqrt{\alpha \binom{n}{2}}}. \tag{3.3}$$

With $dF(t|\Lambda_{n-1})$ the conditional distribution of T given Λ_{n-1}, constructing the pair

$$(T^\dagger, T^\ddagger) \sim (t'' - t')^2 dF(t'|\Lambda_{n-1}) dF(t''|\Lambda_{n-1}) \tag{3.4}$$

on the same space as Λ_{n-1}, and letting $U \sim \mathcal{U}[0,1]$ be independent of $V, T^\dagger$ and $T^\ddagger$, the variable

$$W^* = \frac{V}{\sqrt{\alpha\binom{n}{2}}} + T^* \quad \text{with } T^* = UT^\dagger + (1-U)T^\ddagger \tag{3.5}$$

has the W-zero bias distribution. In fact, the joint distribution on the right hand side of (3.4) can be achieved by running Kerov's growth process twice, conditionally independent on Λ_{n-1}. As shown in Fulman and Goldstein (2011), the resulting variables, say T' and T'', yield the crucial exchangeable Stein pair in (1.3) via (3.3). Again by Fulman and Goldstein (2011), both the conditional mean and variance of T given Λ_{n-1} do not depend on Λ_{n-1}; specifically,

$$\mathbb{E}\{T|\Lambda_{n-1}\} = 0 \quad \text{and} \quad \mathbb{E}\{T^2|\Lambda_{n-1}\} = \frac{2}{n}. \tag{3.6}$$

It is essentially for this reason that we may construct W^* as in (3.5), using V; for details, see Fulman and Goldstein (2011).

Proof of Theorem 1.5. We verify the conditions of Theorem 1.2.

Condition (Z1). Fix an $\varepsilon \in (0,1)$, suppressed in the notation, and let

$$\Theta = \{(n,\alpha) : \alpha > n^{1+\varepsilon}, n \geq 2\} \quad \text{and} \quad \text{☺} = \{(n,\alpha) \in \Theta : r_{n,\alpha} > 2^{1/2-\varepsilon/2}\}, \tag{3.7}$$

where

$$r_{n,\alpha} = \frac{n}{\sqrt{\alpha}}, \tag{3.8}$$

which is positive and measurable. Note that

$$(n,\alpha) \in \text{☺} \quad \Longleftrightarrow \quad n^{1+\varepsilon} < \alpha < \frac{n^2}{2^{1-\varepsilon}}, \tag{3.9}$$

which implies in particular that $n \geq 3$ if $(n,\alpha) \in \text{☺}$.

From Fulman (2004), the mean and variance of the content Y of a tableaux of a partition of n under Jack_α measure is given, respectively, by

$$\mu_{n,\alpha} = 0 \quad \text{and} \quad \sigma_{n,\alpha}^2 = \alpha\binom{n}{2} \quad \text{for all } (n,\alpha) \in \Theta. \tag{3.10}$$

In particular we have that $\text{Var}_{n,\alpha} Y > 0$ for all $(n,\alpha) \in \text{☺}$.

Condition (Z2). The variable Y, given in (3.1) is easily seen to satisfy the needed conditions, and the construction of the zero bias variable W^* is outlined above in (3.3), (3.4) and (3.5).

Condition (Z3). From (3.3) and (3.5) we see that

$$D = T^* - T.$$

For each $(n,\alpha) = \text{☺}$ let $\mathcal{F}_{n,\alpha}$ be the trivial σ-algebra $\{\emptyset, \Omega\}$, let

$$\overline{D} = \frac{10\sqrt{\alpha}}{n\varepsilon} \quad \text{and let} \quad F_{n,\alpha,1} = \left\{\lambda_1 \leq \frac{2}{\varepsilon}\right\}, \tag{3.11}$$

where λ_1 and λ_1' respectively denote the length of the first row and first column of the tableaux Λ_{n-1} produced by Kerov's growth process at time $n-1$. Clearly $\overline{D}$ is $\mathcal{F}_{n,\alpha}$ measurable.

We next argue that $|D| \leq \overline{D}$ on $F_{n,\alpha,1}$ as follows. With $c_\alpha(x_n), c_\alpha(x_n')$ and $c_\alpha(x_n'')$ the contents of the boxes added to Λ_{n-1} by Kerov's growth process, all conditionally independent given Λ_{n-1}, with probability one,

$$\{c_\alpha(x_n), c_\alpha(x_n'), c_\alpha(x_n'')\} \subset [-(\lambda_1' + 1), \alpha(\lambda_1 + 1)],$$

as the extreme values $\alpha(\lambda_1 + 1)$ and $-(\lambda_1' + 1)$ are achieved, respectively, by adding a box at the end of first row, and at bottom of the first column. Scaling by $\sigma_{n,\alpha}$ in (3.10) to obtain T, T' and T'', respectively, with probability one

$$\{T, T', T''\} \subset [-(\lambda_1' + 1)/\sigma_{n,\alpha}, \alpha(\lambda_1 + 1)/\sigma_{n,\alpha}]. \tag{3.12}$$

Now note that by (3.4) the distribution of $(T^\dagger, T^\ddagger)$ is absolutely continuous with respect to that of (T', T''), and hence with probability one

$$\{T, T^\dagger, T^\ddagger\} \subset [-(\lambda_1' + 1)/\sigma_{n,\alpha}, \alpha(\lambda_1 + 1)/\sigma_{n,\alpha}].$$

As T^* is the convex combination $UT^\dagger + (1-U)T^\ddagger$ of $T^\dagger, T^\ddagger$, it too must lie in this same interval, and hence, as the length of the first column of Λ_{n-1} can be no more than n, we obtain

$$\begin{aligned}
|D| = |T^* - T| &\leq \frac{\alpha\lambda_1 + \lambda_1' + \alpha + 1}{\sigma_{n,\alpha}} \\
&\leq \frac{2\alpha/\varepsilon + n + \alpha + 1}{\sigma_{n,\alpha}} \leq \frac{5\alpha}{\varepsilon\sigma_{n,\alpha}} \leq \overline{D} \quad \text{on } F_{n,\alpha,1} \text{ for all } (n,\alpha) \in \text{☺}.
\end{aligned} \tag{3.13}$$

In what follows, we think of $(n,\alpha) \in \text{☺}$ as fixed and suppress the subscript in $\mathbb{E}_{n,\alpha}$. Turning to the moment conditions, we claim that

$$\sqrt{\mathbb{E}D^2} \leq C\left(\frac{1}{\sqrt{n}} + \frac{\sqrt{\alpha}}{n}\right) \leq C\frac{\sqrt{\alpha}}{n}. \tag{3.14}$$

Now,

$$\begin{aligned}
\sqrt{\mathbb{E}D^2} &= \sqrt{\mathbb{E}(T^* - T)^2} \\
&\leq \sqrt{2(\mathbb{E}\{(T^*)^2\} + \mathbb{E}T^2)} \leq \sqrt{2}\left(\sqrt{\mathbb{E}\{(T^*)^2\}} + \sqrt{\mathbb{E}T^2}\right).
\end{aligned} \tag{3.15}$$

To bound the second moment of T^*, by the zero bias formula (1.4) with $f(x) = x^3/3$, and the proof of Theorem 4.1 in Fulman and Goldstein (2011), we obtain

$$3\operatorname{Var}(T)\mathbb{E}\{(T^*)^2\} = \mathbb{E}T^4 \leq \frac{8}{n^2} + \frac{4\alpha}{n^2(n-1)}.$$

Hence, by (3.6),

$$\sqrt{\mathbb{E}\{(T^*)^2\}} = \sqrt{\frac{n}{6}\left(\frac{8}{n^2} + \frac{4\alpha}{n^2(n-1)}\right)} \leq C\sqrt{\frac{1}{n} + \frac{\alpha}{n^2}} \leq C\left(\frac{1}{\sqrt{n}} + \frac{\sqrt{\alpha}}{n}\right).$$

For the second term of (3.15), by (3.6), we obtain $\sqrt{\mathbb{E}T^2} = \sqrt{2/n}$, thus showing first inequality in (3.14). The final inequality in (3.14) holds as $(n,\alpha) \in \text{☺}$ implies $\alpha \geq n$.

To verify the first condition in (1.16), apply the Cauchy Schwarz inequality, (3.8) and (3.14) to obtain

$$r_{n,\alpha}^2 \, \mathbb{E}\{|D|(1 - I_{F_{n,\alpha,1}})\} \le r_{n,\alpha}^2 \sqrt{\mathbb{E}D^2 \mathbb{P}[F_{n,\alpha,1}^c]} \le \frac{Cn}{\sqrt{\alpha}} \sqrt{\mathbb{P}[F_{n,\alpha,1}^c]}. \tag{3.16}$$

To control $\mathbb{P}[F_{n,\alpha,1}^c]$, with $m = n - 1$, we apply the inequality

$$\mathbb{P}[\lambda_1 = l] \le \left(\frac{m}{\alpha}\right)^l \frac{\alpha l}{l!^2}$$

from the proof of Lemma 6.6 in Fulman (2004). Using that $\alpha \ge n^{1+\varepsilon} \ge m^{1+\varepsilon}$ in the third inequality below we obtain

$$\mathbb{P}[F_{n,\alpha,1}^c] \le \mathbb{P}[\lambda_1 \ge 2/\varepsilon] \le \sum_{l \ge 2/\varepsilon} \left(\frac{m}{\alpha}\right)^l \frac{\alpha l}{l!^2} = \frac{\alpha}{m^2} \sum_{l \ge 2/\varepsilon} \frac{m^{l+2}}{\alpha^l} \frac{l}{l!^2}$$

$$\le \frac{\alpha}{m^2} \sum_{l \ge 2/\varepsilon} m^{2-l\varepsilon} \frac{l}{l!^2} \le \frac{\alpha}{m^2} \sum_{l \ge 2/\varepsilon} \frac{l}{l!^2} \le \frac{\alpha}{m^2} \sum_{l \ge 0} \frac{l}{l!^2} \le \frac{e\alpha}{m^2} \le \frac{4e\alpha}{n^2}.$$

Substitution into (3.16) now verifies the first condition in (1.16).

For the second condition in (1.16), using (3.8), the Cauchy Schwarz inequality, that $\mathbb{E}W^2 = 1$, (3.14) and (3.11) we obtain

$$\sup_{(n,\alpha)\in\text{☺}} r_{n,\alpha} \, \mathbb{E}\{|DW| + \overline{D}\} \le \sup_{n,\alpha\in\text{☺}} \frac{n}{\sqrt{\alpha}} \left(\sqrt{\mathbb{E}D^2} + \mathbb{E}|\overline{D}|\right) < \infty.$$

Condition (Z4). For $(n, \alpha) \in \text{☺}$, let

$$\Psi(n, \alpha) = (n - 1, \alpha), \tag{3.17}$$

which is $\mathcal{F}_{n,\alpha}$ measurable, let $F_{n,\alpha,2} = \Omega$, and let V be as in (3.2). The conditional distribution condition (1.17) is satisfied for V with $\theta = (n, \alpha)$ by the properties of Kerov's growth process. Clearly the set $F_{n,\alpha,2}$ is measurable with respect to $\mathcal{F}_{n,\alpha}$. The moment condition (1.18) is trivially satisfied, as $1 - 1_{F_{n,\alpha,2}} = 0$ almost surely.

Condition (Z5). By (3.1) and (3.2) we have that $(Y - V)/\sigma_{n,\alpha} = T$ as in (3.3), the scaled content $c_\alpha(x_n)$ of the box x_n added at time n in Kerov's growth process. Hence, the first part of Condition (1.19) holds with $\overline{B} = \overline{D}$ in (3.11), as by (3.12), and arguing as in (3.13), we have

$$\frac{|Y - V|}{\sigma_{n,\alpha}} = |T| \le \frac{\alpha\lambda_1 + \lambda_1' + \alpha + 1}{\sigma_{n,\alpha}} \le \overline{D} \quad \text{on } F_{n,\alpha,1} \text{ for } (n, \alpha) \in \text{☺}.$$

The second part of this condition holds easily, as

$$r_{n,\alpha}^2 (\overline{D}(\overline{B} + \overline{D})) = 2r_{n,\alpha}^2 \overline{D}^2 = 200/\varepsilon^2 \quad \text{almost surely.}$$

Condition (G7). To verify the variance ratio condition (1.12), recalling $\sigma_{n,\alpha}^2$ from (3.10) and $\Psi(\alpha, n)$ from (3.17), we have

$$\frac{\sigma_{\alpha,n}^2}{\sigma_{\Psi(\alpha,n)}^2} = \frac{\alpha\binom{n}{2}}{\alpha\binom{n-1}{2}} = \frac{n}{n-2} \le 3 \quad \text{for all } (n, \alpha) \in \text{☺},$$

as $n \ge 3$ for all $(n, \alpha) \in \text{☺}$ by the comment after (3.9). For this same reason condition (1.13) holds, as

$$\frac{r_{\alpha,n}}{r_{\Psi(\alpha,n)}} = \frac{r_{\alpha,n}}{r_{\alpha,n-1}} = \frac{n}{n-1} \in [1, 3/2].$$

Conditions (Z1)–(Z5) and (G7) have been verified, and Theorem 1.5 now follows from Theorem 1.2. □

The next result shows that the case when α is taken larger than that in Theorem 1.5 is degenerate; the boundary case $\varepsilon = 1$ is left unresolved.

Theorem 3.1. *For all $\varepsilon > 1$, along any sequence $\{(n, \alpha_n), n \geq 1\}$ for which $\alpha_n \geq n^{1+\varepsilon}$,*

$$\lim_{n \to \infty} \mathbb{P}_{n,\alpha_n}[\lambda_1' = n] = 1.$$

Proof. Note that for all boxes x in the Tableaux with $\lambda_1' = n$ we have $a(x) = 0$ and $l(x)$ takes all values between 0 and $n-1$. Hence, from the Jack$_\alpha$ measure distribution as given in (1.22),

$$\frac{1}{\mathbb{P}_{n,\alpha_n}[\lambda_1' = n]} = \frac{\prod_{l=0}^{n-1}(l+1)(l+\alpha_n)}{\alpha_n^n n!} = \frac{\prod_{l=0}^{n-1}(l+\alpha_n)}{\alpha_n^n}$$

$$= \prod_{l=0}^{n-1}\left(1 + \frac{l}{\alpha_n}\right) \leq \prod_{l=0}^{n-1}\exp\left(\frac{l}{\alpha_n}\right) \leq \exp\left(\frac{n^2}{\alpha_n}\right).$$

Substituting the lower bound on α_n into this inequality yields

$$\mathbb{P}_{n,\alpha_n}[\lambda_1' = n] \geq \exp(-n^{1-\varepsilon}) \to 1 \quad \text{as } n \to \infty. \qquad \square$$

Remark 3.2. The Wasserstein bound in (1.25) suggests that a bound in the Kolmogorov metric should hold with rate function

$$r_{n,\alpha} = \left(\frac{1}{\sqrt{n}} + \frac{\sqrt{\alpha}}{n}\right)^{-1} \quad \text{for all } n \geq 2 \text{ and } \alpha > 0. \tag{3.18}$$

This rate function is equivalent to the one we take in (3.8) for the 'large α' parameter set (3.7), as there $n \leq \alpha$ and $1/\sqrt{n}$ is dominated by $\sqrt{\alpha}/n$. Directly extending the arguments used here to cover the 'small' alpha regime requires that (3.16) hold for some choice of $F_{n,\alpha,1}$. In particular, (3.14) shows that $\mathbb{E}_{n,\alpha}D^2 \leq C/r_{n,\alpha}^2$, with $r_{n,\alpha}$ as in (3.18). Hence, taking this route, one needs to specify $F_{n,\alpha,1}$ as an appropriate restriction on Λ_{n-1} that satisfies $\mathbb{P}_{n,\alpha}[F_{n,\alpha,1}^c] < C/r_{n,\alpha}^2$, and which gives rise to a bounding $\overline{D}$ of the right order. If in this case $\overline{B}$ may be taken to be $\overline{D}$ as in (Z5) above, then $\overline{D}$ needs to be of order $1/r_{n,\alpha}$.

4 Proof of Theorems 1.1 and 1.2

The proofs of Theorems 1.1 and 1.2 ultimately rely on obtaining information about the solution to a certain recursive inequality. In its simplest form, and closely related to the argument in Bolthausen (1984), this inequality becomes

$$a_n \leq q a_{n-1} + c \quad \text{for } n \geq 2 \text{ and } a_1 = 1 \tag{4.1}$$

for some $0 < q < 1$ and $c > 0$. In this simple case, it is not difficult to solve the corresponding equality explicitly to yield

$$a_n = q^{n-1} + c\frac{1 - q^{n-1}}{1 - q} \quad \text{for } n \geq 1.$$

What is important here is not the exact form of the solution but rather that a_n is uniformly bounded over $n \geq 1$. We show below that this property holds in greater generality when we replace n on the left hand side of (4.1) by a generic parameter $\theta \in \Theta$, and average

the right hand side over a randomly chosen parameter $Y \in \Theta$, rather than evaluate at $n-1$. Although, in the general case, there may exist additional solutions to the inequality that are unbounded, it turns out that these solutions must grow exponentially fast along some sequence, which is a behavior that can be excluded in our applications.

Lemma 4.1. *Let $(\Theta, \mathcal{T})$ and $(\Omega, \mathcal{F})$ be measurable spaces. For each $\theta \in \Theta$, let $\mathbb{P}_\theta[\cdot]$ be a probability measure on Ω. Let $X : \Theta \times \Omega \to [0, \infty)$ and $\Psi : \Theta \times \Omega \to \Theta$ be such that, for each $\theta \in \Theta$, both $X(\theta, \cdot)$ and $\Psi(\theta, \cdot)$ are measurable functions. Assume there are constants $0 < q < 1$ and $c > 0$, measurable functions $a : \Theta \to [0, \infty)$ and $r : \Theta \to [0, \infty)$, and a measurable set $\odot \subset \Theta$ such that*

$$(A1)\ \mathbb{E}_\theta X = 1 \text{ for all } \theta \in \odot, \qquad (A2)\ \mathbb{E}_\theta X = 0 \text{ for all } \theta \in \Theta \setminus \odot,$$

$$(A3)\ a(\theta) \le q\, \mathbb{E}_\theta\{X a(\Psi)\} + c < \infty \text{ for all } \theta \in \Theta,$$

$$(A4)\ a(\theta) \le r(\theta) \text{ and } \mathbb{P}_\theta - \operatorname*{ess\,sup}_{\{X>0\}} r(\Psi) \le \frac{r(\theta)}{2q} \text{ for all } \theta \in \odot.$$

Then

$$\sup_{\theta \in \Theta} a(\theta) \le \frac{c}{1-q}.$$

Proof. Note that, for $\theta \in \Theta \setminus \odot$, the variable X must be zero $\mathbb{P}_\theta$-almost surely by (A2), and so (A3) yields that

$$a(\theta) \le c \qquad \text{for } \theta \in \Theta \setminus \odot. \tag{4.2}$$

We may therefore assume that $\odot$ is non-empty, else the claim in trivial. We argue by contradiction; so assume Conditions (A1)–(A4) are satisfied and that the opposite of the conclusion is true. For every $\theta \in \odot$, we can use (A1) and consider the probability measure $\mathbb{P}_\theta^X$ specified by its Radon-Nikodym derivative

$$\frac{d\mathbb{P}_\theta^X}{d\mathbb{P}_\theta} = X, \quad \text{so that} \quad \mathbb{E}_\theta\{X a(\Psi)\} = \mathbb{E}_\theta^X\{a(\Psi)\},$$

where $\mathbb{E}_\theta^X$ denotes expectation with respect to $\mathbb{P}_\theta^X$. We argue by contradiction, assuming that when

$$\sup_{\theta \in \Theta} a(\theta) > \frac{c}{1-q} \tag{4.3}$$

and Conditions (A1)–(A4) hold, there exists a sequence $\{\theta_n\}_{n\ge0} \subset \odot$ and a constant C such that, for all $n \ge 0$,

$$\frac{c}{1-q} + \frac{c\delta}{q^n} \le a(\theta_n) \le r(\theta_n) \le C/(2q)^n, \tag{4.4}$$

which is clearly impossible.

We proceed by induction. For the base case $n = 0$, we note that since $a(\cdot)$ is bounded by c on $\Theta \setminus \odot$ by (4.2), from (4.3) that there is $\theta_0 \in \odot$ such that $a(\theta_0) = c/(1-q) + c\delta$, for some $\delta > 0$; taking also $C = r(\theta_0)$, (4.4) is satisfied.

For the induction step, assume that the lower bound in (4.4) is true for $n - 1 \ge 0$. As $\theta_{n-1} \in \odot$, Condition (A3) yields that $\mathbb{E}_{\theta_{n-1}}^X a(\Psi) \ge (a(\theta_{n-1}) - c)/q$, and so the integrand must be at least this lower bound on a set of positive $\mathbb{P}_{\theta_{n-1}}^X$–measure; that is,

$$A_{n-1} := \big\{\omega \in \Omega\,:\, a\big(\Psi(\theta_{n-1}, \omega)\big) \ge (a(\theta_{n-1}) - c)/q\big\}$$

$$\text{satisfies} \quad \mathbb{P}_{\theta_{n-1}}^X[A_{n-1}] > 0.$$

Moreover, by the definition of essential supremum,

$$B_{n-1} := \{\omega \in \Omega : r(\Psi(\theta_{n-1}, \omega)) \leq \mathbb{P}_{\theta_{n-1}}\text{-}\operatorname{ess\,sup} r(\Psi)\}$$

$$\text{satisfies} \quad \mathbb{P}^X_{\theta_{n-1}}[B_{n-1}] = 1.$$

Hence $\mathbb{P}^X_{\theta_{n-1}}[A_{n-1} \cap B_{n-1}] = \mathbb{P}^X_{\theta_{n-1}}[A_{n-1}] > 0$, and we can find $\theta_n \in \Theta$ satisfying

$$a(\theta_n) \geq \frac{a(\theta_{n-1}) - c}{q} \quad \text{and} \quad r(\theta_n) \leq \mathbb{P}^X_{\theta_{n-1}}\text{-}\operatorname{ess\,sup} r(\Psi(\theta_{n-1}, \cdot)). \tag{4.5}$$

Since $a(\cdot) \leq c$ on $\Theta \setminus \copyright$ we conclude that $\theta_n \in \copyright$ in view of the first inequality of (4.5), which also completes the induction for the lower bound in (4.3). Applying (4.5) and (A4) yields

$$r(\theta_n) \leq \mathbb{P}^X_{\theta_{n-1}}\text{-}\operatorname{ess\,sup} r(\Psi(\theta_{n-1}, \cdot)) \leq \frac{1}{2q} r(\theta_{n-1}),$$

yielding the upper bound in (4.4), and concluding the induction. $\qquad\square$

Proof of Theorem 1.1. Throughout the proof, C denotes a constant that does not depend on θ and can change from formula to formula. Note first that by Condition (G1) the bound (1.14) trivially holds for every $\theta \in \Theta \setminus \copyright$ by taking $C = \bar{r}$. Therefore we need only show that (1.14) holds for all $\theta \in \copyright$. Let

$$\delta(\theta) = \begin{cases} \sup_{z \in \mathbb{R}} |\mathbb{P}_\theta[W \leq z] - \mathbb{P}[Z \leq z]| & \theta \in \copyright \\ 1 & \theta \in \Theta \setminus \copyright. \end{cases} \tag{4.6}$$

Fix $\varepsilon > 0$, whose exact value is to be chosen later, and for $z \in \mathbb{R}$ define

$$h_{z,\varepsilon}(x) = \begin{cases} 1 & \text{if } x \leq z, \\ 1 + (z-x)/\varepsilon & \text{if } z < x \leq z + \varepsilon, \\ 0 & \text{if } z + \varepsilon < x. \end{cases}$$

Let $f_{z,\varepsilon}$ be the unique bounded solution to the Stein equation

$$f'_{z,\varepsilon}(x) - x f_{z,\varepsilon}(x) = h_{z,\varepsilon}(x) - \mathbb{E} h_{z,\varepsilon}(Z).$$

Using a standard smoothing inequality, see e.g. the proof of Theorem 5.1 in Chen, Goldstein and Shao (2011), we have

$$\delta(\theta) \leq \sup_{z \in \mathbb{R}} |\mathbb{E}_\theta\{f'_{z,\varepsilon}(W) - W f_{z,\varepsilon}(W)\}| + \frac{\varepsilon}{\sqrt{2\pi}}. \tag{4.7}$$

For ease of notation, we drop the indices z and ε from f.

Bound on $|\mathbb{E}_\theta\{f'(W) - W f(W)\}|$ Taking an arbitrary $\theta \in \copyright$ and using the definition (1.2) of a Stein coupling in the second line below, we have

$$|\mathbb{E}_\theta\{f'(W) - W f(W)\}|$$

$$= |\mathbb{E}_\theta\{(1 - GD)f'(W) - G \int_0^D (f'(W+t) - f'(W))dt\}|$$

$$\leq |\mathbb{E}_\theta\{f'(W)\mathbb{E}_\theta[1 - GD|W]\}| + \left|\mathbb{E}_\theta\left\{G \int_0^D (f'(W+t) - f'(W))dt\right\}\right|$$

$$=: R_1 + R_2.$$

From (4.6) and (4.7) of Chen and Shao (2004) we have, respectively, that $\|f'\| \le 1$ and

$$|f'(x+t) - f'(x)| \le |t|\Big(1 + |x| + \frac{1}{\varepsilon}\int_0^1 \mathrm{I}[z < x + ut \le z + \varepsilon]du\Big), \qquad (4.8)$$

implying, by the first condition in (1.7), that

$$R_1 \le \frac{C}{r_\theta} \qquad \text{for all } \theta \in \text{☺}, \qquad (4.9)$$

and that

$$\begin{aligned} R_2 &\le \mathbb{E}_\theta\Big\{|G|(1+|W|)\int_{0\wedge D}^{0\vee D}|t|dt\Big\} \\ &\quad + \frac{1}{\varepsilon}\mathbb{E}_\theta\Big\{|G|\int_{0\wedge D}^{0\vee D}\int_0^1 |t|\,\mathrm{I}[z < W+ut \le z+\varepsilon]dudt\Big\} =: R_{2,1} + R_{2,2}. \end{aligned}$$

Using the second condition in (1.7), and that $|t| \le |D|$ in the integral, we have

$$R_{2,1} \le \mathbb{E}_\theta\big\{(1+|W|)|G|D^2\big\} \le \frac{C}{r_\theta} \qquad \text{for all } \theta \in \text{☺}. \qquad (4.10)$$

Let $F_\theta = F_{\theta,1} \cap F_{\theta,2}$. To handle the indicator in $R_{2,2}$, write

$$\begin{aligned} &\mathrm{I}[z < W+ut \le z+\varepsilon] \\ &\qquad \le (1 - I_{F_{\theta,1}}) + (1 - I_{F_{\theta,2}}) + I_{F_\theta}\,\mathrm{I}[z < W+ut \le z+\varepsilon]. \end{aligned} \qquad (4.11)$$

Using (4.11), and again that $|t| \le |D|$, we have

$$\begin{aligned} R_{2,2} &\le \frac{1}{\varepsilon}\mathbb{E}_\theta\Big\{|G|\int_{0\wedge D}^{0\vee D}\int_0^1 |t|(1 - I_{F_{\theta,1}})dudt + |G|\int_{0\wedge D}^{0\vee D}\int_0^1 |t|(1 - I_{F_{\theta,2}})dudt\Big\} \\ &\quad + \frac{1}{\varepsilon}\mathbb{E}_\theta\Big\{|G|\int_{0\wedge D}^{0\vee D}\int_0^1 |t|I_{F_\theta}\,\mathrm{I}[z < W+ut \le z+\varepsilon]dudt\Big\} \\ &\le \frac{1}{\varepsilon}\mathbb{E}_\theta\Big\{|G|D^2(1 - I_{F_{\theta,1}}) + |G|D^2(1 - I_{F_{\theta,2}})\Big\} \\ &\quad + \frac{1}{\varepsilon}\mathbb{E}_\theta\Big\{|GD|\int_{0\wedge D}^{0\vee D}\int_0^1 I_{F_\theta \cap F_\circ}\,\mathrm{I}[z < W+ut \le z+\varepsilon]dudt\Big\} \\ &\quad + \frac{1}{\varepsilon}\mathbb{E}_\theta\Big\{|G|D^2 I_{F_\theta \cap F_\circ^c}\Big\} =: R_{2,2,1} + R_{2,2,2} + R_{2,2,3}, \end{aligned} \qquad (4.12)$$

where $F_\circ = \{\Psi(\theta, \cdot) \in \text{☺}\}$. Now, by (1.10) and the first condition of (1.8)

$$R_{2,2,1} \le \frac{C}{\varepsilon r_\theta^2}. \qquad (4.13)$$

Since $F_\theta \cap F_\circ$ is contained in $F_\circ$ and $\sigma_\theta > 0$ for $\theta \in \text{☺}$, on this intersection we may define

$$\widetilde{W} = \frac{V - \mu_\Psi}{\sigma_\Psi},$$

and thus write

$$W = \frac{\sigma_\Psi}{\sigma_\theta}\widetilde{W} + \frac{Y - V}{\sigma_\theta} - \frac{\mu_\theta - \mu_\Psi}{\sigma_\theta} =: \rho\widetilde{W} + T_1 - T_2,$$

where ρ, T_1 and T_2 are to be understood as random variables on $\odot \times \Omega$. By the first condition in (1.11), we have $|T_1| \leq \overline{B}$ on $F_\theta \cap F_\circ$. Hence,

$$I_{F_\theta \cap F_\circ} I[z < W + ut \leq z + \varepsilon] \tag{4.14}$$

$$= I_{F_\theta \cap F_\circ} I\left[\frac{z - T_1 + T_2 - ut}{\rho} < \widetilde{W} \leq \frac{z - T_1 + T_2 - ut + \varepsilon}{\rho}\right]$$

$$\leq I_{F_\theta \cap F_\circ} I\left[\frac{z - \overline{B} + T_2 - ut}{\rho} < \widetilde{W} \leq \frac{z + \overline{B} + T_2 - ut + \varepsilon}{\rho}\right]$$

$$= I_{F_\theta \cap F_\circ} I\left[Q_{z,ut} - \frac{\overline{B}}{\rho} < \widetilde{W} \leq Q_{z,ut} + \frac{\overline{B} + \varepsilon}{\rho}\right]$$

where

$$Q_{z,y} = \frac{z + T_2 - y}{\rho}$$

is $\mathcal{F}_\theta$ measurable by Condition (G5).

Note that $F_\circ \in \mathcal{F}_\theta$ since $\odot$, given in Condition (G1), is in $\mathcal{T}$ and $\Psi(\theta, \cdot)$ is $\mathcal{F}_\theta$-measurable by Condition (G5) for $\theta \in \odot$. Now using Condition (G4) to bound $|D|$ by $\overline{D}$ on $F_{\theta,1}$, and applying the measurability of $\overline{G}, \overline{D}$ and $F_{\theta,2}$ with respect to $\mathcal{F}_\theta$ by Conditions (G4) and (G5), we obtain

$$R_{2,2,2}$$

$$\leq \frac{1}{\varepsilon} \mathbb{E}_\theta \left\{ \overline{G}\,\overline{D} \int_{-\overline{D}}^{\overline{D}} \int_0^1 I_{F_\theta \cap F_\circ} \, I\left[Q_{z,ut} - \frac{\overline{B}}{\rho} < \widetilde{W} \leq Q_{z,ut} + \frac{\overline{B}+\varepsilon}{\rho}\right] du\, dt \right\} \tag{4.15}$$

$$\leq \frac{1}{\varepsilon} \mathbb{E}_\theta \left\{ \overline{G}\,\overline{D} \int_{-\overline{D}}^{\overline{D}} \int_0^1 I_{F_{\theta,2} \cap F_\circ} \mathbb{P}_\theta\left[Q_{z,ut} - \frac{\overline{B}}{\rho} < \widetilde{W} \leq Q_{z,ut} + \frac{\overline{B}+\varepsilon}{\rho} \Big| \mathcal{F}_\theta\right] du\, dt \right\}.$$

Using (4.6) and (1.9) we obtain

$$\sup_{x \in \mathbb{R}} |\mathbb{P}_\theta[\widetilde{W} \leq x | \mathcal{F}_\theta] - \mathbb{P}[Z \leq x]| \leq \delta(\Psi),$$

and as the normal density is bounded by $1/\sqrt{2\pi}$, using (1.12) we see that the integrand in (4.15) can be no more than

$$I_{F_{\theta,2} \cap F_\circ}\left(2\delta(\Psi) + \frac{2\overline{B} + \varepsilon}{\rho\sqrt{2\pi}}\right) \leq C\, I_{F_{\theta,2} \cap F_\circ}\left(\delta(\Psi) + \overline{B} + \varepsilon\right).$$

Therefore, using the second condition in (1.8) and the second inequality in (1.11) for the fourth inequality below, and then the first condition in (1.13) for the last, we obtain

$$R_{2,2,2} \leq \frac{C}{\varepsilon} \mathbb{E}_\theta \left\{ \overline{G}\,\overline{D}^2 I_{F_{\theta,2} \cap F_\circ}\left(\delta(\Psi) + \overline{B} + \varepsilon\right) \right\}$$

$$\leq \frac{C}{\varepsilon} \mathbb{E}_\theta \left\{ \overline{G}\,\overline{D}^2 I_{F_{\theta,2} \cap F_\circ} \delta(\Psi) \right\} + \frac{C}{\varepsilon} \mathbb{E}_\theta \left\{ \overline{G}\,\overline{D}^2 \overline{B} I_{F_{\theta,2}} \right\} + C\, \mathbb{E}_\theta \left\{ \overline{G}\,\overline{D}^2 \right\}$$

$$\leq \frac{C\, \mathbb{E}_\theta \{ \overline{G}\,\overline{D}^2 \}}{\varepsilon} \mathbb{E}_\theta \left\{ \frac{\overline{G}\,\overline{D}^2 I_{F_{\theta,2} \cap F_\circ}}{\mathbb{E}_\theta \{ \overline{G}\,\overline{D}^2 I_{F_{\theta,2}} \}} \delta(\Psi) \right\} + \frac{C}{\varepsilon} \mathbb{E}_\theta \left\{ \overline{G}\,\overline{D}^2 \overline{B} I_{F_{\theta,2}} \right\} + C\, \mathbb{E}_\theta \left\{ \overline{G}\,\overline{D}^2 \right\}$$

$$\leq \frac{C}{\varepsilon r_\theta} \mathbb{E}_\theta \left\{ \frac{\overline{G}\,\overline{D}^2 I_{F_{\theta,2} \cap F_\circ}}{\mathbb{E}_\theta \{ \overline{G}\,\overline{D}^2 I_{F_{\theta,2}} \}} \delta(\Psi) \right\} + \frac{C}{\varepsilon r_\theta^2} + \frac{C}{r_\theta}$$

$$\leq \frac{C}{\varepsilon r_\theta^2} \mathbb{E}_\theta \left\{ \frac{\overline{G}\,\overline{D}^2 I_{F_{\theta,2} \cap F_\circ}}{\mathbb{E}_\theta \{ \overline{G}\,\overline{D}^2 I_{F_{\theta,2}} \}} \delta(\Psi) r_\Psi \right\} + \frac{C}{\varepsilon r_\theta^2} + \frac{C}{r_\theta}, \tag{4.16}$$

where $R_{2,2,2} = 0$ in the case $\mathbb{E}\{\overline{G}\,\overline{D}^2 I_{F_{\theta,2}}\} = 0$, by the first line of the display above.

In order to bound $R_{2,2,3}$, using that $\delta(\theta) = 1$ for $\theta \in \Theta \setminus ☺$ by (4.6) for the second equality, that $F_\theta \subset F_{\theta,2}$ for the first inequality, the first condition in (1.13) for the second inequality, and the second condition in (1.8) for the last, we have

$$
\begin{aligned}
R_{2,2,3} &= \frac{1}{\varepsilon} \mathbb{E}_\theta \big\{ \overline{G}\,\overline{D}^2 I_{F_\theta \cap F_\circ^c} \big\} = \frac{1}{\varepsilon} \mathbb{E}_\theta \big\{ \overline{G}\,\overline{D}^2 I_{F_\theta \cap F_\circ^c} \delta(\Psi) \big\} \\
&\leq \frac{1}{\varepsilon} \mathbb{E}_\theta \big\{ \overline{G}\,\overline{D}^2 I_{F_{\theta,2} \cap F_\circ^c} \delta(\Psi) \big\} \leq \frac{C}{\varepsilon r_\theta} \mathbb{E}_\theta \big\{ \overline{G}\,\overline{D}^2 I_{F_{\theta,2} \cap F_\circ^c} \delta(\Psi) r_\Psi \big\} \\
&\leq \frac{C\, \mathbb{E}_\theta \big\{ \overline{G}\,\overline{D}^2 \big\}}{\varepsilon r_\theta} \mathbb{E}_\theta \left\{ \frac{\overline{G}\,\overline{D}^2 I_{F_{\theta,2} \cap F_\circ^c}}{\mathbb{E}_\theta \big\{ \overline{G}\,\overline{D}^2 I_{F_{\theta,2}} \big\}} \delta(\Psi) r_\Psi \right\} \\
&\leq \frac{C}{\varepsilon r_\theta^2} \mathbb{E}_\theta \left\{ \frac{\overline{G}\,\overline{D}^2 I_{F_{\theta,2} \cap F_\circ^c}}{\mathbb{E}_\theta \big\{ \overline{G}\,\overline{D}^2 I_{F_{\theta,2}} \big\}} \delta(\Psi) r_\Psi \right\},
\end{aligned}
\tag{4.17}
$$

where $R_{2,2,3} = 0$ when $\mathbb{E}_\theta \big\{ \overline{G}\,\overline{D}^2 I_{F_{\theta,2}} \big\} = 0$, by the first line of the display.

Collecting the bounds (4.9), (4.10), (4.13), (4.16) and (4.17) and using (4.7) we arrive at

$$
\begin{aligned}
\delta(\theta) &\leq R_1 + R_{2,1} + R_{2,2,1} + R_{2,2,2} + R_{2,2,3} + \frac{\varepsilon}{\sqrt{2\pi}} \\
&\leq \frac{C}{\varepsilon r_\theta^2} \mathbb{E}_\theta \left\{ \frac{\overline{G}\,\overline{D}^2 I_{F_{\theta,2}}}{\mathbb{E}_\theta \big\{ \overline{G}\,\overline{D}^2 I_{F_{\theta,2}} \big\}} \delta(\Psi) r_\Psi \right\} + \frac{C}{\varepsilon r_\theta^2} + \frac{C}{r_\theta} + C\varepsilon.
\end{aligned}
\tag{4.18}
$$

Since Condition (G1) implies that $\overline{r}$ is an upper bound on r_θ for $\theta \in \Theta \setminus ☺$, and a lower bound on r_θ for $\theta \in ☺$, we conclude that

$$
\sup_{\theta \in ☺} \mathbb{P}_\theta\text{-} \operatorname*{ess\,sup}_{\omega \in F_{\theta,2} \cap \{\Psi \in \Theta \setminus ☺\}} \frac{r_{\Psi(\theta,\omega)}}{r_\theta} < \infty.
$$

Hence, by the second condition in (1.13),

$$
q = \frac{1}{2} \left(1 \vee \sup_{\theta \in ☺} \mathbb{P}_\theta\text{-} \operatorname*{ess\,sup}_{\omega \in F_{\theta,2}} \frac{r_{\Psi(\theta,\omega)}}{r_\theta} \right)^{-1} \in (0,1).
\tag{4.19}
$$

Choosing $\varepsilon = C/r_\theta q$ with C as in (4.18) and multiplying that inequality by r_θ on both sides and then setting $a(\theta) = \delta(\theta) r(\theta)$ we obtain, for some possibly different constant $c > 0$, which does not depend on θ but may depend on q,

$$
a(\theta) \leq q\, \mathbb{E}_\theta \left\{ \frac{\overline{G}\,\overline{D}^2 I_{F_{\theta,2}}}{\mathbb{E}_\theta \big\{ \overline{G}\,\overline{D}^2 I_{F_{\theta,2}} \big\}} a(\Psi) \right\} + c \qquad \text{for all } \theta \in ☺.
$$

We now verify the hypotheses of Lemma 4.1, with the additional identification

$$
X = \frac{\overline{G}\,\overline{D}^2 I_{F_{\theta,2}}}{\mathbb{E}_\theta \big\{ \overline{G}\,\overline{D}^2 I_{F_{\theta,2}} \big\}} \mathrm{I}[\theta \in ☺].
\tag{4.20}
$$

Conditions (A1) and (A2) follow directly from the definition of X, while (A3) on $☺$ is (4.20), and is satisfied on $\Theta \setminus ☺$ as $\delta(\theta) \leq 1$, and we may replace c by $\max\{\overline{r}, c\}$. Condition (A4) follows from (4.19). The conclusion of Lemma 4.1 now implies that $\delta(\theta) \leq C/r_\theta$ for all $\theta \in \Theta$. $\qquad \square$

Proof of Theorem 1.2. The proof for zero biasing is quite similar, but simpler, than the proof of Theorem 1.1; we only highlight the important differences.

Recalling $D = W^* - W$, applying the bound (4.8), and the zero bias characterization (1.4), we obtain

$$|\mathbb{E}_\theta(f'(W) - Wf(W))| = |\mathbb{E}_\theta(f'(W + D) - f'(W))|$$

$$\leq \mathbb{E}_\theta\left(|D|\left(1 + |W| + \frac{1}{\varepsilon}\int_0^1 1_{[z,z+\varepsilon]}(W + uD)du\right)\right). \quad (4.21)$$

Using (Z3), noting in particular that $|D| \leq |\overline{D}|$ on $F_{\theta,1}$, and the fact that $r_\theta > \overline{r}$ for $\theta \in \smiley$ yields $1/r_\theta^2 \leq C/r_\theta$, for the first two terms in (4.21), we have

$$\mathbb{E}_\theta\{|D| + |DW|\} \leq \mathbb{E}_\theta\{|D|(1 - I_{F_{\theta,1}})\} + \mathbb{E}_\theta\{\overline{D} + |DW|\} \leq \frac{C}{r_\theta}.$$

Following the reasoning in (4.12) and labeling the corresponding terms that arise here in the same manner, for $R_{2,2}$, the only remaining term, by the first condition in (1.16), and (1.18), we obtain the bound

$$R_{2,2,1} \leq \frac{1}{\varepsilon}\mathbb{E}_\theta\big(|D|(1 - I_{F_{\theta,1}}) + |D|(1 - I_{F_{\theta,2}})\big) \leq \frac{C}{\varepsilon r_\theta^2}.$$

For $R_{2,2,2}$, as ut in (4.14) is replaced by uD, separating the term that arises from uD out of $Q_{z,y}$ as defined there, here we obtain

$$I_{F_\theta \cap F_\circ}I[z < W + uD \leq z + \varepsilon] \leq I_{F_\theta \cap F_\circ}I\left[Q_z - \frac{\overline{B} + \overline{D}}{\rho} < \widetilde{W} \leq Q_z + \frac{\overline{B} + \overline{D} + \varepsilon}{\rho}\right],$$

where $Q_z = (z + T_2)/\rho$ is $\mathcal{F}_\theta$ measurable. Now arguing as in (4.16) we obtain

$$R_{2,2,2} \leq \frac{C}{\varepsilon}\mathbb{E}_\theta\left\{\overline{D}I_{F_{\theta,2}\cap F_\circ}\left(\delta(\Psi) + \overline{B} + \overline{D} + \varepsilon\right)\right\}$$

$$\leq \frac{C}{\varepsilon r_\theta^2}\mathbb{E}_\theta\left\{\frac{\overline{D}I_{F_{\theta,2}\cap F_\circ}}{\mathbb{E}\{\overline{D}I_{F_{\theta,2}}\}}\delta(\Psi)r_\Psi\right\} + \frac{C}{\varepsilon r_\theta^2} + \frac{C}{r_\theta}$$

using the second condition of (1.16) and the first one of (1.13) for the first term, and the second conditions of (1.19) and (1.16), respectively, to obtain the last two terms in the bound.

As in (4.17), using the first condition of (1.13) and the second condition of (1.16), we obtain

$$R_{2,2,3} = \frac{1}{\varepsilon}\mathbb{E}_\theta\{\overline{D}I_{F_{\theta,2}\cap F_\circ^c}\} \leq \frac{C}{\varepsilon r_\theta^2}\mathbb{E}_\theta\left\{\frac{\overline{D}I_{F_{\theta,2}\cap F_\circ^c}}{\mathbb{E}_\theta\{\overline{D}I_{F_{\theta,2}}\}}\delta(\Psi)r_\Psi\right\}.$$

Combining terms as in (4.18) yields

$$\delta(\theta) \leq \frac{C}{\varepsilon r_\theta^2}\mathbb{E}_\theta\left\{\frac{\overline{D}I_{F_{\theta,2}}}{\mathbb{E}_\theta\{\overline{D}I_{F_{\theta,2}}\}}\delta(\Psi)r_\Psi\right\} + \frac{C}{\varepsilon r_\theta^2} + \frac{C}{r_\theta} + C\varepsilon.$$

The proof can now be concluded as for Theorem 1.1. $\qquad\square$

Appendix

We illustrate two instances where the conditions in the General Framework of the Introduction are implicitly invoked. First we show that random version of the random variable Y at the (random) 'smaller' parameter value is a random variable. The maps

$$(\theta, \omega) \to (\Psi(\theta, \omega), \omega) \quad \text{and} \quad (\theta, \omega) \to Y\big(\Psi(\theta, \omega), \omega\big)$$

are measurable, the first as each component is measurable, and the second being a composition of measurable maps.

Next, we show that if $f(\theta, \omega)$ is measurable and $\mathbb{P}_\theta$-integrable for all $\theta \in \Theta$, then

$$\theta \to \int_\Omega f(\theta, \omega) dP_\theta(\omega)$$

is a measurable function of θ. Indeed, the collection $\mathcal{M}$ of subsets E of $\Theta \times \Omega$ for which the integral of $f(\omega, \theta) = I_E(\omega, \theta)$ is measurable with respect to $\mathbb{P}_\theta$ is a monotone class. The class $\mathcal{M}$ contains the rectangles which are products of measurable sets A and B, as their indicator

$$f(\theta, \omega) = \mathrm{I}[\theta \in A]\,\mathrm{I}[\omega \in B] \quad \text{has integral} \quad \int_\Omega f(\theta, \omega) d\mathbb{P}_\theta(\omega) = \mathrm{I}[\theta \in A]\mathbb{P}_\theta[B],$$

which is a product of measurable functions of θ. Hence $\mathcal{M}$ contains the algebra of all finite disjoint unions of such rectangles, and hence, by the Monotone Class theorem, the sigma-algebra these rectangle generate, that is, the product sigma-algebra. Given a non-negative integrable function $f(\theta, \omega)$, standard arguments using an approximating sequence of simple functions from below in concert with the Monotone Convergence Theorem yields the measurability of the integral of $f(\theta, \omega)$, and then for real valued functions by breaking up of any given integrable function into positive and negative parts.

References

Bartroff, J. and Goldstein, L. (2013). A Berry-Esseen bound for the uniform multinomial occupancy model. *Electron. J. Probab.* **18**, article 27, 1–29. MR-3035755

Bergström, H. (1944). On the central limit theorem. *Skand. Aktuarietidskr.* **27**, 139–153. MR-0015703

Bolthausen, E. (1984). An estimate of the remainder in a combinatorial central limit theorem. *Z. Wahrsch. Verw. Gebiete* **66**, 379–386. MR-0751577

Chen, L. H. Y., Goldstein, L. and Shao, Q.-M. (2010). *Normal Approximation by Stein's Method*. Springer-Verlag. MR-2732624

Chen, L. H. Y. and Shao, Q.-M. (2004). Normal approximation under local dependence. *Ann. Probab.* **32**, 1985–2028. MR-2073183

Chen, L. H. Y. and Röllin, A. (2010). Stein couplings for normal approximation. *Preprint*, arXiv:1003.6039.

Chen, L. H. Y., Raič, M. and Thành, L. V. (2020+). On the error bound in the normal approximation for Jack measures. *Bernoulli*, to appear.

Englund, G. (1981). A remainder term estimate for the normal approximation in classical occupancy. *Ann. Probab.* **9**, 684–692. MR-0624696

Erdős, P. and Rényi, A. (1960). On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci.* **5** MR-0125031

Fulman, J. (2004). Stein's method, Jack measure, and the Metropolis algorithm. *J. Combin. Theory Ser. A* **108**, 275–296. MR-2098845

Fulman, J. (2006). An inductive proof of the Berry-Esseen theorem for character ratios. *Ann. Comb.* **10**, 319–332. MR-2284273

Fulman, J. (2006). Martingales and character ratios. *Trans. Amer. Math. Soc.* **358**, 4533–4552. MR-2231387

Fulman, J. and Goldstein, L. (2011). Zero biasing and Jack measures. *Combin. Probab. Comput.* **20**, 753–762. MR-2825589

Ghosh, S. (2009). L^p bounds for a central limit theorem with involutions. *Preprint*, arXiv:0905.1150

Ghosh, S. and L. Goldstein (2011). Concentration of measures via size-biased couplings. *Probab. Theory Related Fields* **149**, 271–278. MR-2773032

Goldstein, L. (2007). L^1 Bounds in normal approximation *Ann. Probab.* **35**, 1888–1930. MR-2349578

Goldstein, L. (2013). A Berry-Esseen bound with applications to vertex degree counts in the Erdős-Rényi random graph. *Ann. Appl. Probab.* **23**, 617–636. MR-3059270

Greene, E. and J.A. Wellner (2017). Exponential bounds for the hypergeometric distribution. *Bernoulli* **23**, 1911–1950. MR-3624882

Goldstein, L. and Reinert, G. (1997). Stein's method and the zero bias transformation with application to simple random sampling. *Ann. Appl. Probab.* **7**, 935–952. MR-1484792

Goldstein, L. and Rinott, Y. (1996). Multivariate normal approximations by Stein's method and size bias couplings. *J. Appl. Probab.* **33**, 1–17. MR-1371949

Hora, A. and Obata, N. (2007). *Quantum Probability and Spectral Analysis of Graphs*. Springer. MR-2316893

Kerov, S.V. (1994). The boundary of Young lattice and random Young tableaux. In *Formal Power Series and Algebraic Combinatorics*, 133–158. American Mathematical Society. MR-1363510

Knuth, D.E. (1969). *The Art of Computer Programming. Vol. 2: Seminumerical Algorithms*, Addison-Wesley. MR-0286318

Kordecki, W. (1987). Normal approximation and isolated vertices in random graphs, *Random Graphs '87* (Poznan 1987), 131–139, Wiley, Chichester. MR-1094128

Ross, N. (2011). Fundamentals of Stein's method. *Probab. Surv.* **8**, 210–293. MR-2861132

Stein, C. (1972). A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. *Proc. Sixth Berkeley Symp. Math. Statist. Prob.* **2**, University of California Press. MR-0402873

Stein, C. (1986). *Approximate Computation of Expectations*. Institute of Mathematical Statistics Lecture Notes—Monograph Series, 7. Institute of Mathematical Statistics, Hayward, CA. MR-0882007

Advantages of publishing in EJP-ECP

- Very high standards

- Free for authors, free for readers

- Quick publication (no backlog)

- Secure publication (LOCKSS[1])

- Easy interface (EJMS[2])

Economical model of EJP-ECP

- Non profit, sponsored by IMS[3], BS[4], ProjectEuclid[5]

- Purely electronic

Help keep the journal free and vigorous

- Donate to the IMS open access fund[6] (click here to donate!)

- Submit your best articles to EJP-ECP

- Choose EJP-ECP over for-profit journals

---

[1]LOCKSS: Lots of Copies Keep Stuff Safe http://www.lockss.org/
[2]EJMS: Electronic Journal Management System http://www.vtex.lt/en/ejms.html
[3]IMS: Institute of Mathematical Statistics http://www.imstat.org/
[4]BS: Bernoulli Society http://www.bernoulli-society.org/
[5]Project Euclid: https://projecteuclid.org/
[6]IMS Open Access Fund: http://www.imstat.org/publications/open.htm