Sesqui-type branching processes

Svante Janson* and Oliver Riordan† and Lutz Warnke‡

June 24, 2017

Abstract

We consider branching processes consisting of particles (individuals) of two types (type L and type S) in which only particles of type L have offspring, proving estimates for the survival probability and the (tail of) the distribution of the total number of particles. Such processes are in some sense closer to single- than to multi-type branching processes. Nonetheless, the second, barren, type complicates the analysis significantly. The results proved here (about point and survival probabilities) are a key ingredient in the analysis of bounded-size Achlioptas processes in a recent paper by the last two authors.

1 Introduction

Throughout the paper we consider branching processes in which every particle is of one of two *types*, called (for compatibility with the notation in [22]), 'type L' and 'type S'. Particles of type S may be thought of as barren: they have no children. Each particle of type L will have some random number of children of each type; as usual, we have independence between the children of different particles, but the numbers Y and Z of type-L and type-S children of one particle need *not* be independent. The formal definition is as follows.

Definition 1.1. Let (Y, Z) and (Y^0, Z^0) be probability distributions on $\mathbb{N}^2$. We write $\mathfrak{X}^1 = \mathfrak{X}^1_{Y,Z}$ for the Galton–Watson branching process started with a single particle of type L, in which each particle of type L has Y children of type L and Z of type S. Particles of type S have no children, and the children of different particles are independent. We write $\mathfrak{X} = \mathfrak{X}_{Y,Z,Y^0,Z^0}$ for the branching process defined as follows: start in generation one with Y^0 particles of type L and Z^0 of type S. Those of type L have children according to $\mathfrak{X}^1_{Y,Z}$, independently of each other and of the first generation. Those of type S have no children. We write $|\mathfrak{X}|$ ($|\mathfrak{X}^1|$) for the total number particles in $\mathfrak{X}$ ($\mathfrak{X}^1$).

These branching processes are in some sense essentially single-type: one could first generate the tree of type-L particles as a classical single-type Galton–Watson process, and then consider particles of type S. However, since the numbers of type-S and type-L children are not necessarily independent, this two-stage description does not seem particularly easy to work with.

The motivation for considering such processes (and in particular for allowing a different rule for the first generation) comes from the application to studying the phase transition in Achlioptas processes in [22]. Achlioptas processes are evolving random graph models that have received considerable attention (see, e.g., [1; 19; 4; 24; 14; 20; 15; 3; 21] and the references therein). We shall say nothing further about these random graph processes here, aiming to keep the paper self-contained, and purely about branching processes.

We shall prove two main results. Firstly, in Section 2, we consider an individual branching process of the type above, giving an asymptotic formula for the point probability $p_N = \mathbb{P}(|\mathfrak{X}| = N)$ under certain conditions on the distributions (Y, Z) and (Y^0, Z^0). This formula is proved in Sections 2.1–2.3, which are the heart of the paper. Then, in Section 3, we consider families of processes where the offspring distribution

AMS 2010 Mathematics Subject Classification: 60J80

*Department of Mathematics, Uppsala University, PO Box 480, SE-751 06 Uppsala, Sweden. E-mail: svante.janson@math.uu.se. Partly supported by the Knut and Alice Wallenberg Foundation.

†Mathematical Institute, University of Oxford, Radcliffe Observatory Quarter, Woodstock Road, Oxford OX2 6GG, UK. E-mail: riordan@maths.ox.ac.uk.

‡School of Mathematics, Georgia Institute of Technology, Atlanta GA 30332, USA; *and* Peterhouse, Cambridge CB2 1RD, UK. E-mail: warnke@math.gatech.edu.

varies analytically in an additional parameter t. Roughly speaking, we show that the key quantities in the formula in Section 2 then vary analytically in t. This result (which in particular implies properties of the near-critical case) is needed in [22]. Finally, in Section 4, we prove corresponding results for the survival probability $\mathbb{P}(|\mathfrak{X}| = \infty)$. Here the barren type plays no role, so the results effectively concern single-type processes and are much simpler.

Remark 1.2. Although the definition of sesqui-type branching processes is adapted to the application in [22], the results here are applicable, at least in principle, to a more general class of branching processes. Consider a finite-type Galton–Watson process in which there is one special type (type L), and all other types are 'doomed' (lead to finite trees of descendants a.s.). Such a process may be transformed into a sesqui-type process in a natural way: for each type-L particle replace its children of all doomed types, and their (necessarily doomed) descendents, by type-S children (keeping the same total number of particles). For our results to apply to the transformed process we need further conditions, roughly speaking that the 'doomed' subtrees are not too close to critical; but in outline, all processes with (at most) one type that can potentially survive are covered. Branching processes of this type (with one doomed type) have been studied by several authors, giving various results different from ours; see for example [23; 25; 7].

1.1 Some notation and conventions

Throughout we write $\mathbb{N} := \{0, 1, 2, \dots\}$ for the non-negative integers.

Given a two-dimensional random variable (Y, Z) taking values in $\mathbb{N}^2$, we denote its bivariate probability generating function by

$$g_{Y,Z}(y, z) := \mathbb{E}(y^Y z^Z) = \sum_{k,l\geqslant 0} \mathbb{P}(Y = k, Z = l) y^k z^l, \tag{1.1}$$

for all complex y and z such that the expectation (or sum) converges absolutely. We will also consider the bivariate moment generating function

$$f_{Y,Z}(y, z) := g_{Y,Z}\big(e^y, e^z\big) = \mathbb{E}\big(e^{yY+zZ}\big). \tag{1.2}$$

When considering a particular branching process as in Definition 1.1, we often write $g = g_{Y,Z}$ and $f = f_{Y,Z}$ for brevity.

We denote the coefficient of $y^k z^l$ in a power series $G(y, z)$ by $[y^k z^l]G(y, z)$.

We say that a function f defined on $I \subseteq \mathbb{C}$ is analytic if for every $x_0 \in I$ there is an $r > 0$ and a power series $g(x) = \sum_{j\geqslant 0} a_j(x - x_0)^j$ with radius of convergence at least r such that f and g coincide on $(x_0 - r, x_0 + r) \cap I$. A function f defined on some domain including I is analytic on I if $f|_I$ is analytic. The definitions for functions of several real or complex variables are analogous.

If f is an analytic function of d variables, defined in an open set $U \subseteq \mathbb{C}^d$, we denote its derivative by Df, and its mth derivative by $D^m f$. Note that $D^m f$ is an analytic function from U to the linear space of all (symmetric) m-linear forms $\mathbb{C}^d \to \mathbb{C}$. In particular, for each $z \in U$, $Df(z)$ is a linear form, which can also be regarded as a vector (the usual gradient); we write $D_i f := \frac{\partial}{\partial x_i} f$, so $Df(z) = \big(D_1 f(z), \dots, D_d f(z)\big)$. Similarly, $D^2 f(z)$ is a bilinear form, which may be regarded as a $d \times d$ matrix with entries $D_{ij} f(z)$, where $D_{ij} = D_i D_j$. We denote its determinant by $\mathrm{Det}(D^2 f(z))$. (This is known as the Hessian of f.)

For a vector $x \in \mathbb{C}^d$, let $D^m f(z)[x]$ denote $D^m f(z)(x, \dots, x)$, where the vector x is repeated m times. When using coordinates $x = (u, v)$ in the case $d = 2$, we write $[u, v]$ for $[(u, v)]$, so, regarding $D^2 f$ as a matrix and x as a (column) vector, we have

$$D^2 f(z)[u, v] = (u\ v)\, D^2 f(z) \begin{pmatrix} u \\ v \end{pmatrix}. \tag{1.3}$$

We denote the usual Euclidean norm of vectors by $|\cdot|$. For operators and the multilinear forms $D^m f$ we use $\|\cdot\|$ for the usual norm (any other norm would do as well).

For real symmetric matrices, $A \leqslant B$ means that $B - A$ is positive definite, i.e., that $v^{\mathrm{tr}}(B - A)v \geqslant 0$ for all real vectors v. In particular, if A is a $d \times d$ symmetric matrix and $c \in \mathbb{R}$, then

$$A \geqslant cI \iff v^{\mathrm{tr}} A v \geqslant c|v|^2 \text{ for all } v \in \mathbb{R}^d. \tag{1.4}$$

Remark 1.3. We adopt the following notational convention regarding constants. c and C are used 'locally' (within a single proof), while numbered constants c_1, C_1 etc retain their meaning throughout the paper. The constants c_i, which are numbered in the order they are introduced, obey the inequalities

$$c_1 \leqslant c_8 \leqslant c_7 \leqslant c_6 \leqslant c_5 \quad \text{and} \quad c_4 \leqslant c_2$$

We write y, z, w for complex variables, and u, v, α, β for real variables. All constants c_i, C_i etc are positive.

2 Point probabilities of a single branching process

In this section we study the point probabilities $\mathbb{P}(|\mathfrak{X}| = N)$ of the branching process $\mathfrak{X} = \mathfrak{X}_{Y,Z,Y^0,Z^0}$ from Definition 1.1. To formulate our main result we need some further definitions (which encapsulate fairly mild and natural conditions for the offspring distributions).

Definition 2.1. Suppose that $R > 1$, $M < \infty$, $k_1, k_2 \in \mathbb{N}$ and $\delta > 0$.

(i) Let $\mathcal{K}^0 = \mathcal{K}^0(R, M, \delta)$ be the set of probability distributions ν on $\mathbb{N}^2$ such that if $(Y, Z) \sim \nu$, then

$$\mathbb{E}\, R^{Y+Z} \leqslant M, \tag{2.1}$$
$$\mathbb{E}\, Y \geqslant \delta. \tag{2.2}$$

(ii) Let $\mathcal{K}^1 = \mathcal{K}^1(k_1, k_2, \delta)$ be the set of probability distributions $\nu = (\pi_{i,j})_{i,j\geqslant 0}$ on $\mathbb{N}^2$ such that

$$\pi_{k_1,k_2} \geqslant \delta, \qquad \pi_{k_1+1,k_2} \geqslant \delta, \qquad \pi_{k_1,k_2+1} \geqslant \delta. \tag{2.3}$$

(iii) Let $\mathcal{K} = \mathcal{K}(R, M, k_1, k_2, \delta) := \mathcal{K}^0(R, M, \delta) \cap \mathcal{K}^1(k_1, k_2, \delta)$.

We write $(Y, Z) \in \mathcal{K}^0$ if the distribution of (Y, Z) is in $\mathcal{K}^0$, and similarly for $\mathcal{K}^1$ and $\mathcal{K}$. The key condition here is the (uniform) bound (2.1) on the probability generating functions. The condition (2.3) is needed, roughly speaking, to ensure that (Y, Z) is not essentially supported on a sublattice of $\mathbb{N}^2$. Note that $(Y, Z) \in \mathcal{K}^1$ trivially implies

$$\mathbb{E}\, Z \geqslant \mathbb{P}(Z = k_2 + 1) \geqslant \delta, \tag{2.4}$$

and similarly $\mathbb{E}\, Y \geqslant \delta$.

The following theorem gives the qualitative behaviour of the size–N point probabilities of the branching process $\mathfrak{X} = \mathfrak{X}_{Y,Z,Y^0,Z^0}$ from Definition 1.1. The statement of Theorem 2.2 is not self contained since the parameters Ψ, Φ and x^* are defined (in a rather involved way) from the generating functions of (Y, Z) and (Y^0, Z^0), see (2.43)–(2.44) and Lemma 2.15 in Section 2.3. A key feature of the result is that the estimates and error-terms are uniform over all distributions $(Y^0, Z^0) \in \mathcal{K}^0$ and $(Y, Z) \in \mathcal{K}$, i.e., the explicit and implicit constants depend only on R, M, k_1, k_2 and δ. Note that, from (2.8) below, $\xi = 0$ if and only if $\mathbb{E}\, Y = 1$, and that $\mathbb{P}(|\mathfrak{X}| = N)$ decays exponentially in $\Theta(\varepsilon^2 N)$ in the near-critical case $\mathbb{E}\, Y = 1 \pm \varepsilon$.

Theorem 2.2 (Point probabilities of $\mathfrak{X}$)**.** *Suppose that $R > 1$, $M < \infty$, $k_1, k_2 \in \mathbb{N}$, and $\delta > 0$. Writing $\mathcal{K}^0 = \mathcal{K}^0(R, M, \delta)$ and $\mathcal{K} = \mathcal{K}(R, M, k_1, k_2, \delta)$, there exists a constant $c_1 > 0$ such that if $(Y^0, Z^0) \in \mathcal{K}^0$, $(Y, Z) \in \mathcal{K}$, and $|\mathbb{E}\, Y - 1| \leqslant c_1$, then for all $N \geqslant 1$ we have*

$$\mathbb{P}(|\mathfrak{X}| = N) = N^{-3/2} e^{-N\xi}\big(\theta + O(N^{-1})\big), \tag{2.5}$$

where, defining Ψ and Φ as in (2.43)–(2.44) *and x^* as in Lemma 2.15, we have*

$$\xi = \xi_{Y,Z} := -\Psi(x^*) \geqslant 0, \tag{2.6}$$
$$\theta = \theta_{Y^0,Z^0,Y,Z} := \sqrt{2\pi/|\Psi''(x^*)|}\ \Phi(x^*) = \Theta(1), \tag{2.7}$$

and

$$\xi = \Theta(|\mathbb{E}\, Y - 1|^2). \tag{2.8}$$

Moreover, the implicit constants in (2.5)–(2.8) *depend only on R, M, k_1, k_2 and δ.*

The remainder of this section is devoted to the proof of Theorem 2.2. To this end we fix $R > 1$, $M < \infty$, $k_1, k_2 \in \mathbb{N}$, and $\delta > 0$, and write $\mathcal{K}^0 = \mathcal{K}^0(R, M, \delta)$ and $\mathcal{K} = \mathcal{K}(R, M, k_1, k_2, \delta)$ to avoid clutter. Let $|\mathfrak{X}^L|$ and $|\mathfrak{X}^S|$ denote the total numbers of type-L and type-S particles in $\mathfrak{X}$, so $|\mathfrak{X}| = |\mathfrak{X}^L| + |\mathfrak{X}^S|$, and set

$$p_{n,m} := \mathbb{P}(|\mathfrak{X}^L| = n,\, |\mathfrak{X}^S| = m). \tag{2.9}$$

Of course, $p_{n,m}$ depends on the distributions of (Y, Z) and (Y^0, Z^0). In Section 2.1 we establish a simple integral formula for $p_{n,m}$. Then, in Section 2.2 we use a version of the saddle point method to estimate this integral asymptotically. Finally, in Section 2.3 we prove (2.5) by summing all $p_{n,m}$ with $n + m = N$.

2.1 An integral formula for $p_{n,m}$

In this section we derive an explicit integral formula for $p_{n,m}$, see (2.14). We start with a simple conditional version of the classical Otter–Dwass formula (see e.g. Dwass [8]), which hinges on the random walk representation of a branching process and a well-known random-walk hitting time result.

Lemma 2.3. *For all integers $n \geqslant 1$ and $m, n_0, m_0 \geqslant 0$,*

$$\mathbb{P}\big(|\mathfrak{X}^L| = n,\, |\mathfrak{X}^S| = m \;\big|\; Y^0 = n_0,\, Z^0 = m_0\big) = \frac{n_0}{n}\,\mathbb{P}\Bigg(n_0 + \sum_{1 \leqslant j \leqslant n} Y_j = n,\; m_0 + \sum_{1 \leqslant j \leqslant n} Z_j = m\Bigg). \tag{2.10}$$

Proof. Let $(Y_j, Z_j)_{j \geqslant 1}$ be independent with each pair having the same distribution as (Y, Z). Since particles of type S do not have any children, by exploring the branching process $\mathfrak{X}$ in the usual way (i.e., revealing the offspring of the particles of type L one-by-one until none are left to explore), we have

$$\begin{aligned}
&\mathbb{P}\big(|\mathfrak{X}^L| = n,\, |\mathfrak{X}^S| = m \;\big|\; Y^0 = n_0,\, Z^0 = m_0\big)\\
&\qquad = \mathbb{P}\Bigg(n_0 + \min_{0 \leqslant n' < n} \sum_{1 \leqslant j \leqslant n'} (Y_j - 1) > 0,\; n_0 + \sum_{1 \leqslant j \leqslant n} (Y_j - 1) = 0,\; m_0 + \sum_{1 \leqslant j \leqslant n} Z_j = m\Bigg).
\end{aligned}$$

That the right-hand side of the above expression equals (2.10) is surely folklore (by conditioning on $\sum_{1 \leqslant j \leqslant n} Z_j = m - m_0$ this also follows directly from [17, Theorem 7]); we include a short argument. Namely, by a version of the well-known Cyclic Lemma (sometimes also called Spitzer's combinatorial lemma), see, e.g., [13, Lemma 15.3] or [18, Lemma 6.1], for any sequence $(y_1, \ldots, y_n)$ with $y_i \in \{-1, 0, 1, 2, \ldots\}$ and $n_0 + \sum_{1 \leqslant i \leqslant n} y_i = 0$, there are exactly n_0 cyclic shifts of $(y_1, \ldots, y_n)$ for which all corresponding partial sums $s_i = y_1 + \cdots + y_i$ of length $i \leqslant n - 1$ satisfy $n_0 + s_i > 0$. Hence, taking a uniformly random cyclic shift of the n independent variables $(Y_j - 1, Z_j)$, the formula (2.10) follows. □

Remark 2.4. This two-type version of the Otter–Dwass formula is a simple variation of the usual one-type case; this is because one type is barren and can essentially be ignored. For a much more complicated formula in the general multi-type case, see Chaumont and Liu [5].

The probability on the right-hand side of (2.10) can be expressed using generating functions as

$$[y^{n-n_0} z^{m-m_0}]\big(g(y, z)^n\big) = [y^n z^m]\big(y^{n_0} z^{m_0} g(y, z)^n\big). \tag{2.11}$$

For $n \geqslant 1$ and $m \geqslant 0$, recalling the notation (2.9) and summing (2.10) over all n_0, m_0, we thus obtain

$$\begin{aligned}
np_{n,m} &= \sum_{n_0, m_0 \geqslant 0} \mathbb{P}(Y^0 = n_0, Z^0 = m_0) n_0 [y^n z^m]\big(y^{n_0} z^{m_0} g(y, z)^n\big)\\
&= [y^n z^m]\big(\tilde{g}_0(y, z) g(y, z)^n\big),
\end{aligned} \tag{2.12}$$

where

$$\tilde{g}_0(y, z) := \sum_{n_0, m_0 \geqslant 0} \mathbb{P}(Y^0 = n_0, Z^0 = m_0) n_0 y^{n_0} z^{m_0} = y \frac{\partial}{\partial y} g_{Y^0, Z^0}(y, z).$$

For later use, we also define

$$\tilde{f}_0(y, z) := \tilde{g}_0(e^y, e^z) = \frac{\partial}{\partial y} f_{Y^0, Z^0}(y, z) = \mathbb{E}\big(Y^0 e^{yY^0 + zZ^0}\big). \tag{2.13}$$

Remark 2.5. Let $G(y,z) := \mathbb{E}\big(y^{|\mathfrak{X}^L|}z^{|\mathfrak{X}^S|}\big)$ be the bivariate generating function for the size of the branching process $\mathfrak{X}$, and let $G_1(y,z) := \mathbb{E}\big(y^{|\mathfrak{X}^{1,L}|}z^{|\mathfrak{X}^{1,S}|}\big)$ be the corresponding generating function when starting with a single particle of type L. Then $G(y,z) = g_0(G_1(y,z),z)$ and $G_1(y,z) = yg(G_1(y,z),z)$, and the formula (2.12) can alternatively be obtained by the Lagrange inversion formula in the Bürmann form, see e.g. [9, A.(14)], regarding the generating functions as (formal) power series in y with coefficients that are power series in z. We omit the details.

The extraction of coefficients in (2.12) can be performed by complex integration in the usual way (e.g., using Cauchy's integral formula to evaluate $\frac{\partial^{n+m}}{\partial y^n \partial z^m}\big(\tilde{g}_0(y,z)g(y,z)^n\big)\big|_{y=z=0} = n!\,m!\,np_{n,m}$ as in the textbook proof of Cauchy's estimates), yielding the formula

$$np_{n,m} = \frac{1}{(2\pi \mathrm{i})^2}\oint\oint y^{-n}z^{-m}\tilde{g}_0(y,z)g(y,z)^n\,\frac{\mathrm{d}y}{y}\,\frac{\mathrm{d}z}{z},$$

where we integrate (for example) over two circles with centre 0 and radii such that $\tilde{g}_0(y,z)$ and $g(y,z)$ are defined. In particular, if (Y,Z) and (Y^0,Z^0) are both in $\mathcal{K}^0$, then for any $\alpha,\beta < \log R$ we can integrate over $|y| = e^\alpha$ and $|z| = e^\beta$, and the standard change of variables $y = e^{\alpha+\mathrm{i}u}$, $z = e^{\beta+\mathrm{i}v}$ then yields

$$np_{n,m} = \frac{1}{4\pi^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} e^{-n(\alpha+\mathrm{i}u)-m(\beta+\mathrm{i}v)}\tilde{f}_0(\alpha+\mathrm{i}u,\beta+\mathrm{i}v)f(\alpha+\mathrm{i}u,\beta+\mathrm{i}v)^n\,\mathrm{d}u\,\mathrm{d}v. \tag{2.14}$$

Remark 2.6. Alternatively, (2.14) can be obtained from (2.10) by first considering suitably tilted versions of the random variables (cf. Cramér [6]), and then passing to characteristic functions and making a Fourier inversion.

Remark 2.7. It is not hard to write an integral formula for the final probability $p_N = \sum_{m+n=N} p_{n,m}$ that we are aiming to estimate. For example, multiplying (2.12) by x^n/n and summing we see that $p_{n,m} = [x^ny^nz^m]H(x,y,z)$, where $H(x,y,z) = -\tilde{g}_0(y,z)\log(1-xg(y,z))$. Thus one can find p_N by extracting the coefficient of w^0t^N in $H(w,t/w,t)$. However, the corresponding integral does not obviously lend itself to asymptotic evaluation by methods such as those used here. Still, a direct estimate of p_N may perhaps be possible by appropriate singularity analysis.

2.2 An asymptotic estimate of $p_{n,m}$

In this section we estimate the integral (2.14) asymptotically (see Theorem 2.11 below), using parameters defined in terms of the moment generating function $f(y,z) = f_{Y,Z}(y,z) = \mathbb{E}\big(e^{yY+zZ}\big)$. Whenever f is defined and non-zero, let

$$\varphi(y,z) = \varphi_{Y,Z}(y,z) := \log f_{Y,Z}(y,z) = \log f(y,z), \tag{2.15}$$

taking the principal value of the logarithm; we shall only consider φ on domains on which $|f-1| \leqslant 1/2$. The next lemma simply states that in suitable domains, f, φ and their (partial) derivatives are all bounded.

Lemma 2.8. *There exist constants $0 < c_2 \leqslant (\log R)/2$ and $C_1^{(m)}$, $m \in \mathbb{N}$, such that if $(Y,Z) \in \mathcal{K}^0$ and $m \in \mathbb{N}$, then the following hold.*

(i) *If $\alpha,\beta,u,v \in \mathbb{R}$ with $|\alpha|,|\beta| \leqslant c_2$, then $\|D^m f(\alpha+\mathrm{i}u,\beta+\mathrm{i}v)\| \leqslant C_1^{(m)}$.*

(ii) *If, in addition, $|u|,|v| \leqslant c_2$, then $\varphi(\alpha+\mathrm{i}u,\beta+\mathrm{i}v)$ is defined, and $\|D^m\varphi(\alpha+\mathrm{i}u,\beta+\mathrm{i}v)\| \leqslant C_1^{(m)}$.*

(iii) *If $|\alpha|,|\beta| \leqslant c_2$, then $\frac{\partial}{\partial y}f(\alpha,\beta) \geqslant \delta/2$.*

Proof. (i): When $|y|,|z| \leqslant R$, then $|g(y,z)| = |\,\mathbb{E}(y^Yz^Z)| \leqslant \mathbb{E}(|y|^Y|z|^Z) \leqslant \mathbb{E}\,R^{Y+Z}$, which is at most M by assumption. Thus $|f(y,z)| \leqslant M$ when $\mathrm{Re}(y) \leqslant \log R$ and $\mathrm{Re}(z) \leqslant \log R$. Recall that $R > 1$ by assumption, so $\log R > 0$. For any $c_2 \leqslant (\log R)/2$, say, for suitable $C_1^{(m)} > 0$ statement (i) follows by standard Cauchy estimates

(ii): Let $C = C_1^{(1)}$ denote the constant from the above proof of (i). Set $c_2 := \min\{(\log R)/2, 1/(8C)\}$. Since $f(0,0) = g(1,1) = 1$, it follows from (i) that if $|\alpha|,|\beta|,|u|,|v| \leqslant c_2$, then

$$\big|f(\alpha+\mathrm{i}u,\beta+\mathrm{i}v) - 1\big| \leqslant (|\alpha+\mathrm{i}u| + |\beta+\mathrm{i}v|)C \leqslant 4c_2C \leqslant 1/2, \tag{2.16}$$

so $\varphi(\alpha+\mathrm{i}u, \beta+\mathrm{i}v)$ is defined and bounded. Furthermore, after decreasing c_2 and increasing $C_1^{(m)}$, if necessary, the bounds for the derivatives now again follow by Cauchy's estimates.

(iii): Let $\tilde{f} = \frac{\partial}{\partial y} f$. By our assumption (2.2), $\tilde{f}(0,0) = \mathbb{E}\, Y \geqslant \delta$. Furthermore, $D\tilde{f}(\alpha,\beta) = DD_1 f(\alpha,\beta) = O(1)$ for $|\alpha|, |\beta| \leqslant c_2$ by part (i). Consequently, after reducing c_2 if necessary, we have $\tilde{f}(\alpha,\beta) \geqslant \frac{1}{2}\delta$ for $|\alpha|$, $|\beta| \leqslant c_2$. □

The next lemma expresses, in a quantitative form, the unsurprising fact that if we evaluate the probability generating function $g(y,z) = g_{Y,Z}(y,z) = \mathbb{E}(y^Y z^Z)$ at y, z which are not positive real numbers, then there is significant cancellation, i.e., $|g(y,z)|$ is significantly smaller than $g(|y|,|z|)$. It will be more convenient to write this in terms of the moment generating function $f = f_{Y,Z}$ rather than g.

Lemma 2.9. *There exists a constant $c_3 > 0$ such that if $(Y,Z) \in \mathcal{K}$ and $\alpha, \beta, u, v \in \mathbb{R}$ with $|\alpha|, |\beta| \leqslant c_2$ and $|u|, |v| \leqslant \pi$, then*

$$\big|f(\alpha+\mathrm{i}u, \beta+\mathrm{i}v)\big| \leqslant f(\alpha,\beta) e^{-c_3(u^2+v^2)}. \tag{2.17}$$

Proof. Let $\pi_{k,l} := \mathbb{P}(Y = k, Z = l)$. Then

$$f(\alpha+\mathrm{i}u, \beta+\mathrm{i}v) = \sum_{k,l\geqslant 0} \pi_{k,l} e^{k(\alpha+\mathrm{i}u)+l(\beta+\mathrm{i}v)},$$

and thus $f(\alpha,\beta) > 0$. Then

$$f(\alpha,\beta)^2 - |f(\alpha+\mathrm{i}u, \beta+\mathrm{i}v)|^2 = \sum_{k,l,m,n} \pi_{k,l}\pi_{m,n} e^{(k+m)\alpha+(l+n)\beta}\Big(1 - \operatorname{Re} e^{\mathrm{i}(k-m)u+\mathrm{i}(l-n)v}\Big).$$

Each term on the right-hand side is non-negative, and considering just the cases $(k,l,m,n) = (k_1, k_2, k_1{+}1, k_2)$ and (k_1, k_2, k_1, k_2+1), recalling (2.3) we obtain

$$\begin{aligned} f(\alpha,\beta)^2 - |f(\alpha+\mathrm{i}u, \beta+\mathrm{i}v)|^2 &\geqslant \delta^2 e^{(2k_1+1)\alpha+2k_2\beta}(1-\cos u) + \delta^2 e^{2k_1\alpha+(2k_2+1)\beta}(1-\cos v) \\ &= \Omega(u^2+v^2), \end{aligned}$$

since $1 - \cos x = \Omega(x^2)$ for $|x| \leqslant \pi$. Moreover, by Lemma 2.8(i), $f(\alpha,\beta) = O(1)$. Consequently

$$1 - |f(\alpha+\mathrm{i}u, \beta+\mathrm{i}v)|^2/f(\alpha,\beta)^2 \geqslant 2c_3(u^2+v^2)$$

for some constant $c_3 > 0$, and thus

$$|f(\alpha+\mathrm{i}u, \beta+\mathrm{i}v)|^2/f(\alpha,\beta)^2 \leqslant 1 - 2c_3(u^2+v^2) \leqslant e^{-2c_3(u^2+v^2)},$$

establishing (2.17) since $f(\alpha,\beta) > 0$. □

We next establish that the symmetric bilinear form $D^2\varphi(\alpha,\beta)$ is positive-definite; a variant of the lower bound (2.18) could also be proved by first considering $D^2\varphi(0,0)$ and then using continuity. For the interpretation of $D^2\varphi(\alpha,\beta)[u,v]$, see (1.3).

Lemma 2.10. *If $(Y,Z) \in \mathcal{K}$ and $\alpha, \beta \in \mathbb{R}$ with $|\alpha|, |\beta| \leqslant c_2$, then $D^2\varphi(\alpha,\beta) \geqslant c_3 I$, i.e.,*

$$D^2\varphi(\alpha,\beta)[u,v] \geqslant c_3(u^2+v^2), \qquad u, v \in \mathbb{R}. \tag{2.18}$$

In particular, $\operatorname{Det}(D^2\varphi(\alpha,\beta)) \geqslant c_3^2$.

Proof. We first consider only $|u|, |v| \leqslant c_2$, so Lemma 2.8(ii) applies. Then the estimate (2.17) can be written

$$\operatorname{Re}\varphi(\alpha+\mathrm{i}u, \beta+\mathrm{i}v) \leqslant \varphi(\alpha,\beta) - c_3(u^2+v^2). \tag{2.19}$$

A Taylor expansion yields

$$\varphi(\alpha+\mathrm{i}u, \beta+\mathrm{i}v) = \varphi(\alpha,\beta) + \mathrm{i}D\varphi(\alpha,\beta)[u,v] - \tfrac{1}{2}D^2\varphi(\alpha,\beta)[u,v] + O\big((|u|+|v|)^3\big).$$

Since $\varphi(\alpha,\beta)$ is real for real α and β, all derivatives $D^m\varphi(\alpha,\beta)$ are real. Hence, when taking the real part, the linear term vanishes, and (2.19) implies

$$\tfrac{1}{2}D^2\varphi(\alpha,\beta)[u,v] \geqslant c_3(u^2+v^2) + O\big((|u|+|v|)^3\big).$$

Exploiting bilinearity, by replacing (u,v) with (tu,tv) and letting $t\to 0$, we now obtain (2.18) for all $u,v\in\mathbb{R}$, with room to spare.

Finally, by (1.4), note that (2.18) can be written $D^2\varphi(\alpha,\beta)\geqslant c_3 I$. This says that both eigenvalues are $\geqslant c_3$, and thus the determinant is $\geqslant c_3^2$. □

For $|\alpha|,|\beta|\leqslant c_2$, define

$$\psi(\alpha,\beta) := \varphi(\alpha,\beta) - \alpha D_1\varphi(\alpha,\beta) - \beta D_2\varphi(\alpha,\beta). \tag{2.20}$$

We are now ready to estimate the integral (2.14) for $p_{n,m}$ using a (two-dimensional) version of the saddle point method (see, e.g., [9, Chapter VIII]). We defer the problem of finding suitable (α,β) satisfying equation (2.21) to Section 2.3. Recall that $\tilde{f}_0(y,z) = \frac{\partial}{\partial y} f_{Y^0,Z^0}(y,z) = \mathbb{E}\big(Y^0 e^{yY^0+zZ^0}\big)$, see (2.13).

Theorem 2.11. *Suppose that $(Y^0,Z^0)\in\mathcal{K}^0$ and $(Y,Z)\in\mathcal{K}$. Suppose further that $n\geqslant 1$, $m\geqslant 0$ are integers and that α, β are real numbers with $|\alpha|,|\beta|\leqslant c_2$ such that*

$$D\varphi(\alpha,\beta) = (1, m/n). \tag{2.21}$$

Then

$$p_{n,m} = n^{-2}e^{n\psi(\alpha,\beta)}\Big((2\pi)^{-1}\tilde{f}_0(\alpha,\beta)\operatorname{Det}\big(D^2\varphi(\alpha,\beta)\big)^{-1/2} + O(n^{-1})\Big), \tag{2.22}$$

where the implicit constant depends only on the parameters R, M, k_1, k_2, δ of $\mathcal{K}^0$ and $\mathcal{K}$.

Proof. We write (2.14) as

$$p_{n,m} = \frac{1}{4\pi^2 n} e^{-n\alpha - m\beta} f(\alpha,\beta)^n I_1, \tag{2.23}$$

where

$$I_1 := \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} e^{-\mathrm{i}nu-\mathrm{i}mv}\tilde{f}_0(\alpha+\mathrm{i}u,\beta+\mathrm{i}v)\left(\frac{f(\alpha+\mathrm{i}u,\beta+\mathrm{i}v)}{f(\alpha,\beta)}\right)^n \mathrm{d}u\,\mathrm{d}v. \tag{2.24}$$

Using assumption (2.21) we have $\psi(\alpha,\beta) = \varphi(\alpha,\beta) - \alpha - \beta m/n$, so

$$e^{-n\alpha-m\beta} f(\alpha,\beta)^n = e^{-n\alpha-m\beta+n\varphi(\alpha,\beta)} = e^{n\psi(\alpha,\beta)}. \tag{2.25}$$

We shall estimate (2.24) using Laplace's method (in two dimensions), cf. e.g. [9, Appendix B.6]. Roughly speaking, the idea is as follows. We view the integrand as a product of a term independent of n with a term that is exponential in n. As we shall see, the condition (2.21) ensures that the exponent has a stationary point, in fact a maximum, at $u=v=0$. It turns out that the main contribution is near to this point, and here the exponent may be approximated by a quadratic, leading to a (two-dimensional) Gaussian integral.

Applying Lemma 2.8(i) to (Y^0,Z^0) shows that $\tilde{f}_0(\alpha,\beta) = O(1)$. Since $\operatorname{Det}\big(D^2\varphi(\alpha,\beta)\big) = \Omega(1)$ by Lemma 2.10, and $\psi(\alpha,\beta) = O(1)$ by (2.20) and Lemma 2.8(ii), the conclusion (2.22) holds for any fixed n simply by taking the implicit constant large enough. Thus we may assume that n is at least any given constant n_0, and in particular that $n^{-0.4}\leqslant c_2$.

Applying Lemma 2.8(i) to (Y^0,Z^0) also shows that $\tilde{f}_0(\alpha+\mathrm{i}u,\beta+\mathrm{i}v) = O(1)$. Hence, if $|u|\geqslant n^{-0.4}$ or $|v|\geqslant n^{-0.4}$, then by Lemma 2.9 the integrand in (2.24) is $O\big(e^{-c_3 n\cdot n^{-0.8}}\big) = O\big(e^{-c_3 n^{0.2}}\big) = O\big(n^{-99}\big)$. On the other hand, if $|u|$, $|v|\leqslant n^{-0.4}$ then, since $n^{-0.4}\leqslant c_2$, Lemma 2.8(ii) shows that $\varphi(\alpha+\mathrm{i}u,\beta+\mathrm{i}v)$ is defined and we obtain

$$I_1 = O\big(n^{-99}\big) + I_2, \tag{2.26}$$

with

$$I_2 := \int_{-n^{-0.4}}^{n^{-0.4}}\int_{-n^{-0.4}}^{n^{-0.4}} \tilde{f}_0(\alpha+\mathrm{i}u,\beta+\mathrm{i}v)e^{n(\varphi(\alpha+\mathrm{i}u,\beta+\mathrm{i}v)-\varphi(\alpha,\beta))-\mathrm{i}nu-\mathrm{i}mv}\,\mathrm{d}u\,\mathrm{d}v. \tag{2.27}$$

Considering a Taylor expansion of φ around (α,β), and noting that the linear terms cancel by our assumption (2.21), we have

$$n(\varphi(\alpha+\mathrm{i}u,\beta+\mathrm{i}v)-\varphi(\alpha,\beta))-\mathrm{i}nu-\mathrm{i}mv \\ = -n\tfrac{1}{2}D^2\varphi(\alpha,\beta)[u,v]-n\tfrac{\mathrm{i}}{6}D^3\varphi(\alpha,\beta)[u,v]+O\big(n(|u|+|v|)^4\big), \quad (2.28)$$

where we used Lemma 2.8(ii) to bound the error term. For $|u|$, $|v|\leqslant n^{-0.4}$, note that Lemma 2.8(ii) implies $nD^3\varphi(\alpha,\beta)[u,v]=O(n(|u|+|v|)^3)=O(n^{-0.2})=O(1)$, and $O(n(|u|+|v|)^4)=O(n^{-0.6})=O(1)$. Hence, writing for brevity

$$Q:=D^2\varphi(\alpha,\beta),$$

the exponential factor in (2.27) is

$$e^{-\frac{1}{2}nQ[u,v]}\exp\Big(-n\tfrac{\mathrm{i}}{6}D^3\varphi(\alpha,\beta)[u,v]+O\big(n(|u|+|v|)^4\big)\Big) \\ = e^{-\frac{1}{2}nQ[u,v]}\Big(1-n\tfrac{\mathrm{i}}{6}D^3\varphi(\alpha,\beta)[u,v]+O\big(n(u^4+v^4)\big)+O\big(n^2(u^6+v^6)\big)\Big). \quad (2.29)$$

Recalling $\tilde f_0=\frac{\partial}{\partial y}f_{Y^0,Z^0}$, using Lemma 2.8(i) we also have the Taylor expansion

$$\tilde f_0(\alpha+\mathrm{i}u,\beta+\mathrm{i}v)=\tilde f_0(\alpha,\beta)+\mathrm{i}D\tilde f_0(\alpha,\beta)[u,v]+O\big((|u|+|v|)^2\big). \quad (2.30)$$

Multiplying together (2.29) and (2.30), the integrand in (2.27) is thus

$$e^{-\frac{1}{2}nQ[u,v]}\Big(\tilde f_0(\alpha,\beta)+\mathrm{i}D\tilde f_0(\alpha,\beta)[u,v]-n\tilde f_0(\alpha,\beta)\tfrac{\mathrm{i}}{6}D^3\varphi(\alpha,\beta)[u,v] \\ +O\big(u^2+v^2\big)+O\big(n(u^4+v^4)\big)+O\big(n^2(u^6+v^6)\big)\Big). \quad (2.31)$$

When we integrate, the terms with $D\tilde f_0$ and $D^3\varphi$ are odd functions of (u,v) so their integrals vanish. Hence,

$$I_2=\int_{-n^{-0.4}}^{n^{-0.4}}\int_{-n^{-0.4}}^{n^{-0.4}}e^{-\frac{1}{2}nQ[u,v]}\Big(\tilde f_0(\alpha,\beta)+O\big(u^2+v^2\big)+O\big(n(u^4+v^4)\big)+O\big(n^2(u^6+v^6)\big)\Big)\,\mathrm{d}u\,\mathrm{d}v.$$

Recalling that $Q=D^2\varphi(\alpha,\beta)$, by Lemma 2.10 we have $Q[u,v]=\Omega(u^2+v^2)$. Since for $k\in\{1,2,3\}$ we have $\iint_{\mathbb{R}^2}e^{-a(u^2+v^2)}(u^2+v^2)^k\,\mathrm{d}u\,\mathrm{d}v=O(a^{-(k+1)})$, it follows that

$$I_2=\tilde f_0(\alpha,\beta)\int_{-n^{-0.4}}^{n^{-0.4}}\int_{-n^{-0.4}}^{n^{-0.4}}e^{-\frac{1}{2}nQ[u,v]}\,\mathrm{d}u\,\mathrm{d}v+O(n^{-2}).$$

Since $Q=D^2\varphi(\alpha,\beta)$ is symmetric and positive-definite by Lemma 2.10, we have the following standard Gaussian integral over $\mathbb{R}^2$:

$$\iint_{\mathbb{R}^2}e^{-nQ[u,v]/2}\,\mathrm{d}u\,\mathrm{d}v=n^{-1}\cdot 2\pi(\mathrm{Det}(Q))^{-1/2}. \quad (2.32)$$

Since $Q[u,v]=\Omega(u^2+v^2)$, the contribution of the range $\max\{|u|,|v|\}\geqslant n^{-0.4}$ to the above integral (2.32) is again exponentially small. Hence

$$I_2=\tilde f_0(\alpha,\beta)\cdot n^{-1}2\pi(\mathrm{Det}(Q))^{-1/2}+O(n^{-2}). \quad (2.33)$$

The result follows by combining (2.23), (2.25), (2.26) and (2.33). □

We next estimate the exponent in (2.22), without assuming that equation (2.21) holds.

Lemma 2.12. *There exists a constant $0<c_4\leqslant c_2$ such that if $(Y,Z)\in\mathcal{K}$ and $\alpha,\beta\in\mathbb{R}$ with $|\alpha|,|\beta|\leqslant c_4$, then*

$$\psi(\alpha,\beta)\leqslant -\tfrac{1}{4}c_3(\alpha^2+\beta^2). \quad (2.34)$$

Moreover, $\psi(0,0)=0$, $D\psi(0,0)=0$ and $D^2\psi(0,0)\leqslant -c_3I$.

Proof. We have $\psi(0,0) = \varphi(0,0) = 0$. Furthermore, differentiating (2.20) yields

$$D_1\psi(\alpha,\beta) = -\alpha D_{11}\varphi(\alpha,\beta) - \beta D_{12}\varphi(\alpha,\beta),$$
$$D_2\psi(\alpha,\beta) = -\alpha D_{12}\varphi(\alpha,\beta) - \beta D_{22}\varphi(\alpha,\beta),$$

and thus $D\psi(0,0) = (0,0)$. Differentiating again shows that $D_{ij}\psi(0,0) = -D_{ij}\varphi(0,0)$ for all $i,j \in \{1,2\}$. Hence, using Lemma 2.10,

$$D^2\psi(0,0) = -D^2\varphi(0,0) \leqslant -c_3 I.$$

Moreover, it follows from Lemma 2.8(ii) that $\|D^3\psi(\alpha,\beta)\| = O(1)$ for $|\alpha|, |\beta| \leqslant c_2$. Consequently, a Taylor expansion yields (2.34) for c_4 sufficiently small. □

2.3 Summing $p_{n,m}$: proof of Theorem 2.2

In this section we prove Theorem 2.2 by summing several different estimates of the point probabilities in

$$\mathbb{P}(|\mathfrak{X}| = N) = \sum_{n=0}^{N} p_{n,N-n}. \tag{2.35}$$

Throughout we consider, as in (2.22), only real inputs α, β for the various functions f, φ etc. Thus, all relevant functions are treated as mapping from (subdomains in) $\mathbb{R}^n$ to $\mathbb{R}^m$ for suitable n, m.

An individual of type L has on average $\mathbb{E}\,Y$ children of type L and $\mathbb{E}\,Z$ children of type S. So, in the near-critical case $\mathbb{E}\,Y \approx 1$, we expect that the overall fraction of type L individuals in $\mathfrak{X}$ should be close to

$$x_0 := 1/(1 + \mathbb{E}\,Z).$$

This suggests that the contribution from terms in (2.35) with n/N far from x_0 will be negligible, and we shall later confirm this by standard Chernoff-like estimates. Below our main focus is thus on the terms where n/N is close to x_0. Here the plan is to rewrite the asymptotic estimate (2.22) for $p_{n,N-n}$ using the following version of the inverse function theorem, where we explicitly state uniformity for a set of functions. We define

$$B_r^d := \{x \in \mathbb{R}^d : |x| < r\} \qquad \text{and} \qquad B_r := B_r^2 = \{x \in \mathbb{R}^2 : |x| < r\}.$$

Lemma 2.13 (Inverse function theorem)**.** *Let $d \geqslant 1$ be an integer and $r > 0$ a real number. For every $0 < A < \infty$, there exist $\sigma > 0$ and $0 < r_1 < r$, both depending only on A, r, such that if $F : B_r^d \to \mathbb{R}^d$ is twice continuously differentiable and satisfies*

(i) $F(0) = 0$,

(ii) *$DF(0)$ is invertible and $\|DF(0)^{-1}\| \leqslant A$, and*

(iii) *$\|D^2F(x)\| \leqslant A$ for all $x \in B_r^d$,*

then there exists a twice continuously differentiable function $G : B_\sigma^d \to B_r^d$ with $G(0) = 0$ and $F(G(y)) = y$ for $y \in B_\sigma^d$. Furthermore, for each $y \in B_\sigma^d$, $x = G(y)$ is the unique $x \in \mathbb{R}^d$ with $|x| \leqslant r_1$ such that $F(x) = y$. Moreover, $\|DG(y)\| = O(1)$ and $\|D^2G(y)\| = O(1)$, uniformly for $y \in B_\sigma^d$ and all such F, and if F is infinitely differentiable or (real) analytic, then so is G.

Proof. This follows by a standard proof of the inverse function theorem; we give some details for completeness.

First, let $r_1 := \frac{1}{2}\min\{r, A^{-2}\}$. If $|x| \leqslant r_1$, then by the mean-value theorem $\|DF(x) - DF(0)\| \leqslant A|x| \leqslant Ar_1$. Hence, $\|DF(0)^{-1}DF(x) - I\| \leqslant A^2 r_1 \leqslant \frac{1}{2}$, and thus $DF(0)^{-1}DF(x)$ is invertible and its inverse has norm at most 2 (e.g., by the von Neumann series representation of the inverse). Consequently, $DF(x)$ is invertible and

$$\|DF(x)^{-1}\| \leqslant \|DF(0)^{-1}\| \cdot \|(DF(0)^{-1}DF(x))^{-1}\| \leqslant 2A \qquad \text{for } |x| \leqslant r_1. \tag{2.36}$$

Next, let $\sigma := r_1/(2A)$. If $|y| < \sigma$, define inductively $x_0 := 0$ and $x_{n+1} := \Gamma(x_n)$, where

$$\Gamma(x) := x + DF(0)^{-1}(y - F(x)).$$

Using $\|DF(0)^{-1}\| \leqslant A$ and that $\|D\Gamma(x)\| = \|I - DF(0)^{-1}DF(x)\| \leqslant \frac{1}{2}$ if $|x| \leqslant r_1$, it is easy to show by induction that $|x_n| \leqslant (1-2^{-n})r_1$ and $|x_{n+1} - x_n| \leqslant 2^{-n}A\sigma \leqslant 2^{-n-1}r_1$. Hence x_n is defined for all $n \geqslant 0$, and converges to some x with $|x| \leqslant r_1 < r$. Furthermore, $y - F(x_n) = DF(0)(x_{n+1} - x_n) \to 0$ as $n \to \infty$, and thus by continuity $F(x) = y$. Define $G(y) := x$.

This shows that the inverse function G exists in B^d_σ. The uniqueness statement is immediate, since any $x \in \mathbb{R}^d$ satisfying $F(x) = y$ is a fixed point of $\Gamma(x)$, which is a contraction for $|x| \leqslant r_1$. Differentiability (and analyticity when F is analytic) follows in the usual way (or by appealing to a standard version of the inverse function theorem, locally at $G(y)$). Finally, $DG(y) = DF(x)^{-1}$, and thus $\|DG(y)\| \leqslant 2A$ by (2.36). Another differentiation (using the chain rule) then yields $\|D^2G(y)\| = O(1)$. $\square$

Our next aim is to construct an (implicit) solution $(\alpha,\beta) = h(n/N)$ to equation (2.21) when $N = n+m$ and n/N is close to $x_0 = 1/(1+\mathbb{E}\,Z)$. We start by applying Lemma 2.13 to the function $F : B_{c_4} \to \mathbb{R}^2$ defined by

$$F(\alpha,\beta) := \Big(\mathbb{E}\,Y - D_1\varphi(\alpha,\beta),\ \frac{1}{1+D_2\varphi(\alpha,\beta)} - x_0\Big). \tag{2.37}$$

Note that $D_2\varphi(\alpha,\beta) = D_2f(\alpha,\beta)/f(\alpha,\beta) \geqslant 0$, and thus $F(\alpha,\beta)$ is well-defined. Furthermore, $D\varphi(0,0) = (\mathbb{E}\,Y, \mathbb{E}\,Z)$, and thus $F(0,0) = (0,0)$. Moreover, using matrix form (where the first column is $\frac{\partial}{\partial\alpha}$ of the vector valued function F and the second is $\frac{\partial}{\partial\beta}$), we have

$$DF = \begin{pmatrix} -1 & 0 \\ 0 & -(1+D_2\varphi)^{-2} \end{pmatrix} D^2\varphi. \tag{2.38}$$

It follows from Lemma 2.10 that $\big\|(D^2\varphi(\alpha,\beta))^{-1}\big\| = O(1)$, and then (2.38) together with Lemma 2.8 yields

$$\|DF(\alpha,\beta)^{-1}\| = O(1). \tag{2.39}$$

Lemma 2.8 also implies $\|D^2F(\alpha,\beta)\| = O(1)$. Consequently, Lemma 2.13 applies (with $d = 2$) and yields a constant $\sigma = c_5 > 0$ and a function $G : B_{c_5} \to B_{c_4}$ such that

$$F(G(y)) = y \qquad \text{for } y \in B_{c_5}. \tag{2.40}$$

Recall that $x_0 = 1/(1+\mathbb{E}\,Z)$. Since $\mathbb{E}\,Z = D_2f(0,0) = O(1)$ by Lemma 2.8 and $\mathbb{E}\,Z \geqslant \delta > 0$ by (2.4), there exists a constant $c > 0$ such that $c \leqslant x_0 \leqslant 1-c$. Let

$$c_6 := \tfrac{1}{2}\min\{c_5,\ c\}.$$

Suppose that $|\,\mathbb{E}\,Y - 1| < c_6$. If also $|x - x_0| < c_6$, then $\big(\mathbb{E}\,Y - 1, x - x_0\big) \in B_{c_5}$; we then define

$$h(x) := G(\mathbb{E}\,Y - 1, x - x_0) \in B_{c_4}. \tag{2.41}$$

Furthermore, $|x - x_0| < c_6 \leqslant \frac{1}{2}c \leqslant \frac{1}{2}x_0$ implies $x \geqslant \frac{1}{2}x_0 \geqslant c_6$ and $1 - x \geqslant c_6$. Now suppose that $0 < n \leqslant N$ and that $|n/N - x_0| < c_6$, and let $m := N - n$ and $(\alpha,\beta) := h\big(n/N\big)$. Then, by (2.41) and (2.40),

$$F(\alpha,\beta) = \Big(\mathbb{E}\,Y - 1,\ \frac{n}{N} - x_0\Big) = \Big(\mathbb{E}\,Y - 1,\ \frac{1}{1+m/n} - x_0\Big). \tag{2.42}$$

Definition (2.37) shows that (2.21) holds. Hence, by Theorem 2.11, (2.22) holds. For $|x - x_0| < c_6$ define

$$\Psi(x) := x\psi(h(x)), \tag{2.43}$$

$$\Phi(x) := (2\pi)^{-1}x^{-2}\tilde{f}_0(h(x))\,\mathrm{Det}\big(D^2\varphi(h(x))\big)^{-1/2}. \tag{2.44}$$

Recall that $h(x) \in B_{c_4} \subseteq B_{c_2}$, and note that Lemma 2.10 implies $\mathrm{Det}\big(D^2\varphi(h(x))\big) \geqslant c_3^2$; thus $\Psi(x)$ and $\Phi(x)$ are well-defined. Then, still assuming $|\,\mathbb{E}\,Y - 1| < c_6$, $|n/N - x_0| < c_6$ and $(\alpha,\beta) := h\big(n/N\big)$, we see that (2.22) can be written

$$p_{n,N-n} = N^{-2}e^{N\Psi(n/N)}\big(\Phi(n/N) + O(N^{-1})\big). \tag{2.45}$$

(Here, we use $|n/N - x_0| < c_6 \leqslant x_0/2$ to bound $n \geqslant c_6N$, so an $O(n^{-1})$ error term is $O(N^{-1})$.)

We next show that, in the relevant domains, the functions Φ, Ψ and their (partial) derivatives are all bounded.

Lemma 2.14. *For each $m \geqslant 0$, there exists a constant $C_2^{(m)}$ such that if $|\mathbb{E}\,Y - 1| < c_6$ and $|x - x_0| < c_6$, then $\|D^m\Phi(x)\| \leqslant C_2^{(m)}$ and $\|D^m\Psi(x)\| \leqslant C_2^{(m)}$.*

Proof. We saw in the proof of Lemma 2.13 that $DG(y) = \big(DF(G(y))\big)^{-1}$, which is bounded for $y \in B_{c_5}$ by (2.36). By further differentiations, using the chain rule, Lemma 2.8(ii) and induction, it follows that for each $m \geqslant 0$,

$$\|D^m G(\mathbb{E}\,Y - 1, x - x_0)\| = O(1) \tag{2.46}$$

when $|\mathbb{E}\,Y - 1| < c_6$ and $|x - x_0| < c_6$. Hence the definition (2.41) yields $|D^m h(x)| = O(1)$, and the result follows by (2.43)–(2.44) together with the chain rule and Lemmas 2.8 and 2.10. $\square$

Note for later than since $G(0) = 0$ and $\|DG(y)\| = O(1)$ in B_{c_6}, we have

$$|G(w, x - x_0)| = O(|(w, x - x_0)|) \tag{2.47}$$

if $(w, x - x_0) \in B_{c_6}$.

We now analyze the exponential term $e^{N\Psi(n/N)}$ of the formula (2.45) for $p_{n,N-n}$, which is valid for $|n/N - x_0| < c_6$. The next result in particular implies that $\Psi(x) \leqslant 0$ is a concave function with a unique maximizer x^* close to x_0. As we shall see, this essentially means that the dominant contribution to the sum of the $p_{n,N-n}$ comes from the terms with n/N close to x^*, which is in turn close to x_0.

Lemma 2.15. *There exist constants $c_7, c_8 > 0$ with $c_8 \leqslant c_7 < \frac{1}{3}c_6$ such that if $|\mathbb{E}\,Y - 1| \leqslant c_8$, then the following hold.*

(i) *If $x \in \mathbb{R}$ with $|x - x_0| \leqslant 3c_7$, then*

$$\Psi(x) = -\Omega\big(|\mathbb{E}\,Y - 1|^2 + |x - x_0|^2\big). \tag{2.48}$$

(ii) *There exists $x^* \in \mathbb{R}$ with $|x^* - x_0| = O(|\mathbb{E}\,Y - 1|)$ and $|x^* - x_0| < c_7$ such that $\Psi'(x^*) = 0$.*

(iii) *$\Psi''(x) = -\Omega(1)$ for every x with $|x - x_0| \leqslant 3c_7$.*

(iv) *$\Phi(x^*) = \Omega(1)$.*

As a consequence, x^* is the unique maximum point of $x \mapsto \Psi(x)$ in $[x_0 - 3c_7, x_0 + 3c_7]$.

Proof. For $|w|, |x - x_0| \leqslant c_6$, let

$$\widehat{\Psi}(w, x) := x\psi(G(w, x - x_0)), \tag{2.49}$$

so that $\Psi(x) = \widehat{\Psi}(\mathbb{E}\,Y - 1, x)$. In the proofs below we assume that c_7 and c_8 are positive constants, chosen later, with $c_8 \leqslant c_7 < \frac{1}{3}c_6$, and that $|w| \leqslant c_8$ and $|x - x_0| \leqslant 3c_7$.

(i): Since $|w| + |x - x_0| \leqslant 4c_7 < 2c_6 \leqslant c_5$, and G maps B_{c_5} into B_{c_4}, we have

$$|G(w, x - x_0)| < c_4. \tag{2.50}$$

Since $F(0) = 0$ and $\|DF(y)\| = O(1)$ in B_{c_4}, using $(w, x - x_0) = F(G(w, x - x_0))$ we also have $|(w, x - x_0)| = O(|G(w, x - x_0)|)$. This and Lemma 2.12 imply

$$\psi(G(w, x - x_0)) = -\Omega(|G(w, x - x_0)|^2) = -\Omega(|(w, x - x_0)|^2).$$

Furthermore, as remarked above, $|x - x_0| \leqslant 3c_7 < c_6$ implies $x \geqslant c_6$. Hence, recalling (2.49),

$$\widehat{\Psi}(w, x) = -\Omega(|(w, x - x_0)|^2) = -\Omega\big(w^2 + (x - x_0)^2\big), \tag{2.51}$$

which yields (2.48) since $\Psi(x) = \widehat{\Psi}(\mathbb{E}\,Y - 1, x)$.

(iii): Using $G(0) = 0$, which is shorthand for $G(0, 0) = (0, 0)$, we have

$$\widehat{\Psi}(0, x_0) = x_0\psi(G(0, 0)) = x_0\psi(0, 0) = 0. \tag{2.52}$$

Together with (2.51), it follows that, for some constant $c > 0$,

$$D\widehat{\Psi}(0, x_0) = 0, \tag{2.53}$$

$$D^2\widehat{\Psi}(0, x_0) \leqslant -cI. \tag{2.54}$$

The same proof as for Lemma 2.14 shows that

$$D^m\widehat{\Psi}(w, x) = O(1) \tag{2.55}$$

for every fixed $m \geqslant 0$. Using (2.55) with $m = 3$ and (2.54), we see that if c_7 and hence $c_8 \leqslant c_7$ is small enough, then

$$D^2\widehat{\Psi}(w, x) \leqslant -\tfrac{c}{2}I$$

when $|w| \leqslant c_8$ and $|x - x_0| \leqslant 3c_7$. In particular, recalling $\Psi(x) = \widehat{\Psi}(\mathbb{E}\,Y - 1, x)$, by taking $w = \mathbb{E}\,Y - 1$ we have

$$\Psi''(x) \leqslant -\tfrac{c}{2}. \tag{2.56}$$

(ii): Similarly, (2.53) and (2.55) with $m = 2$ imply that $D\widehat{\Psi}(w, x) = O(|w| + |x - x_0|)$. In particular, $\Psi'(x_0) = D_2\widehat{\Psi}(\mathbb{E}\,Y-1, x_0) = O(|\,\mathbb{E}\,Y-1|)$. Hence we may choose c_8 sufficiently small such that $|\,\mathbb{E}\,Y-1| \leqslant c_8$ implies $|\Psi'(x_0)| \leqslant cc_7/3$. Then the mean value theorem and (2.56) imply $\Psi'(x_0-c_7) > 0$ and $\Psi'(x_0+c_7) < 0$, so $\Psi'(x^*) = 0$ for some $x^* \in (x_0 - c_7, x_0 + c_7)$. Moreover, by the mean value theorem and (2.56) we also have $|x^* - x_0| \leqslant \frac{2}{c}|\Psi'(x_0)|$, so (ii) holds.

(iv): Since $|x^* - x_0| \leqslant c_7 < c_6$, by (2.50) and the definition (2.41) of h we have $|h(x^*)| \leqslant c_4 \leqslant c_2$, so Lemma 2.8(iii) applied to (Y^0, Z^0) gives $\tilde{f}_0(h(x^*)) \geqslant \frac{1}{2}\delta$. The other factors in (2.44) are bounded below, using $x^* \leqslant x_0+c_7 \leqslant 1+c_7$ and Hadamard's inequality together with Lemma 2.8(ii), and thus (iv) follows. □

The following technical lemma will be useful for expanding the sum of the $p_{n,N-n}$ estimates (2.45) around $n/N \approx x^*$ (it is easy to give a much more precise formula for T_{2j}, but we do not need this).

Lemma 2.16. *For $a > 0$, $y \in \mathbb{R}$ and an integer $j \geqslant 0$, let*

$$T_j = T_j(a, y) := \sum_{n\in\mathbb{Z}} (n - y)^j e^{-a(n-y)^2}. \tag{2.57}$$

Then, uniformly for all $0 < a \leqslant 1$ and $y \in \mathbb{R}$,

$$T_0 = \sqrt{\frac{\pi}{a}} + O\bigl(a^{-1/2}e^{-\pi^2/a}\bigr), \tag{2.58}$$

and for every fixed integer $i \geqslant 0$,

$$T_{2i} = O\bigl(a^{-i-1/2}\bigr), \tag{2.59}$$

$$T_{2i+1} = O\bigl(a^{-i-3/2}e^{-\pi^2/a}\bigr). \tag{2.60}$$

Proof. We first consider $T_0 = \sum_{n\in\mathbb{Z}} e^{-a(n-y)^2}$. Applying the well-known Poisson summation formula [26, (II.13.4) or (II.13.14)] and then using the Gaussian integral $\int_{-\infty}^{\infty} e^{-(ax^2+bx+c)}\,\mathrm{d}x = \sqrt{\frac{\pi}{a}}e^{b^2/(4a)-c}$, a short standard calculation yields the identity

$$T_0 = \sum_{n\in\mathbb{Z}} \int_{-\infty}^{\infty} e^{-a(x-y)^2} e^{-2\pi i n x}\,\mathrm{d}x = \sqrt{\frac{\pi}{a}} \sum_{n\in\mathbb{Z}} e^{-\pi^2 n^2/a - 2\pi i n y}, \tag{2.61}$$

which for $a \leqslant 1$, say, implies (2.58). (In fact, (2.61) is equivalent to a well-known identity for the theta function θ_3, see [16, (20.7.32)].)

Moreover, taking the partial derivative of (2.57) with respect to y we obtain

$$\frac{\partial}{\partial y} T_j(a, y) = -jT_{j-1} + 2aT_{j+1}. \tag{2.62}$$

In particular, $2aT_1 = \frac{\partial}{\partial y}T_0$, and termwise differentiation of the right-hand side in (2.61) (noting that the main term, $n = 0$, is constant) yields

$$T_1 = O\big(a^{-3/2}e^{-\pi^2/a}\big).$$

Repeated differentiation of (2.62) and induction now yield (2.59) and (2.60). □

We also have to estimate the sum of the $p_{n,N-n}$ in (2.35) where n/N is far from x_0. Based on simple Chernoff-type arguments, the next result shows that their contribution is negligible.

Lemma 2.17. *If $|n/N - x_0| \geqslant c_7$, then $p_{n,N-n} \leqslant e^{-\Omega(N)}$.*

Proof. For any $u, v > 0$, from (2.12) we have

$$np_{n,m} \leqslant u^{-n}v^{-m}\tilde{g}_0(u,v)g(u,v)^n. \tag{2.63}$$

Take $u = 1$ and $v = e^t$, with $|t| \leqslant \log R$, and define

$$\gamma(t) := e^{-t\,\mathbb{E}\,Z}g(1,e^t) = \mathbb{E}\,e^{t(Z-\mathbb{E}\,Z)}.$$

For any $0 \leqslant n \leqslant N$, (2.63) yields

$$np_{n,N-n} \leqslant e^{t(n-N)+tn\,\mathbb{E}\,Z}\tilde{g}_0(1,e^t)\gamma(t)^n. \tag{2.64}$$

Note that $\gamma(0) = 1$ and $\gamma'(0) = 0$. Since $g(1,e^t) = f(0,t)$ and $\mathbb{E}\,Z = D_2 f(0,0)$, by Lemma 2.8(i) there is a constant $C_3 > 0$ such that $\gamma''(t) \leqslant C_3$ whenever $|t| \leqslant c_2$, and so

$$\gamma(t) \leqslant 1 + C_3t^2 \leqslant e^{C_3t^2}. \tag{2.65}$$

By assumption, $|n - Nx_0| \geqslant c_7N$. Recalling that $x_0 = 1/(1+\mathbb{E}\,Z)$ and $\mathbb{E}\,Z \geqslant 0$, it follows that

$$|t(n - N + n\,\mathbb{E}\,Z)| = |t|\cdot|n(1+\mathbb{E}\,Z) - N| \geqslant |t|\cdot c_7N(1+\mathbb{E}\,Z) \geqslant c_7|t|N. \tag{2.66}$$

We now choose $t = \pm c$ where $c := \min\{\frac{1}{2}c_7/C_3, c_2\}$, and the sign is such that $t(n - N + n\,\mathbb{E}\,Z) < 0$. Using (2.64)–(2.66) and $n \leqslant N$, we infer

$$np_{n,N-n} \leqslant \tilde{g}_0(1,e^t)\cdot e^{-c_7|t|N + C_3t^2N} \leqslant O(1)\cdot e^{-c_7cN/2},$$

completing the proof for $n \geqslant 1$.

Finally, in the remaining case $n = 0$ we have $|\mathfrak{X}^S| = Z^0$, since $|\mathfrak{X}^L| = 0$ if and only if $Y^0 = 0$. Hence

$$p_{0,N} = \mathbb{P}(Y^0 = 0, Z^0 = N) \leqslant g_{Y^0,Z^0}(1,R)\cdot R^{-N} = O(1)\cdot R^{-N},$$

completing the proof (since $R > 1$). □

We are now ready to prove Theorem 2.2.

Proof of Theorem 2.2. We suppose throughout that $c_1 \leqslant c_8$ and that $|\,\mathbb{E}\,Y - 1| \leqslant c_1$.

We start by considering the quantities ξ and θ defined in (2.6) and (2.7). By Lemma 2.15, $\Psi(x)$ has a local maximum point $x^* \in (x_0 - c_7, x_0 + c_7)$. As in (2.6) and (2.7), let

$$\xi := -\Psi(x^*) \qquad \text{and} \qquad \theta := \sqrt{2\pi/|\Psi''(x^*)|}\;\Phi(x^*).$$

By Lemmas 2.14 and 2.15(iii), $\Psi''(x^*) = -\Theta(1)$. By (2.48) we have $\xi = -\Psi(x^*) = \Omega(|\,\mathbb{E}\,Y - 1|^2)$. Recalling that $\Psi(x^*) = \widehat{\Psi}(\mathbb{E}\,Y - 1, x^*)$, see (2.49), by combining (2.52), (2.53) and (2.55) (with $m = 2$) together with Lemma 2.15(ii), it follows that

$$\xi = |\widehat{\Psi}(\mathbb{E}\,Y - 1, x^*)| = O\big(|\,\mathbb{E}\,Y - 1|^2 + |x^* - x_0|^2\big) = O\big(|\,\mathbb{E}\,Y - 1|^2\big). \tag{2.67}$$

Hence $\xi = \Theta\big(|\,\mathbb{E}\,Y - 1|^2\big)$, as claimed. That $\theta = \Theta(1)$ follows from the bound $|\Psi''(x^*)| = \Theta(1)$ above and Lemmas 2.14 and 2.15(iv), which give $\Phi(x^*) = \Theta(1)$.

Since ξ and θ, which do not depend on N, are both $O(1)$, for any fixed N, (2.5) holds trivially simply by taking the implicit constant large enough. Thus we may assume throughout that $N^{-0.4} \leqslant c_7$.

We have $\mathbb{P}(|\mathfrak{X}| = N) = \sum_{n=0}^{N} p_{n,N-n}$. We estimate this sum by Laplace's method, similarly to the argument in the proof of Theorem 2.11, but now for a sum instead of a two-dimensional integral.

We consider first n such that $|n/N - x^*| < N^{-0.4}$, which includes the main terms in the sum. Suppose that $|x - x^*| < N^{-0.4}$. Using Lemma 2.14, a Taylor expansion then yields, cf. (2.28),

$$N\Psi(x) + N\xi = N\Psi(x) - N\Psi(x^*) = N\tfrac{1}{2}\Psi''(x^*)(x-x^*)^2 + N\tfrac{1}{6}\Psi'''(x^*)(x-x^*)^3 + O\big(N|x-x^*|^4\big),$$

which by exponentiation and a Taylor expansion of $\Phi(x)$ yields, cf. (2.31),

$$e^{N\Psi(x)}\Phi(x) = e^{-N\xi + N\frac{1}{2}\Psi''(x^*)(x-x^*)^2}\Big(\Phi(x^*) + \Phi'(x^*)(x-x^*) + N\Phi(x^*)\tfrac{1}{6}\Psi'''(x^*)(x-x^*)^3 + O\big(|x-x^*|^2 + N|x-x^*|^4 + N^2|x-x^*|^6\big)\Big).$$

Similar, but simpler, reasoning also shows that if $|x - x^*| < N^{-0.4}$, then

$$e^{N\Psi(x)}N^{-1} = e^{-N\xi + N\frac{1}{2}\Psi''(x^*)(x-x^*)^2} \cdot O(N^{-1}).$$

Consequently, since $\Psi''(x^*) \leqslant 0$, if we define

$$S_j := \sum_{|n/N - x^*| < N^{-0.4}} \Big(\frac{n}{N} - x^*\Big)^j e^{-\frac{1}{2}|\Psi''(x^*)|(n-Nx^*)^2/N}, \tag{2.68}$$

then (2.45) yields

$$\sum_{|n/N-x^*|<N^{-0.4}} p_{n,N-n} = N^{-2}e^{-N\xi}\Big(\Phi(x^*)S_0 + \Phi'(x^*)S_1 + N\Phi(x^*)\tfrac{1}{6}\Psi'''(x^*)S_3 + O\big(S_2 + NS_4 + N^2S_6\big) + O\big(N^{-1}S_0\big)\Big). \tag{2.69}$$

(The odd sums S_1 and S_3 do not vanish as the corresponding integrals in the proof of Theorem 2.11 do, but we shall see that they are exponentially small.) Recall (from the start of the proof) that $\Psi''(x^*) = -\Theta(1)$. It follows that if we extend the summation in the definition (2.68) to all $n \in \mathbb{Z}$, and denote the result by S'_j, then $S_j - S'_j$ is $O\big(e^{-\Omega(N^{0.2})}\big)$ for each fixed j. Let $a = |\Psi''(x^*)|/2N$. In the notation of Lemma 2.16, $S'_j = N^{-j}T_j(a, Nx^*)$. The error terms of the form $O(a^{-O(1)}e^{-\pi^2/a})$ in the conclusion of Lemma 2.16 are $e^{-\Theta(N)}$ and so negligible. Thus, from Lemma 2.16 and (2.69), recalling the definitions (2.6) and (2.7) of ξ and θ, we find

$$\sum_{|n/N-x^*|<N^{-0.4}} p_{n,N-n} = N^{-2}e^{-N\xi}\big(N^{1/2}\theta + O(N^{-1/2})\big). \tag{2.70}$$

Next, consider n such that $N^{-0.4} \leqslant |n/N - x^*| \leqslant 2c_7$, and recall that $3c_7 < c_6$. If $N^{-0.4} \leqslant |x - x^*| \leqslant 2c_7$, then Lemma 2.15 implies that $|x - x_0| \leqslant 3c_7$ and $\Psi(x) \leqslant \Psi(x^*) - \Omega((x-x^*)^2) \leqslant \Psi(x^*) - \Omega(N^{-0.8}) = -\xi - \Omega(N^{-0.8})$. Hence, by (2.45) and Lemma 2.14, if $N^{-0.4} \leqslant |n/N - x^*| \leqslant 2c_7$, then Lemma 2.15 implies that

$$p_{n,N-n} = N^{-2}e^{-N\xi - \Omega(N^{0.2})} \cdot O(1).$$

The sum over such n is easily absorbed into the error term we are aiming for: we have, say,

$$\sum_{N^{-0.4} \leqslant |n/N-x^*| \leqslant 2c_7} p_{n,N-n} = O(N^{-5/2}) \cdot e^{-N\xi}. \tag{2.71}$$

Finally, since $|x^* - x_0| \leqslant c_7$ by Lemma 2.15(ii) and $0 \leqslant n \leqslant N$, using Lemma 2.17 there exists a constant $c > 0$ such that, say,

$$\sum_{|n/N-x^*|>2c_7} p_{n,N-n} \leqslant \sum_{|n/N-x_0|>c_7} p_{n,N-n} \leqslant O(N) \cdot e^{-2cN} = O(N^{-5/2}) \cdot e^{-cN}. \tag{2.72}$$

Recalling that $|\mathbb{E}\,Y - 1| \leqslant c_1$, by (2.67) we may choose $c_1 \leqslant c_8$ sufficiently small so that $\xi < c$, and then (2.5) follows from (2.70), (2.71) and (2.72). □

3 Application to branching process families

In this section we apply the main result of Section 2 (Theorem 2.2) to a family of branching processes. The goal is to prove Theorem 3.4 below, giving estimates for the point probabilities $\mathbb{P}(|\mathfrak{X}| = N)$ in a form suitable for the application to Achlioptas processes in [22].

3.1 Properties of general parameterized families

By a *branching process family* $(\mathfrak{X}_{Y_u,Z_u,Y_u^0,Z_u^0})_{u\in I}$ we simply mean a family of branching processes of the type in Definition 1.1, one for each u in some interval $I \subset \mathbb{R}$. Given such a family, we write

$$g_u(y,z) := g_{Y_u,Z_u}(y,z) = \mathbb{E}(y^{Y_u} z^{Z_u}) \qquad \text{and} \qquad g_u^0(y,z) := g_{Y_u^0,Z_u^0}(y,z) = \mathbb{E}(y^{Y_u^0} z^{Z_u^0})$$

for the corresponding probability generating functions. Note that the branching process family is fully specified by the interval I and the functions g_u and g_u^0.

The following auxiliary result shows that the associated parameters $\xi_u = \xi_{Y_u,Z_u}$ and $\theta_u = \theta_{Y_u,Z_u,Y_u^0,Z_u^0}$ defined as in Theorem 2.2 vary smoothly in u. This will later allow us to compare the parameters $\xi_{Y,Z}$ and θ_{Y,Z,Y^0,Z^0} resulting from different probability distributions $(Y^0, Z^0) \in \mathcal{K}^0$ and $(Y, Z) \in \mathcal{K}$ (by integrating linear mixtures that interpolate between them); here the extra $|\mathbb{E}\, Y_u - 1| = O(1)$ factor in (3.2) is crucial.

Lemma 3.1. *Suppose that $R > 1$, $M < \infty$, $k_1, k_2 \in \mathbb{N}$, and $\delta > 0$. Set $\mathcal{K}^0 = \mathcal{K}^0(R,M,\delta)$ and $\mathcal{K} = \mathcal{K}(R,M,k_1,k_2,\delta)$. Let $(\mathfrak{X}_u)_{u\in I} = (\mathfrak{X}_{Y_u,Z_u,Y_u^0,Z_u^0})_{u\in I}$ be a branching process family such that, for every $u \in I$, we have $(Y_u^0, Z_u^0) \in \mathcal{K}^0$, $(Y_u, Z_u) \in \mathcal{K}$, and $|\mathbb{E}\, Y_u - 1| \leqslant c_1$, where $c_1 > 0$ is the constant appearing in Theorem 2.2. Suppose that $g_u(y,z)$ and $g_u^0(y,z)$ are analytic as functions of (u,y,z) in the domain*

$$\mathcal{D}_{I,R} := I \times \{(y,z) \in \mathbb{C} : |y|, |z| < R\} \subset \mathbb{R} \times \mathbb{C}^2,$$

and that for some λ,

$$\max\Big\{\Big|\frac{\partial}{\partial u} g_u(y,z)\Big|,\ \Big|\frac{\partial}{\partial u} g_u^0(y,z)\Big|\Big\} \leqslant \lambda \tag{3.1}$$

for all $(u,y,z) \in \mathcal{D}_{I,R}$. Let

$$\xi_u := \xi_{Y_u,Z_u} \qquad \text{and} \qquad \theta_u := \theta_{Y_u,Z_u,Y_u^0,Z_u^0}$$

be defined as in Theorem 2.2. Then ξ_u and θ_u are (real) analytic as functions of $u \in I$. Furthermore,

$$\frac{\mathrm{d}}{\mathrm{d}u}\xi_u = O\big(\lambda |\mathbb{E}\, Y_u - 1|\big), \tag{3.2}$$

$$\frac{\mathrm{d}}{\mathrm{d}u}\theta_u = O(\lambda), \tag{3.3}$$

where the implicit constants in (3.2) *and* (3.3) *depend only on R, M, k_1, k_2, δ.*

Proof. By assumption, the conditions of Theorem 2.2 hold for each $u \in I$. For any of the quantities or functions defined in previous sections for a single branching process, we use a subscript u to denote the corresponding quantity or function associated to $\mathfrak{X}_u$. As in previous sections, α and β always denote real numbers.

The idea of the proof is as follows. For a given u, the functions defined in the previous sections are defined, either explicitly or implicitly, in terms of g_u and g_u^0 (or their reparameterizations f_u and f_u^0). Roughly speaking, since g_u and g_u^0 vary analytically in u by assumption (and with u-derivative $O(\lambda)$), it follows that the same is true for the derived quantities. There are various steps where we must be slightly careful; for example, when taking logs (there is no problem as we stick to the domain $|z-1| \leqslant \frac{1}{2}$), or dividing by the square root of a certain determinant (there is no problem since this determinant is $\Omega(1)$ by Lemma 2.10). We must also be careful with the implicit definitions of G_u and x_u^*; the hardest part of the argument is to establish (3.2) with $O(\lambda |\mathbb{E}\, Y_u - 1|)$ instead of $O(\lambda)$.

Turning to the details, from (3.1) and standard Cauchy estimates we see that for each fixed m we have

$$\Big\|D^m \frac{\partial}{\partial u} g_u(y,z)\Big\| = O(\lambda) \qquad \text{and} \qquad \Big\|D^m \frac{\partial}{\partial u} g_u^0(y,z)\Big\| = O(\lambda) \tag{3.4}$$

whenever $|y|, |z| \leqslant R^{1/2}$, say. (Here and below, D does not include derivatives with respect to u.) Since $c_4 \leqslant c_2 \leqslant (\log R)/2$, the same estimates hold for the derivatives of $f_u(y,z) = g_u(e^y, e^z)$ and $f_u^0(y,z) = g_u^0(e^y, e^z)$ in the domain $B_{c_4} \subset \mathbb{R}^2$; from now on we work over the reals. Recalling the definition (2.37) and $\varphi_u = \log f_u$, from (3.4) it follows that $\|\frac{\partial}{\partial u} F_u(\alpha,\beta)\| = O(\lambda)$ for $(\alpha,\beta) \in B_{c_4}$.

From the definition (2.37), the function $F_u(\alpha,\beta)$ is a (real) analytic function of $(u,\alpha,\beta) \in I \times B_{c_4}$. For each $u \in I$, by (2.40) we have an inverse $G_u : B_{c_5} \to B_{c_4}$ of the 2-variable function F_u. Applying a standard version of the implicit function theorem locally, we see that $G_u(\alpha,\beta)$ is analytic as a function of $(u,\alpha,\beta) \in I \times B_{c_5}$.[1]

Noting $\mathbb{E}\, Y_u = \frac{\partial}{\partial y} g_u(y,z)\big|_{y=z=1}$, by definition (2.41) and $|\mathbb{E}\, Y_u - 1| \leqslant c_1$ it follows that $h_u(x)$ is an analytic function of u, x for $u \in I$ and $|x - x_{0,u}| < c_6$; we consider in the sequel only such u and x. Inspecting the definitions (2.43) and (2.44), using Lemma 2.10 (to ensure that the determinant is not degenerate), we see that $\Psi_u(x)$ and $\Phi_u(x)$ are well-defined compositions of analytic functions, and thus analytic as functions of u, x.

Since $F_u(G_u(y)) = y$ is independent of u, writing $x = G_u(y)$ and differentiating yields $\frac{\partial}{\partial u} F_u(x) + DF_u(x)(\frac{\partial}{\partial u} G_u(y)) = 0$ and thus, recalling (2.39), for $y \in B_{c_5}$,

$$\Big\| \frac{\partial}{\partial u} G_u(y) \Big\| \leqslant \Big\| DF_u(x)^{-1} \Big\| \cdot \Big\| \frac{\partial}{\partial u} F_u(x) \Big\| = O(\lambda). \tag{3.5}$$

Recalling the definition (2.41), note that (3.5) implies $\|\frac{\partial}{\partial u} h_u(x)\| = O(\lambda)$. Since ψ_u is defined in terms of $\varphi_u = \log f_u = \log f_{Y_u,Z_u}$ and its derivatives, see (2.20), using $(Y_u, Z_u) \in \mathcal{K}$ it follows that $\|D\psi_u(\alpha,\beta)\| = O(1)$. Furthermore, since estimates analogous to (3.4) also hold for $f_u = f_{Y_u,Z_u}$, we have $\frac{\partial}{\partial u}\psi_u(y) = O(\lambda)$. Hence, recalling (2.43) and writing $y = h_u(x)$, we have

$$\frac{\partial}{\partial u} \Psi_u(x) = x \left(\frac{\partial}{\partial u} \psi_u(y) + D\psi_u(y) \Big(\frac{\partial}{\partial u} h_u(x) \Big) \right) = O(\lambda). \tag{3.6}$$

Recalling the definitions (2.43) and (2.44), and the estimates in Section 2.3, we similarly deduce $\frac{\partial}{\partial u}\Phi_u(x) = O(\lambda)$, $\frac{\partial}{\partial u}\Psi_u'(x) = O(\lambda)$ and $\frac{\partial}{\partial u}\Psi_u''(x) = O(\lambda)$.

Since x_u^* is defined by $\Psi_u'(x_u^*) = 0$, we have $(\frac{\partial}{\partial u}\Psi_u')(x_u^*) + \Psi_u''(x_u^*)\frac{\mathrm{d}}{\mathrm{d}u} x_u^* = 0$. It follows, using the lower bound $\Psi_u''(x_u^*) = -\Omega(1)$ from Lemma 2.15(iii) and the implicit function theorem, that x_u^* is an analytic function of u, and that

$$\frac{\mathrm{d}}{\mathrm{d}u} x_u^* = \frac{-(\frac{\partial}{\partial u}\Psi_u')(x_u^*)}{\Psi_u''(x_u^*)} = O(\lambda). \tag{3.7}$$

As (3.5) implies $\|\frac{\partial}{\partial u} h_u(x)\| = O(\lambda)$, and $\|h_u'(x)\| = O(1)$ by (2.46), it follows that

$$\Big\| \frac{\mathrm{d}}{\mathrm{d}u} h_u(x_u^*) \Big\| \leqslant \Big\| \Big(\frac{\partial}{\partial u} h_u \Big)(x_u^*) \Big\| + \Big\| h_u'(x_u^*) \Big\| \cdot \Big| \frac{\mathrm{d}}{\mathrm{d}u} x_u^* \Big| = O(\lambda). \tag{3.8}$$

Similarly, from the definitions (2.6) and (2.7), using (3.7) and the bounds above on $\frac{\partial}{\partial u}\Psi_u$, $\frac{\partial}{\partial u}\Psi_u''$ and $\frac{\partial}{\partial u}\Phi_u$ it follows that ξ_u and θ_u are analytic functions of u, with $\frac{\mathrm{d}}{\mathrm{d}u}\xi_u = O(\lambda)$ and $\frac{\mathrm{d}}{\mathrm{d}u}\theta_u = O(\lambda)$. It remains only to establish (3.2).

For this final step, recalling (2.20) and $\varphi_u(\alpha,\beta) = \log f_u(\alpha,\beta) = \log g_u(e^\alpha, e^\beta)$, note that $\|D\frac{\partial}{\partial u}\psi_u(x)\| = O(\lambda)$ follows from (3.4). Since $\frac{\partial}{\partial u}\psi_u(0) = 0$, we thus have $\frac{\partial}{\partial u}\psi_u(x) = O\big(\lambda|x|\big)$. Similarly, as a consequence of Lemma 2.12, $\psi_u(x) = O\big(|x|\big)$ and $\|D\psi_u(x)\| = O\big(|x|\big)$. Writing y_u^* for $h_u(x_u^*)$, it follows from the definition (2.43) and (3.7)–(3.8) that

$$\begin{aligned} -\frac{\mathrm{d}}{\mathrm{d}u}\xi_u = \frac{\mathrm{d}}{\mathrm{d}u}\Psi_u(x_u^*) &= \frac{\mathrm{d}}{\mathrm{d}u}\left(x_u^* \psi_u(y_u^*) \right) \\ &= \left(\frac{\mathrm{d}}{\mathrm{d}u} x_u^* \right) \psi_u(y_u^*) + x_u^* \left(\frac{\partial}{\partial u}\psi_u(y_u^*) + D\psi_u(y_u^*) \Big(\frac{\mathrm{d}}{\mathrm{d}u} h_u(x_u^*) \Big) \right) = O\big(\lambda |y_u^*|\big). \end{aligned}$$

[1]Fix u_0 and $y_0 = (\alpha_0, \beta_0)$ in the relevant domain. By (2.36), DF_u is invertible at $x_0 = G_{u_0}(y_0)$, so there is an analytic function $\widehat{G}_u(y)$ defined in a neighbourhood of (u_0, y_0) such that $F_u(\widehat{G}_u(y)) = y$. By local uniqueness, $G_u(y) = \widehat{G}_u(y)$ near (u_0, y_0), so G is indeed analytic at this point.

Since $y_u^* = h_u(x_u^*)$, from the definition (2.41) of h_u, the bound (2.47), and, for the final step, Lemma 2.15(ii), it follows that

$$\frac{\mathrm{d}}{\mathrm{d}u}\xi_u = O\Big(\lambda\big(|\,\mathbb{E}\,Y_u - 1| + |x_u^* - x_{0,u}|\big)\Big) = O\big(\lambda|\,\mathbb{E}\,Y_u - 1|\big),$$

completing the proof of the lemma. □

3.2 A specific result suitable for application to Achlioptas processes

In this section we use Theorem 2.2 and Lemma 3.1 to prove the case $p_{\mathcal{R}} = 1$, $K = 0$ of Theorem A.10 of [22], used there for the analysis of Achlioptas processes. To formulate this main application, i.e., our point probability result for certain (perturbed) branching process families, we need some some further definitions.

Definition 3.2. Let $t_0 < t_{\mathrm{c}} < t_1$ be real numbers. The branching process family $(\mathfrak{X}_t)_{t\in(t_0,t_1)} = (\mathfrak{X}_{Y_t,Z_t,Y_t^0,Z_t^0})_{t\in(t_0,t_1)}$ is t_{c}*-critical* if the following hold:

(i) There exist $\delta > 0$ and $R > 1$ with $(t_{\mathrm{c}} - \delta, t_{\mathrm{c}} + \delta) \subseteq (t_0, t_1)$ such that the probability generating functions

$$g_t(y,z) := g_{Y_t,Z_t}(y,z) = \mathbb{E}\big(y^{Y_t} z^{Z_t}\big) \qquad \text{and} \qquad g_t^0(y,z) := g_{Y_t^0,Z_t^0}(y,z) = \mathbb{E}\big(y^{Y_t^0} z^{Z_t^0}\big) \tag{3.9}$$

are defined and analytic on the domain

$$\{(t,y,z) \in \mathbb{R}\times\mathbb{C}^2 :\ |t - t_{\mathrm{c}}| < \delta \text{ and } |y|, |z| < R\}.$$

(ii) We have

$$\mathbb{E}\,Y_{t_{\mathrm{c}}} = 1, \qquad \mathbb{E}\,Y_{t_{\mathrm{c}}}^0 > 0 \qquad \text{and} \qquad \frac{\mathrm{d}}{\mathrm{d}t}\,\mathbb{E}\,Y_t\Big|_{t=t_{\mathrm{c}}} > 0. \tag{3.10}$$

(iii) There exists some $k_0 \in \mathbb{N}$ such that

$$\min\Big\{\mathbb{P}\big(Y_{t_{\mathrm{c}}} = k_0,\, Z_{t_{\mathrm{c}}} = k_0\big),\, \mathbb{P}\big(Y_{t_{\mathrm{c}}} = k_0+1,\, Z_{t_{\mathrm{c}}} = k_0\big),\, \mathbb{P}\big(Y_{t_{\mathrm{c}}} = k_0,\, Z_{t_{\mathrm{c}}} = k_0+1\big)\Big\} > 0. \tag{3.11}$$

Definition 3.3. Let $(\mathfrak{X}_t)_{t\in(t_0,t_1)}$ be a t_{c}-critical branching process family, and let δ, R and k_0 be as in Definition 3.2. Given $t, \eta \geqslant 0$ with $|t - t_{\mathrm{c}}| < \delta$, we say that the branching process $\mathfrak{X}_{Y,Z,Y^0,Z^0}$ is *of type* (t,η) (with respect to $(\mathfrak{X}_t)$, δ, R, and k_0) if the following hold:

(i) Writing $\mathcal{N} := \{(y,z) \in \mathbb{C}^2 :\ |y|, |z| < R\}$, the expectations

$$\tilde{g}(y,z) := g_{Y,Z}(y,z) = \mathbb{E}\big(y^{Y} z^{Z}\big) \qquad \text{and} \qquad \tilde{g}^0(y,z) := g_{Y^0,Z^0}(y,z) = \mathbb{E}\big(y^{Y^0} z^{Z^0}\big) \tag{3.12}$$

are defined (i.e., the expectations converge absolutely) for all $(y,z) \in \mathcal{N}$.

(ii) For all $(y,z) \in \mathcal{N}$ we have

$$\big|\tilde{g}(y,z) - g_t(y,z)\big| \leqslant \eta \qquad \text{and} \qquad \big|\tilde{g}^0(y,z) - g_t^0(y,z)\big| \leqslant \eta. \tag{3.13}$$

Note that $\mathfrak{X}_t = \mathfrak{X}_{Y_t,Z_t,Y_t^0,Z_t^0}$ is itself of type (t,η) for any $\eta \geqslant 0$. The following result relates the point probabilities from $\mathfrak{X}_t$ with those from branching processes $\mathfrak{X}$ of type (t,η). A key feature is the form of the uniform $O(\eta|t-t_{\mathrm{c}}|+\eta^2)$ error term in (3.15). In (3.14) and (3.15) below, we have $\xi_{Y_t,Z_t} = \psi(t) = \Theta((t-t_{\mathrm{c}})^2)$ and $\theta_{Y_t,Z_t,Y_t^0,Z_t^0} = \theta(t) = \Theta(1)$ for $\mathfrak{X} = \mathfrak{X}_t$ (using $\eta = 0$), and $\xi_{Y,Z} \sim \psi(t)$ and $\theta_{Y,Z,Y^0,Z^0} \sim \theta(t)$ for any branching process $\mathfrak{X}$ of type (t,η) with $\eta \ll |t - t_{\mathrm{c}}| \leqslant \varepsilon_0$. In the near-critical case $t = t_{\mathrm{c}} \pm \varepsilon$, the size–$N$ point probabilities of $\mathfrak{X}_t$ and $\mathfrak{X}$ thus both decay exponentially in $\Theta(\varepsilon^2 N)$.

Theorem 3.4 (Point probabilities of $\mathfrak{X}$ of type (t,η))**.** *Let $(\mathfrak{X}_t)_{t\in(t_0,t_1)}$ be a t_{c}-critical branching process family. Then there exist constants $\varepsilon_0, \eta_0 > 0$ and analytic functions θ, ψ on the interval $I = [t_{\mathrm{c}} - \varepsilon_0, t_{\mathrm{c}} + \varepsilon_0]$ such that*

$$\mathbb{P}(|\mathfrak{X}| = N) = (1 + O(1/N))N^{-3/2}\theta_{Y,Z,Y^0,Z^0}\ e^{-\xi_{Y,Z}N} \tag{3.14}$$

uniformly over all $N \geqslant 1$, $t \in I$, $0 \leqslant \eta \leqslant \eta_0$ and all branching processes $\mathfrak{X} = \mathfrak{X}_{Y,Z,Y^0,Z^0}$ of type (t,η) (with respect to $(\mathfrak{X}_t)$), where the parameters $\xi_{Y,Z}$ and θ_{Y,Z,Y^0,Z^0}, which depend on the distributions of (Y,Z) and of (Y^0,Z^0), satisfy

$$\xi_{Y,Z} = \psi(t) + O(\eta|t-t_{\mathrm{c}}| + \eta^2) \qquad \text{and} \qquad \theta_{Y,Z,Y^0,Z^0} = \theta(t) + O(\eta). \tag{3.15}$$

Moreover, $\theta(t) > 0$, $\psi(t) \geqslant 0$, $\psi(t_{\mathrm{c}}) = \psi'(t_{\mathrm{c}}) = 0$, and $\psi''(t_{\mathrm{c}}) > 0$.

Proof. Fix a t_{c}-critical branching process family $(\mathfrak{X}_t)_{t\in(t_0,t_1)}$, and let $\delta > 0$ and $R > 1$ be as in the definitions above. We pick $0 < \varepsilon_0 < \delta$, and decrease R slightly, keeping $R > 1$. Then $g(t,y,z) = g_t(y,z)$ and $g^0(t,y,z) = g^0_t(y,z)$ are continuous on the compact domain $|t-t_{\mathrm{c}}| \leqslant \varepsilon_0$, $|y|,|z| \leqslant R$ and so bounded, say by M_1. Let $M = M_1 + 1$. Then, provided $|t-t_{\mathrm{c}}| \leqslant \varepsilon_0$, by (3.13) any $\mathfrak{X}$ of type (t,η) with $\eta \leqslant 1$ satisfies

$$\max\Big\{\big|\tilde{g}(y,z)\big|,\ \big|\tilde{g}^0(y,z)\big|\Big\} \leqslant M \qquad \text{whenever} \quad |y|,|z| \leqslant R.$$

For any integers k,ℓ, we have

$$\mathbb{P}(Y_t = k,\ Z_t = \ell) = \frac{1}{k!\ell!} \cdot \frac{\partial^{k+\ell}}{\partial y^k \partial z^\ell} g_t(y,z)\Big|_{y=z=0}.$$

Since $g_t(y,z)$ is analytic in t,y,z, this probability varies continuously in t. Moreover, since $\mathbb{P}(Y = k,\ Z = \ell)$ can analogously be written as a derivative of $\tilde{g}$ evaluated at $(0,0)$, using standard Cauchy estimates and (3.13) we infer

$$\big|\mathbb{P}(Y = k,\ Z = \ell) - \mathbb{P}(Y_t = k,\ Z_t = \ell)\big| = O(\eta).$$

A similar argument shows that $\mathbb{E}\, Y_t = \frac{\partial}{\partial y} g_t(y,z)\big|_{y=z=1}$ is continuous in t, and that Cauchy's estimates imply

$$\big|\mathbb{E}\, Y_t - \mathbb{E}\, Y\big| = O(\eta). \tag{3.16}$$

Analogous reasoning applies to $\mathbb{E}\, Y^0_t$ and $\mathbb{E}\, Y^0$.

By definition of a t_{c}-critical branching process family, there is some $\delta > 0$ such that for $t = t_{\mathrm{c}}$ all of $\mathbb{E}\, Y^0_t$, $\mathbb{P}(Y_t = k_0, Z_t = k_0)$, $\mathbb{P}(Y_t = k_0+1, Z_t = k_0)$ and $\mathbb{P}(Y_t = k_0, Z_t = k_0+1)$ are at least 2δ, say. Furthermore, at $t = t_{\mathrm{c}}$ we have $\mathbb{E}\, Y_t = 1$. From the argument above these quantities all vary continuously in t, and change by $O(\eta)$ when we move from $\mathfrak{X}_t$ to some $\mathfrak{X}$ of type (t,η). It follows that there is a constant $\eta_0 > 0$ such that, after reducing ε_0 if necessary, whenever $|t-t_{\mathrm{c}}| \leqslant \varepsilon_0$ and $\eta \leqslant \eta_0$, then any $\mathfrak{X}$ of type (t,η) satisfies the conditions of Theorem 2.2, namely that $(Y^0,Z^0) \in \mathcal{K}^0$, $(Y,Z) \in \mathcal{K}$ and $|\,\mathbb{E}\, Y - 1| \leqslant c_1$.

Now, applying Theorem 2.2 to each branching process in the family $(\mathfrak{X}_t)_{t\in[t_{\mathrm{c}}-\varepsilon_0,t_{\mathrm{c}}+\varepsilon_0]}$, and Lemma 3.1 to the family itself, establishes the $\eta = 0$ case of Theorem 3.4 with $\theta(t) = \theta_t$ and $\psi(t) = \xi_t$. Indeed, Theorem 2.2 gives that $\theta = \Theta(1)$, so we do have $\theta(t) > 0$, while (2.8) gives $\psi(t) = \xi_t = \Theta(|\,\mathbb{E}\, Y_t - 1|^2)$, which is $\Theta(|t-t_c|^2)$ since (3.10) implies, after reducing ε_0 if necessary, that

$$\big|\mathbb{E}\, Y_t - 1\big| = \Theta(|t-t_c|). \tag{3.17}$$

It follows that $\psi(t_{\mathrm{c}}) = \psi'(t_{\mathrm{c}}) = 0$ and $\psi''(t_{\mathrm{c}}) > 0$.

To complete the proof, assume now that $\mathfrak{X} = \mathfrak{X}_{Y,Z,Y^0,Z^0}$ is of type (t,η), with $0 \leqslant \eta \leqslant \eta_0$ and $|t-t_{\mathrm{c}}| \leqslant \varepsilon_0$. As noted above, Theorem 2.2 applies to $\mathfrak{X}$, giving (3.14); it remains to establish (3.15). We do this by interpolating between $\mathfrak{X}$ and $\mathfrak{X}_t$, and applying Lemma 3.1. Consider the branching process family $(\bar{Y}_u, \bar{Z}_u, \bar{Y}^0_u, \bar{Z}^0_u)_{u\in[0,1]}$ defined by the mixtures

$$\bar{g}_u(y,z) := (1-u)g_t(y,z) + u\tilde{g}(y,z) \qquad \text{and} \qquad \bar{g}^0_u(y,z) := (1-u)g^0_t(y,z) + u\tilde{g}^0(y,z). \tag{3.18}$$

(As noted earlier, the probability generating functions $\bar{g}_u$, $\bar{g}^0_u$ and the interval $I = [0,1]$ fully specify the family.) Since the assumptions of Theorem 2.2 are preserved by taking mixtures, every branching process in this family satisfies these assumptions. (In fact, they are all clearly of type (t,η) too.) Moreover, the assumption (3.13) implies that (3.1) holds with $\lambda = \eta$, and since $\mathbb{E}\, \bar{Y}_u = \frac{\partial}{\partial y}\bar{g}_u(y,z)\big|_{y=z=1}$ we have

$$|\,\mathbb{E}\, \bar{Y}_u - 1| \leqslant |\,\mathbb{E}\, Y_t - 1| + u|\,\mathbb{E}\, Y - \mathbb{E}\, Y_t| = O(|t-t_c| + \eta) \tag{3.19}$$

by (3.16) and (3.17). Thus we may apply Lemma 3.1, and, by integrating (3.2) with $\xi_u = \xi_{\bar{Y}_u,\bar{Z}_u}$, we infer

$$\xi - \xi_t = \xi_{\bar{Y}_1,\bar{Z}_1} - \xi_{\bar{Y}_0,\bar{Z}_0} = O\big(\eta|t-t_c| + \eta^2\big). \tag{3.20}$$

Finally, $\theta - \theta_t = O(\eta)$ follows similarly by integrating (3.3). □

Theorem 3.4 immediately implies the key case $p_{\mathcal{R}} = 1$, $K = 0$ of Theorem A.10 of [22] with any positive value of the constant c. Indeed, after reducing ε_0 if necessary, the assumption $\eta \leqslant c|t-t_{\mathrm{c}}|$ in the latter theorem implies the assumption $\eta \leqslant \eta_0$ of Theorem 3.4. Moreover, the same assumption $\eta \leqslant c|t-t_{\mathrm{c}}|$ together with (3.15) implies the bound $\xi_{Y,Z} = \psi(t) + O(\eta|t-t_{\mathrm{c}}|)$ in Theorem A.10 of [22].

4 The survival probability

In this section we study the survival probability of the branching process $\mathfrak{X} = \mathfrak{X}_{Y,Z,Y^0,Z^0}$ from Definition 1.1 and the branching process family $(\mathfrak{X}_u)_{u\in I}$ from Section 3. The goal is to prove Theorem 4.5 below, i.e., to give estimates for $\mathbb{P}(|\mathfrak{X}| = \infty)$ suitable for the application to Achlioptas processes in [22].

Our strategy mimics the general approach used in Sections 2–3 for point probabilities, though the technical details are much simpler. In Section 4.1 we first prove a technical result for the survival probability $\mathbb{P}(|\mathfrak{X}| = \infty)$ of a single branching process (Lemma 4.2). Then we show that in a branching process family $(\mathfrak{X}_u)_{u\in I}$ certain parameters related to the survival probability vary smoothly in u (Lemma 4.4). Finally, in Section 4.2 we combine these two auxiliary results to prove Theorem 4.5.

4.1 Properties of a single process and general parameterized families

As far as the survival of $\mathfrak{X}_{Y,Z,Y^0,Z^0}$ is concerned, particles of type S are irrelevant and may be ignored, so we may consider a standard single-type Galton–Watson branching process with offspring distribution Y and initial distribution Y^0, which we henceforth denote by $\mathfrak{X}_{Y,Y^0}$. Thus

$$\mathbb{P}(|\mathfrak{X}_{Y,Z,Y^0,Z^0}| = \infty) = \mathbb{P}(|\mathfrak{X}_{Y,Y^0}| = \infty) =: \rho_{Y,Y^0}. \tag{4.1}$$

Writing $\mathbb{1}$ as shorthand for the distribution with constant value one, it similarly follows that

$$\mathbb{P}(|\mathfrak{X}^1_{Y,Z}| = \infty) = \mathbb{P}(|\mathfrak{X}_{Y,\mathbb{1}}| = \infty) =: \rho_Y. \tag{4.2}$$

Throughout this section, we shall work with the univariate probability generating functions $g_Y(y) := \mathbb{E}\, y^Y = g_{Y,Z}(y,1)$ and $g_{Y^0}(y) := g_{Y^0,Z^0}(y,1)$. By standard branching process arguments (see, e.g., [10, Theorem 5.4.5]), we have

$$1 - \rho_{Y,Y^0} = g_{Y^0}(1-\rho_Y), \tag{4.3}$$

where the extinction probability $1-\rho_Y$ is the smallest non-negative solution to

$$1 - \rho_Y = g_Y(1-\rho_Y). \tag{4.4}$$

Fix $R > 1$, M, k_1, k_2 and $\delta > 0$. We henceforth assume that $(Y,Z) \in \mathcal{K} = \mathcal{K}(R,M,k_1,k_2,\delta)$ and $(Y^0,Z^0) \in \mathcal{K}^0 = \mathcal{K}^0(R,M,\delta)$. Since $(Y,Z) \in \mathcal{K}$, by (2.1) the function $g_Y(y)$ is analytic in $\{y \in \mathbb{C} : |y| < R\}$, with $g_Y(1) = 1$. A Taylor expansion of $g_Y(y)$ at $y = 1$ yields, for $|x| < R-1$,

$$g_Y(1-x) = \mathbb{E}(1-x)^Y = \sum_{n=0}^{\infty} (-1)^n\, \mathbb{E}\binom{Y}{n} x^n = 1 - \mathbb{E}\, Y x + \mathbb{E}\binom{Y}{2}x^2 - \mathbb{E}\binom{Y}{3}x^3 + \cdots. \tag{4.5}$$

Define

$$h_Y(x) := \frac{1-g_Y(1-x)}{x}, \tag{4.6}$$

removing the removable singularity at $x = 0$. Then h_Y is analytic in $\{x \in \mathbb{C} : |x| < R-1\}$, and

$$h_Y(x) = \sum_{n=0}^{\infty} (-1)^n\, \mathbb{E}\binom{Y}{n+1} x^n = \mathbb{E}\, Y - \mathbb{E}\binom{Y}{2}x + \mathbb{E}\binom{Y}{3}x^2 + \cdots. \tag{4.7}$$

Observe that if $\rho_Y > 0$, then (4.4) is equivalent to $h_Y(\rho_Y) = 1$. Furthermore,

$$h_Y(0) = \mathbb{E}\,Y = g'_Y(1), \qquad -h'_Y(0) = \mathbb{E}\binom{Y}{2} = \frac{\mathbb{E}(Y(Y-1))}{2}. \tag{4.8}$$

We next derive bounds on the derivatives of h_Y valid for small x.

Lemma 4.1. *Suppose that $R > 1$, $M < \infty$, $k_1, k_2 \in \mathbb{N}$, and $\delta > 0$. There exist constants $0 < c_9 \leqslant \min\{R-1, 1\}/3$ and $C_4^{(m)}$ such that if $(Y, Z) \in \mathcal{K} = \mathcal{K}(R, M, k_1, k_2, \delta)$, then the following hold.*

(i) *If $m \in \mathbb{N}$ and $|x| \leqslant c_9$, then $|D^m h_Y(x)| \leqslant C_4^{(m)}$.*

(ii) *If $\mathbb{E}\,Y \geqslant 1-\delta$, then $h'_Y(0) \leqslant -\delta$ and $\mathbb{P}(Y \geqslant 2) > 0$.*

(iii) *If $\mathbb{E}\,Y \geqslant 1-\delta$ and $|x| \leqslant c_9$, then $h'_Y(x) \leqslant -\delta/2$.*

Proof. (i): By (4.6) and (2.1), $h(x) = O(1)$ if $|x| = (R-1)/2$, say. Hence the result, with $c_9 := \min\{R-1, 1\}/3$, say, follows by Cauchy's estimates.

(ii): If (2.3) holds with $k_1 \geqslant 1$, then $\mathbb{P}(Y \geqslant 2) \geqslant \mathbb{P}(Y = k_1+1) \geqslant \pi_{k_1+1,k_2} \geqslant \delta$, and thus $\mathbb{E}(Y(Y-1)) \geqslant 2\delta$.

If instead (2.3) holds with $k_1 = 0$, then $\mathbb{P}(Y = 0) \geqslant \pi_{k_1,k_2} + \pi_{k_1,k_2+1} \geqslant 2\delta$. Since $\mathbb{E}\,Y \geqslant 1-\delta$, then $Y \in \mathbb{N}$ implies $\mathbb{E}(\mathbb{1}_{\{Y \geqslant 2\}}(Y-1)) = \mathbb{E}(Y-1) + \mathbb{P}(Y = 0) \geqslant \delta$, and thus $\mathbb{E}(Y(Y-1)) \geqslant 2\,\mathbb{E}(\mathbb{1}_{\{Y \geqslant 2\}}(Y-1)) \geqslant 2\delta$.

In both cases, $h'_Y(0) \leqslant -\delta$ follows by (4.8), and $\mathbb{P}(Y \geqslant 2) > 0$ holds, too.

(iii): Follows by (ii) and (i) (with $m = 2$), replacing c_9 by $\min\{c_9, \delta/2C_4^{(2)}\}$. $\square$

We next characterize the survival probability ρ_Y in terms of the (unique) solution to $h_Y(\hat\rho) = 1$.

Lemma 4.2. *Suppose that $R > 1$, $M < \infty$, $k_1, k_2 \in \mathbb{N}$, and $\delta > 0$. There exists a constant $0 < c_{10} \leqslant \delta$ such that the following holds. If $(Y, Z) \in \mathcal{K}(R, M, k_1, k_2, \delta)$ and $|\mathbb{E}\,Y - 1| < c_{10}$, then there is a unique $\hat\rho = \hat\rho_Y \in \{x \in \mathbb{R} : |x| < c_9\}$ such that*

$$h_Y(\hat\rho) = 1.$$

Furthermore, $\rho_Y = \max\{\hat\rho, 0\}$, $\operatorname{sign}(\hat\rho) = \operatorname{sign}(\mathbb{E}\,Y - 1)$, and $|\hat\rho| = \Theta(|\mathbb{E}\,Y - 1|)$, where the implicit constants depend only on R, M, k_1, k_2 and δ.

Proof. We apply the inverse function theorem, Lemma 2.13, with $d = 1$, $r = c_9$ and

$$F(x) := h_Y(x) - \mathbb{E}\,Y,$$

using (4.8) and Lemma 4.1 to verify the assumptions; we shall ensure that $c_{10} \leqslant \delta$, so $|\mathbb{E}\,Y - 1| < c_{10}$ implies $\mathbb{E}\,Y \geqslant 1-\delta$. Writing $B_r = B_r^1 = \{x \in \mathbb{R} : |x| < r\}$ to avoid clutter (as before), Lemma 2.13 shows the existence of a constant $c_{10} > 0$, which we may assume to be at most δ, and an inverse function $G : B_{c_{10}} \to B_{c_9}$ with $F(G(x)) = x$ and $G(0) = 0$. We define

$$\hat\rho := G(1 - \mathbb{E}\,Y),$$

so that $h_Y(\hat\rho) = F(\hat\rho) + \mathbb{E}\,Y = 1$. Since $\|DG(y)\| = O(1)$ in $B_{c_{10}}$ by Lemma 2.13 and $\|DF(x)\| = O(1)$ in B_{c_9} by Lemma 4.1(i), using $G(0) = F(0) = 0$ we have $|\hat\rho| = |G(1-\mathbb{E}\,Y)| = O(|\mathbb{E}\,Y - 1|)$ and $|\mathbb{E}\,Y - 1| = |F(\hat\rho)| = O(\hat\rho)$, establishing $|\hat\rho| = \Theta(|\mathbb{E}\,Y - 1|)$.

We relate $\hat\rho$ and ρ_Y by a variant of the usual fixed point analysis of $g_Y(x) = x$ in $[0, R]$. Since $\mathbb{P}(Y \geqslant 2) > 0$ by Lemma 4.1(ii), g_Y is strictly convex on $[0, R]$, which implies that $g_Y(x) = x$ has at most two solutions in this interval, and exactly one solution if $\mathbb{E}\,Y = 1$, since $g_Y(1) = 1$ and $g'_Y(1) = \mathbb{E}\,Y$. Now $x = 1$ and $x = 1 - \rho_Y \in [0, 1]$ are solutions. Since $h_Y(\hat\rho) = 1$, $x = 1 - \hat\rho$ is also a solution (see (4.6)); since $|\hat\rho| < c_9 < \min\{R-1, 1\}$, we have $1 - \hat\rho \in (0, R)$.

If $\hat\rho > 0$, then $1 - \hat\rho \in (0, 1)$ and 1 are two distinct solutions; thus $1 - \rho_Y = 1 - \hat\rho$, and $g'_Y(1) > 1$ by strict convexity. Similarly, if $\hat\rho < 0$, then $1 - \hat\rho \in (1, R)$ and thus $1 - \rho_Y = 1$, and $g'_Y(1) < 1$ by strict convexity. Finally, if $\hat\rho = 0$, then $\mathbb{E}\,Y = h_Y(0) = h_Y(\hat\rho) = 1$ by (4.8), so that $1 - \rho_Y = 1$ (since then $x = 1$ is the only solution to $g_Y(x) = x$ in $[0, R]$). Hence $\rho_Y = \max\{\hat\rho, 0\}$ in all cases. It follows also that $\hat\rho$ is unique, and that $\hat\rho$ has the same sign as $\mathbb{E}\,Y - 1 = g'_Y(1) - 1$. $\square$

Remark 4.3. Since $F'(0) = h'_Y(0) = -\mathbb{E}(Y(Y-1))/2$, when $\mathbb{E}\,Y > 1$ it follows easily that $\rho_Y = \frac{2(\mathbb{E}\,Y-1)}{\mathbb{E}(Y(Y-1))} + O(|\,\mathbb{E}\,Y-1|^2)$. In particular $\rho_Y \sim \frac{2(\mathbb{E}\,Y-1)}{\mathbb{E}(Y(Y-1))}$ as $\mathbb{E}\,Y \searrow 1$, assuming, as always here, that $(Y,Z) \in \mathcal{K}$. This holds under much weaker conditions on Y, see [12] and [2] for precise conditions; see also [11, Section 3].

We next consider a branching process family $(\mathfrak{X}_u)_{u\in I} = (\mathfrak{X}_{Y_u,Z_u,Y_u^0,Z_u^0})_{u\in I}$ as in Section 3; as there we indicate the parameter u by subscripts. Thus, for example, $\hat\rho_u = \hat\rho_{Y_u}$ is defined as in Lemma 4.2, with (Y,Z) replaced by (Y_u,Z_u). Furthermore, in analogy to (4.3), we also define

$$1 - \hat\rho_{Y_u,Y_u^0} := g_{Y_u^0}(1-\hat\rho_u). \tag{4.9}$$

Thus, by combining (4.3) with Lemma 4.2, when $\mathbb{E}\,Y_u \geqslant 1$ we have $\hat\rho_u = \rho_{Y_u}$ and $\hat\rho_{Y_u,Y_u^0} = \rho_{Y_u,Y_u^0}$. Mimicking Lemma 3.1, the following auxiliary result shows that $\hat\rho_u$ and $\hat\rho_{Y_u,Y_u^0}$ both vary smoothly in u.

Lemma 4.4. *Suppose that $R > 1$, $M < \infty$, $k_1,k_2 \in \mathbb{N}$, and $\delta > 0$. Set $\mathcal{K}^0 = \mathcal{K}^0(R,M,\delta)$ and $\mathcal{K} = \mathcal{K}(R,M,k_1,k_2,\delta)$. Let $(\mathfrak{X}_u)_{u\in I} = (\mathfrak{X}_{Y_u,Z_u,Y_u^0,Z_u^0})_{u\in I}$ be a branching process family satisfying the assumptions of Lemma 3.1, with $|\,\mathbb{E}\,Y_u-1| \leqslant c_1$ replaced by $|\,\mathbb{E}\,Y_u-1| \leqslant c_{10}$. Let $\hat\rho_u$ and $\hat\rho_{Y_u,Y_u^0}$ be defined as in Lemma 4.2 and (4.9). Then $\hat\rho_u$ and $\hat\rho_{Y_u,Y_u^0}$ are analytic functions of $u \in I$. Furthermore,*

$$\frac{\mathrm{d}}{\mathrm{d}u}\hat\rho_u = O(\lambda), \qquad \frac{\mathrm{d}}{\mathrm{d}u}\hat\rho_{Y_u,Y_u^0} = O(\lambda), \tag{4.10}$$

where the implicit constants depend only on R, M, k_1, k_2 and δ.

Proof. Let $h_u(x) = h_{Y_u}(x) := (1-g_u(1-x))/x$ be the equivalent of (4.6) for $\mathfrak{X}_u$, again removing the removable singularity at $x = 0$. Then $h_u(x)$ is an analytic function of $(u,x) \in I \times \{x \in \mathbb{C} : |x| < R-1\}$. Note that (3.1) implies $|\frac{\partial}{\partial u}h_u(x)| = O(\lambda)$ if $|x| = (R-1)/3$, say. Since $c_9 \leqslant (R-1)/3$, by the maximum modulus principle (applied with u fixed) it follows that

$$\Big|\frac{\partial}{\partial u}h_u(x)\Big| \leqslant C\lambda \tag{4.11}$$

for all $u \in I$ and $|x| \leqslant c_9$.

By Lemma 4.2, for every $u \in I$ there is a unique $\hat\rho_u \in \mathbb{R}$ with $|\hat\rho_u| < c_9$ such that

$$h_u(\hat\rho_u) = 1. \tag{4.12}$$

Since $|h'_u(\hat\rho_u)| \geqslant \delta/2$ by Lemma 4.1(iii) and $|\,\mathbb{E}\,Y_u-1| \leqslant c_{10} \leqslant \delta$, the implicit function theorem shows that $\hat\rho_u$ is an analytic function of $u \in I$. That $\hat\rho_{Y_u,Y_u^0}$ is analytic then follows from (4.9) and the assumption that $g_u^0(y,z)$ is analytic. By differentiating (4.12) we obtain $\frac{\partial h_u}{\partial u}(\hat\rho_u) + h'_u(\hat\rho_u)\cdot\frac{\mathrm{d}}{\mathrm{d}u}\hat\rho_u = 0$. So, using $|h'_u(\hat\rho_u)| \geqslant \delta/2$ and (4.11),

$$\frac{\mathrm{d}}{\mathrm{d}u}\hat\rho_u = -h'_u(\hat\rho_u)^{-1}\cdot\frac{\partial h_u}{\partial u}(\hat\rho_u) = O(1)\cdot O(\lambda) = O(\lambda). \tag{4.13}$$

Finally, $g'_{Y_u^0}(1-\hat\rho_u) = O(1)$ follows from (2.1) and Cauchy's estimates (recall that $|\hat\rho_u| < c_9 \leqslant (R-1)/3$). By differentiating (4.9) and then using (3.1) and (4.13), we obtain

$$\frac{\mathrm{d}}{\mathrm{d}u}\hat\rho_{Y_u,Y_u^0} = -\Big(\frac{\partial}{\partial u}g_{Y_u^0}\Big)(1-\hat\rho_u) + g'_{Y_u^0}(1-\hat\rho_u)\cdot\frac{\mathrm{d}}{\mathrm{d}u}\hat\rho_u = O(\lambda) + O(1)\cdot O(\lambda) = O(\lambda),$$

completing the proof. □

4.2 A specific result suitable for application to Achlioptas processes

We are now ready to prove our main result, concerning the t-dependence of the survival probability of $\mathfrak{X}_t$ when $(\mathfrak{X}_t)_{t\in I}$ is a t_{c}-critical branching process family, as well as the survival probability of branching processes $\mathfrak{X} = \mathfrak{X}_{Y,Z,Y^0,Z^0}$ of type (t,η); see Section 3.2 for the relevant definitions. Two key features are the convergent power series expansion (4.14), and the uniform $O(\eta)$ error term in (4.15). In particular, we have $\tilde\rho \sim \rho(t_{\mathrm{c}}+\varepsilon) = \Theta(\varepsilon)$ for any branching process $\mathfrak{X}$ of type $(t_{\mathrm{c}}+\varepsilon,\eta)$ with $\eta \ll \varepsilon \leqslant \varepsilon_0$. In the supercritical case $t = t_{\mathrm{c}}+\varepsilon$, the survival probabilities of $\mathfrak{X}_t$ and $\mathfrak{X}$ thus both grow linearly in ε.

Theorem 4.5 (Survival probabilities). *Let $(\mathfrak{X}_t)_{t\in(t_0,t_1)} = (\mathfrak{X}_{Y_t,Z_t,Y_t^0,Z_t^0})_{t\in(t_0,t_1)}$ be a t_c-critical branching process family. Then there exist constants $\varepsilon_0, c > 0$ with the following properties. Firstly, the survival probability $\rho(t) := \mathbb{P}(|\mathfrak{X}_t| = \infty)$ is zero for $t_c - \varepsilon_0 \leqslant t \leqslant t_c$, and is positive for $t_c < t \leqslant t_c + \varepsilon_0$. Secondly, $\rho(t)$ is analytic on $[t_c, t_c + \varepsilon_0]$. More precisely, there are constants a_i with $a_1 > 0$ such that*

$$\rho(t_c + \varepsilon) = \sum_{i=1}^{\infty} a_i \varepsilon^i \tag{4.14}$$

for $0 \leqslant \varepsilon \leqslant \varepsilon_0$. Thirdly, for any t, η with $|t - t_c| \leqslant \varepsilon_0$ and $\eta \leqslant c|t - t_c|$, and any branching process $\mathfrak{X} = \mathfrak{X}_{Y,Z,Y^0,Z^0}$ of type (t, η) (with respect to $(\mathfrak{X}_t)$), the survival probability $\tilde{\rho} := \mathbb{P}(|\mathfrak{X}| = \infty)$ is zero if $t \leqslant t_c$, and is positive and satisfies

$$\tilde{\rho} = \rho(t) + O(\eta) \tag{4.15}$$

if $t > t_c$, where the implicit constant depends only on the family $(\mathfrak{X}_t)$, not on t or $\mathfrak{X}$. Moreover, analogous statements hold for the survival probabilities $\rho_1(t) := \mathbb{P}(|\mathfrak{X}^1_{Y_t,Z_t}| = \infty)$ and $\tilde{\rho}_1 := \mathbb{P}(|\mathfrak{X}^1_{Y,Z}| = \infty)$.

Proof. We argue as in the proof of Theorem 3.4. In particular, we may assume that $(Y, Z) \in \mathcal{K}$ and $(Y^0, Z^0) \in \mathcal{K}^0$ for some R, M, k_1, k_2, δ. We shall also assume that $c \leqslant 1$.

We consider only t with $|t-t_c| \leqslant \varepsilon_0$; we may assume that ε_0 is small enough that this implies $t \in (t_0, t_1)$, and, by (3.10), that $\mathbb{E}\, Y_t^0 > 0$, and that

$$\operatorname{sign}(\mathbb{E}\, Y_t - 1) = \operatorname{sign}(t - t_c) \qquad \text{and} \qquad |\,\mathbb{E}\, Y_t - 1| = \Theta(|t - t_c|) < c_{10}. \tag{4.16}$$

By (4.2) and Lemma 4.2 and it follows that $\rho_1(t) = \rho_{Y_t}$ is zero for $t_c - \varepsilon_0 \leqslant t \leqslant t_c$, and positive for $t_c < t \leqslant t_c + \varepsilon_0$. Since $\mathbb{P}(Y_t^0 \geqslant 1) > 0$, now (4.1) and (4.3) imply an analogous statement for $\rho(t) = \rho_{Y_t,Y_t^0}$. Lemmas 4.2 and 4.4 also imply that

$$\rho_1(t) = \hat{\rho}_{Y_t} \qquad \text{and} \qquad \rho(t) = \hat{\rho}_{Y_t,Y_t^0} \tag{4.17}$$

are both analytic for $t_c \leqslant t \leqslant t_c + \varepsilon_0$. Hence (4.14) holds if ε_0 is sufficiently small.

Next, for a branching process of type (t, η), by (3.16) we have $|\,\mathbb{E}\, Y_t - \mathbb{E}\, Y| = O(\eta)$. Since $\eta \leqslant c|t - t_c|$, it follows from (4.16) that if c is small enough, then $\operatorname{sign}(\mathbb{E}\, Y - 1) = \operatorname{sign}(t - t_c)$. Moreover, since $\eta \leqslant c|t - t_c| \leqslant |t - t_c| \leqslant \varepsilon_0$, using (4.16) we also have $|\,\mathbb{E}\, Y - 1| < c_{10}$ if ε_0 is small enough. Mimicking the above reasoning for $\rho_1(t)$ and $\rho(t)$, using (4.1)–(4.3) and Lemma 4.2 it follows for $\eta \leqslant c|t - t_c|$ that $\tilde{\rho}_1 = \rho_Y$ and $\tilde{\rho} = \rho_{Y,Y^0}$ satisfy $\tilde{\rho}_1 = \tilde{\rho} = 0$ if $t_c - \varepsilon_0 \leqslant t \leqslant t_c$, and $\tilde{\rho}_1, \tilde{\rho} > 0$ if $t_c < t \leqslant t_c + \varepsilon_0$; furthermore,

$$\tilde{\rho}_1 = \hat{\rho}_Y \qquad \text{and} \qquad \tilde{\rho} = \hat{\rho}_{Y,Y^0} \tag{4.18}$$

for $t_c \leqslant t \leqslant t_c + \varepsilon_0$ and $\eta \leqslant c|t - t_c|$.

Finally, we consider the interpolating branching process family $(\bar{Y}_u, \bar{Z}_u, \bar{Y}_u^0, \bar{Z}_u^0)_{u\in[0,1]}$ defined by (3.18), for which, as noted in Section 3.2, (3.1) holds with $\lambda = \eta$ and $I = [0, 1]$. Note that (3.19) and $\eta \leqslant |t-t_c| \leqslant \varepsilon_0$ imply $|\,\mathbb{E}\, \bar{Y}_u - 1| < c_{10}$ provided ε_0 is small enough. Integrating (4.10) of Lemma 4.4 over $u \in [0, 1]$ similarly to (3.20) in the proof of Theorem 3.4, using the identities (4.17)–(4.18) we infer $\tilde{\rho}_1 - \rho_1(t) = \hat{\rho}_Y - \hat{\rho}_{Y_t} = O(\eta)$ and $\tilde{\rho} - \rho(t) = \hat{\rho}_{Y,Y^0} - \hat{\rho}_{Y_t,Y_t^0} = O(\eta)$ for $t_c \leqslant t \leqslant t_c + \varepsilon_0$ and $\eta \leqslant c|t - t_c|$, completing the proof. ☐

Theorem 4.5 immediately implies the key case $p_{\mathcal{R}} = 1$, $K = 0$ of Theorem A.11 of [22], used there for the analysis of Achlioptas processes.

Acknowledgement. The last two authors are grateful to Christina Goldschmidt for useful pointers to the local limit theorem literature, which were helpful for the developing parts of the slightly more involved (large deviation based) point probability analysis contained in an earlier version of [22].

References

[1] D. Achlioptas, R.M. D'Souza, and J. Spencer. Explosive percolation in random networks. *Science* **323** (2009), 1453–1455.

[2] K.B. Athreya. Rates of decay for the survival probability of a mutant gene. *J. Math. Biol.* **30** (1992), 577–581.

[3] S. Bhamidi, A. Budhiraja, and X. Wang. The augmented multiplicative coalescent, bounded size rules and critical dynamics of random graphs. *Probab. Theory Related Fields* **160** (2014), 733–796.
[4] T. Bohman and D. Kravitz. Creating a giant component. *Combin. Probab. Comput.* **15** (2006), 489–511.
[5] L. Chaumont and R. Liu. Coding multitype forests: Application to the law of the total population of branching forests. *Trans. Amer. Math. Soc.* **368** (2016), 2723–2747.
[6] H. Cramér. Sur un nouveau théorème-limite de la théorie des probabilités. *Actualités Scientifiques et Industrielles* **736** (1938), 5–23.
[7] M. Drmota and V. Vatutin. Limiting distributions in branching processes with two types of particles. In *Classical and modern branching processes (Minneapolis, MN, 1994)*, IMA Vol. Math. Appl., 84, Springer, New York (1997), 89–110.
[8] M. Dwass. The total progeny in a branching process and a related random walk. *J. Appl. Probab.* **6** (1969), 682–686.
[9] P. Flajolet and R. Sedgewick. *Analytic Combinatorics.* Cambridge Univ. Press, Cambridge, UK (2009).
[10] G. Grimmett and D. Stirzaker. Probability and Random Processes. 3rd ed., Oxford Univ. Press, New York (2001).
[11] R. van der Hofstad, S. Janson and M. Luczak. Component structure of the configuration model: barely supercritical case. Preprint (2016). `arXiv:1611.05728`
[12] F.M. Hoppe. Asymptotic rates of growth of the extinction probability of a mutant gene. *J. Math. Biol.* **30** (1992), no. 6, 547–566.
[13] S. Janson. Simply generated trees, conditioned Galton–Watson trees, random allocations and condensation. *Probab. Surv.* **9** (2012), 103–252.
[14] S. Janson and J. Spencer. Phase transitions for modified Erdős-Rényi processes. *Ark. Math.* **50** (2012), 305–329.
[15] M. Kang, W. Perkins, and J. Spencer. The Bohman–Frieze process near criticality. *Random Struct. Alg.* **43** (2013), 221–250.
[16] *NIST Handbook of Mathematical Functions.* Eds. F. W. J. Olver, D. W. Lozier, R. F. Boisvert and C. W. Clark. Cambridge Univ. Press, Cambridge, UK (2010). Also available as *NIST Digital Library of Mathematical Functions*, `http://dlmf.nist.gov/`.
[17] J. Pitman. Enumerations of trees and forests related to branching processes and random walks. *Microsurveys in Discrete Probability (Princeton, NJ, 1997)*, DIMACS Ser. Discrete Math. and Theoret. Comput. Sci., 41, Amer. Math. Soc., Providence, RI (1998), 163–180.
[18] J. Pitman. *Combinatorial Stochastic Processes.* Lecture Notes in Mathematics, 1875. Springer-Verlag, Berlin (2006).
[19] O. Riordan and L. Warnke. Explosive percolation is continuous. *Science* **333** (2011), 322–324.
[20] O. Riordan and L. Warnke. Achlioptas process phase transitions are continuous. *Ann. Appl. Probab.* **22** (2012), 1450–1464.
[21] O. Riordan and L. Warnke. The evolution of subcritical Achlioptas processes. *Rand. Struct. & Algor.* **47** (2015), 174–203.
[22] O. Riordan and L. Warnke. The phase transition in bounded-size Achlioptas processes. Preprint (2017). `arXiv:1704.08714`
[23] B.A. Sevast'yanov. *Branching processes.* (Russian) Izdat. "Nauka", Moscow (1971).
[24] J. Spencer and N.C. Wormald. Birth control for giants. *Combinatorica* **27** (2007), 587–628.
[25] O.V. V'jugin. Transition phenomena in branching processes with final type. (Russian) *Teor. Veroyatnost. i Primenen* **23** (1978), 169–175. English translation: *Theor. Probability Appl.* **23** (1978), 163–168.
[26] A. Zygmund. *Trigonometric Series.* 2nd ed., Cambridge Univ. Press, Cambridge, UK (1959).