

# Exact solution of a two-type branching process: models of tumor progression

Tibor Antal<sup>1</sup> and P L Krapivsky<sup>2</sup>

<sup>1</sup> School of Mathematics, University of Edinburgh, Edinburgh, EH9 3JZ, UK

<sup>2</sup> Department of Physics, Boston University, Boston, MA 02215, USA

E-mail: [Tibor.Antal@ed.ac.uk](mailto:Tibor.Antal@ed.ac.uk) and [pkrapivsky@gmail.com](mailto:pkrapivsky@gmail.com)

Received 21 May 2011

Accepted 14 July 2011

Published 30 August 2011

Online at [stacks.iop.org/JSTAT/2011/P08018](http://stacks.iop.org/JSTAT/2011/P08018)

[doi:10.1088/1742-5468/2011/08/P08018](https://doi.org/10.1088/1742-5468/2011/08/P08018)

**Abstract.** An explicit solution for a general two-type birth–death branching process with one-way mutation is presented. This continuous time process mimics the evolution of resistance to treatment, or the onset of an extra driver mutation during tumor progression. We obtain the exact generating function of the process at arbitrary times, and derive various large time scaling limits. In the limit of simultaneous small mutation rate and large time scaling, the distribution of the mutant cells develops some atypical properties, including a power law tail and diverging average.

**Keywords:** exact results, stochastic processes (theory), mutational and evolutionary processes (theory), population dynamics (theory)

**ArXiv ePrint:** [1105.1157](https://arxiv.org/abs/1105.1157)

---

**Contents**

<b>1. Introduction</b>	<b>2</b>
<b>2. General case</b>	<b>4</b>
<b>3. Critical cases</b>	<b>8</b>
<b>4. Asymptotic behaviors</b>	<b>9</b>
4.1. Survival probability . . . . .	10
4.2. Critical case . . . . .	11
4.3. Supercritical case . . . . .	13
4.4. Small mutation, large time limit . . . . .	15
<b>5. Conclusions</b>	<b>16</b>
<b>Acknowledgments</b>	<b>18</b>
<b>Appendix A. Mutation at cell divisions</b>	<b>18</b>
<b>Appendix B. Case of integer <math>b</math></b>	<b>18</b>
B.1. $b = 1$ . . . . .	19
B.2. $b = 2$ . . . . .	19
B.3. Arbitrary integer $b$ . . . . .	20
<b>Appendix C. Useful formulas</b>	<b>20</b>
<b>References</b>	<b>21</b>

---

**1. Introduction**

Mathematical modeling has a long history in cancer research [1, 2]. Stochastic models helped to establish the concept of multiple mutations in tumor progression [3], and led to the understanding of tumor suppressor genes [4, 5]. These models described the transitions between various stages of cancer [6, 7] in an effective way, without considering population genetics. A more ‘microscopic’ approach is to model the stochastic evolution of a population of individual cells [8]. The simplest such models are branching processes [9, 10], with countless biological applications [11, 12] ranging from bacterial evolution [13, 14] to cell homeostasis [15]. Branching processes have also been used to model several aspects of tumor progression [2, 11], [16]–[21].

Can, however, simple branching processes provide a quantitative description of such complex biological phenomena as homeostasis—the maintenance of healthy tissues? A promising positive answer appeared in [22], where inducible genetic labeling was used to analyze the stochastic fate of the progenies of a single initial marked cell in the basal layer of epidermis in mice. The probabilities of finding clones of any given sizes at various times were fitted remarkably well by a simple constant rate two-type branching process

of progenitor ( $A$ ) and post-mitotic cells ( $C$ ). The scheme of the process is as follows:



This model seemed unsolvable [22] in the sense that there was no known explicit formula for the probability  $P_{m,n}(t)$  of finding  $m$  number of  $A$  cells and  $n$  number of  $C$  cells for any finite time. We showed in [23] how to obtain an exact expression for the generating function of  $P_{m,n}(t)$ . Once the generating function is known, the probability  $P_{m,n}(t)$  can be efficiently obtained numerically by fast Fourier transformation.

The simplicity of the above model lies in the trivial behavior of  $C$  cells—which can only die but not proliferate—and in the criticality of the process, that is the symmetry of birth and death rates of  $A$  cells. Is it possible to solve more general two-type birth–death processes explicitly where the second cell type can both die and proliferate? In this paper we present an explicit solution of such a general continuous time birth–death process, proposed by Kendall [16] to model the onset of mutations during tumor progression. In this model,  $A$  cells can mutate irreversibly into  $B$  cells according to



We calculate the generating function that encapsulates the probability distribution  $P_{m,n}(t)$  for finding  $m$  copies of  $A$  and  $n$  copies of  $B$  cells at time  $t$ , provided we have a single initial  $A$  cell. Without loss of generality, we set the division rate of  $A$  cells to one,  $\alpha_1 \equiv 1$ , since this can always be achieved by simply rescaling time. We shall use the following shorthand notations for the rate differences:

$$\lambda_1 = 1 - \beta_1 - \nu, \quad \lambda_2 = \alpha_2 - \beta_2 \tag{3}$$

which can be considered as the fitness values of the corresponding cell types. Note that there is no restriction on the signs of the parameters  $\lambda_1$  and  $\lambda_2$ .

In Kendall's original interpretation [16] two specific mutations are necessary to initiate tumor growth, in agreement with the tumor suppressor gene hypothesis [4]. An initial  $A$  cell acquires only the first mutation, hence  $A$  cells represent a benign tumor ( $\lambda_1 \leq 0$ ). When an  $A$  cell obtains the second mutation, it turns into a malignant  $B$  cell ( $\lambda_2 > 0$ ). There can also be other interpretations of model (2). It can describe the onset of a chemotherapy resistant mutation in cancer [18, 19], corresponding to the case  $\lambda_1 = \lambda_2 > 0$ , or provide a minimal model of metastasis formation, where  $B$  cells play the role of metastasized cells [24]–[26]. When an advantageous (driver) mutation appears in a supercritical process ( $\lambda_2 > \lambda_1 > 0$ ), model (2) provides a detailed description of the progression of a tumor toward malignancy. This scenario occurs in models that follow the onsets of multiple mutations [3, 21], [27]–[29]. Such a multistage model was recently found to fit clinical data (on the number of driver versus passenger mutations in

cancer) remarkably well [20]. This demonstrates that branching processes can provide a quantitative description of certain aspects of cancer.

Note that in model (2) mutations happen at arbitrary times, which can be either a reasonable simplifying assumption, or a realistic feature when, for instance, mutations are caused by UV radiation [30]. More often, however, mutations happen during cell divisions [7]; this case is described by the scheme



instead of the third process of (2). The solution of this alternative process goes completely analogously to that of (2), and is discussed in appendix A.

The two-type process (2) and other multi-type branching processes have been extensively studied [9]–[11], but the main focus was on the large time behavior. Our results, however, are exact for any finite time, which can be relevant when comparing them with experiments. For example in the experiments [22] that inspired model (1), it takes several months to get sufficiently close to the asymptotic regime, but data are already available from the first week. One can hypothesize that similar experiments could be designed to test the validity of model (2), by simply measuring the sizes of  $A$  and  $B$  clones in a tumor initiation study [16, 30, 31].

Multi-type branching processes were generally considered intractable as even in the simplest cases one arrives at (generally unsolvable) Riccati equations. Some notable exceptions are the two-type Luria–Delbruck models, which describe bacteria growth in a dish, where cell death can be neglected. Indeed, in the case of no death,  $\beta_1 = \beta_2 = 0$ , model (2) with (4) can be turned into a much easier Bernoulli equation, which leads to a simple solution [11]. Recently, however, we solved [23] a two-type branching process with cell death (1), and this renewed the hope that Kendall’s model of tumor growth is also solvable. We shall show that the processes (2) and (4) are also analytically tractable.

The rest of the paper is organized as follows. In section 2 we obtain the explicit solution of the process (2) in terms of generating functions. Special cases of the solution when either  $A$  or  $B$  cells behave critically are discussed in section 3. Knowing the explicit solution allows us to calculate  $P_{m,n}(t)$  for arbitrary times, but we also derive various asymptotic limits of the solution in section 4. Finally, we draw conclusions in section 5, and relegate some details to the appendices.

## 2. General case

Let us first try to develop an intuition for the system (2) based on exact results regarding average densities. The average number of  $A$  cells obeys the rate equation

$$\frac{d\langle m \rangle}{dt} = \lambda_1 \langle m \rangle. \quad (5)$$

Hence, starting with one  $A$  cell we get

$$\langle m \rangle = e^{\lambda_1 t}. \quad (6)$$

Similarly the average number of  $B$  cells evolves according to the rate equation

$$\frac{d\langle n \rangle}{dt} = \nu \langle m \rangle + \lambda_2 \langle n \rangle. \quad (7)$$

Using (6) and  $\langle n \rangle(0) = 0$  we solve (7) to yield

$$\langle n \rangle = \nu \frac{e^{\lambda_2 t} - e^{\lambda_1 t}}{\lambda_2 - \lambda_1}. \quad (8)$$

Higher order densities satisfy more complicated equations, but their solution is just a combination of exponentials.

The full description of the dynamics of the two-type branching process (2) is provided by the probabilities  $P_{m,n}(t)$  of having  $m$  copies of  $A$  and  $n$  copies of  $B$  at time  $t$ . These probabilities satisfy both the forward and the backward Kolmogorov equations [10]. In this paper we employ these latter equations solely for the reason that they are technically somewhat easier to tackle in the present case. The backward equations can be simply derived by accounting for all possible events that can happen to a single initial  $A$  or  $B$  cell. Let us have an  $A$  cell at time  $-\tau$  (with  $\tau \ll 1$ ) and ask for the probability  $P_{m,n}^A(t+\tau)$  of having  $(m, n)$  cells at time  $t$  (the superscript refers to the initial cell). We write this probability as a sum of the probabilities of all possible outcomes at  $t = 0$ . For example, with probability  $\alpha_1 \tau \equiv \tau$  the initial cell turns into two  $A$  cells until  $t = 0$ , hence it has a contribution  $P_{m,n}^{AA}(t)$ : the probability that two initial  $A$  cells evolved into  $(m, n)$  cells during time  $t$ . By similarly accounting for all possible transitions that can happen to an initial  $A$  or  $B$  cell and by taking the  $\tau \rightarrow 0$  limit, we arrive at the rate equations

$$\begin{aligned} \frac{dP_{m,n}^A}{dt} &= P_{m,n}^{AA} + \beta_1 \delta_{m,0} \delta_{n,0} + \nu P_{m,n}^B - (1 + \beta_1 + \nu) P_{m,n}^A \\ \frac{dP_{m,n}^B}{dt} &= \alpha_2 P_{m,n}^{BB} + \beta_2 \delta_{m,0} \delta_{n,0} - (\alpha_2 + \beta_2) P_{m,n}^B \end{aligned} \quad (9)$$

(where we omitted the  $t$  arguments). The terms containing Kronecker delta functions represent cell death which results in zero offspring at time  $t$ . We introduce the generating function

$$\mathcal{P}(x, y, t) = \sum_{m,n \geq 0} x^m y^n P_{m,n}(t). \quad (10)$$

The following shorthand notation is used to specify the initial conditions:  $\mathcal{P}(x, y, t)$  is denoted by  $\mathcal{A}(x, y, t)$  for a single initial  $A$  cell, and by  $\mathcal{B}(x, y, t)$  for a single initial  $B$  cell. We can rewrite the backward equations (9) in terms of generating functions by multiplying both sides by  $x^m y^n$  and summing over all values of  $m, n \geq 0$ . We also need to use the independence of the progenies of two initial cells, hence for example the generating function of  $P_{m,n}^{AA}$  simply becomes  $\mathcal{A}^2$ . In this way the (infinite number of) backward Kolmogorov equations (9) turn into a pair of coupled non-linear ordinary differential equations

$$\partial_t \mathcal{A} = \mathcal{A}^2 + \beta_1 + \nu \mathcal{B} - (1 + \beta_1 + \nu) \mathcal{A} \quad (11a)$$

$$\partial_t \mathcal{B} = \alpha_2 \mathcal{B}^2 + \beta_2 - (\alpha_2 + \beta_2) \mathcal{B} \quad (11b)$$

with initial conditions

$$\mathcal{A}(x, y, t = 0) = x \quad (12a)$$

$$\mathcal{B}(x, y, t = 0) = y. \quad (12b)$$

For a more thorough derivation of these types of equations see [9]–[11]. Both the positive (gain) terms and the negative (loss) terms in (11a) and (11b) correspond to the individual processes in (2) in a straightforward way, reflecting what happens to the initial cell. For example, since the initial  $A$  cell divides into two  $A$  cells at rate one, we find an  $\mathcal{A}^2$  term with coefficient one on the right-hand side of (11a). Note that the forward Kolmogorov equation is a first-order hyperbolic partial differential equation and hence one can analyze it using the method of characteristics. The equations for the characteristics are mathematically identical to the backward Kolmogorov equations (11a) and (11b), so the analytical framework (forward or backward Kolmogorov equations) is the matter of choice; the latter description is a little more convenient as it does not involve intermediate steps related to the characteristics.

Equation (11b) contains only  $\mathcal{B}$ . It is tractable and its solution, subject to (12b), reads

$$\mathcal{B} = 1 - \frac{\lambda_2}{\alpha_2(1-z)}, \quad z = \left[ 1 - \frac{\lambda_2}{\alpha_2(1-y)} \right] e^{-\lambda_2 t}. \quad (13)$$

Plugging (13) into (11a) we find that  $X \equiv 1 - \mathcal{A}$  satisfies a first-order non-linear differential equation

$$\frac{dX}{dt} = -X^2 + \lambda_1 X + \frac{\nu \lambda_2}{\alpha_2(1-z)}, \quad (14)$$

the Riccati equation. A useful way to treat Riccati equations is based on the standard trick [32] that turns a Riccati equation into a linear Sturm–Liouville equation. In the present case, this trick suggests that we write

$$X \equiv \frac{d}{dt} \log Z = \frac{1}{Z} \frac{dZ}{dt}. \quad (15)$$

One finds that  $Z$  satisfies the following Sturm–Liouville equation:

$$\frac{d^2 Z}{dt^2} = \lambda_1 \frac{dZ}{dt} + \frac{\nu \lambda_2}{\alpha_2(1-z)} Z. \quad (16)$$

Let us consider  $\lambda_2 > 0$  for an instant. In the  $t \rightarrow \infty$  limit, we have  $z \rightarrow 0$  and hence equation (16) turns into a linear equation with constant coefficients which admits an exponential asymptotic solution

$$Z \propto e^{-\omega t}, \quad \omega^2 + \lambda_1 \omega - \frac{\nu \lambda_2}{\alpha_2} = 0. \quad (17)$$

Using (13) we can rewrite the above asymptotic solution as  $Z \propto z^{\omega/\lambda_2}$  when  $z \rightarrow 0$ . This suggests that we seek the solution of equation (16) in the form

$$Z(t) \equiv z^{\omega/\lambda_2} \Phi(z). \quad (18)$$

We shall use this ansatz for arbitrary  $\lambda_2$ . Plugging (18) into (16) we obtain

$$z(1-z)\Phi'' + \left( 1 + \frac{2\omega + \lambda_1}{\lambda_2} \right) (1-z)\Phi' = \frac{\nu}{\alpha_2 \lambda_2} \Phi \quad (19)$$

where the prime denotes the derivative with respect to  $z$ . Equation (19) admits a solution in terms of the hypergeometric function. Indeed, the canonical hypergeometric equation

reads [32]

$$z(1-z)\Phi'' + [c - (a+b+1)z]\Phi' - ab\Phi = 0 \quad (20)$$

and it has two linearly independent solutions

$$F(a, b; c; z), \quad z^{1-c}F(a-c+1, b-c+1; 2-c; z). \quad (21)$$

Equation (19) coincides with (20) if the parameters  $a, b, c$  satisfy

$$a+b = \frac{2\omega + \lambda_1}{\lambda_2}, \quad ab = \frac{\nu}{\alpha_2\lambda_2}, \quad c = a+b+1. \quad (22)$$

Using (17) and (21) we arrive at

$$\Phi(z) = F(a, b; c; z) + Cz^{1-c}F(-b, -a; 2-c; z) \quad (23)$$

with parameters

$$a = \frac{\omega}{\lambda_2}, \quad b = \frac{\omega + \lambda_1}{\lambda_2}, \quad c = 1 + \frac{2\omega + \lambda_1}{\lambda_2}. \quad (24)$$

To complete the solution (18), (23)–(24) we need to know  $\omega$  and  $C$ . Recall that  $\omega$  is found from (17); the proper root is

$$\omega = -\frac{\lambda_1}{2} + \sqrt{\left(\frac{\lambda_1}{2}\right)^2 + \frac{\nu\lambda_2}{\alpha_2}}. \quad (25)$$

Recalling previous definitions we have

$$\mathcal{A} = 1 - \frac{1}{Z} \frac{dZ}{dt} = 1 + \omega + \frac{\lambda_2 z}{\Phi} \frac{d\Phi}{dz}.$$

We now compute the derivative of the hypergeometric function using identity (C.4) given in appendix C. The original generating function  $\mathcal{A}$  becomes

$$\mathcal{A} = 1 + \omega + \lambda_2\Psi(z) \quad (26)$$

where we use the shorthand notation

$$\Psi(z) = \frac{z^c F_3(z) + C(1-c)F_2(z) + CzF_4(z)}{z^{c-1}F_1(z) + CF_2(z)} \quad (27)$$

with

$$\begin{aligned} F_1(z) &= F(a, b; c; z), & F_2(z) &= F(-a, -b; 2-c; z), \\ F_3(z) &= \frac{ab}{c}F(1+a, 1+b; 1+c; z), & F_4(z) &= \frac{ab}{2-c}F(1-a, 1-b, 3-c; z). \end{aligned} \quad (28)$$

From (12a) we determine the value of the parameter

$$C = z_0^c \frac{\kappa F_1(z_0) - F_3(z_0)}{(1-c-\kappa z_0)F_2(z_0) + z_0 F_4(z_0)} \quad (29)$$

with

$$\kappa = \frac{x-1-\omega}{\lambda_2 z_0}, \quad z_0 = 1 - \frac{\lambda_2}{\alpha_2(1-y)}. \quad (30)$$

Now that we know the complete generating function  $\mathcal{A}(x, y, t)$ , we immediately know e.g. the generating function of the distribution of the number of  $B$  cells,  $\mathcal{A}(1, y, t)$ , and that of the total number of  $A$  and  $B$  cells,  $\mathcal{A}(x, x, t)$ . We also know the probability of having no  $B$  cells,  $\mathcal{A}(1, 0, t)$ , or having no cells at all,  $\mathcal{A}(0, 0, t)$ . We can extract  $P_{m,n}(t)$  from this solution by using Cauchy's integral formula

$$P_{m,n}(t) = \frac{1}{(2\pi i)^2} \oint \frac{dx}{x^{m+1}} \oint \frac{dy}{y^{n+1}} \mathcal{A}(x, y, t). \quad (31)$$

Numerically, this inverse transformation can be efficiently performed via the fast Fourier transform algorithm [23]. The solution of the alternative process (4), where mutations happen at cell divisions, is given in appendix A.

### 3. Critical cases

Here we consider the special cases of critical dynamics of  $A$  cells ( $\lambda_1 = 0$ ), of  $B$  cells ( $\lambda_2 = 0$ ), or both. If only  $A$  cells are critical, all formulas of section 2 can be simply taken in the  $\lambda_1 \rightarrow 0$  limit. If, however,  $B$  cells are critical ( $\lambda_2 = 0$ ), expressions like (24) are ill defined. This  $\lambda_2 = 0$  case for general  $\lambda_1$  is a bit cumbersome, but also implausible, as it describes a super- or subcritical population producing a completely neutral mutant. Hence we only consider here the bi-critical case  $\lambda_1 = \lambda_2 = 0$ , that is when the birth and death rates of both cell types are balanced:  $1 - \nu = \beta_1$ ,  $\alpha_2 = \beta_2$ .

The behavior of critical  $B$  cells can be obtained from (13) by taking the  $\lambda_2 \rightarrow 0$  limit

$$\mathcal{B} = \frac{y + \alpha_2 t(1 - y)}{1 + \alpha_2 t(1 - y)}. \quad (32)$$

We substitute this back into (11a), and for  $X = 1 - \mathcal{A}$  we arrive at

$$\frac{dX(t)}{dt} = -X^2 + \frac{\nu(1 - y)}{1 + \alpha_2 t(1 - y)}$$

or equivalently

$$\frac{dX(\tau)}{d\tau} = -\frac{2\alpha_2\tau}{\nu} X^2 + \frac{2\nu}{\alpha_2\tau} \quad (33)$$

in terms of the new variable

$$\tau = \sqrt{\frac{\nu}{\alpha_2} \left[ t + \frac{1}{\alpha_2(1 - y)} \right]}. \quad (34)$$

Again, by writing  $X(\tau) = (\nu/2\alpha_2 t)(d/d\tau) \log Z(\tau)$ , we simplify (33) to  $Z'' - Z'/\tau - 4Z = 0$ , which is solved by the modified Bessel functions

$$Z = \tau [I_1(2\tau) + CK_1(2\tau)] \quad (35)$$

up to an arbitrary constant. By using identities (C.1) for the modified Bessel functions, solution (35) can be transformed back to the original generating function

$$\mathcal{A}(x, y, t) = 1 - \frac{\nu}{\alpha_2\tau} \frac{I_0(2\tau) - CK_0(2\tau)}{I_1(2\tau) + CK_1(2\tau)}. \quad (36)$$

The parameter  $C$  is fixed by the initial condition (12a):

$$C = \frac{I_0(2\tau_0) - \alpha_2\nu^{-1}\tau_0(1-x)I_1(2\tau_0)}{K_0(2\tau_0) + \alpha_2\nu^{-1}\tau_0(1-x)K_1(2\tau_0)} \quad (37)$$

with

$$\tau_0 = \frac{1}{\alpha_2} \sqrt{\frac{\nu}{1-y}}. \quad (38)$$

This bi-critical case solution is somewhat simpler than the general solution (26), as it only contains low index modified Bessel functions.

#### 4. Asymptotic behaviors

Now that we have the exact solution (26), let us derive some asymptotic scaling behaviors. The first task is to determine the right (or interesting) scaling limit. In the case of studying, e.g., only the distribution  $P_m(t)$  of one type of cell (either  $A$  or  $B$ ), we first guess the leading order large time asymptotic of the typical number of surviving cells  $\chi(t)$ , and we study the  $t \rightarrow \infty$ ,  $m \rightarrow \infty$  scaling limit, with  $\hat{m} = m/\chi(t)$  kept constant. We define the scaling limit for the distribution  $P_m(t)$  of the number of cells as  $\chi(t)P_m(t) \rightarrow p(\hat{m})$ , and note that the generating function becomes a Laplace transform, if we also take the  $x \rightarrow 1$  limit, while keeping  $\xi = \chi(t) \log(1/x) \sim \chi(t)(1-x)$  constant:

$$\begin{aligned} \mathcal{A}(x, t) &= \sum_{m \geq 0} x^m P_m(t) = \sum_{m \geq 0} e^{-\hat{m}\xi} P_m(t) \\ &\rightarrow \int_0^\infty e^{-\hat{m}\xi} p(\hat{m}) d\hat{m}. \end{aligned} \quad (39)$$

The scaled density  $p(\hat{m})$  has both a singular part describing the extinction of cells and a regular part describing the surviving cells. The distribution of cells conditioned on survival is by definition  $P_m^*(t) = P_m(t)/S(t)$ , and we define its scaling limit as  $\chi(t)P_m^*(t) \rightarrow p^*(\hat{m})$ . Now by treating the first term of the sum in (39) separately, up to first order we obtain

$$\mathcal{A}(x, t) \rightarrow 1 - s(t) + s(t)\mathcal{A}^*(\xi) \quad (40)$$

with the Laplace transform

$$\mathcal{A}^*(\xi) = \int_0^\infty e^{-\hat{m}\xi} p^*(\hat{m}) d\hat{m} \quad (41)$$

where the survival probability is asymptotically  $S(t) \sim s(t)$ . Consequently,  $p(\hat{m}) = [1 - s(t)]\delta(\hat{m}) + s(t)p^*(\hat{m})$ . Our scaling describes the fate of all surviving cells if  $\mathcal{A}^*(\xi)$  is the Laplace transform of a valid probability density, i.e. if  $\mathcal{A}^*(0) = 1$ . But to obtain the right scaling of the surviving cell distribution, we first need to understand the asymptotic behavior of the survival probability.

#### 4.1. Survival probability

From the exact expressions (36) or (26) for  $\mathcal{A}(x, y, t)$ , we know the survival probability of  $A$  cells,  $S_A(t) = 1 - \mathcal{A}(0, 1, t)$ , that of  $B$  cells,  $S_B(t) = 1 - \mathcal{A}(1, 0, t)$ , and that of any cell,  $S(t) \equiv 1 - \mathcal{A}(0, 0, t)$ , at time  $t$ . Let us now use these exact expressions to understand their large time asymptotic behavior.

We first consider the bi-critical case  $\lambda_1 = \lambda_2 = 0$ . It is immediate that the survival probability of  $A$  cells  $S_A(t) \sim 1/t$  as  $t \rightarrow \infty$ . What is the survival probability of any cells  $S(t) = 1 - \mathcal{A}(0, 0, t)$ ? As  $t \rightarrow \infty$ , also  $\tau \sim \sqrt{\nu t / \alpha_2} \rightarrow \infty$ , while  $C$  from (37) remains constant. Now using the large argument limits (C.2) of the Bessel functions in (36), we find that terms containing  $C$  are asymptotically negligible and

$$S(t) \sim \sqrt{\frac{\nu}{\alpha_2 t}}. \tag{42}$$

In this bi-critical case, of course, all survival probabilities  $S_A$ ,  $S_B$  and  $S$  tend to zero, that is the eventual extinction is certain. But since  $S$  is asymptotically much larger than  $S_A$ , it means that surviving cells are typically  $B$  cells. In other words  $S_B(t) \sim S(t)$ , which is true in general for any parameter values due to the finite mutation rate  $\nu$ . This strange behavior, that neutral mutants are more likely to survive than non-mutants, can be generalized for multiple type processes, with successive mutations  $A_1 \rightarrow A_2 \rightarrow A_3 \dots$ . It turns out that the survival probability of  $A_i$  is asymptotically  $S_i \propto t^{-1/2^{i-1}}$  [33].

In the supercritical case the survival probability is positive, hence the leading order asymptotic is constant,  $S(t) \rightarrow s_\infty \equiv 1 - \mathcal{A}(0, 0, \infty)$ . Let us first assume that  $B$  cells are supercritical  $\lambda_2 > 0$ . In this case  $z(t = \infty) = 0$ , and since from (24) and (25)

$$c - 1 = \sqrt{\left(\frac{\lambda_1}{\lambda_2}\right)^2 + \frac{4\nu}{\alpha_2 \lambda_2}} > 0$$

we find that equation (26) reduces to

$$\mathcal{A}(0, 0, \infty) = 1 + \omega + \lambda_2(1 - c) = 1 - s_\infty. \tag{43}$$

Using (24) and (25) we can rewrite the probability to end up in the state without cells as

$$s_\infty = \frac{\lambda_1}{2} + \sqrt{\left(\frac{\lambda_1}{2}\right)^2 + \frac{\nu \lambda_2}{\alpha_2}} = \omega + \lambda_1. \tag{44}$$

For  $\lambda_2 \rightarrow 0$  we recover the classical single type result  $s_\infty = 1 - \beta_1 - \nu$  for  $\beta_1 + \nu < 1$ , and  $s_\infty = 0$  otherwise. For  $\lambda_2 < 0$  there is no effect of  $B$  cells on the survival probability, hence the result is the same as for  $\lambda_2 \rightarrow 0$ . Note again that in the leading order  $S_B(t) \sim S(t)$  for  $t \rightarrow \infty$  due to the finite mutation rate  $\nu$ . Note also that the asymptotic survival probability  $s_\infty$  can be obtained without the explicit knowledge of  $\mathcal{A}(x, y, t)$ , by solving the equations for  $\mathcal{A}(0, 0, t \rightarrow \infty)$  derived from (11). The above calculation therefore provides a check of self-consistency.

## 4.2. Critical case

Many interesting limits are worth investigating in the simplest bi-critical case,  $\lambda_1 = \lambda_2 = 0$ , that is when the birth and death rates of both cell types are balanced,  $\alpha_2 = \beta_2$  and  $\beta_1 = 1 - \nu$ . We are going to describe the large time scaling behavior of  $P_{m,n}(t)$  in three naturally arising regions of the  $(m, n)$  plane: (i) corresponds to extinction,  $m = n = 0$ ; (ii) corresponds to the survival of only  $B$  cells,  $m = 0, n > 0$ ; and finally (iii) corresponds to the survival of both types,  $m, n > 0$ . For all cases of this bi-critical process, the average number of cells behaves as

$$\langle m \rangle = 1, \quad \langle n \rangle = \nu t, \quad (45)$$

but these averages include extinct lineages as well. To find the correct scaling for each region we need to know the corresponding typical behaviors.

We understand the behavior in region (i) since in (42) we already obtained that all cells go extinct asymptotically with probability  $1 - \sqrt{\nu/\alpha_2 t}$ . Hence the main weight of  $P_{m,n}$  is asymptotically concentrated at the origin.

Before exploring the  $(m, n)$  plane further, as a warmup let us first consider the behavior of  $A$  cells alone, which sheds some light on the behavior of region (iii). Type  $A$  cells are not affected by  $B$  cells, and behave as a simple critical process, with generating function

$$\mathcal{A}(x, 1, t) = \frac{t(1-x) + x}{t(1-x) + 1}. \quad (46)$$

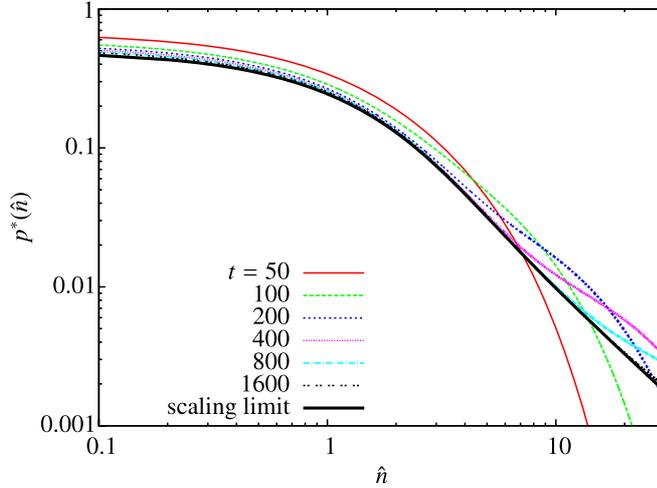
The average (and typical) number of surviving cells is asymptotically  $\chi(t) = t$ , hence following the above method we need to take the  $t, m \rightarrow \infty$  and  $x \rightarrow 0$  limits with  $\hat{m} = m/t$  and  $\xi = t(1-x)$  constant. In this limit we recover the scaling (40) with  $\mathcal{A}^*(\xi) = 1/(1+\xi)$ , with  $S_A \sim 1/t$ . By an inverse Laplace transform we obtain an exponential density  $p^*(\hat{m}) = e^{-\hat{m}}$  for the scaled number of surviving cells.

Now let us look at  $B$  cells as well. As we have shown in section 4.1, for large times if there are surviving cells they are typically  $B$  cells only, hence the next largest weight corresponds to region (ii), that is to  $m = 0, n > 0$ . Therefore, we are interested in the distribution  $P_n(t)$  of  $B$  cells, irrespective of  $A$  cells. The generating function of  $P_n(t)$  is simply  $\mathcal{A}(1, y, t)$ . Next we need to guess the typical number of surviving  $B$  cells in the asymptotic limit, which we set to  $\chi(t) = \alpha_2 t$ . The reason is that before the extinction of  $A$  cells a few  $B$  cells are produced, and they behave as a single type critical branching process of rate  $\alpha_2$ , which we have just discussed above. This guess will be justified by the scaling function we obtain, i.e. by  $\mathcal{A}^*(0) = 1$ . Alternatively, we could have just assumed a general algebraic ansatz for  $\chi(t)$ , and established its linear time dependence from the scaling.

Hence we take the scaling limit  $t \rightarrow \infty, y \rightarrow 1$ , with  $\eta = \alpha_2 t(1-y)$  kept constant, and from (34) and (38) we find that both

$$\tau_0 = \sqrt{\frac{\nu t}{\alpha_2 \eta}} \quad \text{and} \quad \tau = \tau_0 \sqrt{1 + \eta} \quad (47)$$

diverge as  $t \rightarrow \infty$ . Note that although  $\tau_0$  depends only on  $y$  in (38), we formally turned that into a  $t$  dependence from  $\eta = \alpha_2 t(1-y)$ . Therefore we use the large argument asymptotic of the modified Bessel functions (C.2), and note that  $x = 1$ , to obtain from (37)



**Figure 1.** Scaling of the probability  $P_n(t)$  of having  $n$  copies of  $B$  cells at time  $t$  in the bi-critical case, that is when death and birth rates are balanced for both cell types,  $\alpha_1 = \alpha_2 = \beta_2 = 1$ , while  $\beta_1 = 1 - \nu$ . Here we present plots for mutation rate  $\nu = 0.1$  at several different times, calculated numerically from the exact generating function (36). In the simultaneous limit  $t, n \rightarrow \infty$  and  $\hat{n} = n/\chi(t)$  being constant with  $\chi(t) = \alpha_2 t$ , the scaled distributions converge to a scaling limit  $\chi(t)P_n(t)/S_B(t) \rightarrow p^*(\hat{n})$  given by (49), where  $S_B \sim \sqrt{\nu/\alpha_2 t}$ .

the asymptotic behavior of the parameter  $C \sim e^{4\tau_0}/\pi$ . We see from (47) that  $\tau > \tau_0$  in the scaling limit, hence in (36) the terms proportional to  $C$  become negligible, and the generating function simplifies to  $A(1, y, t) = 1 - \nu/\alpha_2 \tau$ . From this we recover the scaling form (40) with

$$\mathcal{A}^*(\eta) = 1 - \sqrt{\frac{\eta}{1 + \eta}} \quad (48)$$

and with  $S_B \sim \sqrt{\nu/\alpha_2 t}$ . Note that  $\mathcal{A}^*(0) = 1$  as required. Now, by taking the inverse Laplace transform of the above expression (e.g. by integrating around the branch cut between  $\eta = -1$  and 0), we find that the surviving  $B$  cells are distributed as

$$p^*(\hat{n}) = e^{-\hat{n}/2} [I_0(\hat{n}/2) - I_1(\hat{n}/2)] \quad (49)$$

with the scaling variable  $\hat{n} = n/\alpha_2 t$ . Interestingly, the scaling function is independent of the mutation rate  $\nu$ . Note also that the above distribution has an algebraic tail  $p^*(\hat{n}) \propto \hat{n}^{-3/2}$  for  $\hat{n} \rightarrow \infty$ , which implies an infinite average number of cells. This is understandable, since the average number of surviving  $B$  cells grows as  $\propto t^{3/2}$  from (42) and (45), which is faster than the typical value  $\propto t$  we used for the scaling. The convergence to the scaling function  $p^*(\hat{n})$  for large times is illustrated in figure 1.

What is the distribution of cells when both types of cells are present, that is in the ‘bulk’ region (iii)? Note that the survival probability that cells of both types are still present at time  $t$  scales asymptotically as  $1/t$ . Since  $A$  cells behave critically, the number of surviving  $A$  cells grows linearly with time, and consequently the typical number of  $B$  cells grows as  $\nu t^2$ . This suggests that we keep  $m/t$  and  $n/\nu t^2$  constant as  $t \rightarrow \infty$ . Hence,

in the generating function (36) we should take the  $t \rightarrow \infty, x, y \rightarrow 1$  limit, with  $\xi = t(1-x)$  and  $\eta = \nu t^2(1-y)$  constant. In this limit, from (34) and (38), we find again that both

$$\tau_0 \sim \frac{t\nu}{\alpha_2\sqrt{\eta}} \quad \text{and} \quad \tau = \tau_0 + \frac{\sqrt{\eta}}{2} + \mathcal{O}\left(\frac{1}{t}\right) \quad (50)$$

diverge as  $t \rightarrow \infty$ . Consequently, all terms of parameter  $C$  of (37) have the same leading order behavior, and we arrive at

$$C \sim \frac{e^{4\tau_0} \sqrt{\eta} - \xi}{\pi \sqrt{\eta} + \xi}. \quad (51)$$

Therefore all terms containing Bessel functions in (36) have the same order, and we finally find that

$$\mathcal{A}^*(\xi, \eta) = 1 - \sqrt{\eta} \frac{\xi + \sqrt{\eta} \operatorname{th}\sqrt{\eta}}{\sqrt{\eta} + \xi \operatorname{th}\sqrt{\eta}}. \quad (52)$$

The scaling function of surviving  $A$  and  $B$  cells  $p^*(m/t, n/\nu t^2)$  is the inverse Laplace transform of (52) in both variables  $\xi$  and  $\eta$ , which we could not obtain explicitly.

### 4.3. Supercritical case

There are many possible limits to take in a non-critical system, where  $\lambda_1, \lambda_2 \neq 0$ . For example, one could consider all six possible orderings of  $0, \lambda_1, \lambda_2$ . Here we restrict our attention to the most interesting case, when an already advantageous cell lineage produces an even fitter mutant, i.e.  $\lambda_2 > \lambda_1 > 0$ . Again, we are interested in the distribution  $P_n(t)$  of  $B$  cells, irrespective of  $A$  cells, which is given by  $\mathcal{A}(1, y, t)$ .

First, let us provide an intuitive reasoning for the scaling, which will be justified later. Since  $\lambda_2 > \lambda_1$ , the first surviving mutant produces the dominant portion of  $B$  cells. A single  $B$  cell eventually survives with probability  $\lambda_2/\alpha_2$ , and the average number of  $B$  cells (including extinct lineages too) grows as  $e^{\lambda_2 t}$ . Consequently, the average number of surviving offspring of a single initial  $B$  cell grows as  $\chi(t) = (\alpha_2/\lambda_2)e^{\lambda_2 t}$  for  $t \rightarrow \infty$ , which we set as the typical number of surviving cells for the two-type process.

Hence, we study again  $\mathcal{A}(1, y, t)$  in the scaling limit  $t \rightarrow \infty, y \rightarrow 1$ , with  $\eta = \chi(t) \log 1/y \sim \chi(t)(1-y)$  being constant. Let us first determine the asymptotic value of the parameter  $C$  given in (29). In the scaling limit the parameter  $z_0$  of (30) diverges:  $|z_0| = |1 - e^{\lambda_2 t}/\eta| \rightarrow \infty$  for  $\lambda_2 > 0$ . Thus we need to determine the asymptotics of functions  $F_j(z_0)$  which appear in (29). Since  $\kappa z_0 = -a$  (this follows from (30) and (24); also recall that we consider  $x = 1$ ), the parameter  $C$  is actually given by

$$C = -z_0^{c-1} \frac{aF_1(z_0) + z_0F_3(z_0)}{-bF_2(z_0) + z_0F_4(z_0)}. \quad (53)$$

Note that  $\lambda_1 > 0$  implies  $a < b$  from (24). However, here we obtain the limit of  $C$  for the case  $a > b$ , since this derivation is simpler, and the result is the same due to

the analyticity of  $C(a, b)$ . Using the identities (C.5) and (C.6) for the hypergeometric functions one finds

$$\begin{aligned}
 F_1(z_0) &\sim (-z_0)^{-b} \frac{\Gamma(c)\Gamma(a-b)}{\Gamma(1+a)\Gamma(a)}, & F_2(z_0) &\sim (-z_0)^a \frac{\Gamma(2-c)\Gamma(a-b)}{\Gamma(-b)\Gamma(1-b)}, \\
 F_3(z_0) &\sim -\frac{b}{z_0} F_1(z_0), & F_4(z_0) &\sim -\frac{a}{z_0} F_2(z_0)
 \end{aligned}
 \tag{54}$$

as  $|z_0| \rightarrow \infty$ . Substituting (54) into (53) we obtain a neat expression for the parameter

$$C = (-1)^{a+b} \frac{b}{a} \frac{\Gamma(c)}{\Gamma(2-c)} \left[ \frac{\Gamma(-b)}{\Gamma(a)} \right]^2.
 \tag{55}$$

Note that this result is valid for arbitrary values of  $a$  and  $b$ .

It is clear from (26) that the generating function  $\mathcal{A}(1, y, t)$  of  $P_n(t)$  depends on  $t$  and  $y$  only through  $C$  and  $z$ . We have just obtained  $C$  in the scaling limit (55), and from (13) we find that  $z \sim -1/\eta$ . Hence analogously to the limiting form of (40), from the definition (26) we obtain the Laplace transform of the surviving  $B$  cell distribution in the scaling limit

$$\mathcal{A}^*(\eta) = 1 + \frac{\omega + \lambda_2 \Psi(-1/\eta)}{s_\infty}
 \tag{56}$$

where  $S(t) \rightarrow s_\infty = \omega + \lambda_1$  of (44) is the leading order constant survival probability of  $B$  cells,  $\Psi(z)$  is given by (27) with  $C$  of (55). Therefore technically all we should do is to invert the Laplace transform (56) to obtain the scaling function  $p^*(\hat{n})$  for the surviving  $B$  cells, with

$$\hat{n} = \frac{n}{\chi(t)} = n \frac{\lambda_2}{\alpha_2} e^{-\lambda_2 t}.
 \tag{57}$$

Although this is easy to do numerically, analytically it is quite challenging, since the function  $\Psi(z)$  involves the ratio of hypergeometric functions. This provides an example of a simple process where even the large time limit behavior is very complicated.

One general comment concerns the behavior of the scaling function in the limit when the scaled variable  $\hat{n}$  is large. We anticipate a generic exponential tail,  $p^*(\hat{n}) \sim U e^{-u\hat{n}}$  as  $\hat{n} \rightarrow \infty$ , which implies that the Laplace transform has a simple pole at  $\hat{n} = -u$  with residue  $U$ , that is

$$\mathcal{A}^*(\eta) = \frac{U}{u + \eta} + \dots.
 \tag{58}$$

To determine  $u, U$  we must therefore find a real and positive zero of the function  $z^{c-1} F_1(z) + C F_2(z)$ . For special cases of integer  $b$ , we can explicitly compute the scaling function  $p^*(\hat{n})$ , and verify the exponential tail. We relegate these details to appendix B.

#### 4.4. Small mutation, large time limit

In many biological applications the mutation rate refers to a single base-pair change, which is typically very small (in human DNA the probability of such a mutation is around  $5 \times 10^{-10}$  per base-pair per cell division [34]). Hence in this subsection we study the joint large time small mutation rate limit.

We consider again the case where both types are advantageous, and also  $\lambda_2 > \lambda_1 > 0$ . We take the  $\nu \rightarrow 0$  and  $t \rightarrow \infty$  limits. As  $\nu \rightarrow 0$ , we have  $\lambda_1 \rightarrow 1 - \beta_1$ , and using (24)–(25) we obtain

$$a \sim a_1\nu, \quad a_1 = \frac{1}{\alpha_2(1 - \beta_1)}, \quad b \rightarrow b_0 = \frac{1 - \beta_1}{\lambda_2} \quad (59)$$

where  $0 < b_0 < 1$ . Also, we anticipate the corresponding scaled number of  $B$  cells  $\hat{n} \rightarrow 0$ , then  $\eta \rightarrow \infty$ , hence also  $z \sim -1/\eta \rightarrow 0$ . Therefore we take a simultaneous  $\nu \rightarrow 0, \eta \rightarrow \infty$  limit, while keeping  $\eta(a_1\nu)^{1/b_0}$  constant. The reason for keeping this particular combination constant will become clear later.

The smallest order terms of (28) are particularly simple due to (C.3)

$$F_1(z) \sim F_2(z) \sim 1, \\ F_3(z) \sim \frac{a_1 b_0}{1 + b_0} \nu, \quad F_4(z) \sim \frac{a_1 b_0}{1 - b_0} \nu$$

while from (55)

$$C \sim a_1\nu B_0 e^{\pi b_0}, \quad B_0 = \frac{\pi}{\sin \pi b_0}.$$

Substituting these limits into (27) we obtain

$$\Psi(-1/\eta) = \frac{-b_0}{1 + B_0^{-1} [\eta(a_1\nu)^{1/b_0}]^{-b_0}}. \quad (60)$$

It is clear that we obtain the only nontrivial limit when the new scaling variable  $\zeta = (a_1\nu)^{1/b_0}\eta = \chi(t, \nu)(1 - y)$  is kept constant, with

$$\chi(t, \nu) = \frac{\alpha_2}{\lambda_2} e^{\lambda_2 t} (a_1\nu)^{1/b_0}. \quad (61)$$

Since for  $\nu \rightarrow 0$  the survival probability (44) is  $s_\infty \rightarrow 1 - \beta_1$ , and  $\omega \sim \lambda_2 a_1\nu$ , the scaling function  $\mathcal{A}^*$  from (56) becomes

$$\mathcal{A}^*(\zeta) = \frac{1}{1 + B_0 \zeta^{b_0}}. \quad (62)$$

This scaling function depends only on the single parameter  $0 < b_0 < 1$ . The corresponding scaled number of  $B$  cells is  $\tilde{n} = n/\chi(t, \nu)$ , with the scaling function  $p^*(\tilde{n})$ . This generating function has been studied numerically in [18], and it has already been derived in a very appealing approximate model in [21].

Let us explore the properties of  $p^*(\tilde{n})$ . For the special case  $b_0 = 1/2$  we can evaluate the inverse Laplace transform of  $\mathcal{A}^*(\zeta)$  from (62) explicitly:

$$p^*(\tilde{n}) = \frac{1}{\pi^2} \left[ \sqrt{\frac{\pi}{\tilde{n}}} - e^{\tilde{n}/\pi^2} \operatorname{Erfc} \frac{\sqrt{\tilde{n}}}{\pi} \right]. \quad (63)$$

For general  $b_0$ , from the large  $\eta$  limit of  $\mathcal{A}^*(\zeta)$ , we can obtain the small  $\tilde{n}$  series of the scaling function

$$p^*(\tilde{n}) = \sum_{j \geq 1} \frac{-1}{(-B_0)^j \Gamma(b_0 j)} \tilde{n}^{b_0 j - 1} \quad (64)$$

where  $B_0 > 0$  is given by (60). Indeed, by taking the Laplace transform of the formal series expansion of  $p^*(\tilde{n})$  around  $\tilde{n} = 0$  and equating it term by term to the series expansion of  $\mathcal{A}^*(\zeta)$  around  $\zeta = \infty$ , we obtain (64). Conversely, from the small  $\eta$  asymptotic of  $\mathcal{A}^*(\zeta)$  we obtain the large  $\tilde{n}$  series

$$p^*(\tilde{n}) = \sum_{j \geq 1} \frac{(-B_0)^j}{\Gamma(-b_0 j)} \tilde{n}^{-b_0 j - 1}. \quad (65)$$

This is achieved similarly to the small  $\tilde{n}$  series, but some care is needed. To avoid certain singularities, the series expansion of the derivative  $(d/d\zeta)\mathcal{A}^*(\zeta)$  should be compared to the Laplace transform of the expanded  $\tilde{n}p^*(\tilde{n})$ . Note that the small  $\tilde{n}$  series (64) has an infinite radius of convergence. Conversely, the large  $\tilde{n}$  series (65) has zero radius convergence, but it is asymptotic and hence truncating it after a finite number of terms provides an excellent approximation [32].

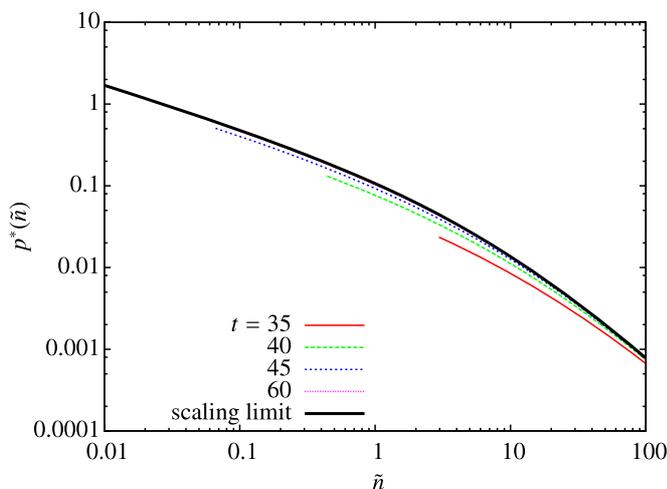
There are two surprising features of the scaling function  $p^*(\tilde{n})$ . First, from the small  $\tilde{n}$  series (64) we find that the density is singular at  $\tilde{n} = 0$  as  $p^*(\tilde{n}) \propto \tilde{n}^{b_0 - 1}$ , where  $-1 < b_0 - 1 < 0$ . Second, from the large  $\tilde{n}$  series (65) we find that the density has a power law tail  $p^*(\tilde{n}) \propto \tilde{n}^{-(1+b_0)}$ , with the exponent being  $1 < 1 + b_0 < 2$ . This means that this density function has an infinite average. The scaling for large times is demonstrated in figure 2, while the scaling function  $p^*(\tilde{n})$  is depicted in figure 3 for several values of its only parameter  $b_0$ .

The power law tail of  $p^*(\tilde{n})$  has been observed numerically in [18]. Note also that we have already encountered a limit density with infinite average in section 4.2, for a critical process for arbitrary mutation rate. Here, however, a power law density appeared in a non-critical situation, where we expect exponential tails in general (see appendix B).

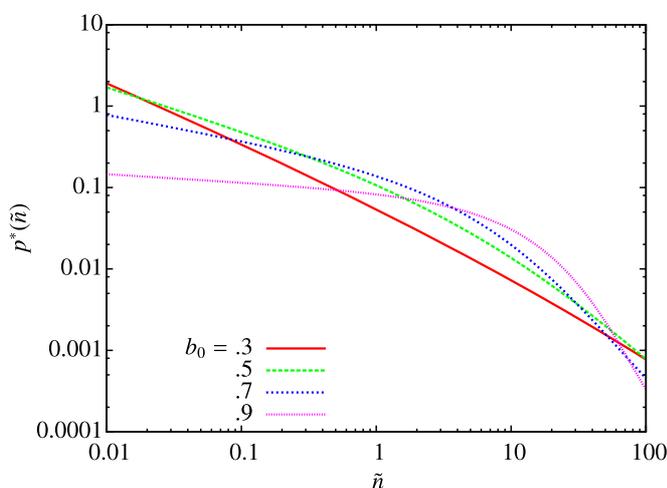
## 5. Conclusions

In this paper we presented an explicit solution for a general two-type birth–death process with one-way mutations. This process was proposed by Kendall [16] to model the onset of a driver mutation in an evolving cell population. We computed the generating function (26)–(28) that encapsulates the probability distribution  $P_{m,n}(t)$ . The generating function can be turned into the exact probabilities  $P_{m,n}(t)$  by a fast Fourier transform [23].

Our explicit result (26)–(28) is not easy to grasp, as it involves the ratio of hypergeometric functions. It becomes somewhat simpler, equation (36), in the bi-critical case, that is when both cell types behave critically,  $\lambda_1 = \lambda_2 = 0$ . The complete solutions, however, contain all information; for instance one immediately knows the survival probability of cells, or any order moments of the cell distribution. We have also extracted several interesting limits from the exact solutions. In the bi-critical case we showed in section 4.2 how different types of scaling apply in different regions of the  $(m, n)$  plane. For the onset of an advantageous mutation,  $\lambda_2 > \lambda_1 > 0$ , we derived the



**Figure 2.** Scaling of the probability  $P_n(t)$  of having  $n$  copies of  $B$  cells at time  $t$  in the supercritical case for the simultaneous limit of large time, large number of cells and small mutation rate,  $t, n \rightarrow \infty, \nu \rightarrow 0$  with  $\tilde{n} = n/\chi(t, \nu)$  constant and  $\chi(t, \nu)$  given by (61). We chose  $\alpha_1 = 1, \beta_1 = 0.81, \alpha_2 = 1.29, \beta_2 = 0.91$ , which correspond to the scaling parameter  $b_0 = 1/2$ . The plots are for mutation rate  $\nu = 0.0001$  at different times, calculated numerically from the exact generating function (26). The scaled distributions converge to the limit function  $\chi(t, \nu)P_n(t)/S_B(t) \rightarrow p^*(\tilde{n})$  given by (63), where  $S_B \rightarrow 1 - \beta_1$ .



**Figure 3.** Scaling function  $p^*(\tilde{n})$  of  $B$  cells for the supercritical case in the large time small mutation limit evaluated from (64). We plotted this function at four values of the single parameter  $b_0$ . Note the explicit formula (63) for  $b_0 = 1/2$ . The small and large  $\tilde{n}$  asymptotic power law behavior can be observed.

large time scaling limit, which limit function still involved hypergeometric functions. This demonstrates that the large time limit behavior can still be very complicated. For certain special cases though, we derived the explicit scaling functions, with generic exponential tails. Conversely, in the simultaneous large time and small mutation limit, the scaling function has a power law tail, with infinite average value. As a consequence, an enormous

number of samples are needed in simulations to recover the exact (finite time) average values.

There are, of course, numerous interesting open questions for two-type branching processes which deserve further attention. The long term description of tumor development though must involve more than two cell types corresponding to multiple stages of tumors [3, 27, 35]. These multi-type branching processes are likely too cumbersome for explicit solutions, and one needs simplifying assumptions to deduce the relevant asymptotic behavior [21, 33]. Nevertheless, some effective models have recently proved to be remarkably successful in connecting theory and clinical data [20].

### Acknowledgments

We thank Ivana Bozic for useful suggestions. PLK gratefully acknowledges financial support from NSF grant CCF-0829541.

### Appendix A. Mutation at cell divisions

Here we briefly discuss a version of the two-type process (2), where mutations happen at cell divisions (4). In (11a) the term  $\nu\mathcal{A}$  should be replaced by  $\nu\mathcal{AB}$ ; everything else in (11a)–(11b) remains the same. We follow the exact same steps as for the original model, and we arrive at the same solution (26)–(30), but with parameters

$$a = \frac{\omega}{\lambda_2}, \quad b = \frac{\omega + 1 - \beta_1}{\lambda_2}, \quad c = 1 + a + b - \frac{\nu}{\alpha_2}$$

and

$$\omega = - \left( \frac{\lambda_1}{2} + \frac{\nu\beta_2}{2\alpha_2} \right) + \sqrt{\left( \frac{\lambda_1}{2} + \frac{\nu\beta_2}{2\alpha_2} \right)^2 + \frac{\nu\lambda_2}{\alpha_2}}$$

instead of (24) and (25). Also, while  $F_1(z)$  and  $F_3(z)$  are unchanged from (28), we need  $F_2(z)$  and  $F_4(z)$  in their general forms:

$$F_2(z) = F(1 + a - c, 1 + b - c; 2 - c; z)$$

$$F_4(z) = \frac{(1 + a - c)(1 + b - c)}{2 - c} F(2 + a - c, 2 + b - c, 3 - c; z)$$

since in (28) we simplified these general formulas by using  $c = 1 + a + b$ , which is not true in the present case.

### Appendix B. Case of integer $b$

In section 4.3 we conjectured that the scaled distribution  $p^*(\hat{n})$  has an exponential tail in general. In this appendix we review a few concrete examples for the special case of integer  $b$  values. Recall that we assumed  $\lambda_2 > \lambda_1 > 0$  in section 4.3, which implies  $b > a > 0$  from (24). We do not impose any other restriction on  $a$ .

Equation (55) shows that  $C$  diverges when  $b$  is a non-negative integer. This case corresponds to particular values of the mutation rate

$$\nu = \frac{1 - \beta_1 - b\lambda_2}{1 - (b\alpha_2)^{-1}}$$

from the definitions (17) and (24), provided that  $0 < \nu < 1 - \beta_1$ . Our sole reason to consider these special mutation rates is to make the problem more tractable, and we anticipate that the behavior is qualitatively similar for arbitrary mutation rates. Indeed, when  $C = \infty$ , equation (27) simplifies to

$$\Psi = 1 - c + z \frac{F_4(z)}{F_2(z)}$$

and hence (56) becomes

$$\mathcal{A}^*(-1/z) = 1 + \frac{1}{s_\infty} \left[ \omega + \lambda_2(1 - c) + \lambda_2 z \frac{F_4(z)}{F_2(z)} \right].$$

Note that  $\eta = -1/z$  in this limit. The constant term on the right-hand side vanishes, see (43), and therefore we finally arrive at

$$\mathcal{A}^*(-1/z) = \frac{\lambda_2 z F_4(z)}{s_\infty F_2(z)} = \frac{z}{b} \frac{d \log F_2}{dz}. \tag{B.1}$$

Here we also used that  $s_\infty/\lambda_2 = b$  from (24) and (44). Note that this relation provides another interpretation for integer  $b$ . Let us consider now concrete examples. Since  $b > 0$ , the simplest possible choice is  $b = 1$ .

**B.1.  $b = 1$**

Since  $c = a + b + 1 = a + 2$  we conclude that

$$F_2 = F(-1, -a; -a; z) = 1 - z, \quad F_4 = -1.$$

Therefore equation (B.1) becomes  $\mathcal{A}^*(-1/z) = z/(z - 1)$ . In other words,  $\mathcal{A}^*(\eta) = 1/(1 + \eta)$ , from which  $p^*(\hat{n}) = \exp(-\hat{n})$ . Thus in this case the scaled distribution of  $B$  cells is a pure exponential.

**B.2.  $b = 2$**

We have  $c = a + b + 1 = a + 3$  and

$$F_2 = F(-2, -a; -1 - a; z) = 1 - \frac{2a}{a + 1} z + \frac{a - 1}{a + 1} z^2$$

$$F_4 = \frac{dF_2}{dz} = -\frac{2a}{a + 1} + 2 \frac{a - 1}{a + 1} z.$$

Equation (B.1) becomes

$$\mathcal{A}^*(\eta) = \frac{a - 1 + a\eta}{a - 1 + 2a\eta + (a + 1)\eta^2}$$

and its inverse Laplace transform

$$p^*(\hat{n}) = \frac{1}{2} \left[ e^{-\hat{n}} + \frac{a - 1}{a + 1} \exp \left( -\hat{n} \frac{a - 1}{a + 1} \right) \right]$$

is a combination of two exponents. It is possible to obtain explicit expressions for larger  $b$  values in particular, but let us consider the problem now in general.

### B.3. Arbitrary integer $b$

In this case  $F_2 = F(-b, -a; 1 - b - a; z)$  is a polynomial of the order  $b$ ; hence  $F_4$  is a polynomial of order  $b - 1$ . It is easy to see that  $\mathcal{A}^*(\eta)$  is the ratio of a polynomial of  $\eta$  of the order  $b - 1$  to the polynomial of  $\eta$  of the order  $b$ . Therefore one can write an exact expansion

$$\mathcal{A}^*(\eta) = \sum_{j=1}^b \frac{U_j}{u_j + \eta}.$$

Thus when  $b$  is integer, the scaling function is a combination of  $b$  exponentials:

$$p^*(\hat{n}) = \sum_{j=1}^b U_j e^{-u_j \hat{n}}.$$

### Appendix C. Useful formulas

Here we list a few identities and limits of special functions from [36, 37]. The modified Bessel functions satisfy

$$\begin{aligned} 2I_1'(z) &= I_0(z) + I_2(z), & -2K_1'(z) &= K_0(z) + K_2(z), \\ 2I_1(z)/z &= I_0(z) - I_2(z), & -2K_1(z)/z &= K_0(z) - K_2(z) \end{aligned} \tag{C.1}$$

and their  $z \rightarrow \infty$  asymptotic behavior is

$$I_a(z) \sim \frac{1}{\sqrt{2\pi z}} e^z, \quad K_a(z) \sim \sqrt{\frac{\pi}{2z}} e^{-z}. \tag{C.2}$$

The hypergeometric function has the series expansion

$$F(a, b; c; z) = 1 + \frac{ab}{c} \frac{z}{1!} + \frac{a(a+1)b(b+1)}{c(c+1)} \frac{z^2}{2!} + \dots \tag{C.3}$$

and the derivative

$$\frac{d}{dz} F(a, b; c; z) = \frac{ab}{c} F(1 + a, 1 + b; 1 + c; z). \tag{C.4}$$

We use the transformation formulas

$$\begin{aligned} F(a, b; c; z) &= (1 - z)^{-a} F(a, c - b; c; 1 - 1/z) \\ &= (1 - z)^{-b} F(b, c - a; c; 1 - 1/z) \end{aligned} \tag{C.5}$$

and the identity

$$F(a, b; c; 1) = \frac{\Gamma(c)\Gamma(c - a - b)}{\Gamma(c - a)\Gamma(c - b)} \tag{C.6}$$

for  $\text{Re } c > \text{Re}(a + b)$ .

## References

- [1] Frank S A, 2007 *Dynamics of Cancer: Incidence, Inheritance, and Evolution* (Princeton, NJ: Princeton University Press)
- [2] Attolini C S O and Michor F, *Evolutionary theory of cancer*, 2009 *Ann. NY Acad. Sci.* vol 1168, p 23
- [3] Armitage P and Doll R, *The age distribution of cancer and a multi-stage theory of carcinogenesis 1954*, 2004 *Int. J. Epidemiol.* **33** 1174
- [4] Knudson A, *Mutation and cancer—statistical study of retinoblastoma*, 1971 *Proc. Nat. Acad. Sci.* **68** 820
- [5] Kern S E, *Research philosophy—whose hypothesis? CIPHERING, sectorials, d lesions, freckles and the operation of stigler’s law*, 2002 *Cancer Biol. Ther.* **1** 571
- [6] Weinberg R A, 2007 *The Biology of Cancer* (New York: Garland Sciences, Taylor and Francis)
- [7] Vogelstein B and Kinzler K W, 2002 *The Genetic Basis of Human Cancer* 2nd edn (New York: McGraw-Hill, Medical Pub. Division)
- [8] Nowak M A, 2006 *Evolutionary Dynamics: Exploring the Equations of Life* (Cambridge, MA: Harvard University Press)
- [9] Harris T E, 1963 *The Theory of Branching Processes* (New York: Dover)
- [10] Athreya K B and Ney P, 1972 *Branching Processes* (Berlin: Springer)
- [11] Kimmel M and Axelrod D E, 2002 *Branching Processes in Biology* (New York: Springer)
- [12] Haccou P, Jagers P and Vatutin V A, 2005 *Branching Processes: Variation, Growth, and Extinction of Populations* (Cambridge: Cambridge University Press)
- [13] Bartlett M S, 1955 *An Introduction to Stochastic Processes: with Special Reference to Methods and Applications* (Cambridge: University Press)
- [14] Zheng Q, *Progress of a half century in the study of the Luria–Delbruck distribution*, 1999 *Math. Biosci.* **162** 1
- [15] Pharr P N, Nedelman J, Downs H P, Ogawa M and Gross A J, *A stochastic model for mast cell proliferation in culture*, 1985 *J. Cell. Physiol.* **125** 379
- [16] Kendall D G, *Birth-and-death processes, and the theory of carcinogenesis*, 1960 *Biometrika* **47** 13
- [17] Luebeck E and Moolgavkar S, *Multistage carcinogenesis and the incidence of colorectal cancer*, 2002 *Proc. Nat. Acad. Sci.* **99** 15095
- [18] Iwasa Y, Nowak M and Michor F, *Evolution of resistance during clonal expansion*, 2006 *Genetics* **172** 2557
- [19] Komarova N, *Stochastic modeling of drug resistance in cancer*, 2006 *J. Theor. Biol.* **239** 351
- [20] Bozic I, Antal T, Ohtsuki H, Carter H, Kim D, Chen S, Karchin R, Kinzler K W, Vogelstein B and Nowak M A, *Accumulation of driver and passenger mutations during tumor progression*, 2010 *Proc. Nat. Acad. Sci.* **107** 18545
- [21] Durrett R and Moseley S, *Evolution of resistance and progression to disease during clonal expansion of cancer*, 2010 *Theor. Popul. Biol.* **77** 42–48
- [22] Clayton E, Doupe D P, Klein A M, Winton D J, Simons B D and Jones P H, *A single type of progenitor cell maintains normal epidermis*, 2007 *Nature* **446** 185
- [23] Antal T and Krapivsky P L, *Exact solution of a two-type branching process: clone size distribution in cell division kinetics*, 2010 *J. Stat. Mech.* P07028
- [24] Michor F, Nowak M and Iwasa Y, *Stochastic dynamics of metastasis formation*, 2006 *J. Theor. Biol.* **240** 521
- [25] Dingli D, Michor F, Antal T and Pacheco J M, *The emergence of tumor metastases*, 2007 *Cancer Biol. Ther.* **6** 383
- [26] Yachida S *et al.*, *Distant metastasis occurs late during the genetic evolution of pancreatic cancer*, 2010 *Nature* **467** 1114
- [27] Fearon E R and Vogelstein B, *A genetic model for colorectal tumorigenesis*, 1990 *Cell* **61** 759
- [28] Beerenwinkel N, Antal T, Dingli D, Traulsen A, Kinzler K W, Velculescu V E, Vogelstein B and Nowak M A, *Genetic progression and the waiting time to cancer*, 2007 *PLoS Comput. Biol.* **3** 2239
- [29] Durrett R, Foo J, Leder K, Mayberry J and Michor F, *Evolutionary dynamics of tumor progression with random fitness values*, 2010 *Theor. Popul. Biol.* **78** 54
- [30] Klein A M, Brash D E, Jones P H and Simons B D, *Stochastic fate of p53-mutant epidermal progenitor cells is tilted toward proliferation by uv b during preneoplasia*, 2010 *Proc. Nat. Acad. Sci.* **107** 270
- [31] Klein A M, Doupe D P, Jones P H and Simons B D, *Kinetics of cell division in epidermal maintenance*, 2007 *Phys. Rev. E* **76** 021910
- [32] Bender C M and Orszag S A, 1978 *Advanced Mathematical Methods for Scientists and Engineers (International Series in Pure and Applied Mathematics)* (New York: McGraw-Hill)
- [33] Antal T and Krapivsky P L, *Multiple-type critical branching processes*, 2011 in preparation

- [34] Jones S *et al*, *Comparative lesion sequencing provides insights into tumor evolution*, 2008 *Proc. Nat. Acad. Sci.* **105** 4283
- [35] Fearon E R, Hamilton S R and Vogelstein B, *Clonal analysis of human colorectal tumors*, 1987 *Science* **238** 193
- [36] Gradshteyn I S, Ryzhik I M and Jeffrey A, 2000 *Table of Integrals, Series, and Products* (San Diego, CA: Academic)
- [37] Olver F W J, 2010 *NIST Handbook of Mathematical Functions* (Cambridge: Cambridge University Press)