HEAVY-TAILED DISTRIBUTIONS IN BRANCHING PROCESS MODELS OF SECONDARY CANCEROUS TUMORS

By PHILIP A. ERNST, MAREK KIMMEL, MONIKA KURPAS AND QUAN ZHOU

Abstract

Recent progress in microdissection and in DNA sequencing has facilitated the subsampling of multi-focal cancers in organs such as the liver in several hundred spots, helping to determine the pattern of mutations in each of these spots. This has led to the construction of genealogies of the primary, secondary, tertiary, and so forth, foci of the tumor. These studies have led to diverse conclusions concerning the Darwinian (selective) or neutral evolution in cancer. Mathematical models of the development of multi-focal tumors have been devised to support these claims. We offer a model for the development of a multi-focal tumor: it is a mathematically rigorous refinement of a model of Ling *et al.* (2015). Guided by numerical studies and simulations, we show that the rigorous model, in the form of an infinite-type branching process, displays distributions of tumor size which have heavy tails and moments that become infinite in finite time. To demonstrate these points, we obtain bounds on the tails of the distributions of the process and an infinite series expansion for the first moments. In addition to its inherent mathematical interest, the model is corroborated by recent literature on apparent super-exponential growth in cancer metastases.

Keywords: Branching process; mutation; cancer cell; heterogeneity; heavy tail; Yule–Simon distribution; infinite moment

2010 Mathematics Subject Classification: Primary 60J80
Secondary 60J85

1. Introduction

Growth patterns and heterogeneity of cancer metastases are not well understood, although it seems clear that they are a product of mutation, genetic drift, migration, selection, and perhaps other population genetics and population dynamics mechanisms. Recently, Baratchart *et al.* [3] reported that in some animal models, metastases exhibit growth patterns which appear to be super-exponential, and in a series of careful experiments and computations, they provided an explanation for this phenomenon involving rather complex biological mechanisms. Our aim in this paper is to show that such behavior can be generated by a fairly simple growth and mutation model.

Our work is also motivated by recent work [7] that presents an analysis of a sequencing experiment using nearly 300 samples taken from a section of hepatocellular carcinoma tumor. The purpose of the analysis has been to compare the Darwinian and non-Darwinian theories of development of human solid cancers. The Darwinian model in [7] involves a primary tumor growing and shedding secondary foci with distributed growth rates, which may seem to be an attractive way of modeling competition among the secondary foci. However, when examining the details of this approach, we found that the 'Darwinian' model introduced in [7] (and also treated in [10]) displays peculiar behavior, characterized by the existence of outlier trajectories and the explosion of expected cell counts in finite time. We attribute this behavior to the exponential model used for the distribution of growth rates.

Similar behavior is exhibited by a simple 'toy model' that involves exponential growth curves with Malthusian parameter (growth rate) that itself is an exponentially distributed random variable; where convenient, $Z \sim \text{Exp}(\lambda)$ means that the random variable Z is exponentially distributed with mean $1/\lambda$. In the toy model the explosions are related to the distributions of population size being of Pareto type with coefficients changing in time. However, neither the model of [7] nor the toy model is truly stochastic in the sense of describing cell divisions and mutations as stochastic events that occur as the cell population evolves in time. This led us to build a stochastic population model as a branching process, in which (for simplicity) lifelengths of cells are assumed to be exponentially distributed. At each division, one progeny cell may mutate and acquire a new lifelength distribution which is exponential with the parameter sampled from an exponential distribution. This process may be classified as an age-dependent Markov branching process with a nondenumerable type space. We show that the process exhibits finite-time explosions of expected values, while simulations indicate Pareto-like tails, with exponents changing in time and becoming equal to 1 at the time the expectations explode. We develop a set of bounds that are consistent with the simulation findings and prove the finite-time explosion of the expected values of the process.

2. Baseline model

Our baseline model is in principle the deterministic model considered in [7] and [10] (see Figure 1(a)). A primary tumor is generated from a single cell born at time $t = 0$ and grows at the rate $g(x_p) = bx_p$, where b is a constant and $x_p := x_p(t)$ is the size of the tumor at time t; think of the size as the total number of cells in the tumor, and note that, from $\mathrm{d}x_p/\mathrm{d}t = g(x_p) = bx_p$ and $x_p(0+) = 1$, $x_p(t) = \mathrm{e}^{bt}$. This growing tumor emits secondary tumor cells at the rate $\beta(x_p) := mx_p^\alpha$. Each secondary tumor grows at a rate $g_s(x) = ax$, where the rate a may vary from one cell to another, and emits further secondary tumor cells: similarly to the primary cell, when a given secondary tumor is of size x_s, it gives birth to further secondary tumors at rate $\beta(x_s)$. Note that a secondary tumor born at time u is of size $\mathrm{e}^{a(t-u)}$ at time $t \geq u$. (In numerical work, we take $\alpha = \frac{2}{3}$ to reflect the shedding of secondary tumors from the surface of the primary tumor of 'size' (volume) x.)

Consistent with the deterministic convention for structured populations, in this section we use a deterministic nonnormed size distribution density function $\rho(x, t)$ to be interpreted as follows. At time $t > 0$ the count of colonies with sizes between x and $x + \Delta x$ equals $\rho(x, t)\Delta x + o(\Delta x)$, where $o(\cdot)$ is small. Following [5], the dynamics of the secondary cell colony size distribution density are given by the von Förster-type equation for $(x, t) \in [1, \infty) \times [0, \infty)$,

$$\frac{\partial \rho(x, t)}{\partial t} + \frac{\partial g(x)\rho(x, t)}{\partial x} = 0, \tag{1}$$

subject to the initial condition $\rho(x, 0) = 0$ and the nonlocal boundary conditions

$$g(1)\rho(1, t) = \int_1^{\mathrm{e}^{at}} \beta(x)\rho(x, t)\,\mathrm{d}x + \beta(x_p(t)), \qquad \beta(x) := mx^\alpha.$$

As shown in Appendix A, if the growth rates a and b are constant, the solution of (1) with the stated boundary conditions is of the form

$$\rho(x, t) = \begin{cases} \dfrac{m[\alpha(a - b)\mathrm{e}^{\alpha bt}x^{-(\alpha b/a+1)} + m\mathrm{e}^{(m+a\alpha)t}x^{-(\alpha+m/a+1)}]}{a[m + \alpha(a - b)]} & \text{if } x < \mathrm{e}^{at}, \\ 0 & \text{otherwise.} \end{cases}$$
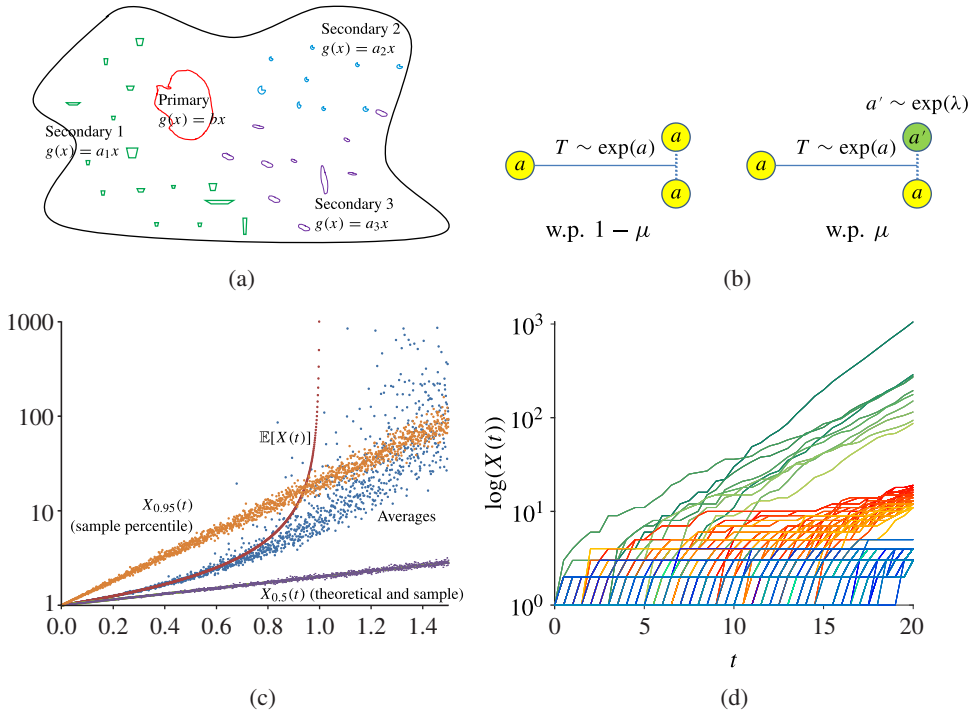
FIGURE 1: (a) Ideogram representation of the baseline process. Within the 'tumor field', the primary tumor grows exponentially at rate b, and then sheds secondary tumors, which may shed further secondary tumors. Secondary tumors grow at rates generally different from that of the primary tumor. (b) Hypotheses underlying the modified Goldie–Coldman model. (c) Toy model with $\lambda = 1$: expected value $\mathbb{E}[X(t)]$ of the process defined in Section 3.1 and which explodes at $t = 1$, averages of 1000 realizations of $X(t)$, and 0.5 and 0.95 quantiles of $X(t)$, all in semi-logarithmic scale. (d) Summary of 10 000 simulated trajectories of the modified Goldie–Coldman process with parameters $\mu = 0.5$, $a = 0.01$, and $\lambda = 10$. Depicted are realizations of the process ranking 1–10 (green), 51–100 (red), and 301–400 (blue) at time $t = 20$.

Then the tail of the distribution corresponding to this density $\rho(x, t)$ has the form

$$
\begin{aligned}
G(x) &:= G(x; a, b) \\
&= \int_x^{e^{at}} \rho(\xi, t)\, d\xi \\
&= \begin{cases} \dfrac{m[((a-b)/b)(e^{\alpha bt}x^{-\alpha b/a} - 1) + (m/(a\alpha + m))(e^{(m+\alpha a)t}x^{-\alpha - m/a} - 1)]}{m + \alpha(a - b)} & \text{if } x \in [1, e^{at}], \\ 0 & \text{if } x > e^{at}. \end{cases}
\end{aligned}
\tag{2}
$$

This function increases exponentially at rate $\max(\alpha b, \alpha a + m)$. However, if the growth rate of secondary tumors a is random and follows the exponential distribution with mean $1/\lambda$ then instead of $G(\cdot)$ at (2) we have

$$
\widetilde{G}(x; b) = \int_{\ln(x)/t}^{\infty} G(x; a, b)\lambda e^{-\lambda a}\, da.
$$

Substitution of the expression for $G(x; a, b)$ leads to intractable integrals, except for the case $x = 1$ (total count of secondary foci), when it leads to the relation

$$\widetilde{G}(1; b) = \frac{m(\mathrm{e}^{b\alpha t} - 1)}{b\alpha} + (2 - \mathrm{e}^{b\alpha t})\frac{m^2\lambda}{b\alpha^2}\mathrm{e}^{\lambda(m/\alpha - b)}\Gamma\left(0, \lambda\left(\frac{m}{\alpha - b}\right)\right)$$

$$+ \frac{m^2\lambda}{b\alpha^2}\left\{-\mathrm{e}^{mt}\left[\mathrm{e}^{(\lambda - \alpha t)(m/\alpha - b)}\Gamma\left(0, (\lambda - \alpha t)\left(\frac{m}{\alpha - b}\right)\right)\right.\right.$$

$$\left.\left. - \mathrm{e}^{(\lambda - \alpha t)(m/\alpha)}\Gamma\left(0, \frac{(\lambda - \alpha t)m}{\alpha}\right)\right] - \mathrm{e}^{\lambda m/\alpha}\Gamma\left(0, \frac{\lambda m}{\alpha}\right)\right\},$$

where $\Gamma(z, w) = \int_w^\infty \mathrm{e}^{-t}t^{z-1}\,\mathrm{d}t$ is the incomplete gamma function for which $\Gamma(0, w) \sim -\log(w)$ as $w \downarrow 0$. Thus, the solution increases to infinity as $t \uparrow \lambda/\alpha$. This highly irregular behavior of the 'quasi-stochastic' version of the baseline model inspired us to seek a fully stochastic model with analogous behavior. We present this model in the following section.

3. Branching process model

3.1. Stochastic toy model

Can a truly stochastic model display the same behavior as the baseline model? We begin with a toy model, as follows. In place of the growth rate $x_s(t) = \mathrm{e}^{at}$ let secondary tumors grow exponentially at rate a, which itself is a random variable, so that

$$X(t \mid a) = \exp(at), \quad t \geq 0, \qquad a \sim \exp(\lambda).$$

For $t > 0$, it now has Pareto tail

$$\mathbb{P}[X(t) > x] = \begin{cases} 1 & \text{if } 0 \leq x < 1, \\ x^{-\lambda/t} & \text{if } x \geq 1. \end{cases}$$

We integrate the above to obtain

$$\mathbb{E}[X(t)] = \int_0^\infty \mathbb{P}[X(t) > x]\,\mathrm{d}x = 1 + \int_1^\infty x^{-\lambda/t}\,\mathrm{d}x = \begin{cases} \lambda(\lambda - t)^{-1} & \text{for } t < \lambda, \\ \infty & \text{for } t \geq \lambda. \end{cases}$$

What of ∞ here? Will tumors really *explode*? We turn our attention to this matter.

3.2. Modified Goldie–Coldman model

Consider a modified Goldie–Coldman (G–C) model (Figure 1(b)) of which the classical version can be found in [6, Section 4.2].

1. Cells are organized in proliferating clones; each clone is characterized by its (random) division rate $a \sim \mathrm{Exp}(\lambda)$; within each clone cells proliferate according to a continuous-time Markov branching process with perfect binary fission after lifetimes that are i.i.d. $\mathrm{Exp}(a)$ for all cells.

2. At each binary fission epoch (and for each cell such epochs occur at the characteristic rate a of the clone to which the dividing cell belongs), one of the offspring cells inherits the characteristic rate a, while for the other cell, either with probability $1 - \mu$ it also inherits the characteristic rate a or with probability μ it is the first cell in a new clone with division rate $a' \sim \mathrm{Exp}(\lambda)$, this rate being determined independently of the division rates for all other clones and independently for each binary fission epoch.

3. The process is originated by an ancestor cell of given type a.

4. A variant of the last assumption is to have the type of the ancestor cell randomly distributed as $a \sim \text{Exp}(\lambda)$.

Formally, consider one ancestor tumor cell with division rate a at time 0. At each division, with probability μ, it can divide into one cell with rate a and another cell with division rate a', where $a' \sim \text{Exp}(\lambda)$. The new type of tumor cells (with rate a') have the same mutation rate, μ, and can continue mutating into new subtypes with random division rate generated from $\text{Exp}(\lambda)$. All the tumor cells are assumed to be independent of each other. Let $X_k(a, t)$ be the number of tumor cells that are generated by $k - 1$ mutations. Accordingly, $X_1(a, t)$ denotes the number of primary tumor cells, i.e. the cells with division rate a; $X_2(a, t)$ denotes the number of cells of types that directly mutated from primary tumor cells; $X_3(a, t), X_4(a, t), \dots$ are defined analogously.

The resulting model is a continuum-type continuous-time Markov branching process. A backward Kolmogorov-type ordinary differential equation can be written for the joint probability generating function (PGF) of the distribution of total cell counts in all clones.

We begin by presenting simulation results, which motivate the more mathematical study that follows. We then perform some asymptotic calculations to characterize the tail distribution of the cell counts of different types. In particular, we show that the tail probability of these cell counts can be bounded from below by a power law with exponent $-\lambda/[(1 - \mu)t]$.

Finally, we derive equations for the PGFs of the total counts of cells in the process. We proceed to derive an infinite series solution for the expected counts of cells and show that it explodes in finite time, almost exactly as occurs in the toy model.

3.3. Simulation results

We begin with the toy model because it provides guidance concerning the behavior of the branching process model. For the version of the toy model with $\lambda = 1$, Figure 1(c) depicts the true expected value $\mathbb{E}[X(t)]$ of the process (which explodes at $t = 1$), averages of 1000 realizations of $X(t)$, and 0.5 and 0.95 quantiles of $X(t)$, all in semi-logarithmic scale. Note that the averages increase faster than any exponential, while the quantiles grow exponentially. The explosion at $t = 1$ is analogous to the behavior exhibited by the baseline model.

Consider next the modified G–C model. We conducted extensive simulations with it, assuming a wide range of parameter values. Selected results are shown in Figure 1(d) and Figures 2–5. Figure 1(d) is based on 10 000 simulated trajectories of the modified G–C process with parameters $\mu = 0.5$, $a = 0.01$, and $\lambda = 10$. Depicted are realizations of the process ranking 1–10 (green), 51–100 (red), and 301–400 (blue) at time $t = 20$. The distribution of trajectories exhibits strong right skewness and suggests heavy tails. Figures 2–4 depict averages of the simulated trajectories of the modified G–C process in three different cases: $\mu = 0.5$, $a = 0.01$, $\lambda = 10$ (Figure 2), $\mu = 0.5$, $a = 0.01$, $\lambda = 100$ (Figure 3), and $\mu = 0.1$, $\lambda = 100$ (Figure 4), based on 200, 1000, and 10 000 trajectories, with expectations $M(a, t)$ computed by numerically solving the integral equation (17) for $\varphi(t)$ and using (16). The averages are convex in semi-log coordinates; this suggests super-exponential growth, but they underestimate the growth of the expectation, which explodes to infinity at $t = \lambda/(1 - \mu)$.

Figure 5 shows simulated tail behavior of the modified G–C process. We depict estimated power exponents of the tail of $X(a, t)$, $t \in [0, \lambda/(1-\mu)]$, and examples of empirical tails at the expected value explosion times, in log-log coordinates, approximated by straight lines. These simulations seemingly indicate that the expectation of $X(a, t)$ tends to infinity as $t \uparrow \lambda/(1-\mu)$. We formally confirm this intuition in the sequel.
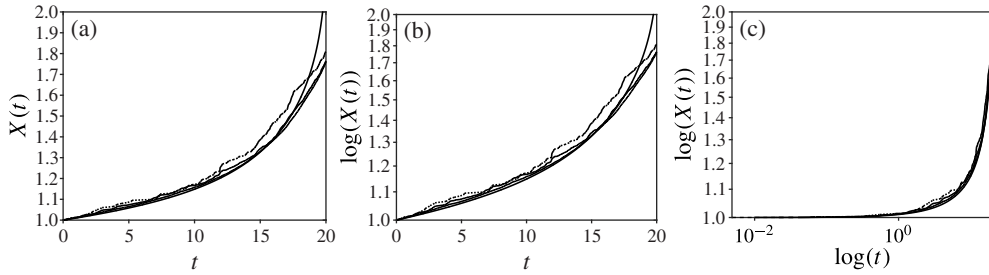
FIGURE 2: Averages of the simulated trajectories of the modified G–C process with parameters $\mu = 0.5$, $a = 0.01$, and $\lambda = 10$, based on 200, 1000, and 10 000 trajectories (dotted, dashed–dotted, and dashed lines, respectively), with the expectation $M(a, t)$, computed by numerically solving the integral equation (17) for $\varphi(t)$ and using (16). (a) t and $X(t)$ in linear scale, (b) t in linear scale and $X(t)$ in logarithmic scale, (c) t and $X(t)$ in logarithmic scale.
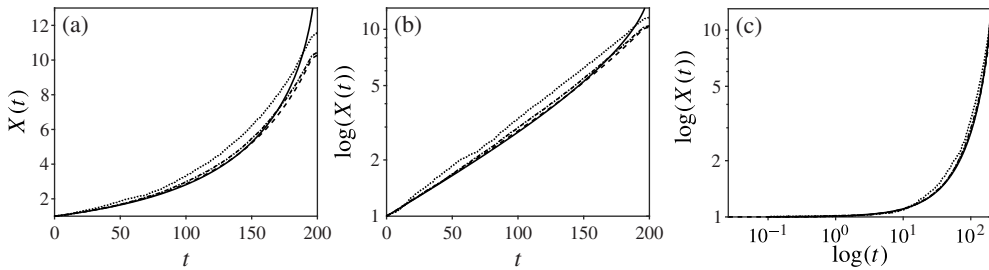

FIGURE 3: Averages of the simulated trajectories of the modified G-C process with parameters $\mu = 0.5$, $a = 0.01$, and $\lambda = 100$. Details as in Figure 2.
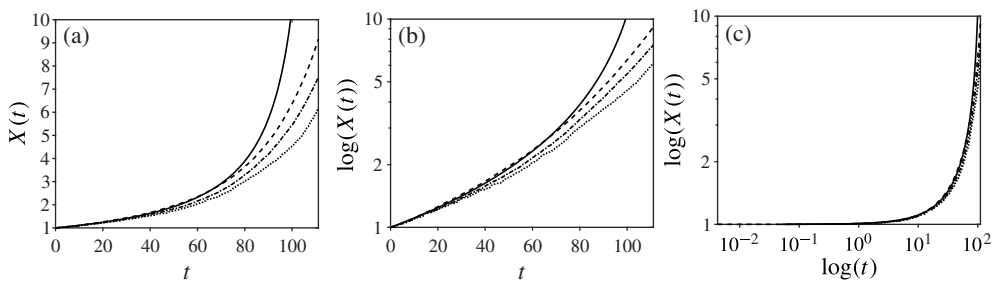

FIGURE 4: Averages of the simulated trajectories of the modified G–C process starting from a cell with randomly selected parameter of lifetime distribution, with parameters $\mu = 0.1$ and $\lambda = 100$, based on 200, 1000, and 10 000 trajectories (dotted, dashed–dotted, and dashed lines, respectively), with the expectation $\varphi(t)$, computed by numerically solving the integral equation (17). (a) t and $X(t)$ in linear scale, (b) t in linear scale and $X(t)$ in logarithmic scale, (c) t and $X(t)$ in logarithmic scale.
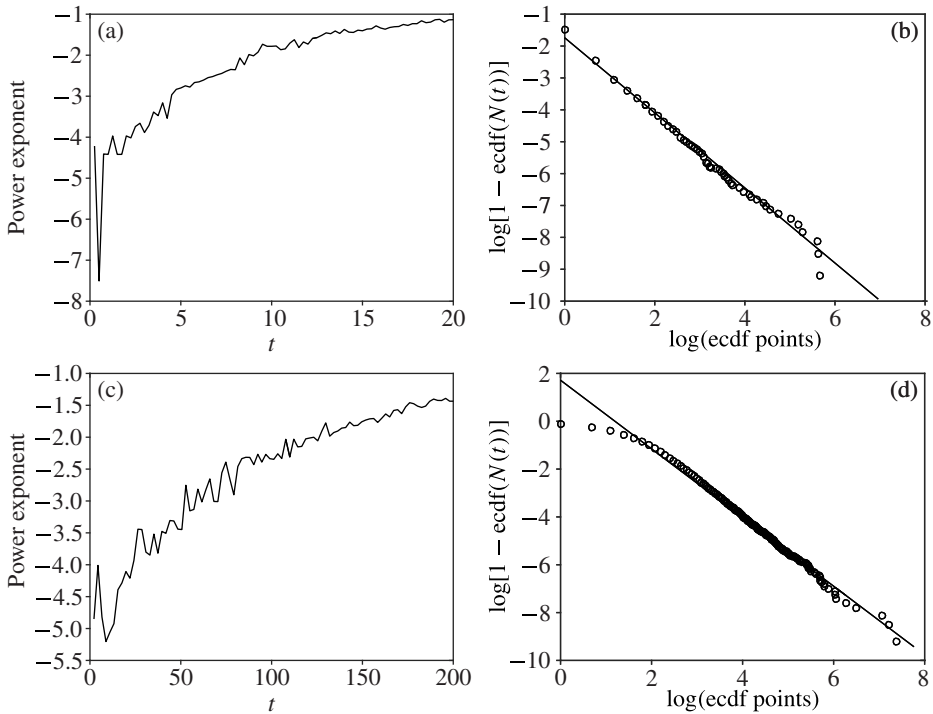
FIGURE 5: Simulated tail behavior of the modified G-C process with parameters $\mu = 0.5$, $a = 0.01$, $\lambda = 10$ (a, b) and $\lambda = 100$ (c, d). Depicted are estimated power exponents of the tail of $X(a, t)$, $t \in [0, \lambda/(1 - \mu)]$, (*left*) and examples of empirical tails at the expected value explosion times, in log-log coordinates, approximated by straight lines (*right*).

3.4. Asymptotic bounds

In this section we consider the distributions of $X_1(a, t)$, $X_2(a, t)$, and, in general, $X_k(a, t)$.

3.4.1. *Distribution of $X_1(a, t)$*. By the independence assumption, the distribution of the primary tumor cells, $X_1(a, t)$, is not affected by the behavior of subtypes that mutated from the primary type. Standard results for Yule's binary fission model gives

$$F_1(s, a, t) = \frac{s\,e^{-a(1-\mu)t}}{1 - s(1 - e^{-a(1-\mu)t})}, \qquad s \in [0, 1],\ t \geq 0,$$

where $F_1(s, a, t)$ is the probability generating function of $X_1(a, t)$. This is a geometric distribution with success probability $e^{-a(1-\mu t)}$. Hence,

$$\mathbb{E}[X_1(a, t)] = e^{a(1-\mu)t}, \qquad \mathbb{P}[X_1(a, t) > n] = (1 - e^{-a(1-\mu)t})^n.$$

We now introduce a result that will be very useful for studying the distribution of $X_2(a, t)$, $X_3(a, t), \ldots$. If we integrate over $a \sim \text{Exp}(\lambda)$, the marginal distribution of X_1 is known as the Yule–Simon distribution [9], [11]. Define

$$\nu := \nu(t) \equiv \frac{\lambda}{(1 - \mu)t}. \tag{3}$$

The probability mass function and tail probability of $X_1(t)$ are given by

$$\mathbb{P}[X_1(t) = n] = \nu B(\nu + 1, n), \qquad \mathbb{P}[X_1(t) > n] = n B(\nu + 1, n),$$

where B is the beta function, so, for sufficiently large n, the tail probability follows a power law

$$\mathbb{P}[X_1(t) > n] = \frac{\Gamma(n+1)\Gamma(\nu+1)}{\Gamma(\nu+n+1)} \sim \frac{\Gamma(\nu+1)}{n^\nu} \quad \text{as } n \to \infty.$$

The first two moments of $X_1(t)$ are

$$\mathbb{E}[X_1(t)] = \begin{cases} \dfrac{\nu}{\nu-1} & \text{if } \nu > 1, \\ \infty & \text{if } \nu \le 1, \end{cases} \qquad \text{var}(X_1(t)) = \begin{cases} \dfrac{\nu^2}{(\nu-1)^2(\nu-2)} & \text{if } \nu > 2, \\ \infty & \text{if } \nu \le 2. \end{cases} \tag{4}$$

This is essentially the same as the result we obtained for the toy model introduced at the beginning of this section.

3.4.2. *Distribution of $X_2(a, t)$.* Let $K(a, t)$ denote the number of tumor types generated by one and only one mutation. Denote the division rates of these subtypes by $a'_1, \ldots, a'_{K(a,t)}$, and let $Y_i(a, t)$ be the number of cells of type a'_i. Thus, $X_2(a, t) = \sum_{i=1}^{K(a,t)} Y_i(a, t)$. Recall that a is the division rate of the ancestor tumor cell. Hence, the notation $Y_i(a, t)$ implies that a'_i is integrated out. Clearly, if a subtype a'_i is born at time $t_0 < t$, the distribution of $Y_i(a, t)$ is the same as the marginal distribution of $X_1(t - t_0)$. The expected value of $X_2(a, t)$ can be computed as

$$\mathbb{E}[X_2(a, t)] = \begin{cases} \displaystyle\int_0^t \frac{\lambda a \mu e^{a_1(1-\mu)s}}{\lambda - (1-\mu)(t-s)} \, ds & \text{if } (1-\mu)t < \lambda, \\ \infty & \text{if } (1-\mu)t \ge \lambda. \end{cases} \tag{5}$$

Consider the tail probabilities $\mathbb{P}[X_2(a, t) > n]$; these can be bounded by

$$\mathbb{P}[X_2(a, t) > n] = \mathbb{P}\left[\sum_{i=1}^{K(a,t)} Y_i(a, t) > n\right] \ge \mathbb{P}\left[\bigcup_{i=1}^{K(a,t)} \{Y_i(a, t) > n\}\right]. \tag{6}$$

We pause to comment on why this bound is useful. For a tumor model, a is typically small and λ is typically large to ensure that both the primary tumor type and most secondary tumor types cannot grow too quickly. From its biological meaning, the mutation rate μ is small. Since, by (5), the number of tumor cells eventually explodes, our primary interest is in the case that t is moderate. Consequently, the event $\{K(a, t) \ge 2\}$ has small probability. More importantly, the tail probability of Y_i is a power law. Thus, we are much more likely to observe one very large Y_i than to observe two or more 'moderately large' Y_i. The left-hand side of (6) can be computed as

$$P_n(a, t) \equiv \mathbb{P}\left[\bigcup_{i=1}^{K(a,t)} \{Y_i(a, t) > n\}\right] = a\mu \int_0^t n e^{a(1-\mu)(t-s)} B\left(\frac{\lambda}{(1-\mu)s} + 1, n\right) ds. \tag{7}$$

To simplify notation, define $\tilde{\lambda} \equiv \lambda/(1-\mu)$. Choosing $\varepsilon > 0$ and omitting the exponential term, we obtain

$$\frac{P_n(a, t)}{a\mu} \ge \int_\varepsilon^t n B\left(1 + \frac{\tilde{\lambda}}{s}, n\right) ds = \int_\varepsilon^t \frac{\Gamma(1 + \tilde{\lambda}/s)\Gamma(n+1)}{\Gamma(n+1+\tilde{\lambda}/s)} \, ds. \tag{8}$$

On $\mathbb{R}^+$, $\Gamma(x)$ attains the minimum ≈ 0.885 at $x \approx 1.46$. Thus we may bound $\Gamma(1 + \widetilde{\lambda}/s)$ by 0.885 or $\Gamma(1 + \widetilde{\lambda}/t)$ if $\widetilde{\lambda}/t > 0.46$. For simplicity, assume henceforth that $\widetilde{\lambda}/t > 0.46$ and obtain

$$P_n(a,t) > a\mu\Gamma\left(1 + \frac{\widetilde{\lambda}}{t}\right)\int_\varepsilon^t \frac{\Gamma(n+1)}{\Gamma(n+1+\widetilde{\lambda}/s)}\,\mathrm{d}s,$$

$$\sim a\mu\Gamma\left(1 + \frac{\widetilde{\lambda}}{t}\right)\int_\varepsilon^t \frac{1}{n^{\widetilde{\lambda}/s}}\,\mathrm{d}s$$

$$= a\mu\Gamma\left(1 + \frac{\widetilde{\lambda}}{t}\right)\widetilde{\lambda}\log n\int_{\widetilde{\lambda}\log n/t}^{\widetilde{\lambda}\log n/\varepsilon} \frac{\mathrm{e}^{-x}}{x^2}\,\mathrm{d}x.$$

We can let n go to infinity since $\widetilde{\lambda}/s \geq \widetilde{\lambda}/\varepsilon$. The exponential integral is not an elementary function, but can be bounded (see [1]) by

$$\frac{\mathrm{e}^{-u}}{u^{n-1}(u+n)} < \int_u^\infty \frac{\mathrm{e}^{-x}}{x^n}\,\mathrm{d}x < \frac{\mathrm{e}^{-u}}{u^{n-1}(u+n-1)} \leq \frac{\mathrm{e}^{-u}}{u^n}, \qquad u > 0, \; n = 1, 2, \ldots. \quad (9)$$

Hence,

$$\widetilde{\lambda}\log n \int_{\widetilde{\lambda}\log n/t}^{\widetilde{\lambda}\log n/\varepsilon} \frac{\mathrm{e}^{-x}}{x^2}\,\mathrm{d}x > \frac{t^2 n^{-\widetilde{\lambda}/t}}{\widetilde{\lambda}\log n + 2t} - \frac{\varepsilon^2 n^{-\widetilde{\lambda}/\varepsilon}}{\widetilde{\lambda}\log n}. \quad (10)$$

Note that in (8) we omitted the integral on $(0, \varepsilon)$, as it is of less interest to us. But using the inequality for the beta function given in [4] and (9), we can show that

$$\int_0^\varepsilon n\mathrm{e}^{a(1-\mu)(t-s)}\mathrm{B}\left(1 + \frac{\widetilde{\lambda}}{s}, n\right)\mathrm{d}s > \frac{\mathrm{e}^{a(1-\mu)(t-\varepsilon)}\varepsilon^3}{(\widetilde{\lambda}+\varepsilon)(\widetilde{\lambda}\log n + 3\varepsilon)}n^{-\widetilde{\lambda}/\varepsilon},$$

which grows at a slower rate (with respect to n) than (10). Since $\log n$ is a slowly varying function, for sufficiently large n, we have

$$P_n(a,t) > \frac{a\mu t^2\Gamma(1 + \widetilde{\lambda}/t)}{\widetilde{\lambda}\log n + 2t}n^{-\widetilde{\lambda}/t} \equiv L_n(a,t). \quad (11)$$

Finally, if we integrate over $a \sim \mathrm{Exp}(\lambda)$ and recall the definition in (3), we obtain

$$\mathbb{P}[X_2(t) > n] > P_n(t) > \frac{\mu\Gamma(\nu+1)}{\nu(\nu\log n + 2)}n^{-\nu}. \quad (12)$$

3.5. Numerical examples

We chose $a = 0.1$, $\mu = 0.2$, and $\lambda = 10$, and simulated 10^6 trajectories of $X_2(a,t)$. For the sample mean of $X_2(a,t)$, we obtained 0.112 at $t = 4$ and 0.339 at $t = 8$; these equal the theoretical values computed using (5). We show the tail probabilities of $X_2(a,t)$ for $t = 8, 15$ in Table 1. Recall that our estimate $P_n(a,t)$ defined in (7) is a lower bound for $\mathbb{P}[X_2(a,t) > n]$, and $L_n(a,t)$ defined in (11) is an asymptotic lower bound for $P_n(a,t)$. Observe that in Table 1, both $P_n(a,t)$ and $L_n(a,t)$ can at least correctly estimate the order of the tail probabilities of $X_2(a,t)$. In fact, $P_n(a,t)$ is very close to the sample average for large n; this is most likely due to the heavy tail of the distribution of $X_2(a,t)$. Furthermore, assuming that the tail probability has the form $n^x/\log n$, we estimate the exponent to be -1.45 for $t = 8$ and -0.81 for $t = 15$; these are very close to the theoretical values -1.56 for $t = 8$ and -0.83 for $t = 15$. Thus, our estimate of the exponent, $\nu = \lambda/(1-\mu)t$, is useful, though it tends to be slightly conservative.

TABLE 1: Results of simulating the model of primary and secondary tumors (with parameters $a = 0.1$, $\mu = 0.2$, and $\lambda = 10$). $\widehat{\mathbb{P}}[X_2(a, t) > n]$ is the observed frequency in 10^6 simulated trajectories; $P_n(a, t)$ is defined in (7), computed by numerical integration; $L_n(a, t)$ is defined in (11).

					n				
		5	10	15	20	25	50	100	200
	$\widehat{\mathbb{P}}(X_2(t, a) > n) \times 10^3$	5.38	1.49	0.69	0.40	0.28	0.07	0.02	0.009
$t = 8$	$P_n(a, t) \times 10^3$	4.54	1.37	0.67	0.40	0.27	0.08	0.02	0.007
	$L_n(a, t) \times 10^3$	3.98	1.09	0.52	0.31	0.21	0.06	0.02	0.005
	$\widehat{\mathbb{P}}(X_2(t, a) > n) \times 10^3$	44.9	19.0	11.6	8.30	6.45	3.00	1.45	0.71
$t = 15$	$P_n(a, t) \times 10^3$	34.3	16.1	10.4	7.59	5.97	2.87	1.40	0.70
	$L_n(a, t) \times 10^3$	22.1	10.6	6.94	5.17	4.12	2.06	1.04	0.53

3.5.1. *Tail probabilities of $X_k(a, t)$.* The asymptotic analysis leading to (12) is readily extended to $X_k(a, t)$ for $k = 3, 4, \ldots$. For example, when analyzing $X_3(a, t)$, we can treat the secondary tumor cells described by X_2 as primary tumor cells and apply our previous result. By both (9) and (12), we obtain, for sufficiently large n,

$$
\begin{aligned}
\mathbb{P}[X_3(a, t) > n] &> a\mu^2 \int_0^t \frac{e^{a(1-\mu)(t-s)}}{\widetilde{\lambda}(\widetilde{\lambda}\log n + 2s)}\Gamma\left(1 + \frac{\widetilde{\lambda}}{s}\right)s^2 n^{-\widetilde{\lambda}/s}\, ds \\
&> \frac{a\mu^2 \Gamma(\nu + 1)}{\widetilde{\lambda}(\widetilde{\lambda}\log n + 2t)}\int_0^t s^2 n^{-\widetilde{\lambda}/s}\, ds \\
&> \frac{a\mu^2 t^4 \Gamma(\nu + 1)}{\widetilde{\lambda}(\widetilde{\lambda}\log n + 2t)(\widetilde{\lambda}\log n + 4t)}n^{-\widetilde{\lambda}/s}.
\end{aligned}
$$

We can repeat this calculation and obtain the general expression of the tail probability of $X_k(a, t)$. Assuming that $(\widetilde{\lambda}\log n + kt) \sim \widetilde{\lambda}\log n$, we find that

$$
\mathbb{P}[X_k(a, t) > n] > Ca\left\{\frac{\mu(1 - \mu)^2 t^2}{\lambda^2 \log n}\right\}^{k-1} n^{-\nu} \quad \text{as } n \to \infty \text{ for } k = 2, 3, \ldots,
$$

where C is a chosen constant. This expression provides insight into the dynamics of the tumor cells. First, the power law exponent $-\nu$ is the same for all tumor cells except the primary ones, but the growth rate of $X_k(a, t)$ is penalized by $(\log n)^{1-k}$. The exponent ν is equal to 1 exactly when the expected value of the number of tumor cells explodes (recall (4) and (5)). Second, for small t, the tumor population is dominated by $X_1(a, t)$ and $X_2(a, t)$, but for large t, the cell populations $X_k(a, t)$ with large k eventually dominate. Finally, given a moderate value of t, the value of μ determines which of $X_1, X_2, \ldots$, dominates: if μ is too small then there can be no mutation to give rise to new subtypes, but if μ is close to 1 then no tumor subtypes can flourish since most divisions do not increase the total number of cells of that subtype.

3.6. Towards general theory

3.6.1. *Branching process with infinite type space.* We return to the modified G–C model specified at the beginning of Section 3.2. Following the hypotheses of the model and under the usual conditional independence assumptions, an ordinary differential equation can be written for the PGF of the distribution of total cell count in all clones

$$
F(s; a, t) = \mathbb{E}[s^{X(a,t)}], \qquad s \in [0, 1],
$$

where $X(a, t)$ denotes the number of cells in the process started by an ancestor of type a. The backward Kolmogorov-type equations for $t \geq 0$ and $a \geq 0$,

$$\frac{\partial F(s; a, t)}{\partial t} = -aF(s; a, t) + a[(1 - \mu)F(s; a, t)^2 + \mu F(s; a, t)\Phi(s; t)], \qquad (13)$$

$$F(s; a, 0) = s,$$

are analogous to the equations of the G–C model of clonal resistance ([6, Section 4.2]), except that the PGF $\Phi(s, t)$ of the cell count of the clone started by a mutant of exponentially distributed type is equal to

$$\Phi(s, t) = \int_0^\infty F(s; a', t) \cdot \lambda e^{-\lambda a'} \, da', \qquad (14)$$

as follows from the second hypothesis of the modified G–C model. Equation (13) can be solved and using (14) compressed into a single integral equation for $\Phi(s, t)$ (see Appendix A). It is also straightforward to obtain

$$M(a, t) = \mathbb{E}[X(a, t)] = \frac{\partial F(s; a, t)}{\partial s}\bigg|_{s \uparrow 1},$$

$$\frac{\partial M(a, t)}{\partial t} = a(1 - \mu)M(a, t) + a\mu\varphi(t), \qquad (15)$$

where

$$\varphi(t) = \int_0^\infty M(a', t)\lambda e^{-\lambda a'} \, da'$$

is also equal to $\partial \Phi(s, t)/\partial s \mid_{s \uparrow 1}$. We can represent the solution of (15) using the variation of constant formula

$$M(a, t) = g(t) + a\mu g(t) \overset{(t)}{*} \varphi(t), \qquad (16)$$

where ' $\overset{(t)}{*}$ ' is the convolution operator of functions on $[0, \infty)$ and

$$g(t) = e^{a(1-\mu)t}.$$

On multiplying (16) by $\lambda e^{-\lambda a}$ and integrating on a over $[0, \infty)$, we obtain

$$\varphi(t) = f_1(t) + \frac{\mu}{\lambda} f_2(t) \overset{(t)}{*} \varphi(t), \qquad (17)$$

where, for $t \in [0, \lambda/(1 - \mu))$,

$$f_1(t) = \int_0^\infty g(t)\lambda e^{-\lambda a} \, da = \frac{1}{1 - t(1 - \mu)/\lambda},$$

$$f_2(t) = \int_0^\infty ag(t)\lambda^2 e^{-\lambda a} \, da = \frac{1}{[1 - t(1 - \mu)/\lambda]^2}.$$

Accordingly,

$$\varphi(t) = f_1(t) + f_1(t) \overset{(t)}{*} \sum_{i \geq 1} \left(\frac{\mu}{\lambda}\right)^i f_2^{\overset{(t)}{*}i}(t), \qquad (18)$$

where the infinite series of convolution powers converges uniformly for t in any closed subset of the interval $[0, \lambda/(1 - \mu))$ (this can be proved by an argument akin to Picard iteration).

Equations for the expected values have solutions expressed as a series of convolution powers. They explode at finite time $t = \lambda/(1-\mu)$. In addition, the function $\varphi(t) = \varphi(t, \lambda)$, is understood to be a function of two variables depending only on t/λ, i.e. it has the following scaling property

for $t \in [0.k\lambda/(1-\mu))$:

$$\varphi(t, k\lambda) = \varphi\left(\frac{t}{k}, \lambda\right).$$

Numerical solutions based on the power series of (18) are depicted as thick continuous lines in Figures 2, 3, and 4, along with simulation averages. The relationship between these averages and the exact expected values $M(a, t)$ and $\varphi(t)$ is analogous to that of the toy model (Figure 1(d)).

4. Conclusions

In this paper we presented three models: a quasi-stochastic baseline model, a stochastic toy model, and a branching process model. Each of these models proposes a mechanism for generating heavy tail and 'explosive' super-exponential growth of a population of secondary tumors under very parsimonious assumptions. Our approach generates somewhat unexpected results without invoking new biological mechanisms. Of course, the finite-time 'explosions' of expected values that we obtained cannot occur in the real world, in which cell proliferation rates a cannot be arbitrarily high as required by the exponential distribution.

Equally important are the statistical and model building consequences. Our analysis demonstrates that averages of empirical trajectories may be quite meaningless when building models of evolutionary phenomena such as cancer, in which heterogeneity plays a major role. Based on the toy model analysis, it is much more realistic to follow quantiles and deduce the growth law of the process from the different growth exponents of the quantiles. Since the branching process model seems to behave very similarly to the toy model, this conclusion is likely to hold for it as well.

The behavior of the models arises from two distributional assumptions: exponentially distributed lifetimes and exponentially distributed growth rates. The assumption of exponential lifetimes is made for simplicity. While the assumption is not realistic in some contexts, it is often true in the theory of branching processes that many results which are exact under this assumption hold asymptotically for more general distributions. The second assumption of exponential growth rates is less obvious. In reality, we deal with a distribution of growth rates with bounded support. It would be interesting to investigate how the behavior of the process changes if we truncate the exponential distributions of growth rates. Without getting into details, which are elementary, this can be clarified via the toy model. If a truncated exponential is adopted as the distribution of growth rates, then asymptotically the expected values become exponential, with rates depending on the truncation. The distribution tails become truncated Pareto, with the truncation point growing exponentially with time. As a result, the expected value remains useless as a measure of central tendency.

Mathematically, our analysis of the branching process model is quite preliminary. We can only conjecture the nature of the asymptotics of the modified G–C model. We do not know which properties of the process persist if cell death is allowed. Finally, we do not know the mathematical structure of the Markov operator semigroup involved in such a process. These questions certainly warrant further research.

Appendix A. Solution of the baseline model

We follow the approach of Iwata [5]; this involves a transport-type partial differential equation with nonlocal boundary conditions, of the type considered in [2] and [8], and which can be used to derive the distribution of the sizes of recurrent secondary tumors shed by a

growing primary tumor. In the simplest cases, we obtain closed-form expressions, while when the growth rates of the metastases are exponentially distributed, we obtain expressions that involve incomplete gamma functions that explode in finite time.

A.1. Deriving the distribution density from the transport equation

Case 1. Primary and metastatic tumors grow at the same rate $a = b$. From (1) and the exponential growth rate hypothesis $g(x) = ax$, we derive the transport equation

$$\frac{\partial \rho}{\partial t} + ax \frac{\partial \rho}{\partial x} = -a\rho.$$

Equivalently, for $x > 0$,

$$\frac{\partial \rho}{\partial x} + \frac{1}{ax} \frac{\partial \rho}{\partial t} = -\frac{\rho}{x}.$$

Regarding x as the independent variable, we apply the method of characteristics

$$\widetilde{\rho}(x) = \rho(x, \tau(x)),$$

where $\widetilde{\rho}$ denotes the distribution density ρ parameterized along characteristics. Integrating the equation $d\tau/dx = (ax)^{-1}$ from 1 to x, we obtain

$$\tau(x) - \tau(1) = \frac{\ln x}{a}, \tag{19}$$

leading to a solution of the form $\rho(x, \tau(x)) = \rho(1, \tau(1))/x$, which, with $t = \tau(x)$ and (19), yields

$$\rho(x, t) = \frac{\rho(1, t - (\ln|x|)/a)}{x}, \tag{20}$$

implying that $\rho(x, t) = 0$ for $x > e^{at}$.

Assume that at time $t = 0$ no metastatic tumor exists, so the initial condition is

$$\rho(x, 0) = 0.$$

The boundary condition at $x = 1$ has the nonlocal form given in (2). Equation (2) indicates that the number of metastatic single cells newly created per unit time at time t (the left-hand side term) is the sum of the total rate of occurrence of metastases due to both metastatic tumors and the primary tumor (corresponding respectively to the first and second terms of the right-hand side of (2)) [5]. $x_p(t)$ represents the number of cells in the primary tumor at time t and is as described in Section 2: the number of cells in the primary tumor as a function of time is

$$x_p(t) = e^{at}. \tag{21}$$

Write $\rho_1(t) = \rho(1, t)$, and substitute (20) and (21) into (2); then

$$a\rho_1(t) = \int_1^{e^{at}} mx^\alpha \frac{\rho_1(t - (\ln|x|)/a)}{x} \, dx + me^{a\alpha t}. \tag{22}$$

Following the change of variables $x = e^{a(t-u)}$, $dx = -ax \, du$, (22) can be written as $a\rho_1(t) = ame^{at} * \rho_1(t)$, where $(f * g)(t) = \int_0^t f(t - \tau)g(\tau) \, d\tau$. Taking Laplace transforms $\widehat{\rho}_1(\cdot)$ gives

$$a\widehat{\rho}_1(s) = a\frac{m\widehat{\rho}_1(s)}{s - a\alpha} + \frac{m}{s - a\alpha}, \tag{23}$$

and

$$\widehat{\rho}_1(s) = \frac{m/a}{s - (a\alpha + m)},$$

leading to $\rho_1(t) = (m/a)e^{(a\alpha+m)t}$. Using (20), we obtain

$$\rho(x,t) = \frac{(m/a)e^{(a\alpha+m)(t-\ln x/a)}}{x} = \frac{m}{a}e^{(a\alpha+m)t}x^{-(\alpha+m/a+1)} \qquad (x \le e^{at}). \qquad (24)$$

Let $G(x)$ be the number of migrant clones with more than x cells at time t; then

$$G(x) = \int_x^{e^{at}} \rho(\xi,t)\,d\xi = \frac{m}{a\alpha+m}(e^{(a\alpha+m)t}x^{-(\alpha+m/a)} - 1) \qquad (x < e^{at}). \qquad (25)$$

Case 2. Metastatic tumors grow at a different rate, $a \ne b$. Equation (25) can be extended to include growth advantage, the newly seeded tumor having a growth rate which can be higher or lower than the growth rate of the parent tumor. Using (2) and the equation for the number of cells in the primary tumor $x_p(t) = e^{bt}$ yields

$$a\rho_1(t) = \int_1^{e^{at}} mx^{\alpha}\frac{\rho_1(t-(\ln|x|)/a)}{x}\,dx + me^{\alpha bt}.$$

Much as before, after a change of variables, $a\rho_1(t) = a\int_1^t me^{a\alpha(t-u)}\rho_1(u)\,du + me^{\alpha bt}$, and taking Laplace transforms gives

$$\widehat{\rho_1}(s) = \frac{m}{a}\frac{s-a\alpha}{(s-\alpha b)(s-a\alpha-m)}.$$

Inverting this Laplace transform yields

$$\rho_1(t) = \frac{m[(a\alpha-\alpha b)e^{\alpha bt} + me^{(a\alpha+m)t}]}{a(a\alpha+m-\alpha b)}. \qquad (26)$$

For $x \le e^{at}$, (20) gives

$$\rho(x,t) = \frac{1}{x}\left(\frac{m}{a(a\alpha+m-\alpha b)}[(a\alpha-\alpha b)e^{\alpha b(t-\ln|x|/a)} + me^{(a\alpha+m)(t-\ln|x|/a)}]\right)$$
$$= \frac{m[(a\alpha-\alpha b)e^{\alpha bt}x^{-(\alpha b/a+1)} + me^{(a\alpha+m)t}x^{-(\alpha+m/a+1)}]}{a(a\alpha+m-\alpha b)}.$$

Correspondingly, $G(x) = G(x; a, b)$, denoting the number of migrant clones with more than x cells, is found by integrating this density over $x \in [1, e^{at}]$ and is given by

$$G(x; a, b) = \begin{cases} \dfrac{m}{a\alpha+m-\alpha b}\left[\dfrac{a-b}{b}(e^{\alpha bt}x^{-\alpha b/a} - 1) \right. \\ \qquad\qquad \left. + \dfrac{m}{a\alpha+m}(e^{(a\alpha+m)t}x^{-\alpha-m/a} - 1)\right] & \text{if } x < e^{at}, \\ 0 & \text{otherwise.} \end{cases}$$

Case 3. Metastatic growth rate has exponential distribution with parameter λ. In this case, we obtain $\widetilde{G}(x,b) = \int_0^{\infty} G(x; a, b)\lambda e^{-\lambda a}\,da$. Since $G(x; a, b) = 0$ for $x > e^{at}$ (equivalently, for $a < (\ln x)/t$), we can write

$$\widetilde{G}(x,b) = \int_{(\ln x)/t}^{\infty} G(x; a, b)\lambda e^{-\lambda a}\,da. \qquad (27)$$

In general, this integral seems analytically intractable, though the special case $x = 1$ can be expressed in the terms of incomplete gamma functions (IGFs), and $\widetilde{G}(1, b)$ is important because it is equal to the total load of metastases at time t. In its evaluation to which we proceed shortly, there arises the IGF $\Gamma(c, x)$ defined for positive x and complex c by $\Gamma(c, x) = \int_x^{\infty} e^{-t}t^{c-1}\,dt$.

For the function $G(1; a, b)$, we can write, using (27) with $x = 1$, hence $(\ln x)/t = 0$,

$$G(1; a, b) = \frac{m/\alpha}{a + m/\alpha - b}\left[\frac{a - b}{b}(e^{\alpha bt} - 1) + \frac{m}{a\alpha + m}(e^{(a\alpha + m)t} - 1)\right],$$

where, since the factor $[\cdots]$ is the algebraic sum of three different functions of a, substitution in (27) implies that $\widetilde{G}(1; b)$ is the sum of the following three terms:

$$(e^{\alpha bt} - 1)\int_0^\infty \frac{m/\alpha}{a + m/\alpha - b}\frac{a - b}{b}\lambda e^{-\lambda a}\,\mathrm{d}a$$

$$= \frac{m(e^{\alpha bt} - 1)}{\alpha b}\int_0^\infty\left[1 - \frac{m/\alpha}{u/\lambda + m/\alpha - b}\right]e^{-u}\,\mathrm{d}u$$

$$= \frac{m(e^{\alpha bt} - 1)}{\alpha b}\left[1 - \frac{\lambda m e^{\lambda(m/\alpha - b)}}{\alpha}\int_{\lambda(m/\alpha - b)}^\infty e^{-v}v^{-1}\,\mathrm{d}v\right]$$

$$= \frac{m(e^{\alpha bt} - 1)}{\alpha b}\left[1 - \frac{\lambda m}{\alpha}e^{\lambda(m/\alpha - b)}\Gamma\left(0, \lambda\left(\frac{m}{\alpha} - b\right)\right)\right], \tag{28}$$

$$-\int_0^\infty \frac{m/\alpha}{a + m/\alpha - b}\frac{m/\alpha}{a + m/\alpha}\lambda e^{-\lambda a}\,\mathrm{d}a$$

$$= -\frac{m^2}{\alpha^2 b}\int_0^\infty\left[\frac{1}{u/\lambda + m/\alpha - b} - \frac{1}{u/\lambda + m/\alpha}\right]e^{-u}\,\mathrm{d}u$$

$$= -\frac{\lambda m^2}{\alpha^2 b}\int_0^\infty\left[\frac{1}{u + \lambda(m/\alpha - b)} - \frac{1}{u + \lambda m/\alpha}\right]e^{-u}\,\mathrm{d}u$$

$$= -\frac{\lambda m^2}{\alpha^2 b}\left[e^{\lambda(m/\alpha - b)}\Gamma\left(0, \lambda\left(\frac{m}{\alpha} - b\right)\right) - e^{\lambda m/\alpha}\Gamma\left(0, \lambda\frac{m}{\alpha}\right)\right], \tag{29}$$

$$\int_0^\infty \frac{m/\alpha}{a + m/\alpha - b}\frac{m/\alpha}{a + m/\alpha}e^{a\alpha t}e^{mt}\lambda e^{-\lambda a}\,\mathrm{d}a$$

$$= \frac{m^2 e^{mt}}{\alpha^2 b}\int_0^\infty\left[\frac{1}{u/\lambda + m/\alpha - b} - \frac{1}{u/\lambda + m/\alpha}\right]e^{-u(1 - \alpha t/\lambda)}\,\mathrm{d}u$$

$$= \frac{\lambda m^2 e^{mt}}{\alpha^2 b(\lambda - \alpha t)}\int_0^\infty\left[\frac{1}{v/(\lambda - \alpha t) + m/\alpha - b} - \frac{1}{v/(\lambda - \alpha t) + m/\alpha}\right]e^{-v}\,\mathrm{d}v$$

$$= \frac{\lambda m^2 e^{mt}}{\alpha^2 b}\int_0^\infty\left[\frac{1}{v + (\lambda - \alpha t)(m/\alpha - b)} - \frac{1}{v + (\lambda - \alpha t)m/\alpha}\right]e^{-v}\,\mathrm{d}v$$

$$= \frac{\lambda m^2 e^{mt}}{\alpha^2 b}\left[e^{(\lambda - \alpha t)(m/\alpha - b)}\Gamma\left(0, (\lambda - \alpha t)\left(\frac{m}{\alpha} - b\right)\right) - e^{(\lambda - \alpha t)m/\alpha}\Gamma(0, (\lambda\alpha - t)m)\right]. \tag{30}$$

Gathering the three expressions at (28), (30) and (29) yields, for $b < m/\alpha$ and $\lambda > \alpha t$,

$$\widetilde{G}(1; b) = \frac{m(e^{b\alpha t} - 1)}{\alpha b}\left[1 - \frac{\lambda m e^{\lambda(m/\alpha - b)}}{\alpha}\Gamma\left(0, \lambda\left(\frac{m}{\alpha} - b\right)\right)\right]$$

$$+ \frac{\lambda m^2 e^{mt}}{\alpha^2 b}\left[e^{(\lambda - \alpha t)(m/\alpha - b)}\Gamma\left(0, (\lambda - \alpha t)\left(\frac{m}{\alpha} - b\right)\right)\right.$$

$$\left. - e^{(\lambda - \alpha t)m/\alpha}\Gamma\left(0, (\lambda - \alpha t)\frac{m}{\alpha}\right)\right]$$

$$- \frac{\lambda m^2}{\alpha^2 b}\left[e^{\lambda(m/\alpha - b)}\Gamma\left(0, \lambda\left(\frac{m}{\alpha} - b\right)\right) - e^{\lambda m/\alpha}\Gamma\left(0, \lambda\frac{m}{\alpha}\right)\right].$$

Acknowledgements

References

[1] ABRAMOWITZ, M. AND STEGUN, I. A. (1964). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables* (Nat. Bureau Stand. Appl. Math. Ser. **55**). US Government Printing Office, Washington, DC.

[2] ARINO, O. AND KIMMEL, M. (1993). Comparison of approaches to modeling of cell population dynamics. *SIAM J. Appl. Math.* **53,** 1480–1504.

[3] BARATCHART, E. *et al.* (2015). Computational modelling of metastasis development in renal cell carcinoma. *PLoS Comput. Biol.* **11,** 23pp.

[4] CERONE, P. (2007). Special functions: approximations and bounds. *Appl. Anal. Discrete Math.* **1,** 72–91.

[5] IWATA, K., KAWASAKI, K. AND SHIGESADA, N. (2000). A dynamical model for the growth and size distribution of multiple metastatic tumors. *J. Theoret. Biol.* **203,** 177–186.

[6] KIMMEL, M. AND AXELROD, D. (2015). *Branching Processes in Biology*. Springer, New York.

[7] LING, S. *et al.* (2015). Extremely high genetic diversity in a single tumor points to prevalence of non-Darwinian cell evolution. *Proc. Nat. Acad. Sci. USA* **112,** E6496–E6505.

[8] METZ, J. A. AND DIEKMANN, O. (eds) (2014). *The Dynamics of Physiologically Structured Populations* (Lecture Notes Biomath. **68**). Springer, Berlin.

[9] SIMON, H. A. (1955). On a class of skew distribution functions. *Biometrika* **42,** 425–440.

[10] TAO, Y. *et al.* (2015). Further genetic diversification in multiple tumors and an evolutionary perspective on therapeutics. Preprint. Available at https://www.biorxiv.org/content/early/2015/08/25/025429.

[11] YULE, G. U. (1925). A mathematical theory of evolution, based on the conclusions of Dr. J. C. Willis, F.R.S. *Phil. Trans. Royal Soc. London B* **213,** 21–87.

PHILIP A. ERNST, *Rice University*

Department of Statistics, Rice University, MS-138, 6100 Main Street, Houston, TX 77005, USA.

MAREK KIMMEL, *Rice University and Silesian University of Technology*

Departments of Statistics and Bioengineering, Rice University, MS-138, 6100 Main Street, Houston, TX 77005, USA.
Email address: kimmel@stat.rice.edu

MONIKA KURPAS, *Silesian University of Technology*

Systems Engineering Group, Silesian University of Technology, Akademicka 16, 44-100, Glinice, Poland.

QUAN ZHOU, *Rice University*

Department of Statistics, Rice University, MS-138, 6100 Main Street, Houston, TX 77005, USA.