RARE EVENTS IN CANCER RECURRENCE TIMING

Jasmine Foo

University of Minnesota
206 Church St
Minneapolis, MN 55455 USA

Kevin Leder

University of Minnesota
111 Church St.
Minneapolis, MN 55455 USA

ABSTRACT

Tumor recurrence due to acquired resistance to anti-cancer treatments poses a major clinical problem in treating cancer. One major cause of drug-resistance is the acquisition of random point mutations in the genomic sequence of cancer cells which confer resistant phenotypes. Despite an initial response to treatment, emergent drug-resistant subpopulations often eventually drive the recurrence and regrowth of the tumor. The timing of such cancer recurrence is highly variable in patient populations, and is governed by a balance between several factors such as initial tumor size, mutation rates, and growth kinetics of drug-sensitive and resistant cells. To better understand patterns of cancer progression in patient populations, we are interested in the mechanisms driving early or late cancer recurrences. In previous work, we modeled the dynamics of recurrence by considering escape from a subcritical branching process, where the establishment of a clone of escape mutants can lead to total population growth after the initial decline. Here, we study and characterize the rare events leading to early or late *crossover time*, defined as the time at which the total cancer population first becomes dominated by the emerging resistant cell population. In particular, using this model we investigate algorithms for estimating the probability of early crossover events, which are correlated with early tumor recurrence.

1 INTRODUCTION

The past decade has seen a revolution in the discovery of targeted anti-cancer therapies (e.g. Iressa - non-small cell lung cancer; Gleevec - chronic myeloid leukemia; Herceptin - HER2+ breast cancer). Many of these drugs lead to dramatic reductions in tumor burden with little toxicity. However, a major limitation to their efficacy is the eventual emergence of drug resistance, leading to a rebound in tumor size and progression of disease. Drug resistance is commonly caused by random point mutations in the genomic sequence of cells, which can confer drug-resistant phenotypes to cells by interfering with the binding site of the drug, up-regulating drug efflux pumps, or activating alternate signaling pathways, for example. Once a viable resistant subpopulation is established, its growth drives the recurrence of the tumor in the presence of therapy. Although tumor recurrence due to acquired resistance is inevitable for many therapies (e.g. Iressa), the timing of recurrence can vary significantly between patients (Demicheli and et al 2008; Demicheli and et al 2012). The source of this variability - e.g. why some patients experience shorter or longer disease-free periods than others, is largely unknown. Gaining a better understanding of the evolutionary paths leading to early or late recurrence will provide insight into the relationship between recurrence timing and the composition of the recurrent tumor, which may eventually aid in clinical treatment strategies post-recurrence.

Markov process population models provide a useful tool to understand the complex evolutionary dynamics of cancer and the emergence of drug-resistant subpopulations (see, e.g. references (Coldman and Goldie 1986; Durrett, Schmidt, and Schweinsberg 2009; Frank 2003; Haeno, Iwasa, and Michor 2007; Komarova 2006)). In this work we analyze an evolutionary branching process model of tumor recurrence due to the emergence of drug resistance mediated by a random point mutation. In particular, we develop

computational tools for effective estimation of the probability of tumor recurrence occurring much earlier than expected. These tools will be used in later work to study the most likely evolutionary paths leading to early tumor recurrence.

2 MODEL DESCRIPTION

In the classic Galton-Watson process each individual lives for one unit of time and then is replaced by a random number of offspring that is characterized by its probability generating function (pgf) $f(s) = \sum_{n=0}^{\infty} p_n s^n$, where $p_n \geq 0$ and $\sum_n p_n = 1$. A natural generalization of this model is to assume that each individual lives for a random amount of time. If we assume that these lifetimes are independent and identically distributed exponential random variables then the total population process is a continuous time Markov process, such a process is called a continuous time Markovian branching process (for further results on these processes see Chapter 3 of Athreya and Ney 2004). These processes are characterized by the pgf of the offspring distribution and the rate at which individuals are replaced by their offspring. If we insist that the offspring distribution has all of its mass at that points 0 and 2 then we have a binary Markovian branching process. We can characterize these processes by their growth and death rates, since this will tell us both the offspring distribution and the total rate at which replacement events occur. Since the somatic cells of the body can either divide in two or die a binary branching process is a natural tool to model the population dynamics of these cells.

We assume a large initial population of n drug-sensitive cancer cells. The evolution of this population is modeled as a binary Markovian branching process, $Z_0(t)$, with birth rate d_0, death rate r_0, net growth rate $\lambda_0 < 0$ and decay rate $r = |\lambda_0|$. The sensitive cell population produces resistance mutants at rate $\mu_n Z_0(t)$ where $\mu_n = \mu n^{-\alpha}$ for $\mu > 0$ and $\alpha \in [0,1)$. Each resistance mutation initiates the creation of a Markovian binary branching process with birth rate r_1, death rate d_1 and net growth rate $\lambda_1 > 0$, we denote total number of resistant cells at time t by $Z_1(t)$. Thus, Z_1 is a binary branching process with immigration from mutations arising in the Z_0 process.

2.1 Previous Results

Approximations of the cross-over time, or first time at which the resistant cells dominate the total population

$$\xi_n = \inf\{t \geq 0 : Z_1(t) \geq Z_0(t)\} \tag{1}$$

were studied in Foo and Leder 2012a. There it was shown that we can approximate the large n behavior of ξ_n by

$$u_n = \frac{\log\left(1 + n^{\alpha}(\lambda_1 + r)/\mu\right)}{(\lambda_1 + r)}. \tag{2}$$

In particular, we have the result that for any $\varepsilon > 0$

$$\lim_{n \to \infty} P(|\xi_n - u_n| > \varepsilon) = 0.$$

In addition, the following weak limit is shown in Foo and Leder 2012b,

$$n^{(1-\alpha)/2}(\xi_n - u_n) \Rightarrow Z, \tag{3}$$

where $Z \sim N(0, \sigma^2)$ and $\sigma^2 = 2r_1 r^2/(r_0 \lambda_1 (2\lambda_1 + r))$. Based on these result we see that the events $|\xi_n - u_n| > y$ for a fixed y become increasingly unlikely as $n \nearrow \infty$.

3 RARE EVENTS IN CROSSOVER TIME

We are interested in rare events in the crossover time ξ_n, which can be driven by deviations by either Z_0 or Z_1. In order to get a sense of which process is more important in terms of deviations, we first consider

their variances. The variance of Z_0 is well known (see e.g. Athreya and Ney 2004) as

$$\text{Var}(Z_0(t)) = n\left(\frac{r_0+d_0}{r}\right)\left(e^{-rt}-e^{-2rt}\right) = C_0 n e^{-rt}(1-e^{-rt}).$$

The variance of Z_1 can be shown (see Foo and Leder 2012a) to be

$$\text{Var}(Z_1(t)) = \left(\frac{\mu}{n^\alpha}\right)^2 \int_0^t \int_0^t \text{Cov}(Z_0(s), Z_0(y)) e^{\lambda_1(t-s)} e^{\lambda_1(t-y)} ds\, dy$$
$$+ \frac{\mu}{\lambda_1 n^\alpha} \int_0^t E Z_0(s) E[\tilde{Z}_1(t-s)^2|\tilde{Z}_1(0)=1] ds,$$

where $\tilde{Z}_1$ is a binary branching process with birth rate r_1 and death rate d_1. $\tilde{Z}_1$ possesses same growth behavior as Z_1 except that no further immigration is allowed. It is easily seen that both integrands in the formula for the variance of Z_1 are $O(n)$, and therefore $\text{Var}\, Z_1(t) = \Theta(e^{2\lambda_1 t} n^{1-\alpha})$. Note that if $\alpha = 0$ then $\text{Var} Z_1(t) > \text{Var} Z_0(t)$ for all t, and if $\alpha \in (0,1)$ the variance of Z_0 is larger for small t. However, equation (2) shows that if $\alpha \in (0,1)$ then our event of interest occurs at large times, i.e., $t \approx t_n = \frac{1}{\lambda_1+r} \log n^\alpha$. If we consider the variances at these time scales we see that

$$\text{Var}(Z_0(t_n)) = k_0 n^{1-r\alpha/(\lambda_1+r)}(1+o(1))$$
$$\text{Var}(Z_1(t_n)) = k_1 n^{1-\alpha+2\lambda_1\alpha/(\lambda_1+r)}(1+o(1))$$

where k_0 and k_1 are constants that depend only on system parameters. Comparing the power of n in the two expressions we see that

$$1 - \alpha + \frac{2\lambda_1\alpha}{\lambda_1+r} - 1 + \frac{r\alpha}{\lambda_1+r} > 0,$$

and thus the variance of Z_1 asymptotically dominates the variance of Z_0 at the crossover time ξ_n if $\alpha \in (0,1)$. Therefore for $a > 1$ we consider

$$p_n(a) = P(Z_1(\xi_n) > a E Z_1(u_n)). \tag{4}$$

Alternatively we could consider the probability $P(Z_1(\xi_n) > aEZ_1(\xi_n))$, however since it is possible to get an explicit expression for $EZ_1(u_n)$ we prefer the current formulation of $p_n(a)$. In order to analyze $p_n(a)$, we must first analyze the behavior of the processes Z_0 and $\tilde{Z}_1$. Let us first consider the moment generating function of $\tilde{Z}_1$

$$\phi_t(\theta) = E\exp(\theta\tilde{Z}_1(t)) = \begin{cases} \frac{d_1(e^\theta-1)-e^{-\lambda_1 t}(r_1 e^\theta - d_1)}{r_1(e^\theta-1)-e^{-\lambda_1 t}(r_1 e^\theta - d_1)}, & \theta < \bar{\theta}_t \\ \infty, & \theta \geq \bar{\theta}_t \end{cases} \tag{5}$$

where

$$\bar{\theta}_t \doteq \log\left(\frac{r_1 e^{\lambda_1 t} - d_1}{r_1 e^{\lambda_1 t} - r_1}\right), \tag{6}$$

see Athreya and Ney 2004 page 109.

To study the subcritical process Z_0 we consider the generating function of the offspring distribution

$$f(s) = \sum_{j=0}^\infty p_j s^j,$$

the infinitesimal generating function $\phi(s) = a(f(s)-s)$ (a is the total rate of cell deaths) and the moment generating function

$$F(s,t) = \sum_{n=0}^\infty s^n P(Z_0(t) = n|Z_0(t) = 1). \tag{7}$$

The function $F(s,t)$ satisfies the initial condition $F(s,0) = s$ and the forward and backward equations (see Athreya and Ney 2004)

$$\frac{\partial}{\partial t}F(s,t) = \phi(s)\frac{\partial}{\partial s}F(s,t) \tag{8}$$

$$\frac{\partial}{\partial t}F(s,t) = \phi(F(s,t)).$$

Note that for our binary branching process $f(s) = \frac{d_0 + r_0 s^2}{d_0 + r_0}$, and $\phi(s) = r_0 s^2 - (r_0 + d_0)s + d_0$. Notice that since Z_0 is subcritical $f'(1) < 1$ and thus by continuity there exists a $\delta > 0$ such that for $s \in (1, 1+\delta)$, we have $\phi(s) < 0$ and $\phi'(s) < 0$. We must address the question of the existence of $F(s,t)$ for $s > 1$. In the case of offspring distribution with finite support (such as our binary branching process), we see that $\phi(s)$ is Lipschitz for s in a compact set. By considering the backward equation in (8), observe that if $F(s_0, 0) = s_0$ is sufficiently close to 1 then the negativity of ϕ near 1 will constrain F to a compact set. Then the local Lipschitz continuity of ϕ gives the existence of F for all time. This is in contrast to the supercritical case where for any fixed s the solution will blow up in finite time.

We now consider upper bounds on the probability of Z_1 exceeding its mean at the crossover time. In order to do this we need the following assumption for $k > 0$

$$P(|z_n| > kn^{(1-\alpha)/2}) \leq \exp\left[-k^2 n^{1-\alpha}/2\right], \tag{9}$$

where $z_n = (\xi_n - u_n)n^{(1-\alpha)/2}$. This result will be established in future work.

Theorem 1 For $\alpha \in (0,1)$ and $a > 1$ there exists a $\gamma(a) > 0$ such that

$$p_n(a) \leq \exp\left(-n^{1-\alpha}(\gamma(a) + o(1))\right),$$

where $o(1)$ is a term that goes to 0 as $n \to \infty$.

Proof. First consider the following decomposition

$$P(Z_1(\xi_n) > aEZ_1(u_n)) \leq P(Z_1(u_n + z_n n^{(\alpha-1)/2}) > aEZ_1(u_n), z_n \in [-M_n, M_n]) + P(|z_n| > M_n),$$

where $M_n = kn^{(1-\alpha)/2}$ for a positive constant k. Due to the assumption (9) it suffices to study the first probability in the previous display. Next notice that

$$P(Z_1(u_n + z_n n^{(\alpha-1)/2}) > aEZ_1(u_n), z_n \in [-M_n, M_n]) \leq P(\sup_{t \in [-M_n, M_n]} Z_1(u_n + tn^{(\alpha-1)/2}) > aEZ_1(u_n)). \tag{10}$$

Then notice that Z_1 is a submartingale; thus if we choose $\theta_n > 0$ appropriately and define $t_n = u_n + k$ we have from Doob's inequality that the RHS of display (10) can be bounded by

$$E\exp(\theta_n Z_1(t_n) - a\theta_n EZ_1(u_n)) = \frac{1}{\exp(a\theta_n EZ_1(u_n))}E\exp\left(\frac{\mu}{n^\alpha}\int_0^{t_n} Z_0(s)(\phi_{t-s}(\theta_n) - 1)ds\right). \tag{11}$$

The equality in the previous display is a basic result for the moment generating function of birth-death processes with immigration, see e.g. Grimmett and Stirzaker 2001, Theorem 11, section 6.12. For $\theta \in (0,1)$, define $\theta_n = \theta\lambda_1 e^{-\lambda_1 t_n}/r_1$, from (6) we can observe that

$$\frac{\lambda_1}{re^{\lambda_1 t}} \leq \bar{\theta}_t + \frac{\lambda_1}{r_1 e^{\lambda_1 t}(e^{\lambda_1 t} - 1)}.$$

Thus for any $\theta < 1$, $\theta_n < \bar{\theta}_{t_n}$ for n sufficiently large n.

A useful decomposition of the expected value on the RHS of (11) is

$$E \exp\left(\frac{\mu}{n^\alpha}\int_0^{t_n} Z_0(s)(\phi_{t_n-s}(\theta_n)-1)ds\right) \tag{12}$$

$$= \exp\left(\frac{\mu}{n^{\alpha-1}}\int_0^{t_n} e^{-rs}(\phi_{t_n-s}(\theta_n)-1)ds\right) E \exp\left(\frac{\mu}{n^\alpha}\int_0^{t_n}(\phi_{t_n-s}(\theta_n)-1)(Z_0(s)-ne^{-rs})ds\right).$$

Roughly, this decomposes the moment generating function into two terms: one dealing solely with the growth of Z_1, and a term that considers deviations of Z_0 away from its mean behavior. A cursory analysis of these two expressions shows that the first term should be of higher order. More specifically, we compare $e^{n^{1-\alpha}}$ versus $e^{n^{1/2-\alpha}}$ since $Z_0(s)-ne^{-rs}$ is of order $n^{1/2}$. We now develop upper bounds on this latter term. First from Jensen's inequality we have

$$E \exp\left(\frac{\mu}{n^\alpha}\int_0^{t_n}(\phi_{t_n-s}(\theta_n)-1)(Z_0(s)-ne^{-rs})ds\right) \le \frac{1}{t_n}\int_0^{t_n} E \exp\left(\frac{\mu t_n}{n^\alpha}(\phi_{t_n-s}(\theta)-1)(Z_0(s)-ne^{-rs})\right)ds.$$
$$\tag{13}$$

To analyze the previous expression define $\psi_t(\theta) = E[\exp(\theta(Z_0(s)-e^{-rs})|Z_0(s)=1]$, and $\omega_n(s,t_n) = (\frac{\mu t_n}{n^\alpha}(\phi_{t_n-s}(\theta_n)-1))$. Then from the independence of the n initial cells we have

$$E \exp\left(\omega_n(s,t_n)(Z_0(s)-ne^{-rs})\right) = (\psi_s(\omega_n(s,t_n)))^n.$$

From the definition of θ_n and $\phi_t(\theta)$ it is possible to show that $\phi_{t_n-s}(\theta_n)$ is bounded for all n and $s \in (0,t_n)$, i.e., there exists $k_0 > 0$ such that $\omega_n(s,t_n) \le k_0 n^{-\alpha}\log n$. Therefore we conclude ω_n is small for n large and perform a Taylor expansion of ψ around 0. Then for $z \in (0,\omega_n(s,t_n))$

$$\psi_s(\omega_n(s,t_n)) = 1 + \frac{\omega_n(s,t_n)^2}{2}\psi_s''(z) \le 1 + \frac{\omega_n(s,t_n)^2}{2}\psi_s''(\omega_n(s,t_n)),$$

and the inequality follows from the monotonicity of $\psi_s''(\theta)$ in θ. It thus remains to study ψ_t''. In particular we want to show that this function is bounded in t for θ sufficiently small. In order to do this first recall $F(s,t)$ from (7) and observe that

$$h_\theta(t) \doteq \psi_t''(\theta) = e^{2\theta}\frac{\partial^2}{\partial s^2}F(e^\theta,t) + e^\theta\frac{\partial}{\partial s}F(e^\theta,t).$$

Our goal will be to show that for θ sufficiently close to 0, $\frac{d}{dt}h_\theta(t) < 0$. From the backward equation in (8) we get that

$$\frac{\partial^2}{\partial s\partial t}F(s,t) = \frac{1}{\phi(s)}\frac{\partial^2}{\partial t^2}F(s,t) \quad \text{and} \quad \frac{\partial^2}{\partial t^2}F(s,t) = \phi'(F(s,t))\frac{\partial}{\partial t}F(s,t).$$

Recall that there exists a $\delta > 0$ such that for $s \in (1,1+\delta)$, $\phi(s) < 0$ and $\phi'(s) < 0$. From the continuity of F and the backward equation we see that $\frac{\partial^2}{\partial t^2}F(s,t) > 0$ and therefore $\frac{\partial^2}{\partial s\partial t}F(s,t) < 0$. From this it also follows that

$$\frac{\partial^3}{\partial s^2\partial t}F(s,t) = \frac{-1}{\phi(s)^2}\frac{\partial^2}{\partial t^2}F(s,t) < 0.$$

Therefore we conclude that for θ sufficiently small $h_\theta(t) \le K_\theta$, and therefore there exists $k_1 > 0$ such that

$$\psi_s(\omega_n(s,t)) \le 1 + \frac{k_1\log n}{n^{2\alpha}}.$$

We can then conclude that

$$\exp\left[n\log\psi_n\left(\omega_n(s,t)\right)\right] \le \exp\left[n\log\left(1+\frac{k_1\log n}{2n^{2\alpha}}\right)\right] \le \exp\left(k_1 n^{1-2\alpha}\log n\right).$$

Since this bound is independent of s, we can apply this the bound to (13) to see that

$$E\exp\left(\frac{\mu}{n^\alpha}\int_0^{t_n}(\phi_{t_n-s}(\theta_n)-1)(Z_0(s)-ne^{-rs})ds\right) \le \exp(k_1 n^{1-2\alpha}\log n). \tag{14}$$

We now consider the deviations of Z_1. First define

$$\Psi_n = \frac{\mu}{n^{\alpha-1}}\int_0^{t_n}e^{-rs}(\phi_{t_n-s}(\theta_n)-1)ds - a\theta_n EZ_1(t_n) + a\theta_n(EZ_1(t_n)-EZ_1(u_n)).$$

Recalling that θ_n is small we look at a Taylor expansion of ϕ_t around 0 and use the fact that $\phi_t''(\cdot)$ is non-decreasing to see that

$$\phi_{t_n-s}(\theta_n) \le 1 + \theta_n\phi'_{t_n-s}(0) + \frac{\theta_n^2}{2}\phi''_{t_n-s}(\theta_n)$$

$$= 1 + \theta_n E\tilde{Z}_t(t_n-s) + \frac{\theta_n^2}{2}\phi''_{t_n-s}(\theta_n).$$

The previous inequality and the fact that

$$EZ_1(t_n) = \frac{\mu}{n^{\alpha-1}}\int_0^{t_n}e^{-rs}E\tilde{Z}_1(t_n-s)ds,$$

gives the bound

$$\Psi_n \le \frac{\mu}{2n^{\alpha-1}}\int_0^{t_n}e^{-rs}\theta_n^2\phi''_{t_n-s}(\theta_n)ds + (1-a)\theta_n EZ_1(t_n) + a\theta_n(EZ_1(t_n)-EZ_1(u_n)). \tag{15}$$

Since $u_n = t_n - k$ we see that

$$(1-a)\theta_n EZ_1(t_n) + a\theta_n(EZ_1(t_n)-EZ_1(u_n)) = \frac{\theta\lambda_1\mu}{r_1(\lambda_1+r)}(1-ae^{-\lambda_1 k}). \tag{16}$$

Using the fact that ϕ_t is the moment generating function of a binary branching process we can explicitly evaluate its second derivative as

$$\phi''_{t_n-s}(\theta_n) = \frac{e^\theta\lambda_1^2 e^{-\lambda_1(t_n-s)}(r_1-d_1 e^{-\lambda_1(t_n-s)})}{\left(r_1(1-e^{\theta_n})+e^{-\lambda_1(t_n-s)}(r_1 e^{\theta_n}-d_1)\right)^3} \le \frac{r_1 e^{2\lambda_1(t_n-s)}}{\lambda_1(1-\theta-\theta_n)^3}$$

$$\le r_1 e^{2\lambda_1(t_n-s)}\left(\frac{1}{\lambda_1(1-\theta)^3}+\frac{\theta_n}{\lambda_1(1-\theta)^3}\right).$$

Plugging the previous bound for $\phi_{t_n-s}(\theta_n)$ into (15) and using (16), we see

$$\Psi_n \le n^{1-\alpha}\left(\frac{\mu\theta^2\lambda_1}{2r_1(r+2\lambda_1)(1-\theta)^3}+\frac{\mu\theta\lambda_1(1-ae^{-\lambda_1 k})}{r_1(r+\lambda_1)}+O(n^{-\lambda_1 c})\right).$$

Looking back at (11), we see that to get an upper bound on tail probabilities of Z_1 we need to bound an exponential integral. In (12) we decomposed this exponential integral into two terms. Combining the previous display with (14) we see that

$$\log P(Z_1(\xi_n) > aEZ_1(u_n)) \leq k_1 n^{1-2\alpha} \log n + n^{1-\alpha} \left(\frac{\mu \theta^2 \lambda_1}{2r_1(r+2\lambda_1)(1-\theta)^3} + \frac{\mu \theta \lambda_1 (1-ae^{-\lambda_1 k})}{r_1(r+\lambda_1)} + O(n^{-\lambda_1 c}) \right)$$

$$= n^{1-\alpha} \left(\frac{\mu \theta^2 \lambda_1}{2r_1(r+2\lambda_1)(1-\theta)^3} + \frac{\mu \theta \lambda_1 (1-ae^{-\lambda_1 k})}{r_1(r+\lambda_1)} + o(1) \right). \tag{17}$$

The exponential decay at rate $n^{1-\alpha}$ follows by observing that the expression

$$\frac{\mu \theta \lambda_1}{2r_1(r+2\lambda_1)(1-\theta)^3} + \frac{\mu \lambda_1 (1-ae^{-\lambda_1 k})}{r_1(r+\lambda_1)}$$

is strictly negative for $k < \frac{1}{\lambda_1} \log a$ and θ sufficiently close to 0. $\quad\square$

4 IMPORTANCE SAMPLING

The previous result establishes that $p_n(a)$ decays exponentially in $n^{1-\alpha}$; thus, estimating these probabilities via Monte Carlo simulation requires the use of variance reduction techniques. We will focus on the method of Monte Carlo with importance sampling. Recall that importance sampling works by sampling from an alternative distribution that reduces the variance of the Monte Carlo estimator. For example, suppose we wish estimate $z = E[f(Z)]$ where Z has cdf F, then we construct a sampling measure $\tilde{F}$ and generate N i.i.d. samples $Z_1, \ldots, Z_N$ according to $\tilde{F}$. The estimate for z is

$$\hat{z} = \frac{1}{N} \sum_{i=1}^{N} f(Z_i) L(Z_i),$$

where L is a likelihood ratio between dF and $d\tilde{F}$ that insures $\hat{z}$ is an unbiased estimator for z. The challenge of importance sampling is choosing $\tilde{F}$ so that the variance of $\hat{z}$ is sufficiently reduced.

An approach that has been successful in constructing sampling measures is to consider a scaled version of the system and look at solutions for a relevant differential inequality. In the following we outline this approach for the construction of estimators of $p_n(a)$. This will be addressed in greater detail in future work.

Define scaled versions of our processes as

$$Z_0^n(t) = \frac{1}{n} Z_0(t) \ , \ Z_1^n(t) = n^{\alpha-1} Z_1(t) \ \text{ and } \ Z^n(t) = (Z_0^n(t), Z_1^n(t)).$$

Then for $a > 1$ define the function

$$p_n(x; a) = P(Z_1^n(\xi_n) > aEZ_1^n(u_n) | Z^n(0) = x).$$

From the previous theorem we know that this probability decays exponentially, so we consider a logarithmic transform

$$W_n(x; a) = \log\left(-n^{1-\alpha} p_n(x; a)\right),$$

where $x \in \mathbb{R}_+^2$. Since $p_n(x; a)$ is a time-homogeneous probability we will focus on the embedded discrete time process for the remainder, which can be written as

$$Z_{k+1} = Z_k + Y_{k+1}(Z_{k+1}), \quad k \geq 0$$

where Y_k represents the result of the kth event, i.e. a birth, death, or mutation, and Z_k is the number of resistant and sensitive cells immediately after the kth event. If we define $e_0 = (1,0)$ and $e_1 = (0,1)$ the possible values of Y_k are $v_1 = e_0, v_2 = -e_0, v_3 = e_1$, and $v_4 = -e_1$.

We can renormalize the sensitive and resistant populations and arrive at a scaled version of the embedded process

$$Z_{k+1}^n = Z_k^n + Y_{k+1}^n(Z_{k+1}^n), \quad k \geq 0.$$

The possible values of Y_k^n are

$$v_i(n) = \begin{cases} \frac{v_i}{n}, & 1 \leq i \leq 2 \\ \frac{v_i}{n^{1-\alpha}}, & 3 \leq i \leq 4 \end{cases}$$

Given the state of the scaled process is $x = (x_0, x_1)$, the rate at which jumps occur is given by

$$R_n(x) = n^{1-\alpha}(x_1(r_1 + d_1) + x_0\mu) + nx_0(r_0 + d_0).$$

and the jump probabilities are given by

$$q_x(v_1) = \frac{nx_0r_0}{R_n(x)}, \quad q_x(v_2) = \frac{nx_0d_0}{R_n(x)}, \quad q_x(v_3) = \frac{n^{1-\alpha}(x_1r_1 + x_0\mu)}{R_n(x)}, \quad \text{and} \quad q_x(v_4) = \frac{n^{1-\alpha}x_1d_1}{R_n(x)},$$

i.e. for $j = 1, \ldots, 4$, $P(Y_1^n(x) = v_j) = q_x(v_j)$.

The following function is useful for the construction and analysis of importance sampling changes of measure

$$\psi(y,x) = \log \sum_j q_x(v_j)e^{y \cdot v_j}.$$

We will use the sampling measure defined by

$$\tilde{q}_x(v) = q_x(v)\exp(-\partial_x g(x) \cdot v - \psi(-\partial_x g(x), x))$$

for a smooth function g that satisfies $g(x) \leq 0$ for x such that $x_1 \geq n^\alpha x_0$ and $x_1 \geq aEZ_1^n(u_n)$. Based on this sampling measure we form the estimator

$$\hat{z}_n(a) = 1_{\{Z_1(N_n) > aEZ_1(u_n)\}} \prod_{j=1}^{N_n} \exp\left[\partial_x g(Z_{j-1}^n) \cdot Y_j^n + \psi(\partial_x(g(Z_{j-1}^n), Z_{j-1}^n))\right], \tag{18}$$

where N_n is the discrete time analog of ξ_n.

We will now discuss how to choose g so that we can estimate the second moment of the associated importance sampling estimator. Define the function $v(x) = \exp(-n^{1-\alpha}g(x))$. Then a condition on g that enables control over the second moment of the importance sampling estimator (in terms of $g(z_0)$ where $z_0 = (1,0)$) is

$$E\left[\frac{v(x + Y^n(x))}{v(x)}\exp\left(\partial_x g(x) \cdot Y + \psi(-\partial_x g(x), x)\right)\right]$$

$$= \sum_{j=1}^4 q_x(v_j)\exp\left[n^{1-\alpha}\left(g(x) - g(x + v_j(n))\right) + \partial_x g(x) \cdot v_j + \psi(-\partial_x g(x), x)\right] \leq 1. \tag{19}$$

This inequality basically implies a supermartingale property for the likelihood ratio and thus allows one to use the optional sampling theorem to control the variance of the estimator, see Blanchet and Glynn 2008; Blanchet, Glynn, and Leder 2012. Satisfying this inequality is in general difficult, so one looks at satisfying a limiting version of this inequality. This leads to a search for sub solutions to the so-called

'Isaac's Equation' (Dupuis and Wang 2009). The structure of the large n limit of (19) breaks down into two cases:

Case 1: $x_0 > 0$

$$E\left[\frac{v(x+Y^n(x))}{v(x)}\exp\left(\partial_x g(x)\cdot Y + \psi(-\partial_x g(x),x)\right)\right]$$
$$\approx \sum_{j=1}^{2}\exp\left[\partial_x g(x)\cdot v_j(1-n^{-\alpha}) + \psi(-\partial_x g(x),x)\right]q_x(v_j).$$

Case 2: $x_0 = 0$

$$E\left[\frac{v(x+Y^n(x))}{v(x)}\exp\left(\partial_x g(x)\cdot Y + \psi(-\partial_x g(x),x)\right)\right]$$
$$= \sum_{j=3}^{4}\exp\left[n^{1-\alpha}\left(g(x)-g(x+v_j(n))\right) + \partial_x g(x)\cdot v_j + \psi(-\partial_x g(x),x)\right]q_x(v_j).$$

Looking at the limit as $n \to \infty$, if $x_0 > 0$, we need g to satisfy the following differential inequality

$$\psi(-\partial_x g(x),x) + \psi(\partial_x g(x),x) \leq 0, \tag{20}$$

and if $x_0 = 0$ we require

$$\psi(-\partial_x g(x),x) \leq 0, \tag{21}$$

with the boundary condition $\lim_{x_1 \to \infty} g(x) \leq 0$ for $x_0 = 0$. Based on the results of Dupuis and Wang 2007 we expect that if g satisfies the equations outlined above then

$$\liminf_{n\to\infty} -n^{\alpha-1}\log E\hat{z}_n(a) \geq g(z_0).$$

Therefore, our goal is to create a function g that satisfies (20) and (21) with the maximum possible value at the origin. Note that due to the convexity of ψ and the fact that $\psi(0) = 0$, (20) in fact imposes that one use no importance sampling in the region $x_0 > 0$.

Interestingly, the limiting differential equation has distinct behavior along the axis $x_0 = 0$ versus the interior $x_0 > 0$. This is similar to what is seen in rare event simulation problems involving queueing systems (Dupuis, Sezer, and Wang. 2007). In queueing problems the separate behavior along the axis is due to discontinuous behavior in the pre-limit. In contrast in this problem the separate behavior along the axis arises because of separation of scales between Z_0 and Z_1 in the limit. In future work we plan to investigate solutions to the differential inequality from (20) and (21). Following work by Dupuis and Wang, we will look at state dependent mixtures between solutions to the two differential equations. Specifically, as $x_0 \to 0$ we will sample more frequently according to a solution of (21) and for x_0 far away from 0 we will sample from the original dynamics.

REFERENCES

Athreya, K., and P. Ney. 2004. *Branching Processes*. New York: Dover.

Blanchet, J., and P. Glynn. 2008. "Efficient rare-event simulation for the maximum of a heavy-tailed random walk.". *Ann. of Appl. Probab.* 18:1351–1378.

Blanchet, J., P. Glynn, and K. Leder. 2012. "On Lyapunov Inequalities and Subsolutions for Efficient Importance Sampling". *ACM TOMACS* Forthcoming.

Coldman, A., and J. Goldie. 1986. "A stochastic model for the origin and treatment of tumors containing drug-resistant cells". *Bull. Math. Biol.* 48 (3/4): 279–292.

Demicheli, R., and et al. 2008. "Recurrence dynamics does not depend on the recurrence site". *Breast Cancer Research* 10:R83.

Demicheli, R., and et al. 2012. "Recurrence Dynamics for NonSmall-Cell Lung Cancer: Effect of Surgery on the Development of Metastases". *Journal of Thoracic Oncology* 7:723–730.

Dupuis, P., A. Sezer, and H. Wang.. 2007. "Dynamic importance sampling for queueing networks.". *Ann. Appl. Probab.* 17:1306–1346.

Dupuis, P., and H. Wang. 2007. "Subsolutions of an Isaacs equation and efficient schemes of importance sampling". *Mathematics of Operations Research* 32:723–757.

Dupuis, P., and H. Wang. 2009. "Importance Sampling for Jackson Networks". *Queueing Systems: Theory and Applications* 62:113–157.

Durrett, R., D. Schmidt, and J. Schweinsberg. 2009. "A waiting time problem arising from the study of multi-stage carcinogenesis". *Annals of Applied Probability* 19:676–718.

Foo, J., and K. Leder. 2012a. "Dynamics of Cancer Recurrence". *Annals of Applied Probability* To Appear.

Foo, J., and K. Leder. 2012b. "Weak Limits of Crossover Times in Multi-type Branching Processes". *Forthcoming*.

Frank, S. 2003. "Somatic mosaicism and cancer: inference based on a conditional Luria-Delbrück distribution". *Journal of Theoretical Biology* 223:405–412.

Grimmett, G., and D. Stirzaker. 2001. *Probability and Stochastic Processes*. Oxford: Oxford University Press.

Haeno, H., Y. Iwasa, and F. Michor. 2007. "The Evolution of Two Mutations Durng Clonal Expansion". *Genetics* 177:2209–2221.

Komarova, N. 2006. "Stochastic modeling of drug resistance in cancer". *Journal of Theoretical Biology* 239:351–366.

AUTHOR BIOGRAPHIES

JASMINE FOO is Assistant Professor in the School of Mathematics at the Univeristy of Minnesota Twin-Cities. She received her PhD in Applied Mathematics from Brown University and was previously a post-doc in bio-statistics at the Harvard School of Public Health. Her research interests are in stochastic population models with applications to cancer modelling. Her e-mail address is jyfoo@math.umn.edu.

KEVIN LEDER is an assistant professor in the ISyE program at the Univeristy of Minnesota Twin-Cities. Prior to joining the University of Minnesota he was a post-doctoral researcher in the Department of Biostatistics at Harvard University, and a post-doctoral researcher in the IEOR department at Columbia University. He received his bachelors degree in applied mathematics from University of Colorado, and his PhD from the Division of Applied Mathematics at Brown University. His research interests include stochastic modelling and simulation, with a particular interest in cancer modelling. His email address is kevin.leder@isye.umn.edu.