Heavy-tail distributions in cell proliferation models

Marek Kimmel
with Philip Ernst, Monika Kurpas and Quan Zhou

Dec 5, 2024

Abstract

Mutation processes in proliferating cancer cells lead to populations with highly diversified genomes. A natural simple tool to describe the resulting stochastic phenomena are branching processes (bp) with countable collections of types. I will review several such models I worked on over a number of years. These includes processes such as (i) gene amplification (increase in copy count) in response to chemotherapy (Kimmel and Stivers, 1994), (ii) uneven replication and segregation of chromosomes (Kimmel, 1997), (iii) bp's with randomly changing lifetime distributions (Ernst et al. 2019), and (iv) Tug-of-War model of competition of advantageous and deleterious mutations embedded in proliferating cells (Wang and Kimmel 2023). Although these are biologically and mathematically different models, they share some "exotic" properties, which result from nontrivial interaction between branching and type transitions. Among other, they result in limit growth laws that can be exponential modified by negative fractional power, and unexpected relations between criticality and finite time moment explosions consistent with increasingly heavy-tail distributions of cell counts. Other interesting properties are present. This review illustrates variety of behaviors of such models.

Nonexponential moments, heavy distribution tails, and criticality conditions in models of secondary tumors

Outline

Extremely high genetic diversity in a single tumor points to prevalence of non-Darwinian cell evolution [*Ling et al. PNAS 2015; 112 : E6496-E6505*]

The prevailing view that the evolution of cells in a tumor is driven by Darwinian selection has never been rigorously tested.

Because selection greatly affects the level of intra-tumor genetic diversity with profound consequences for treatment outcomes, it is important to assess whether intra-tumor evolution follows the Darwinian or the non-Darwinian mode of evolution.

To provide statistical power, many regions in a single tumor need to be sampled.

This account is mostly based on *[Ernst et al. Advances in Applied Probability, 2018]*.

Observations

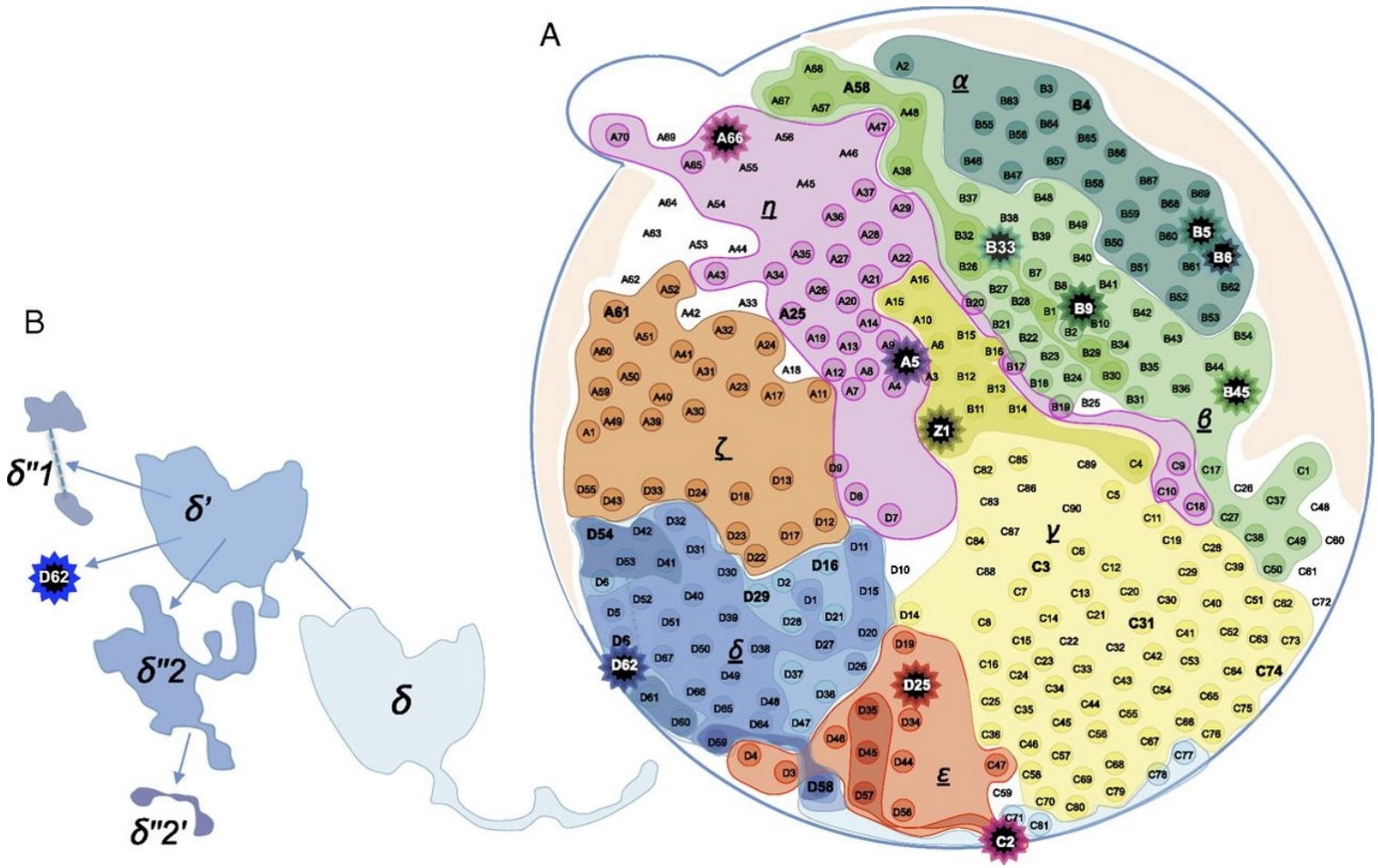


Figure: From a hepatocellular carcinoma (HCC) tumor, multiregional samples from the tumor were evaluated, using either whole-exome sequencing (WES) (n=23 samples) or genotyping (n=286).

Tumor field model

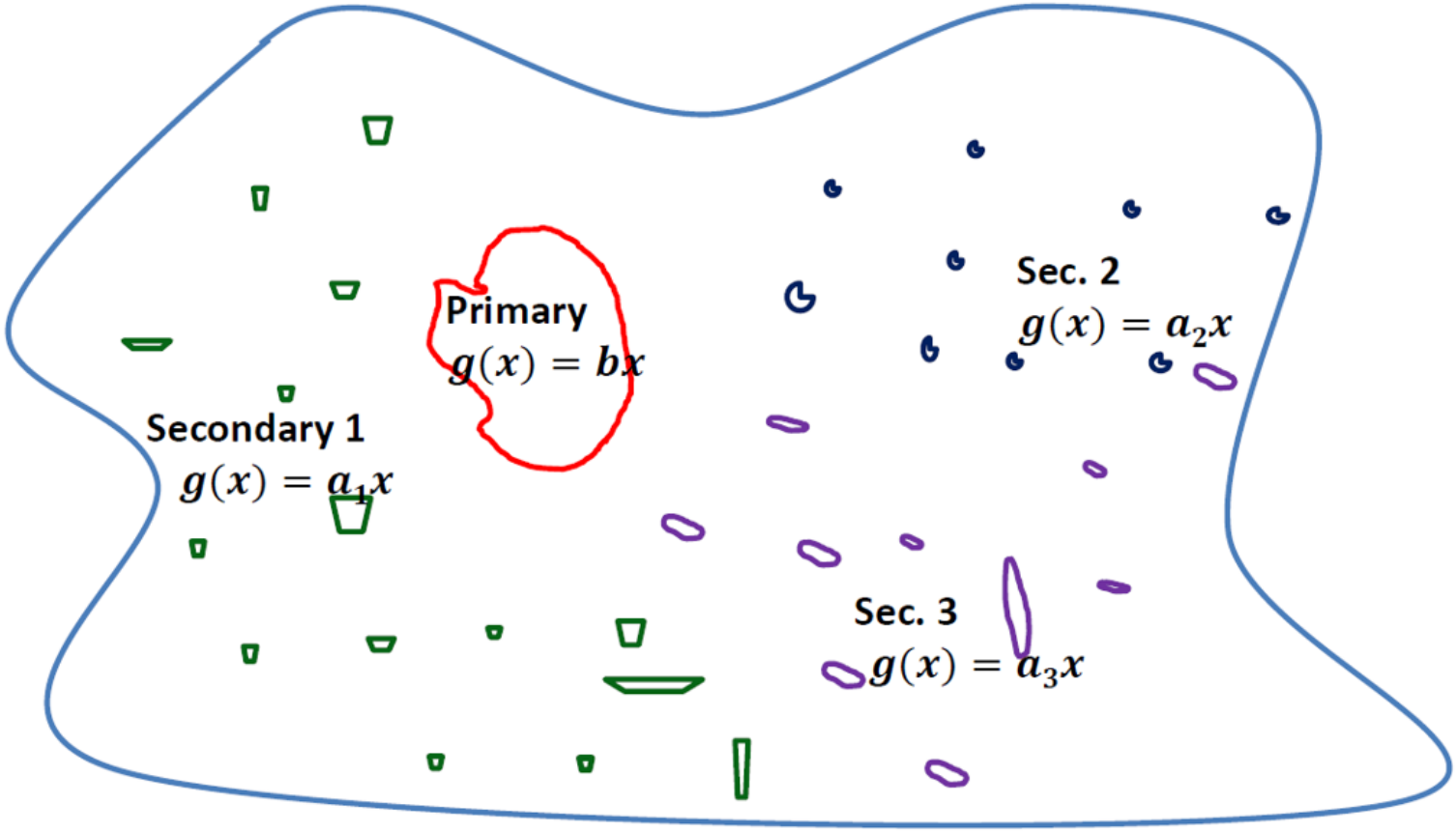


Figure: Wide distribution of secondary tumor growth rates a_i

Hypotheses of the model

A primary tumor is generated from a single cell at time $t = 0$ and grows at rate $g(x)$, where x denotes the number of cells in the tumor.

$$g(x) = bx$$

The growing tumor emits transformed single cells at rate $\beta(x)$.

$$\beta(x) = mx^{\alpha}$$

Each transformed cell develops into a new tumor, which grows at a generally different rate $g(x)$ and emits new transformed cells just as the primary does.

$$g(x) = ax$$

Growth rate of secondary tumors is **a random variable with exponential distribution**

$$a \sim \exp(\lambda)$$

PDE model with randomized growth rates

Solving the PDE (v. Foerster type) and randomizing the growth rate we obtain the total count of secondary foci

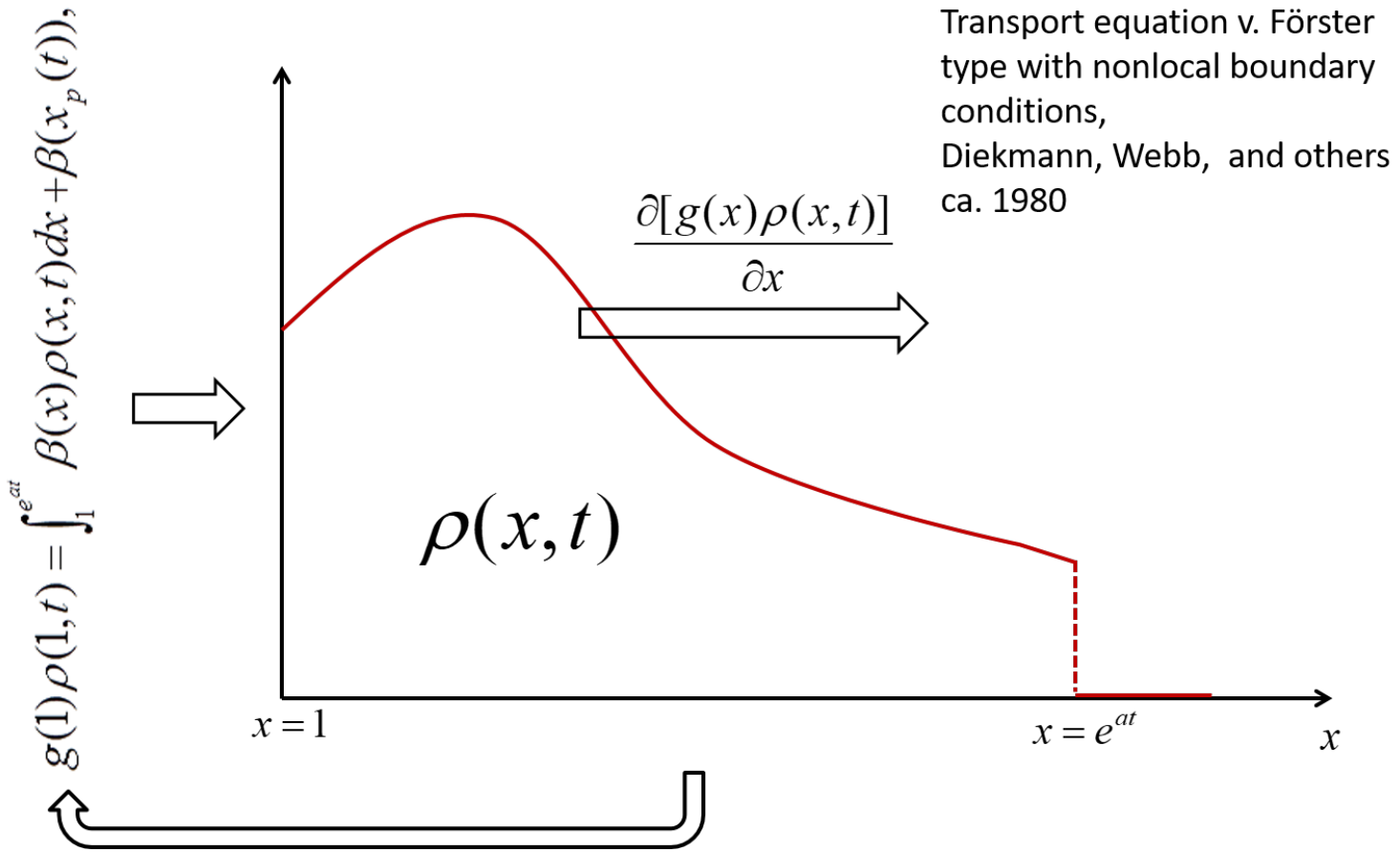
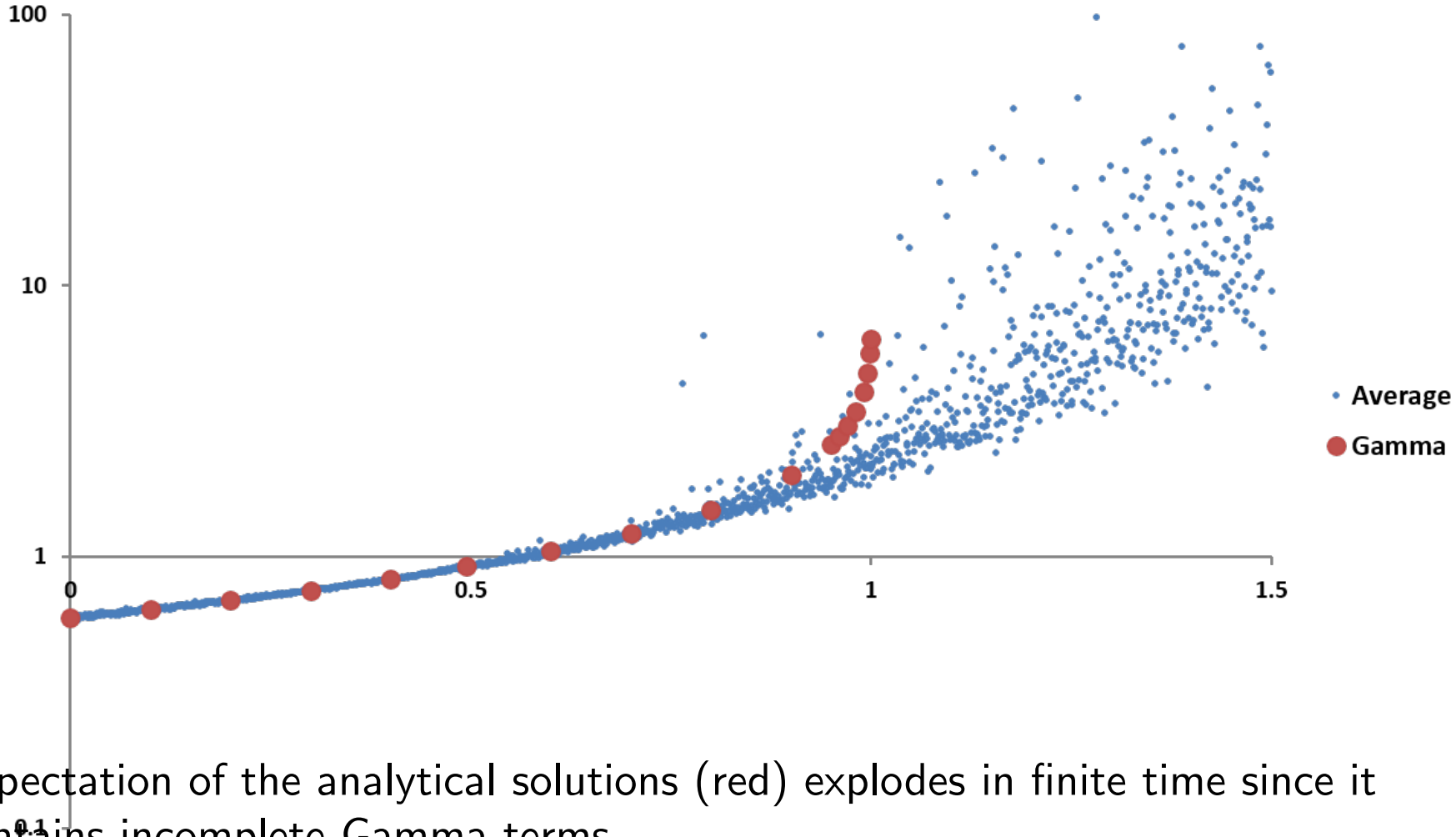


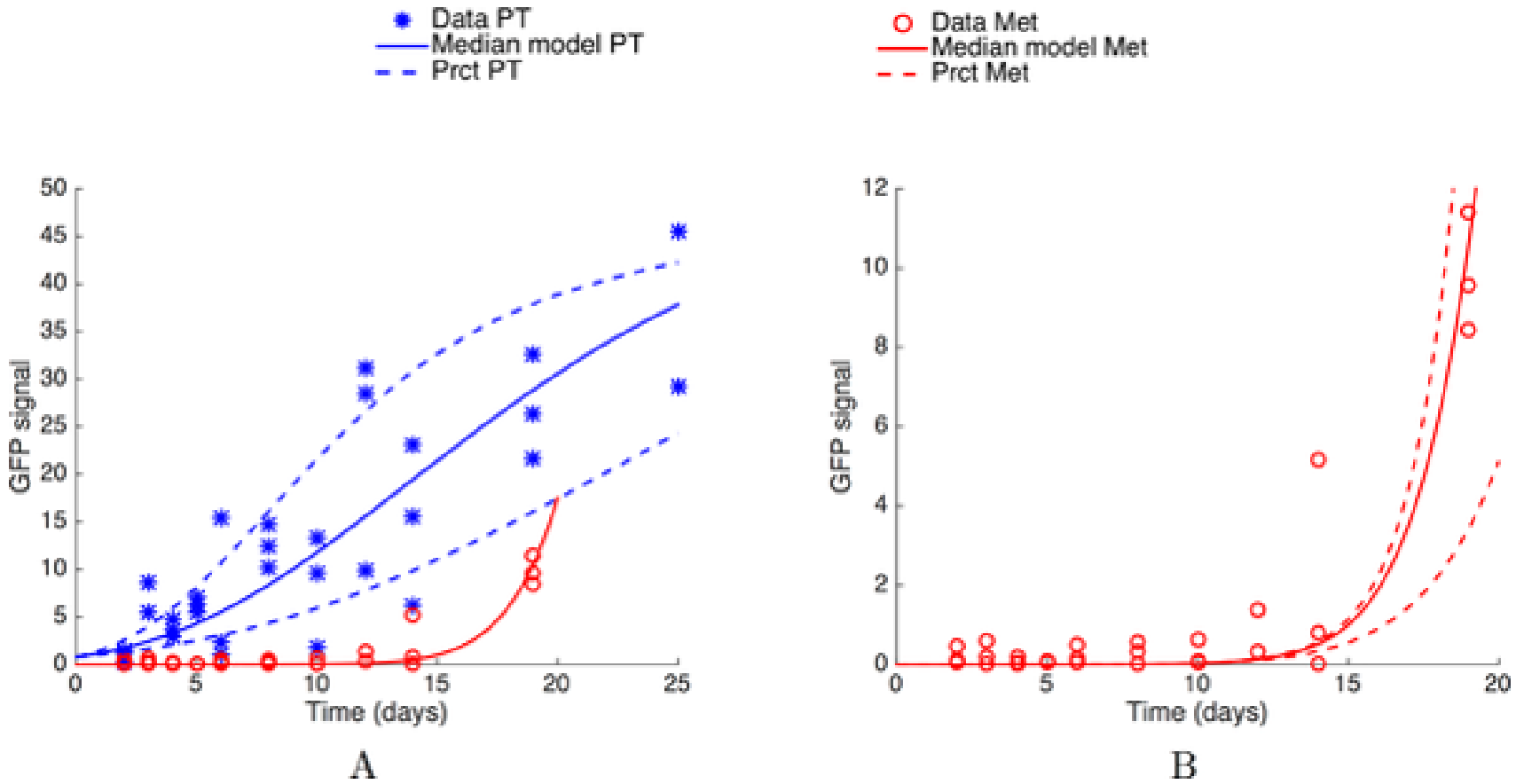
Figure: Schematic of transport PDE with non-local boundary conditions

Analytical solution explodes in finite time



Expectation of the analytical solutions (red) explodes in finite time since it contains incomplete Gamma terms.
Randonly sampled trajectories (blue) do not explode but grow faster than exponential

A piece of data on rapid metastasis lab model growth from Baratchart et al



right) but is only quasi-stochastic Model captures some features of the Baratchart et al. experiment.

Toy Model

Secondary tumors grow exponentially at rate a, which itself is a random variable

$$X(t|a) = \exp(at),\ t \geq 0,\ a \sim \exp(\lambda)$$

It now has Pareto tail

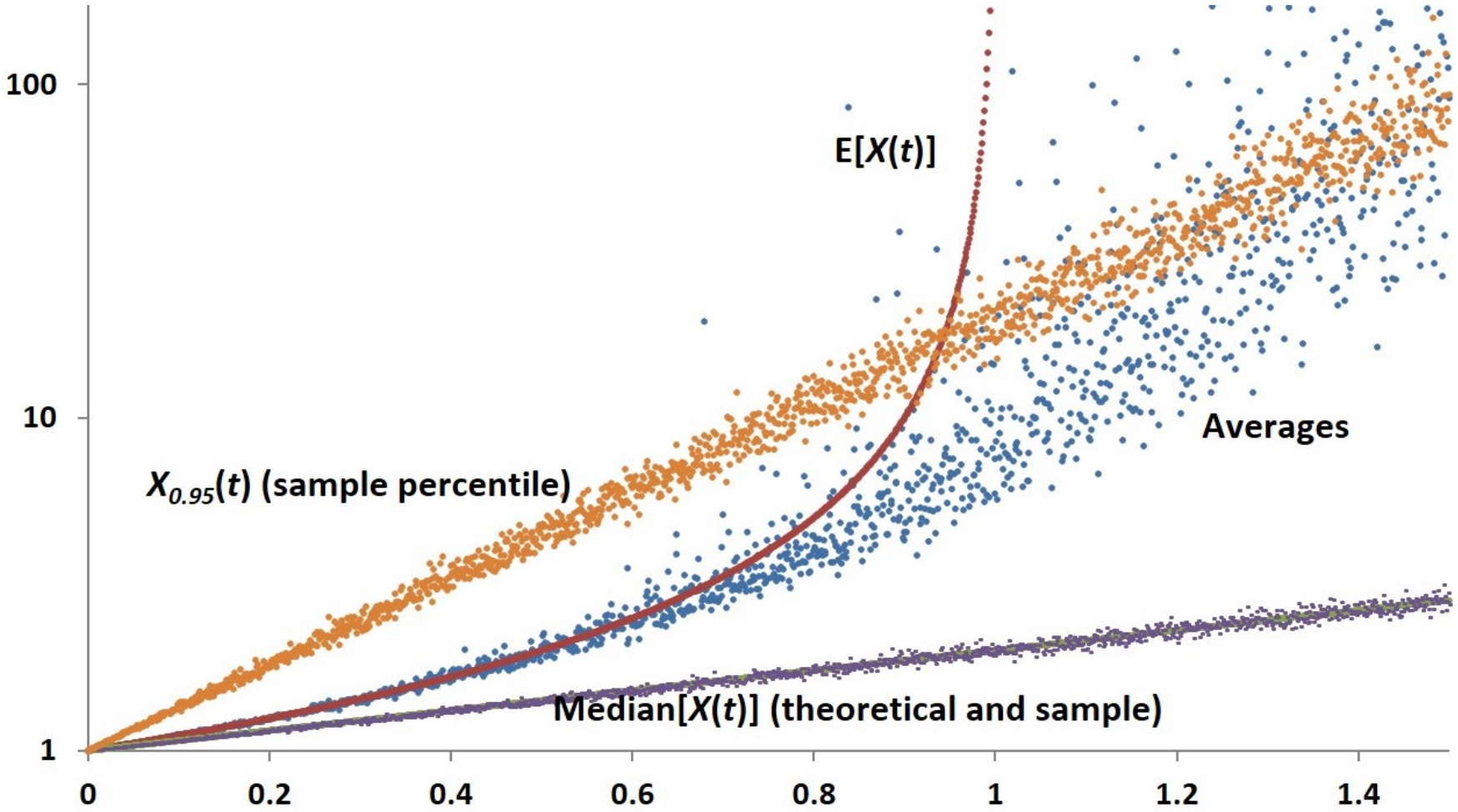
$$\mathbb{P}[X(t) > x] = \min\{1, x^{-\lambda/t}\},\ t \geq 0$$

$$\mathbb{E}[X(t)] = \int_0^\infty \mathbb{P}[X(t) > x]dx = 1 + \int_1^\infty x^{-\lambda/t}dx$$

$\mathbb{E}[X(t)] = \lambda(\lambda - t)^{-1}$ explodes when $t \uparrow \lambda$

Higher moments explode faster.

Simulations of the Toy Model



Conclusions from the Toy Model

Conclusions:

- Averages do not explode but grow faster than exponential.
- Expected value may not be useful as the central tendency of growth after certain time.
- Median is a good model in this case, but have you heard of anybody using medians to model?
- Quantiles of increasing order increase exponentially at increasing rates

We need a truly stochastic model.
We start with something really simple

Coldman-Goldie model

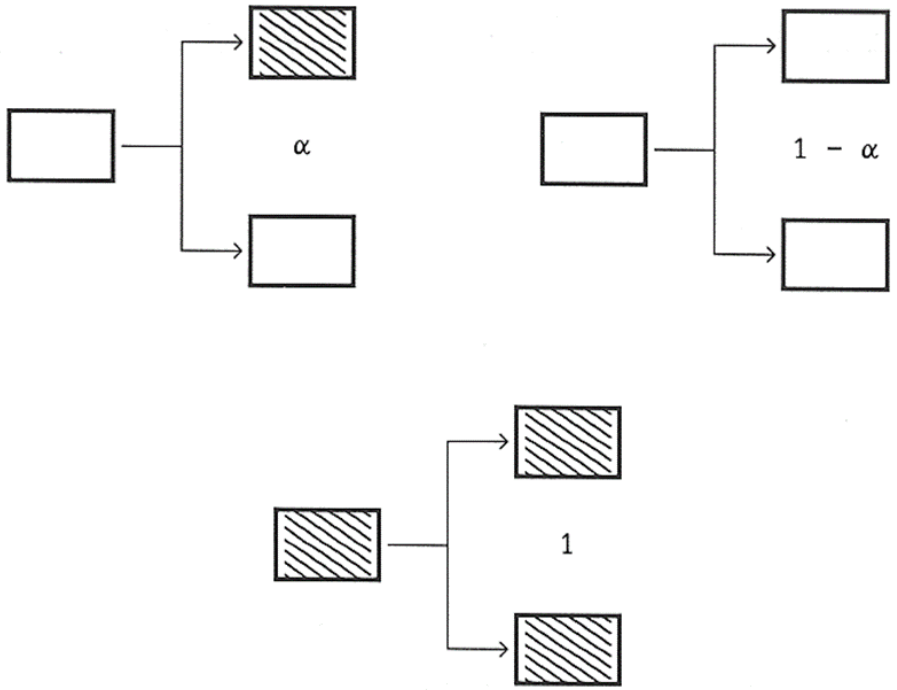
[Goldie and Coldman, 1979; Coldman and Goldie, 1985]
Based on *[Kimmel and Axelrod 2015]* book

Branching process approach to a theory of resistance, which has become influential in the cancer research community.

The assumptions of the theory are as follows

1. Cancer cell population is initiated by a single cell which is sensitive to the cytotoxic (chemotherapeutic) agent. The population proliferates without losses.
2. Interdivision time of cells is a random variable with a given distribution.
3. At each division, with given probability, a single progeny cell mutates and becomes resistant to the cytotoxic agent.
4. Mutations are irreversible.

Coldman-Goldie model



We wish to compute the probability that when the tumor is discovered, it does not contain any resistant cells.

Only in such a situation is the use of a cytotoxic agent effective.

If even a small subpopulation of resistant cells exists, the cancer cell population will eventually re-emerge despite the therapy.

Coldman-Goldie model

We translate the hypotheses of clonal resistance into the language of branching processes.

1. In the process, there exist two types of particles, labeled 0 (sensitive) and 1 (resistant).
2. The process is initiated by a single type 0 particle.
3. The life spans of particles are independent random variables, distributed exponentially with parameter λ.
4. Each particle, at death, divides into exactly two progeny particles:
   - 0-particle produces either two 0-particles, wp $1-\alpha$, or one 0− and one 1-particle, wp α.
   - 1-particle produces two 1-particles.

Thus, we have a $two-type$ time continuous Markov branching process.

Coldman-Goldie model

- $F_0(s_0, s_1; t)$ is the joint probability generating function of the numbers of cells of both types, present at time t in the process initiated at time 0 by a type 0 cell.
- $F_1(s_1; t)$ is the pgf of the numbers of cells of type 1, present at time t in the process initiated at time 0 by a type 1 cell.

Theorem

The solution of the differential equation

$$\frac{dF(t)}{dt} = f(t)F(t) + hF(t)^2, \tag{1}$$

where $f \in C[0, \infty)$, with initial condition $F(0)$, is a uniquely defined function $F \in C^1[0, \infty)$

$$F(t) = \frac{F(0)e^{\int_0^t f(u)du}}{1 - hF(0)\int_0^t e^{\int_0^u f(v)dv}du}. \tag{2}$$

Coldman-Goldie model

In our application, $f_0(s) = (1-\alpha)s_0^2 + \alpha s_0 s_1$, $f_1(s) = s_1^2$, and $\lambda_0 = \lambda_1 = \lambda$. In consequence,

$$\frac{dF_0}{dt} = -\lambda F_0 + \lambda[(1-\alpha)F_0^2 + \alpha F_0 F_1], \tag{3}$$

$$\frac{dF_1}{dt} = -\lambda F_1 + \lambda F_1^2. \tag{4}$$

$$F_1(s;t) = \frac{s_1}{s_1 + (1-s_1)e^{\lambda t}}. \tag{5}$$

Substituting (5) into Eqn. (3) and employing Theorem 1, we obtain

$$F_0(s;t) = \frac{s_0 e^{-\lambda t}[e^{-\lambda t}s_1 + (1-s_1)]^{-\alpha}}{1 + s_0\{[e^{-\lambda t}s_1 + (1-s_1)]^{1-\alpha} - 1\}s_1^{-1}}. \tag{6}$$

Coldman-Goldie model

Differentiating $F_0(s;t)$ with respect to s_0 and s_1 we obtain the expressions for the expected counts of the sensitive and resistant cells

$$M_0(t) = \frac{\partial F(1,1;t)}{\partial s_0} = e^{\lambda(1-\alpha)t}, \quad t \geq 0,$$

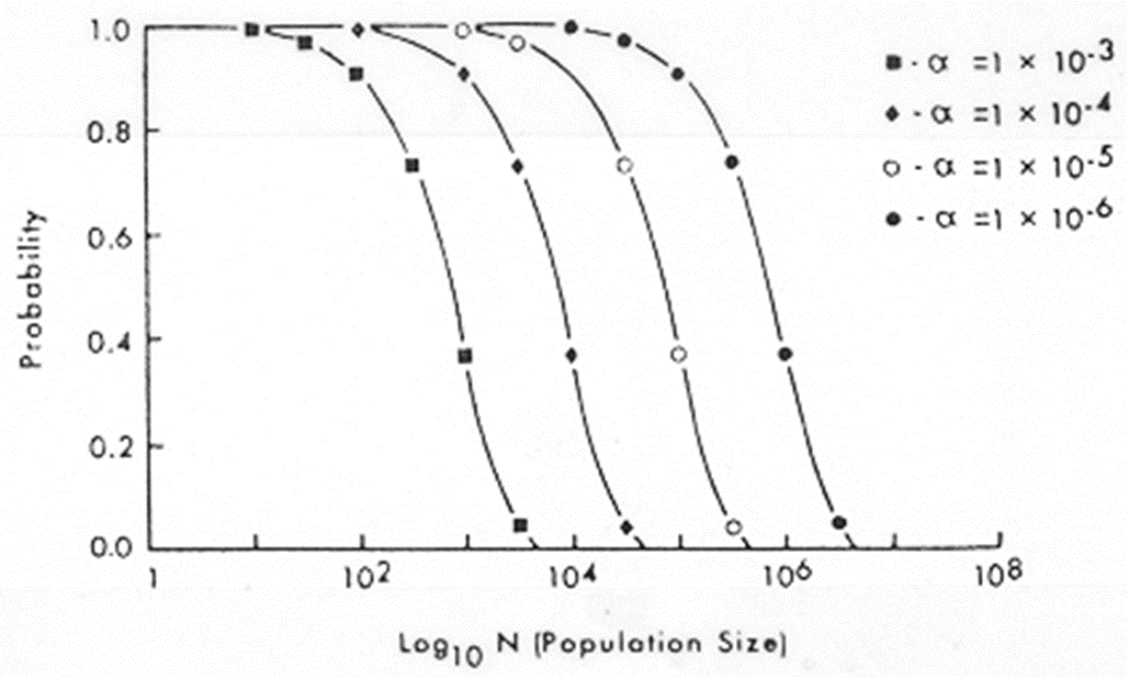
$$M_1(t) = \frac{\partial F(1,1;t)}{\partial s_1} = e^{\lambda t} - e^{\lambda(1-\alpha)t}, \quad t \geq 0.$$

The conclusion is that in absence of intervention the resistant cells eventually outgrow the sensitive ones.

Coldman-Goldie model

The probability of no resistant cells at time t is also easy to obtain

$$P(t) = \lim_{s_0 \uparrow 1} \lim_{s_1 \downarrow 0} F_0(s; t) = \frac{1}{(1-\alpha) + \alpha e^{\lambda t}} = \frac{1}{(1-\alpha) + \alpha[M_0(t) + M_1(t)]}. \tag{7}$$

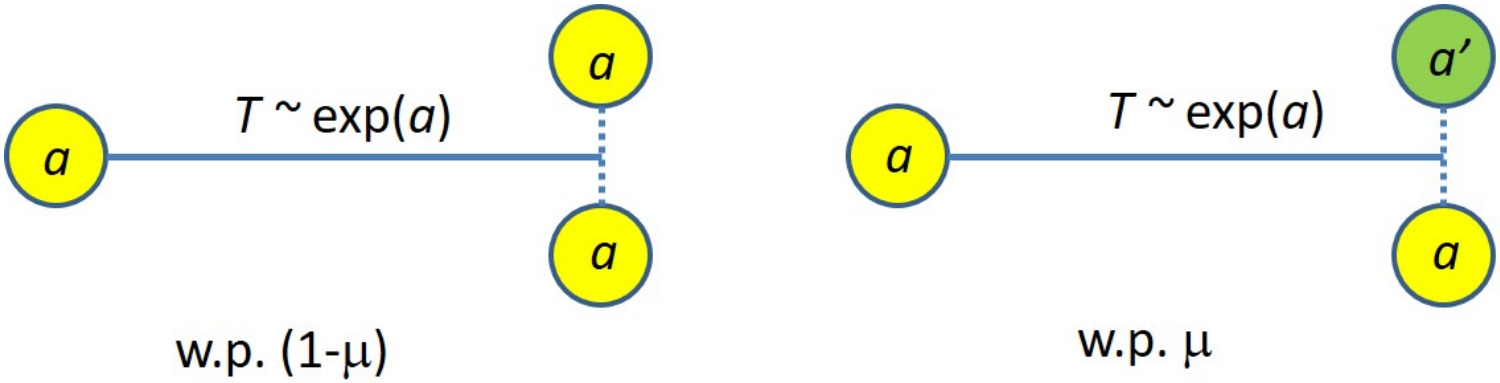


Depending on the mutation rate α, we obtain a window of opportunity for a sledgehammer therapy.

Back to our problem

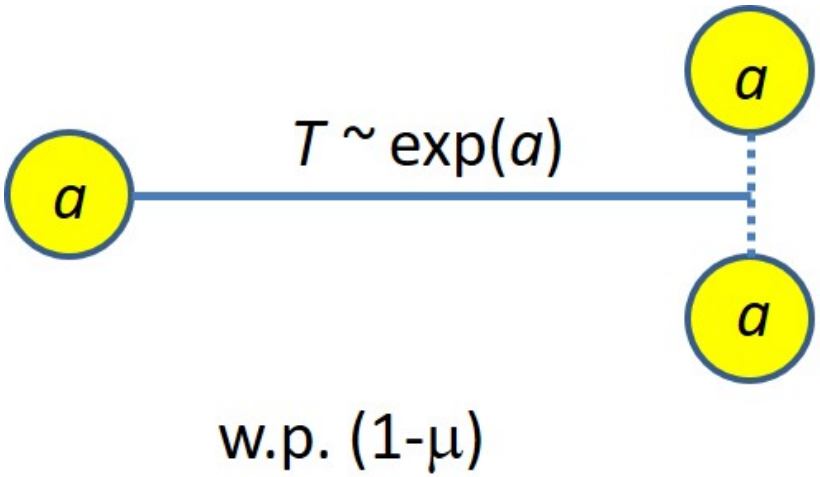
We need to modify the model so it can handle the diversity of cancer cell populations.

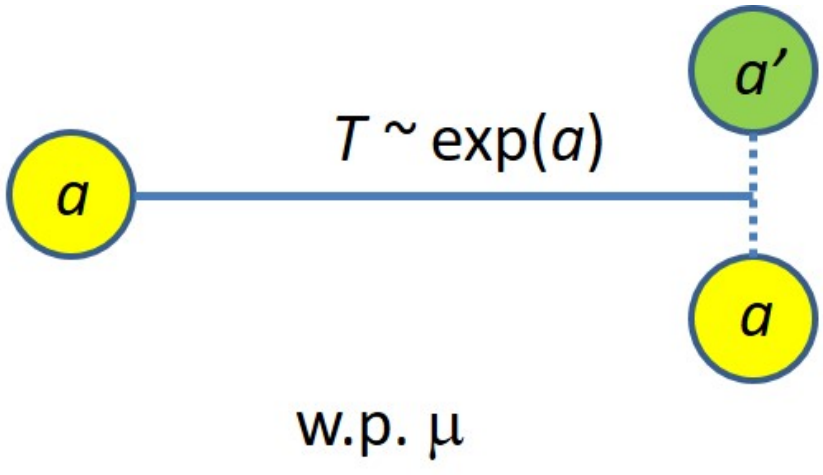
Coldman-Goldie model with a twist



- Cells are organized in proliferating clones characterized by division rates
- At each division, with probability , one cell mutates and assumes random division rate
- This means that mutant clones arising may be frequently quite sluggish (depending on) but sometimes very fast (passengers or drivers)
- Resulting model is a continuum-type time-continuous Markov branching process
- Infinite systems of ODEs can be written for the pgfs of the distribution of total cell count in all clones started by a mutant with division rate

Infinite Riccati Equation system





$X(a,t) = \#\{$ cells in the process started by cell of type $a\}$

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

Infinite Riccati Equation system

We can write the infinite Riccati Equation system

$$\frac{\partial F(s;a,t)}{\partial t} = -aF(s;a,t)+a[(1-\mu)F(s;a,t^2)+\mu F(s;a,t)\Phi(s;\lambda,t)], \quad t \geq 0, a > 0$$

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

$$\Phi(s;\lambda,t) = \int_0^\infty F(s;a',t)\lambda \exp(-\lambda a')d\lambda$$

We can use the Theorem for equation that has the form of

$$\frac{dF(t)}{dt} = f(t)F(t) + hF(t)^2, \tag{8}$$

with $f(t) = a\mu\Phi(s;\lambda,t)$

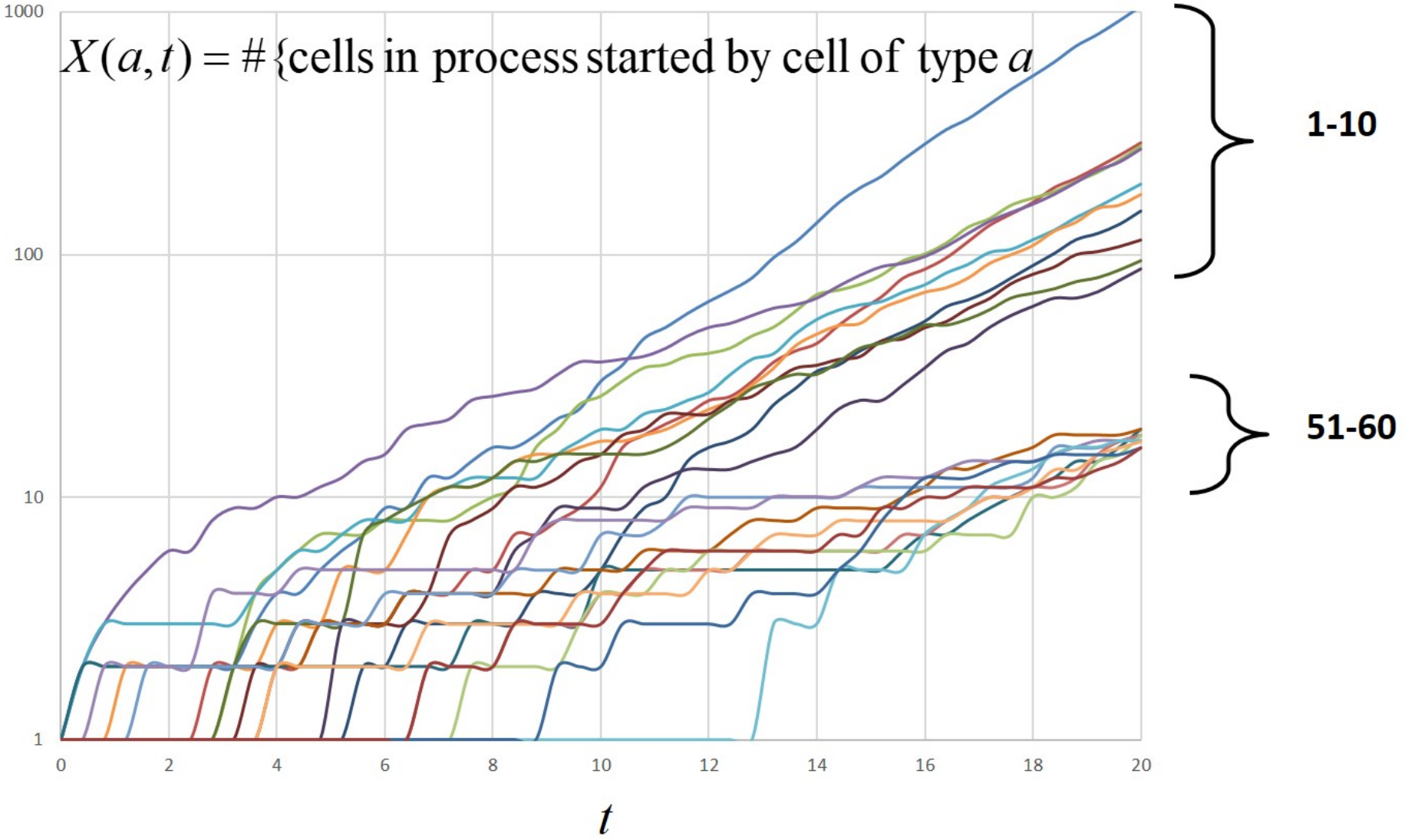
$$F(t) = \frac{F(0)e^{\int_0^t f(u)du}}{1 - hF(0)\int_0^t e^{\int_0^u f(v)dv}du}. \tag{9}$$

to see that pgf satisfies a nonlinear integral equation ...

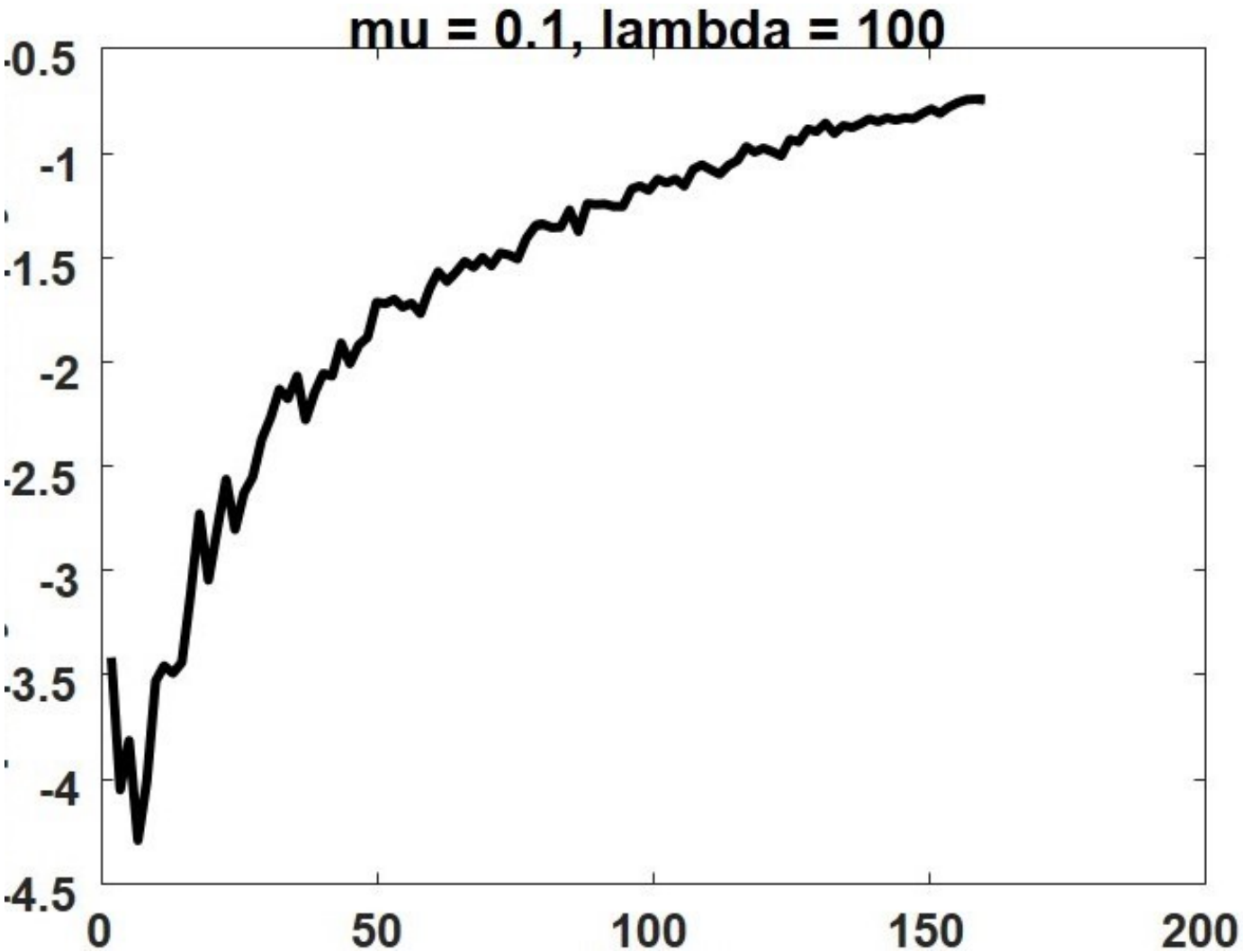
Variability of simulated stochastic trajectories

Trajectories ranked **1-10** and **51-60** out of 10^4 simulated

Parameter values: $\mu = 0.5$, $a = 0.01$, and $\lambda = 10$

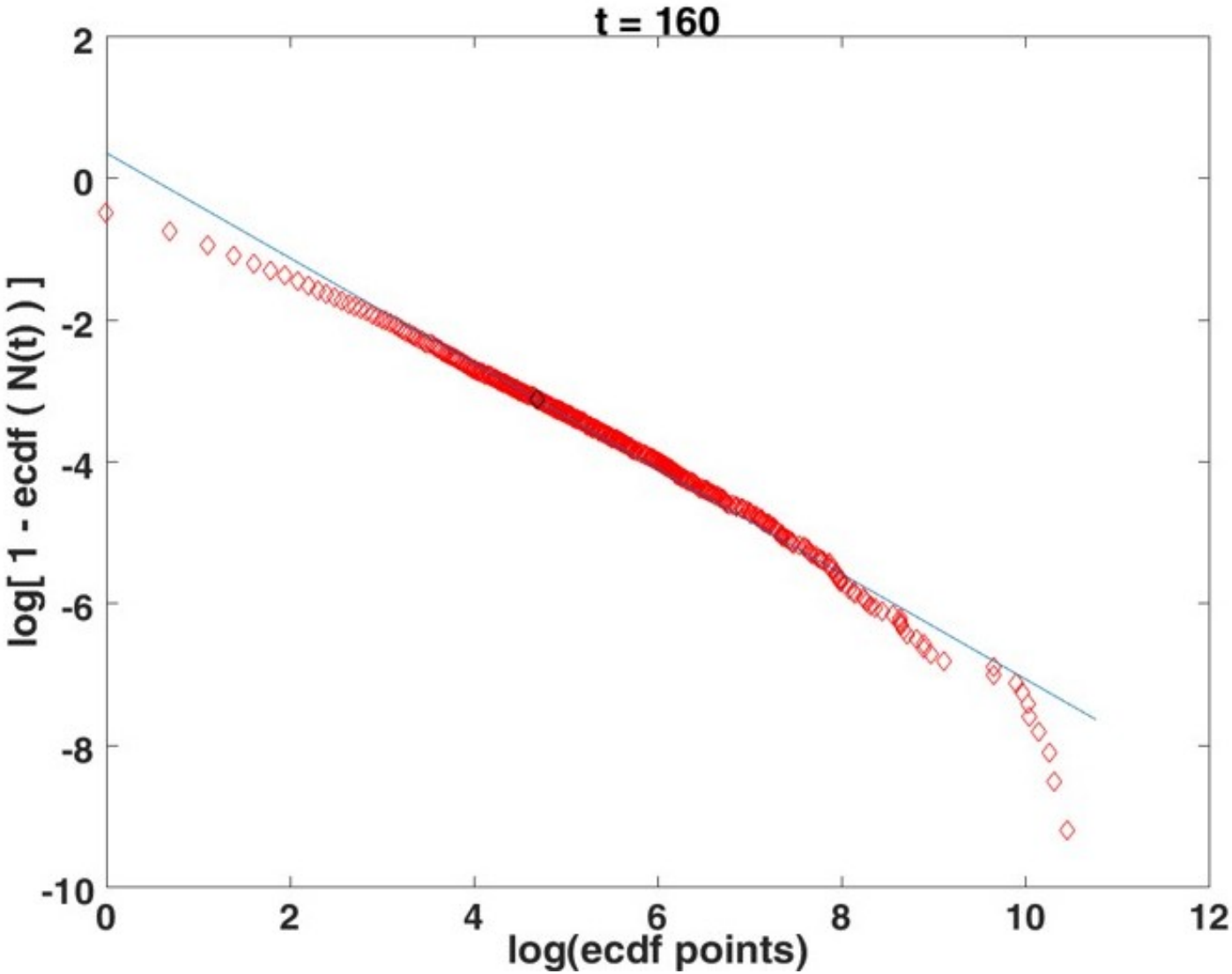


Power exponents



Tails of the distributions of cell counts seem to obey a power law, with estimated exponent close to -1 at the time when $\mathbb{E}[X(a,t)]$ explodes

Fit at the expectation explosion time point



Some Estimates (more in the Ernst et al.)

Let $X_k(a,t)$ be the number of cells generated by $k-1$ mutations
$X_1(a,t)$ denotes the number of primary cells (division rate a)
$X_2(a,t)$ denotes the number of cells that directly mutated from primary cells
$X_3(a,t)$
$X_4(a,t)$
.....

Solutions for $X_1(a,t)$ cells

Primary cells follow Yule's binary fission model

$$X_1(a,t) \sim F_1(s,a,t) = \frac{se^{-a(1-\mu)t}}{1-s(1-e^{-a(1-\mu)t})}$$

If we integrate over $a \sim \exp(\lambda)$, we obtain Yule-Simon distribution for $X_1(t)$

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

with $\nu(t) = \frac{\lambda}{(1-\mu)t}$.

For large n

$$\mathrm{P}[X_1(t)>n] \sim \frac{\Gamma(\nu+1)}{n^{\nu}}$$

with

$$E[X_1(t)] = \begin{cases} \dfrac{\nu}{\nu-1}; & \nu > 1 \\ \infty; & \nu \le 1 \end{cases}$$

As in the toy model and simulations.

Direct equations for expectations

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

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

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

$$M(a,t) = g(t) + a\mu g(t) \overset{(t)}{*} \varphi(t) \quad | \times \lambda e^{-\lambda a} | \int_0^\infty (\cdot) da$$

$$g(t) = \exp(a(1-\mu)t)$$

$$\Rightarrow \varphi(t) = f_1(t) + (\mu/\lambda) f_2(t) \overset{(t)}{*} \varphi(t) \quad \Rightarrow \varphi(t) = f_1(t) + f_1(t) \overset{(t)}{*} \sum_{i\geq 1} (\mu/\lambda)^i f_2^{\overset{(t)}{*} i}(t)$$

$$f_1(t) = \frac{1}{1 - t(1-\mu)/\lambda}, t \in [0, \lambda/(1-\mu)), f_2(t) = f_1(t)^2$$

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

Cell death and criticality

Suppose that each of the progeny cells may die with probability d.

Then expectation $\varphi(\lambda, t)$ explodes at $t = \lambda/c$, where

$c = (2 - \mu)(1 - d) - 1$

Time at explosion is becoming infinite if

$c = 0 \Leftrightarrow d = d^* = \frac{1-\mu}{2-\mu} < 0.5$

At $c = 0$, expectation $\varphi(t) = \varphi(0) \exp\left(\frac{\mu t}{\lambda(2-\mu)}\right)$.

For $d \in [0, d^*]$ solutions do not explode. Nothing special at $d = 0.5$?

End of Presentation

Thank you!