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

Philip Ernst, Marek Kimmel, Monika Kurpas and Quan Zhou

Rice University, Houston, TX, USA

Silesian University of Technology, Gliwice, Poland

kimmel@rice.edu

Outline

- Deterministic model of secondary tumors and incomplete gamma functions
- Will tumors explode? Or, how to model **truly** stochastic population processes?
- Goldie-Coldman with a twist, or thoughtless mutations and experimental math of heavy tails

Extremely high genetic diversity in a single tumor points to prevalence of non-Darwinian cell evolution Ling et al. PNAS (2015)

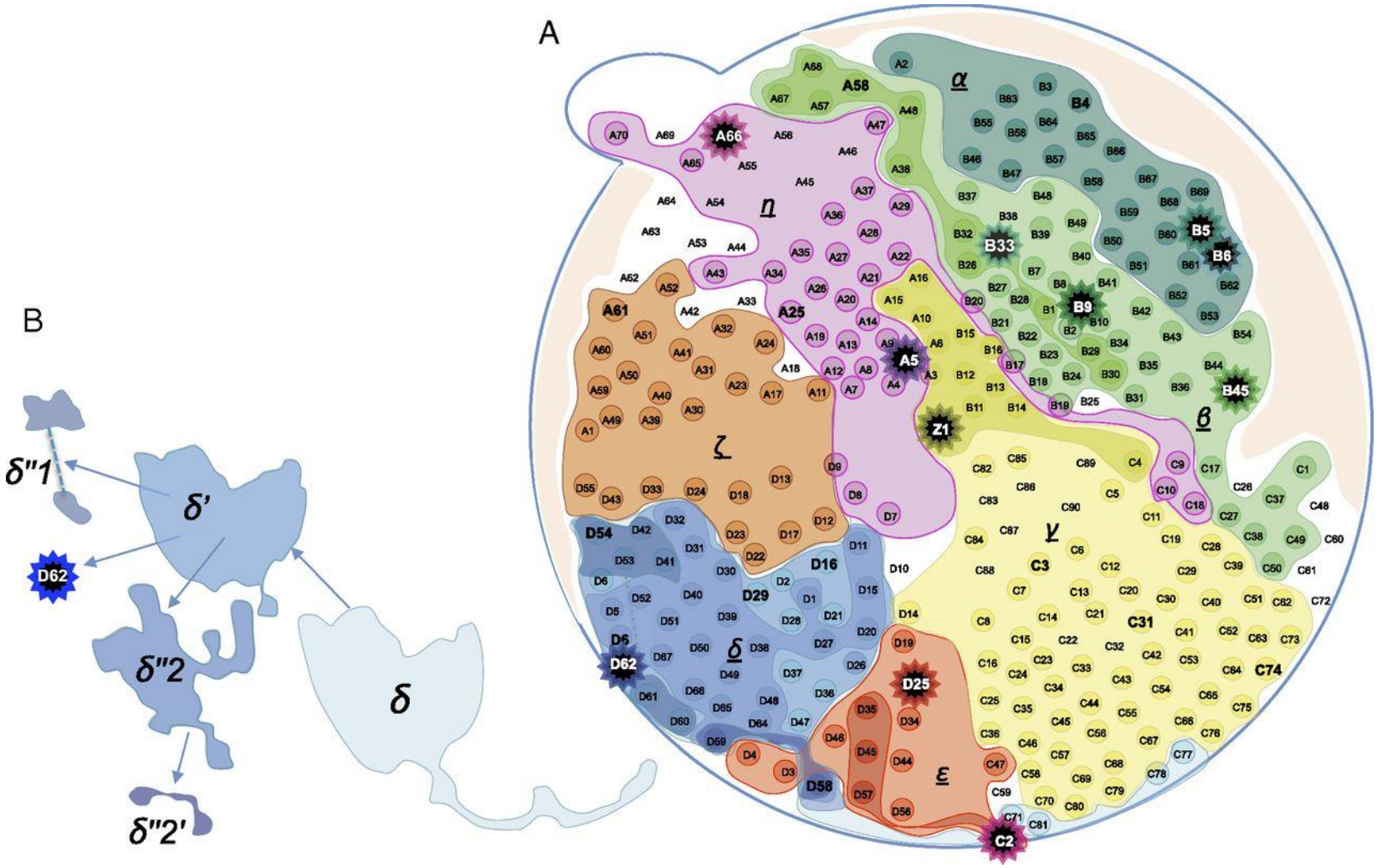
- The prevailing view that the evolution of cells in a tumor is driven by Darwinian selection has never been rigorously tested.

- 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.

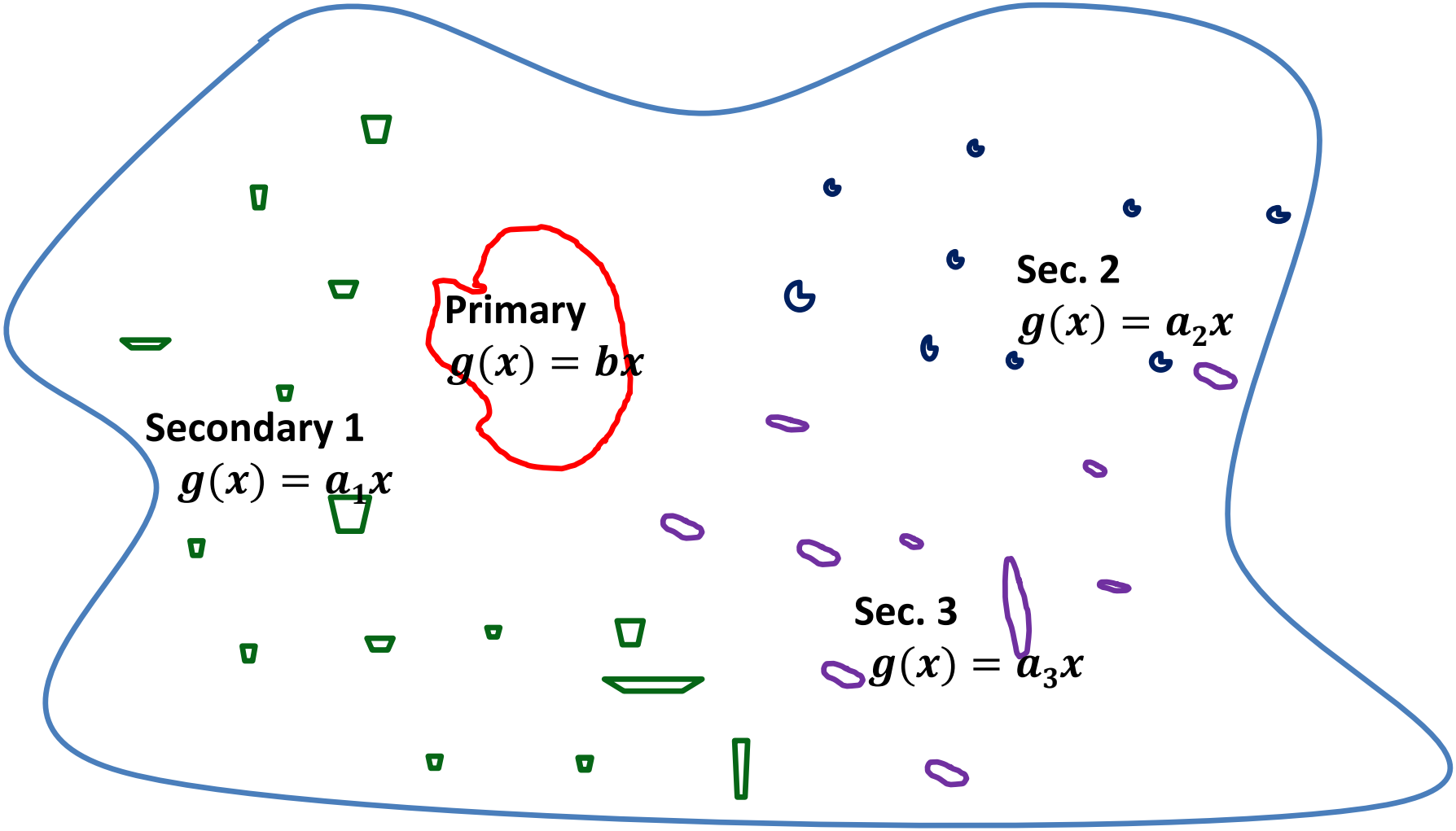
- From **a hepatocellular carcinoma (HCC)** tumor, **multiregional samples** from the tumor were evaluated, using either whole-exome sequencing **(WES)** (*n*=23 samples) or **genotyping** (*n*=286).

Map of the mutation clones of HCC-15.



Shaoping Ling et al. PNAS 2015;112:E6496-E6505

Tumor field model



Wide distribution of 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

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

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

$$\tilde{G}(1;b)$$

$$= \frac{m(e^{b\alpha t}-1)}{b\alpha} + (2-e^{b\alpha t})\frac{m^2\lambda}{b\alpha^2} e^{\lambda(m/\alpha - b)}\Gamma(0,\lambda(m/\alpha - b))$$

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

$$- e^{\lambda m/\alpha}\Gamma(0,\lambda m/\alpha)\}$$

where

$$\Gamma(z,w) = \int_w^\infty e^{-t}t^{z-1}dt$$

is the incomplete Gamma function, which behaves as

$$-\log(w), \text{ as } w \downarrow 0$$

Analytical solution (red) explodes in finite time

Simulated trajectories grow faster than exponential

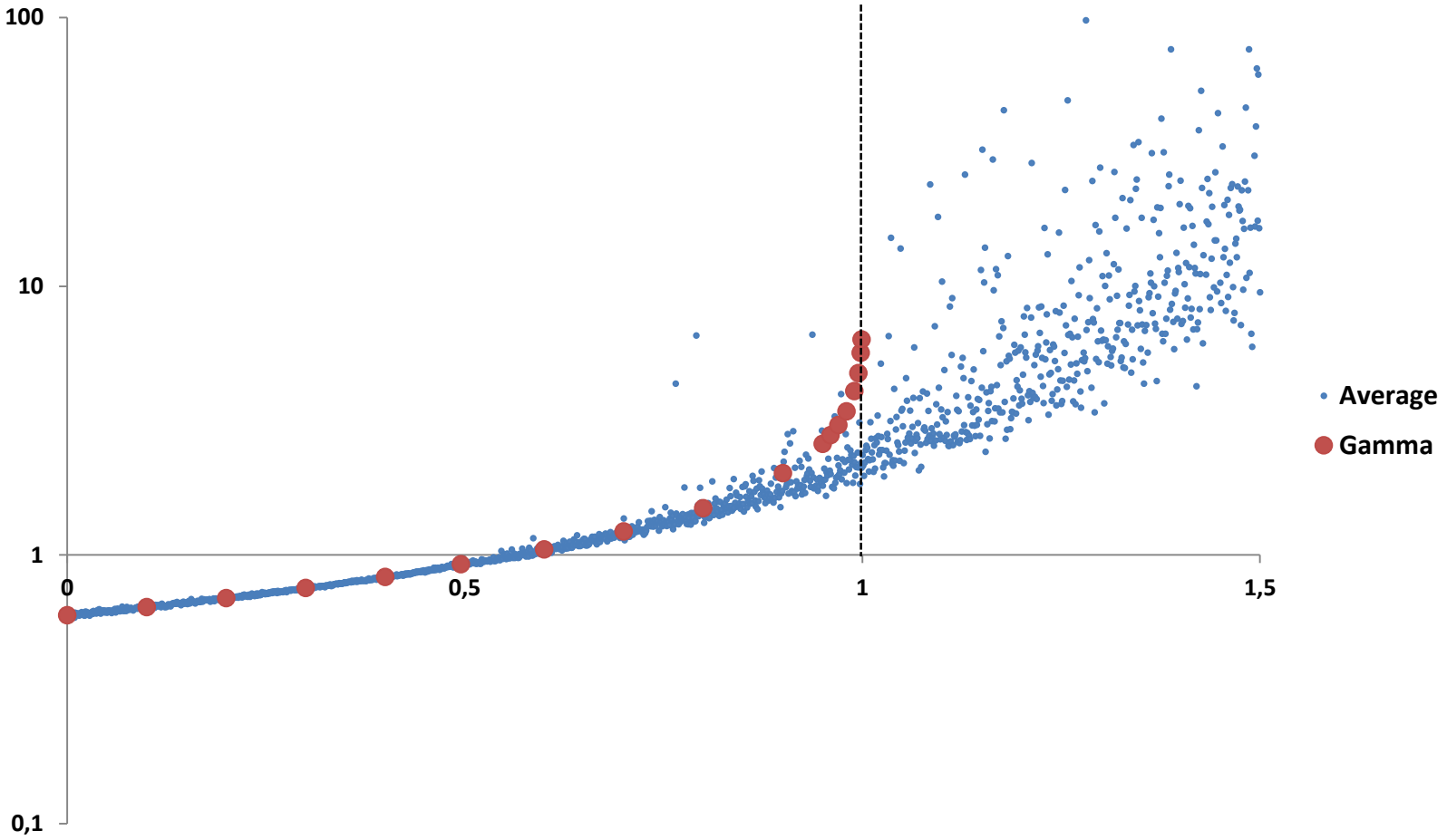
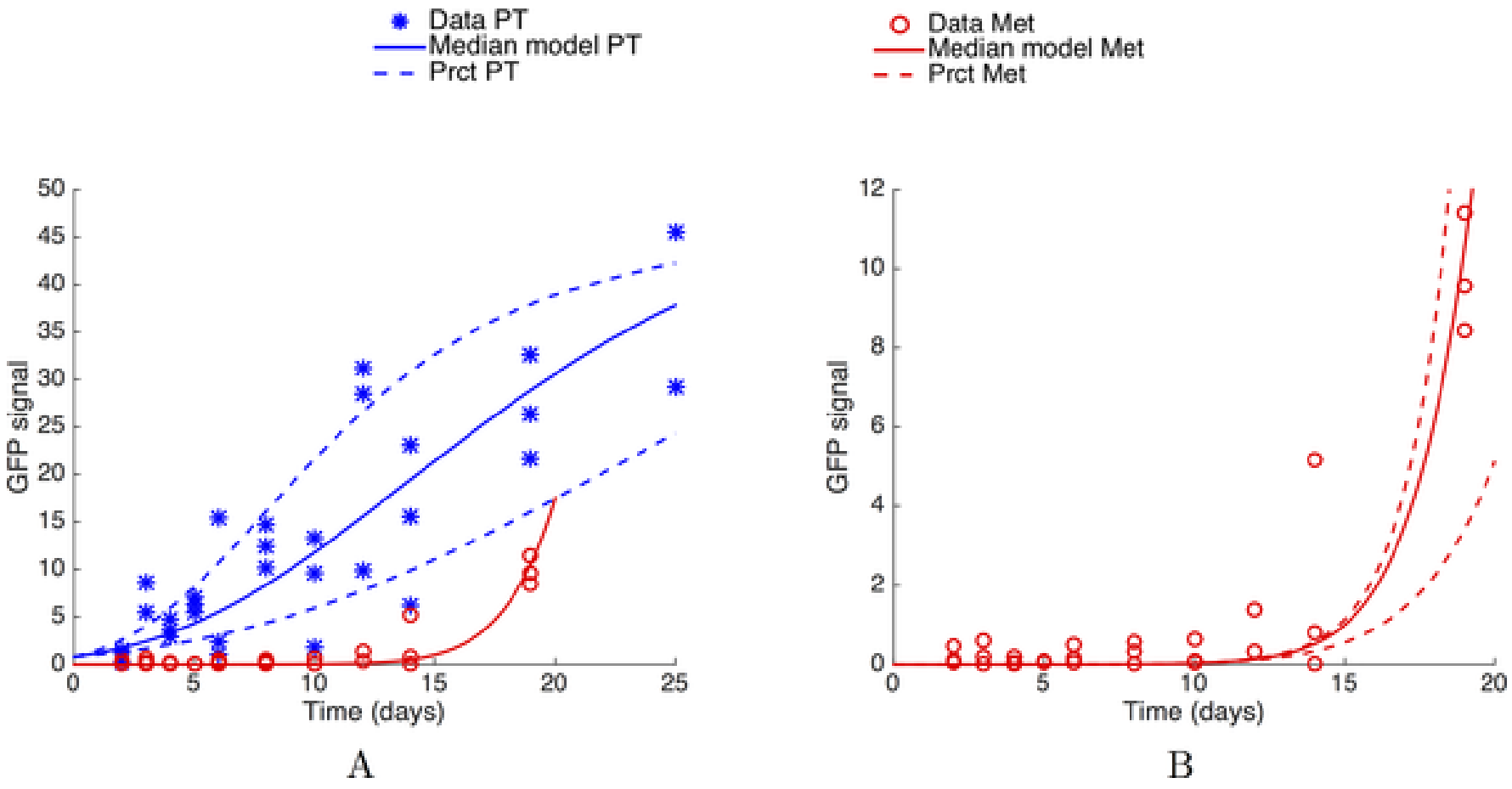


Fig 2. The standard theory: Primary tumour and metastatic burden dynamics fitting.



Baratchart E, Benzekry S, Bikfalvi A, Colin T, Cooley LS, et al. (2015) Computational Modelling of Metastasis Development in Renal Cell Carcinoma. PLOS Computational Biology 11(11): e1004626. https://doi.org/10.1371/journal.pcbi.1004626
http://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1004626

Some comments

- Ling et al. Model captures some essentials but it is only quasi-stochastic
- Does a truly stochastic model display the same expected-value behavior?
  - Let us try a toy model
  - We can then try to build a Ling et al. - like branching-process model

Toy model

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

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

It now has Pareto tail

$$\Pr[X(t) > x] = \begin{cases} 1, & 0 \leq x < 1, \\ x^{-\lambda/t}, & x \geq 1, \end{cases} \quad t \geq 0,$$

…. which can be integrated to get the expectation

$$E[X(t)] = \int_0^\infty \Pr[X(t) > x]dx = 1 + \int_1^\infty x^{-\lambda/t}\,dx = \begin{cases} \lambda(\lambda - t)^{-1}, & t < \lambda, \\ \infty, & t \geq \lambda, \end{cases}$$

Success ! But what about the ∞ ? Will tumors really **explode** ?!

Let us simulate growth with random rates

- Blue dots are averages of 1000 simulations with the same rate
- Red line is expected value

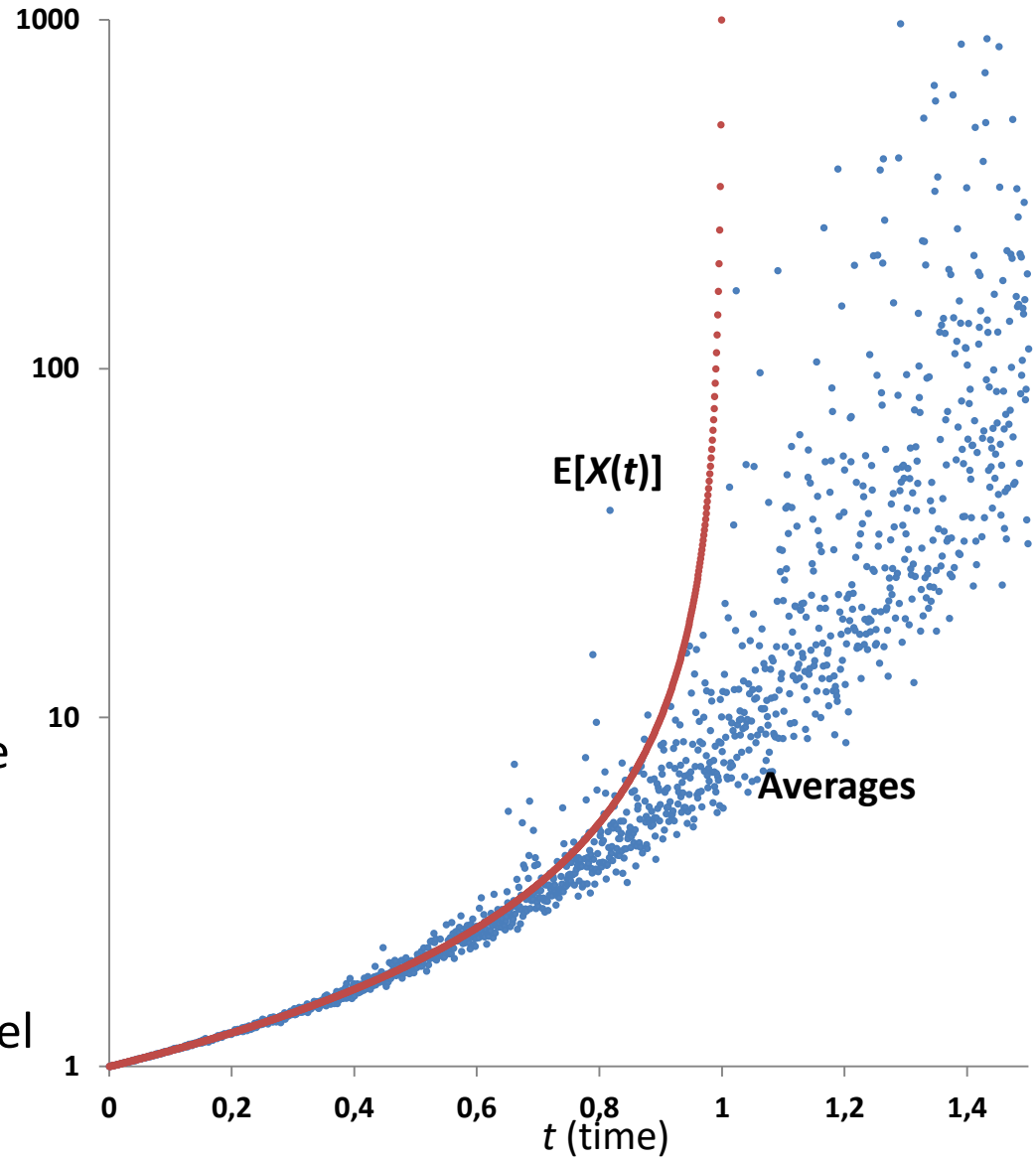
Averages grow faster than exponential, but do not explode.

Theoretical expectation does

Conclusion:

Expected value is not useful as the central tendency of growth after certain time.

Inference from deterministic model is inaccurate (*really ?*)



But we forgot the medians ….

By definition:

$$\Pr[X(t) > Median\{X(t)\}] = \frac{1}{2}$$

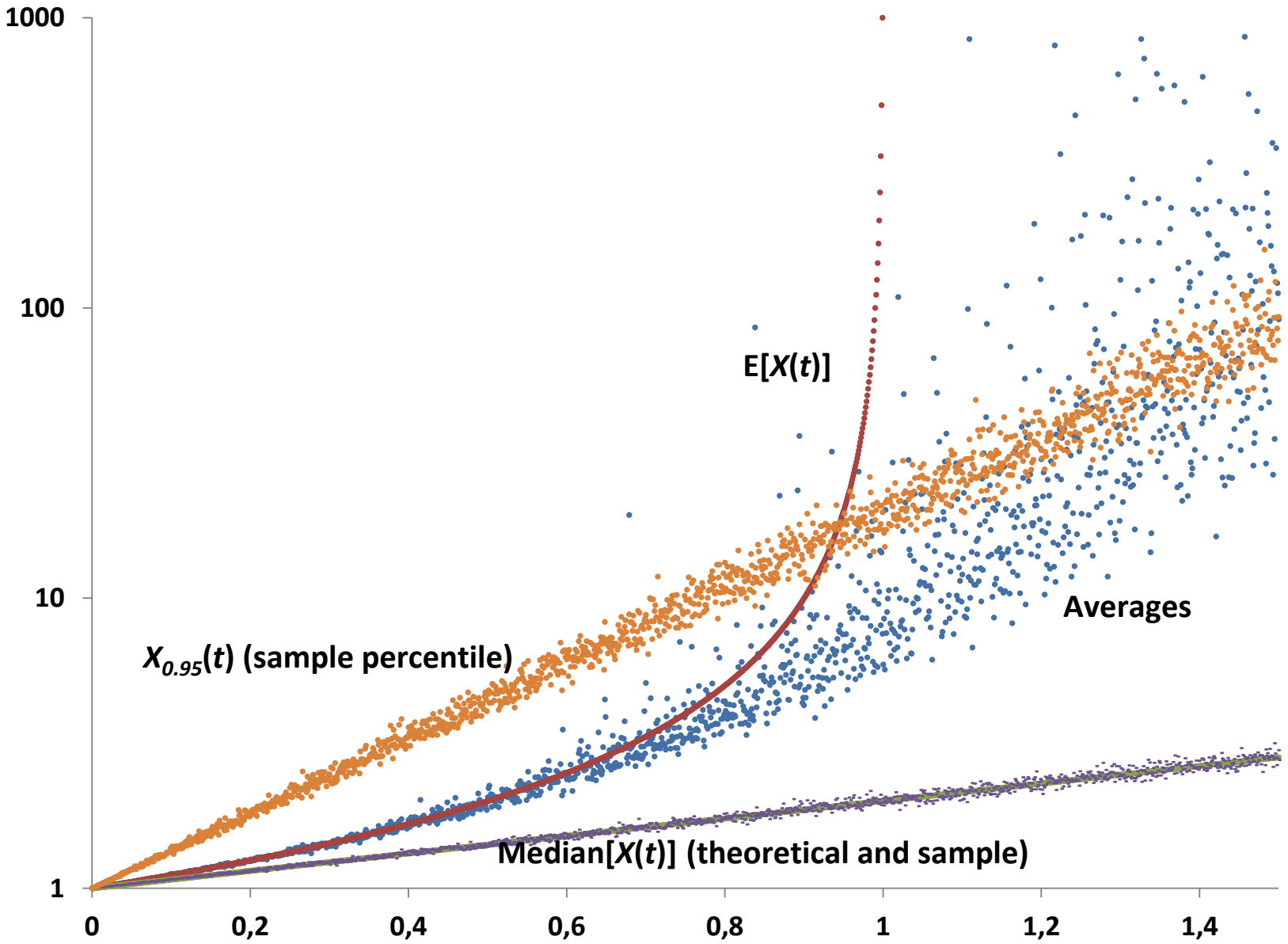
So

$$Median\{X(t)\} = 2^{t/\lambda}$$

and it is well-matched by its sample counterpart

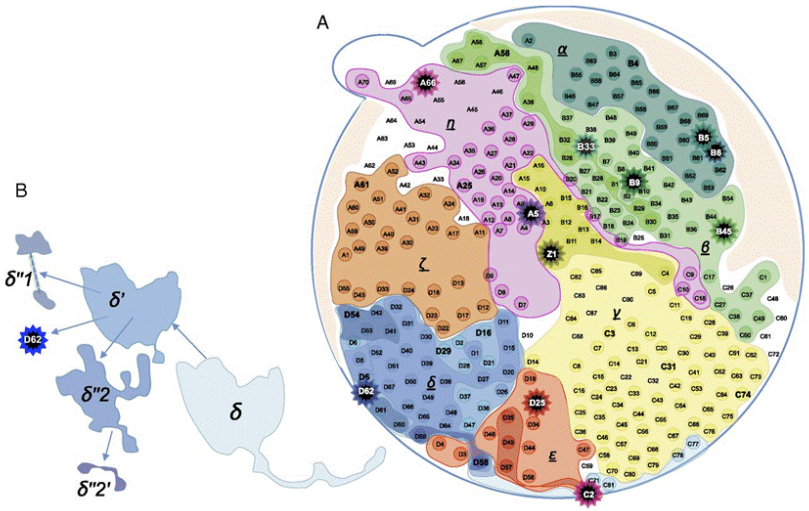
Median is a good model in this case

[Have you heard of ***anybody*** using medians to model?]



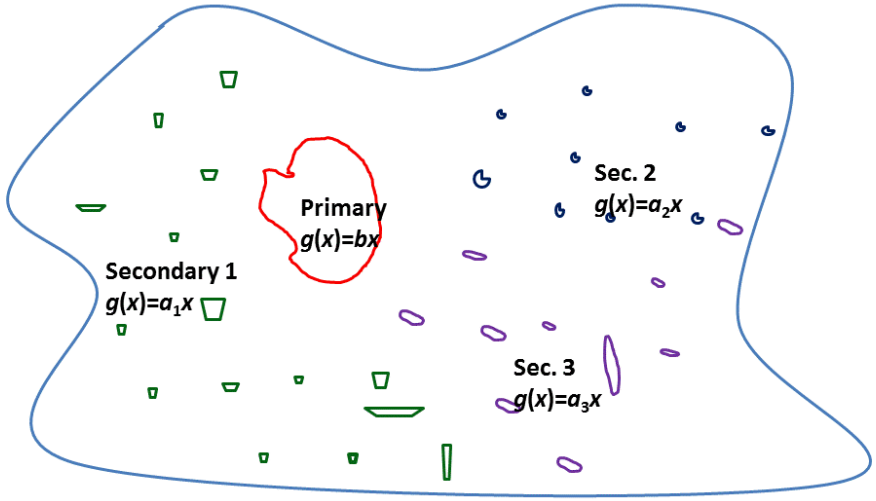
Can we reproduce such behavior in a truly stochastic model that is not "toy" but still simple?

Map of the mutation clones of HCC-15.



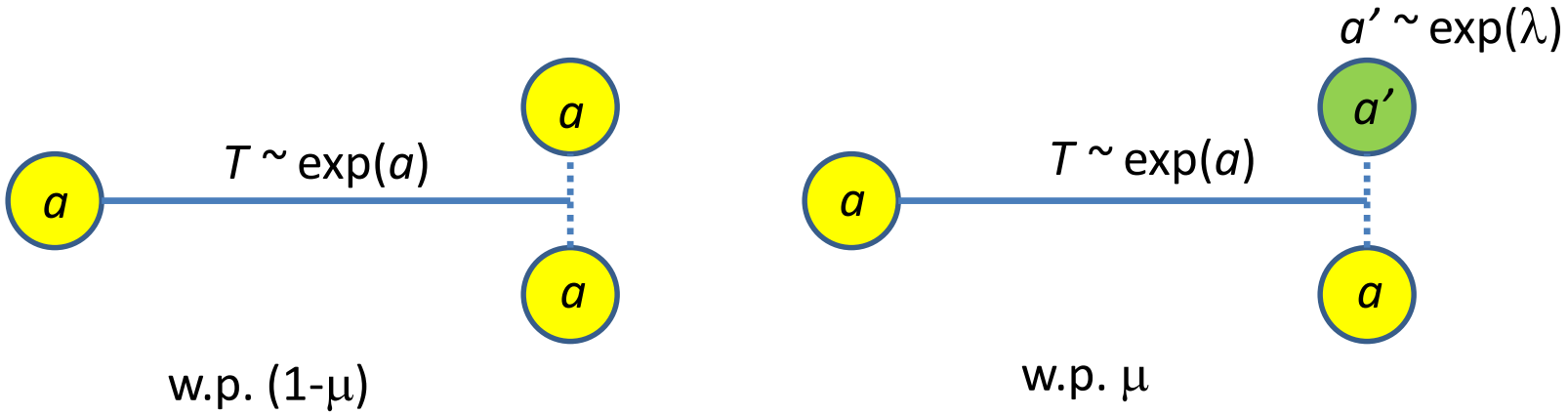
Shaoping Ling et al. PNAS 2015;112:E6496-E6505

Tumor field model



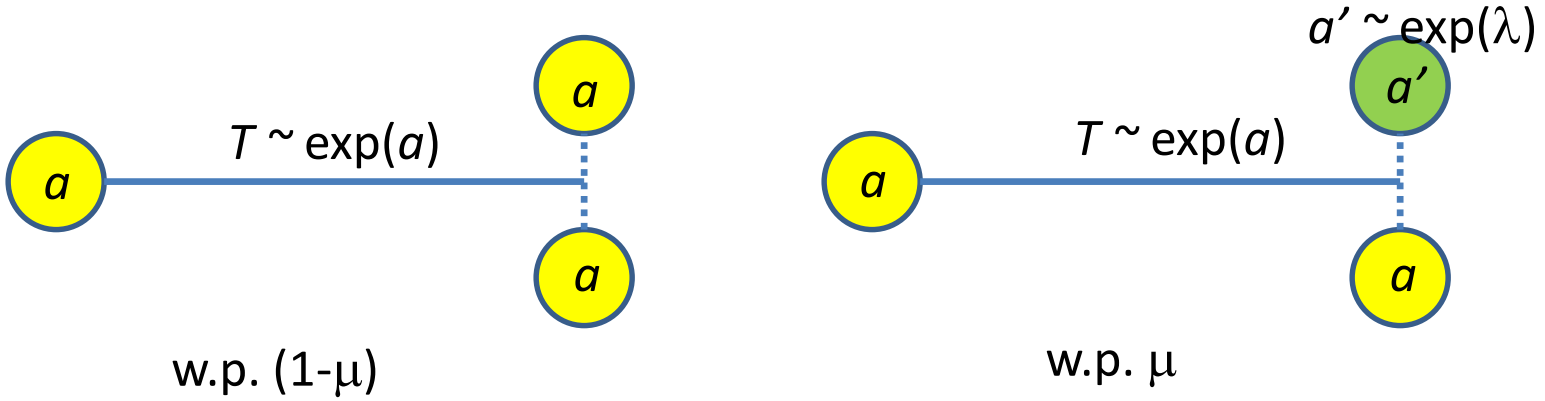
Wide distribution of growth rates a_i

Coldman-Goldie model with a twist



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

Coldman-Goldie model with a twist



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

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

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

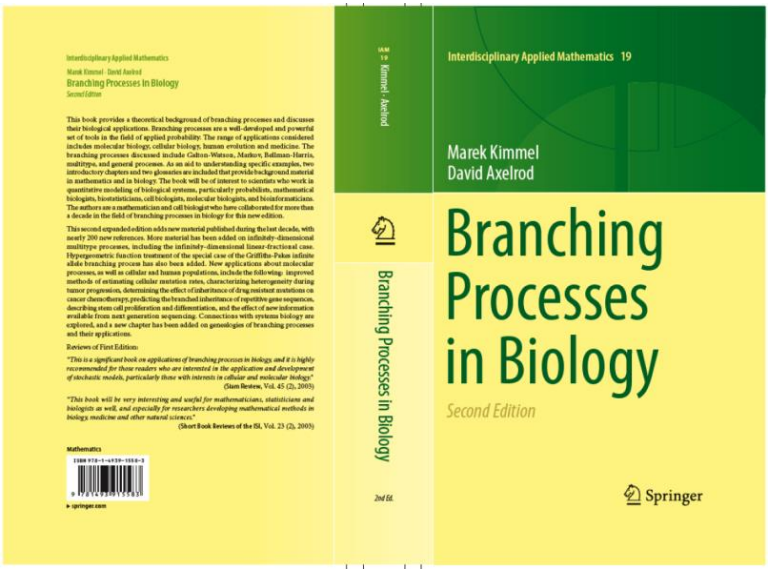
$$\Phi(s;\lambda,t) = \int_0^\infty F(s;a',t)\cdot\lambda\exp(-\lambda a')d\lambda$$

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

We can solve the Riccati equation!

Only to see that …

pgf $\Phi(s,t;\ \lambda)$ satisfies
a really hairy integral equation …



However …

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

Expected value equations …

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

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

have solutions expressed as series of convolution powers …
And they explode at finite times …

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

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

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

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

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

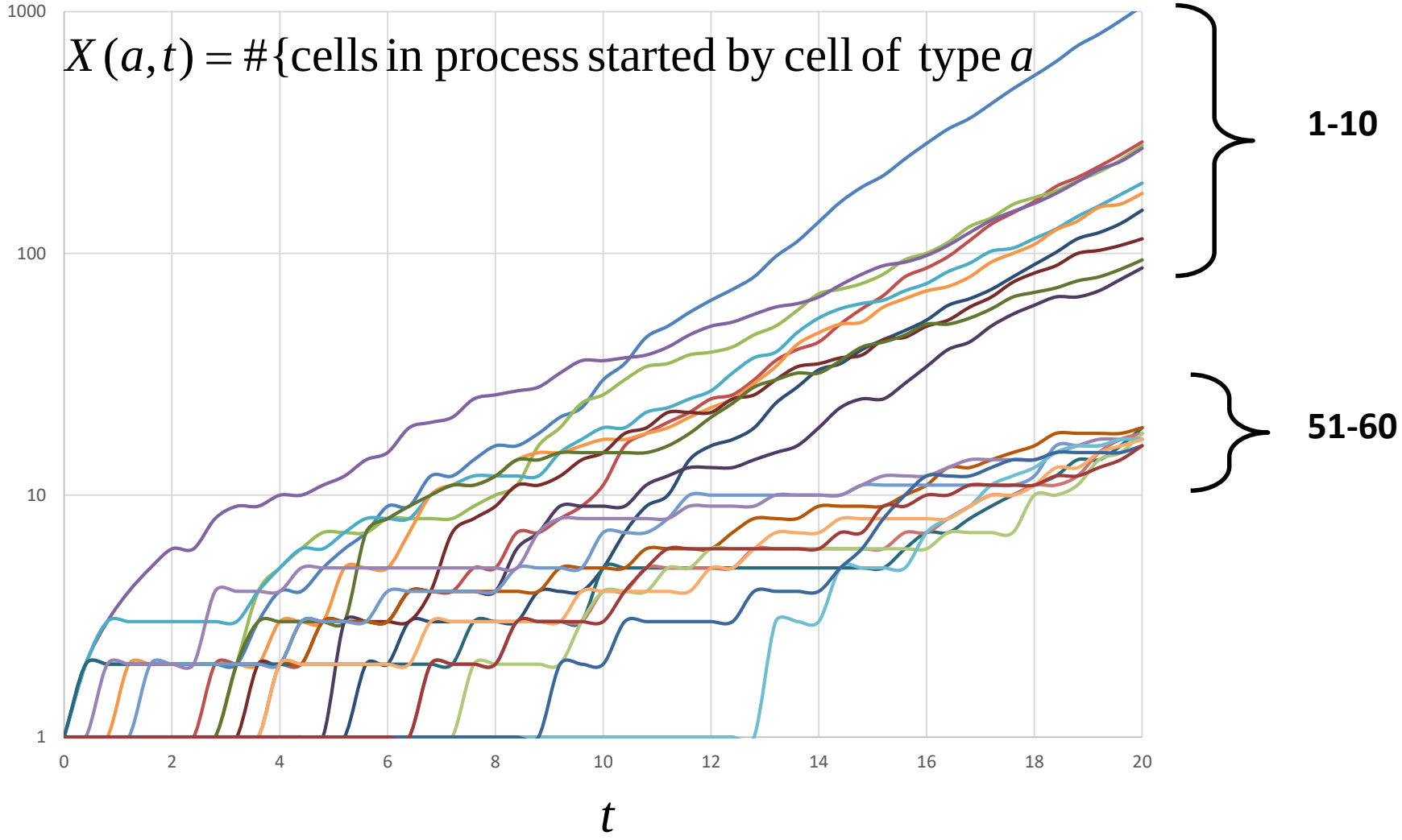
Also a scaling property …

$$\varphi(t, k\lambda) = \varphi(t/k, \lambda), t \in [0, k\lambda/(1-\mu))$$

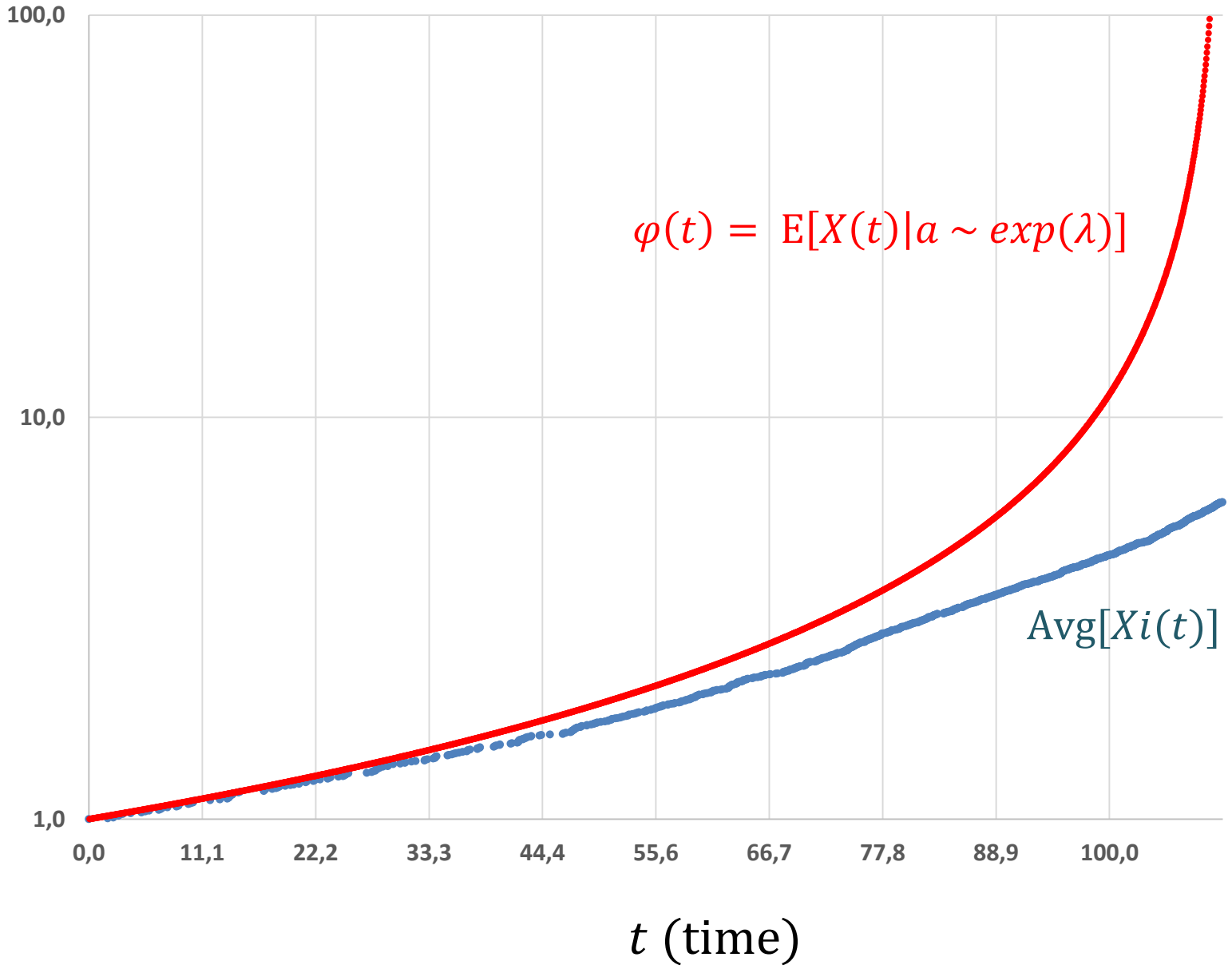
Variability of simulated stochastic trajectories

Trajectories ranked **1-10** and **51-60** out of $\mathbf{10^4}$ simulated

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



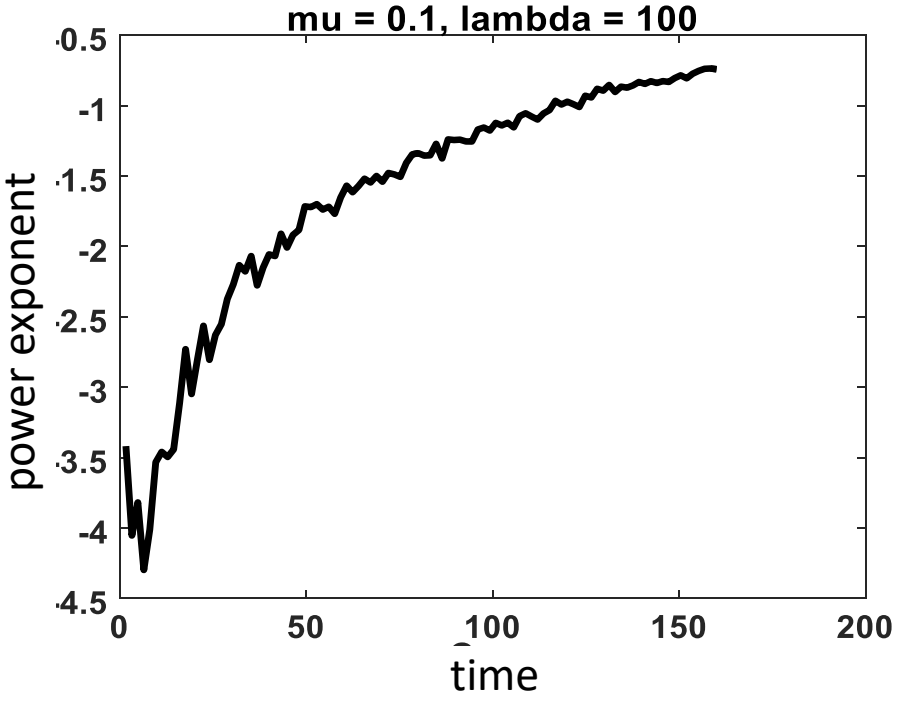
$\lambda = 100, \qquad \mu = 0.1$

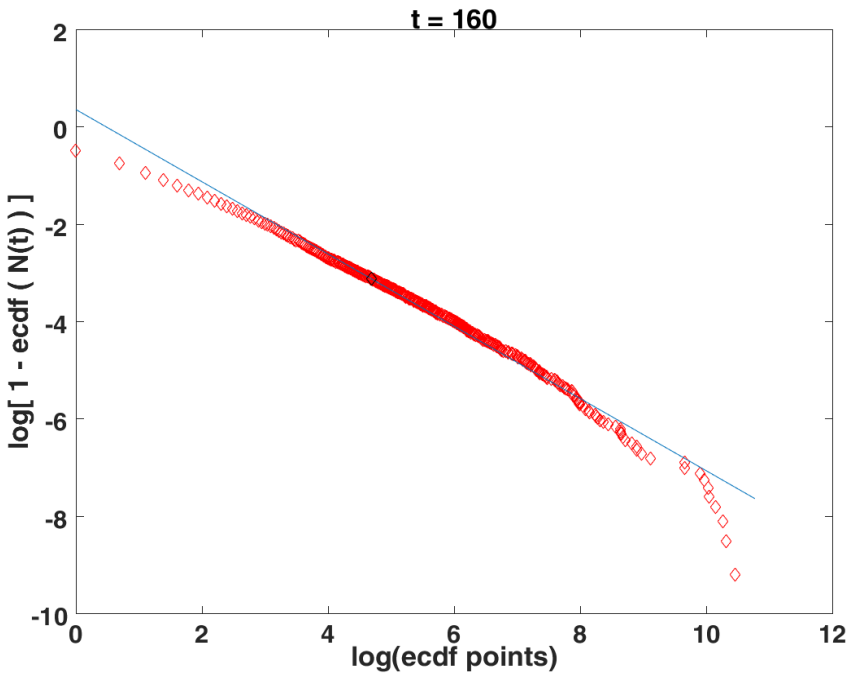


Tails of the distributions of cell counts $X(t,a)$

seem to obey a power law,

with exponent estimated close to -**1** at the time when the $E[X(t)]$ explodes





Some math …

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

$X_3(a,t)$ …

$X_4(a,t)$ …

…

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

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

As in the toy model and simulations.

Another partial result for $k = 2, 3, \ldots$

$$\mathrm{P}[X_k(a,t) > n] > Ca \left\{ \frac{\mu(1-\mu)^2 t^2}{\lambda^2 \ln n} \right\}^{k-1} n^{-\nu}, \qquad n \to \infty$$

suggests that

- The power law exponent is the same for all tumor cells
- But, the growth rate is penalized by $(\ln n)^{1-k}$
- Eventually, for large times, $X_k(a,t)$ with large k will eventually dominate

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 $d = 0.5$?

Conclusions

- Mechanisms of growth and mutation of malignant cell clones in tumors are gradually uncovered due to technological revolution in biology
- Secondary tumors display highly variable growth rates
- Understanding differences between stochastic processes and their central tendency characteristics helps understand new data.
- Truncating the distribution of division rates eliminates literal explosions but tails can be quite long.

2018 q-bio Summer School

2018 q-bio Summer School

June 11-25, 2018
Rice Unversity | University of New Mexico
q-bio.org/wp/qbss/

Overview

The q-bio Summer School is an annual event intended to advance predictive modeling of cellular regulatory systems by exposing participants to a survey of work in quantitative biology and by providing in-depth instruction in selected techniques, with an emphasis on techniques useful for modeling cellular regulatory networks, although data analysis techniques and experimental methods will also be covered.

Students will each work on mentored projects and will have opportunities to present a talk and a poster about their work and receive feedback on their presentation skills. Participants will attend daily core lectures, chalk talks, journal clubs, expert panel discussions, and hands-on computer labs.

The summer school is designed for graduate students, postdocs, or anyone with a quantitative background who is new to modeling cellular regulatory systems/networks. Students attending the school are strongly encouraged to also attend the affiliated q-bio Conference, which will take place at Rice University from June 26 to 29.

Separate registration is required for the school and conference.

Locations

- Gulf Coast Campus, Rice University, Houston, Texas
- Southwest Campus, University of New Mexico, Albuquerque, New Mexico

Courses

Gulf Coast Campus, Rice University
- Stochastic Cell Regulation
- Cancer Dynamics

Southwest Campus, University of New Mexico
- Cell Signaling and Rule-Based Modeling
- Membrane Biology and Quantitative Microscopy

Important Dates

- Application Deadline: March 30, 2018

Cost of Attendance

- $500 for tuition
- $1500 for lodging and meals

Financial assistance is available to help defer tuition and lodging costs. Please refer to the application website for more information.

This event is brought to you by the Center for Nonlinear Studies at Los Alamos National Laboratory. This event has been declared "Open to the Public" with Registration".

Course Leaders

Marek Kimmel
(Rice University)

Tomasz Lipniacki
(Polish Academy of Science)

William S. Hlavacek
(Los Alamos National Laboratory)

Mara P. Steinkamp
(University of New Mexico School of Medicine)

Contact Information

For inquiries about the summer school, please contact: **cnls-conferences@lanl.gov**

For more information, please visit the school website at: **q-bio.org/wp/qbss/**

THE UNIVERSITY of NEW MEXICO

RICE
Rice University Creative Ventures Funds

Colorado State University

NSF

STMC
SpatioTemporal Modeling Center

New Mexico CONSORTIUM

Los Alamos NATIONAL LABORATORY
EST 1943

Center for Nonlinear Studies

What has happenned

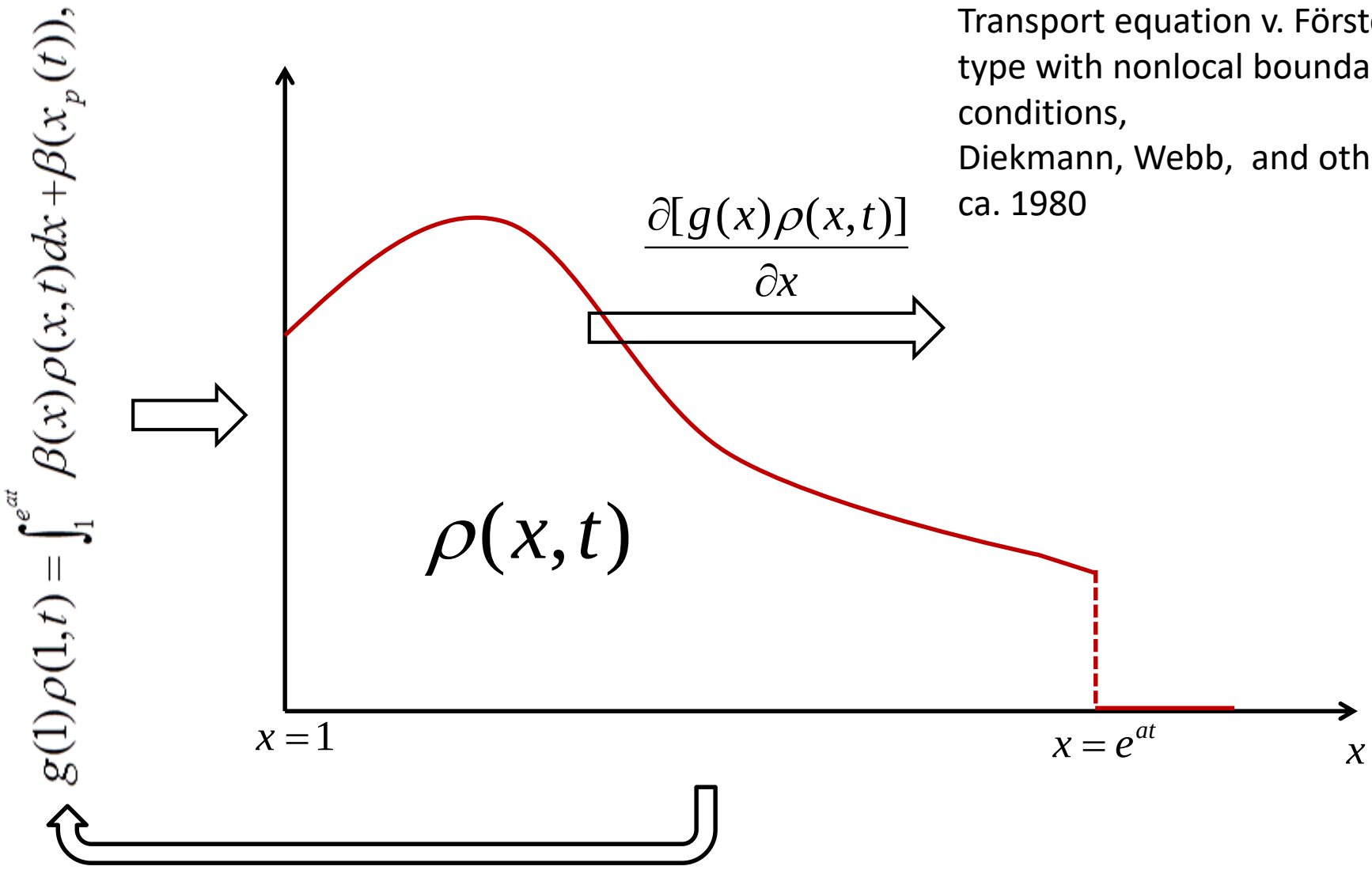
- Ling et al (2015) is a comprehensive treatment of several aspects of secondary tumors, using for example Durrett (2013) generalization of the Ewens sampling formula for growing populations, based among other on Polanski and Kimmel (2003).

- They also formulated a mathematical model of the primary tumor shedding secondary foci and used it to conclude about the distribution of the sizes of secondary foci in HCC.

- We solved the model equations based on the methods reported by Iwata (2000) and showed that the rigorous solution is different from Ling et al. (2015) (and later Tao et al. (2015)) intuitive solution.

Time evolution of the size distribution of secondary tumors

Transport equation v. Förster type with nonlocal boundary conditions,
Diekmann, Webb, and others ca. 1980



Equations of the model

Dynamics of the secondary cell colony size distribution density is described by the following von Förster-type equation:

$$\frac{\partial \rho(x,t)}{\partial t} + \frac{\partial [g(x)\rho(x,t)]}{\partial x} = 0, \quad (x,t) \in [1,\infty) \times [0,\infty)$$

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

with nonlocal boundary conditions

$$g(1)\rho(1,t) = \int_1^{e^{at}} \beta(x)\rho(x,t)dx + \beta(x_p(t)),\ \beta(x) = mx^{\alpha}$$

Solution in the "selective randomized" case

- In this case
  - the primary tumor grows at rate bx^p
  - each new secondary originating from the primary grows at a different rate, sampled from the exponential distribution with parameter λ
  - Secondaries originating from other secondaries grow at unchanged rates

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

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

- Substitution of the expression for $G(x;a,b)$ leads to intractable integrals …

Assumptions a bit unrealistic?

So … Try a model with expected lifetimes of mutants ≥ 1/*A*

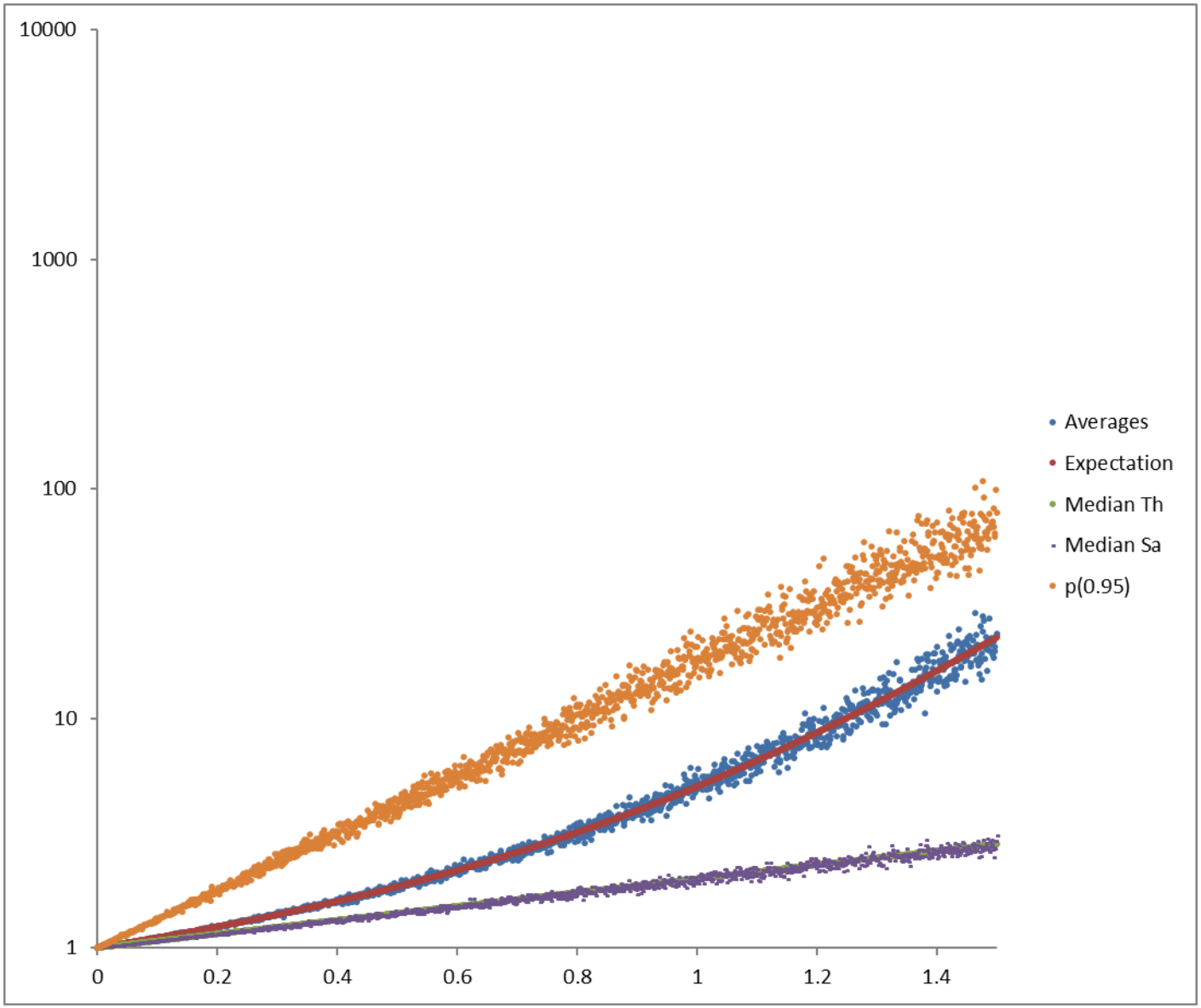
$$X(t|a) = \exp(at),\ t \geq 0,\ a \sim \exp(\lambda; A)$$

$$\Pr[X(t) > x] = \frac{x^{-\lambda/t} - e^{-\lambda A}}{1 - e^{-\lambda A}}\ x \in [1, e^{At}]$$

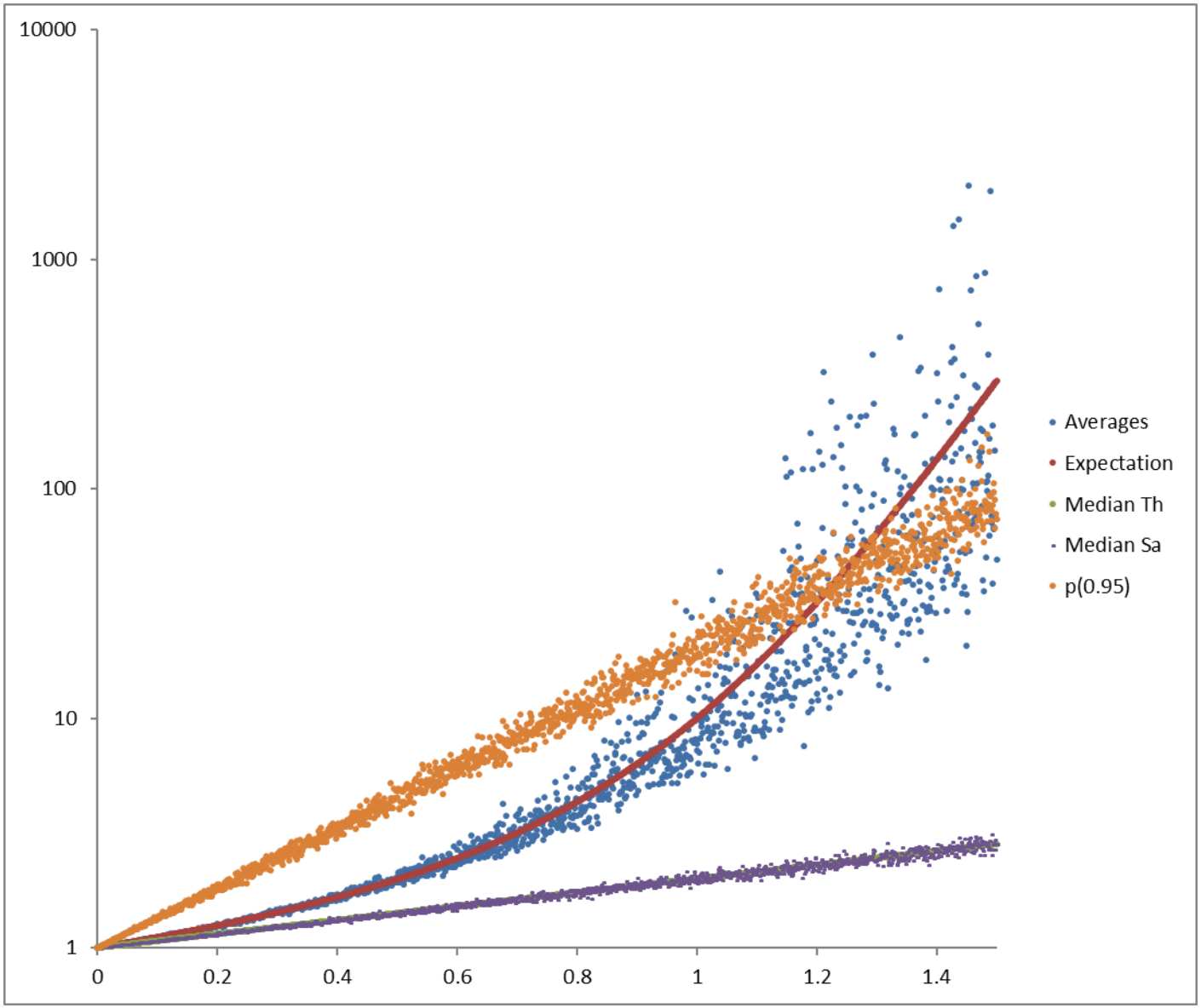
$$E[X(t)] = \int_0^\infty \Pr[X(t) > x]dx = 1 + (1 - e^{-\lambda A})^{-1}\left[\frac{e^{A(t-\lambda)}}{1 - \lambda/t} - e^{-\lambda A}(e^{At} - 1)\right]$$

How do expectations, averages and quantiles depend on *A*?

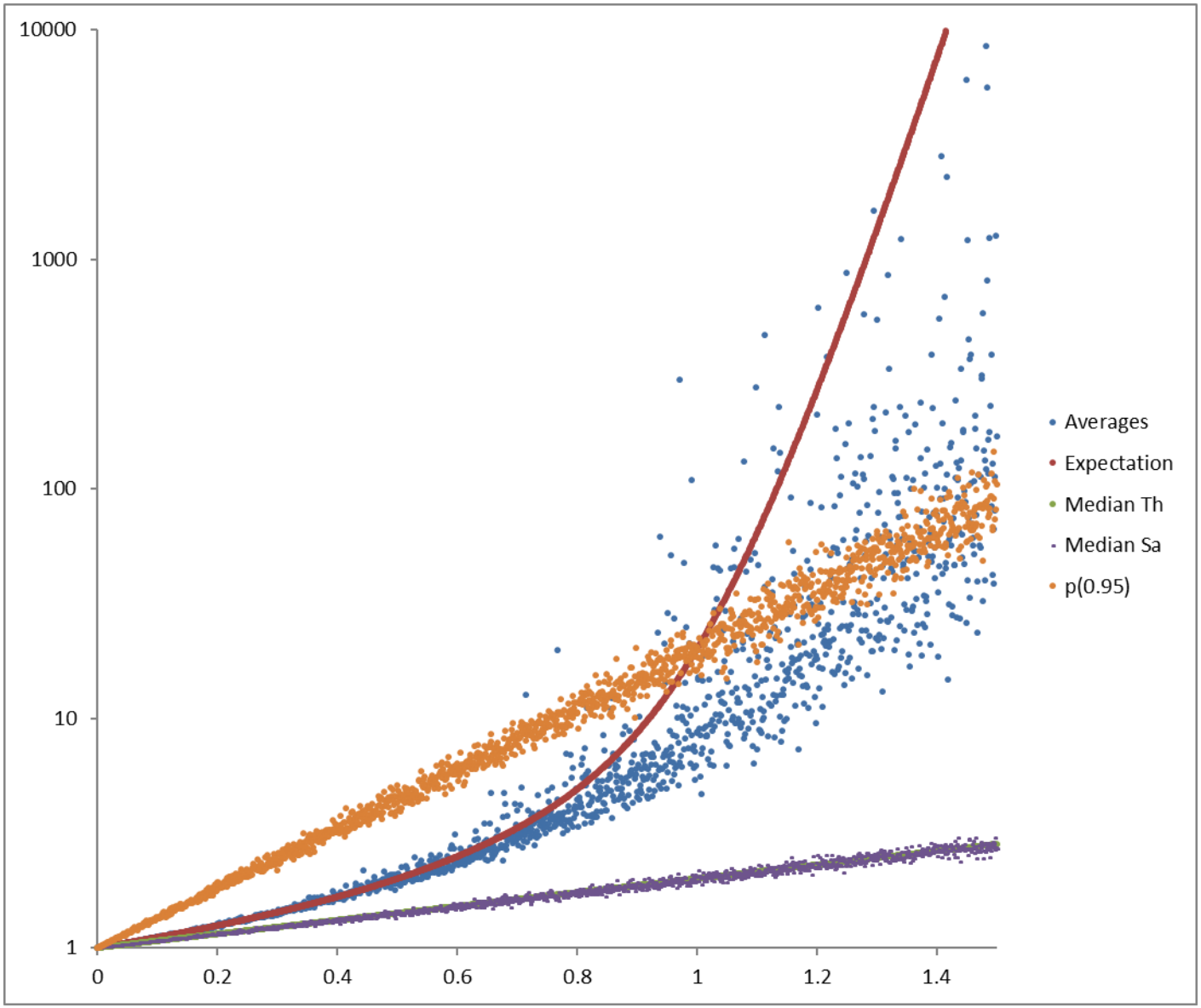
$A = 5$


10000
1000
100
10
1
0
0.2
0.4
0.6
0.8
1
1.2
1.4
Averages
Expectation
Median Th
Median Sa
p(0.95)

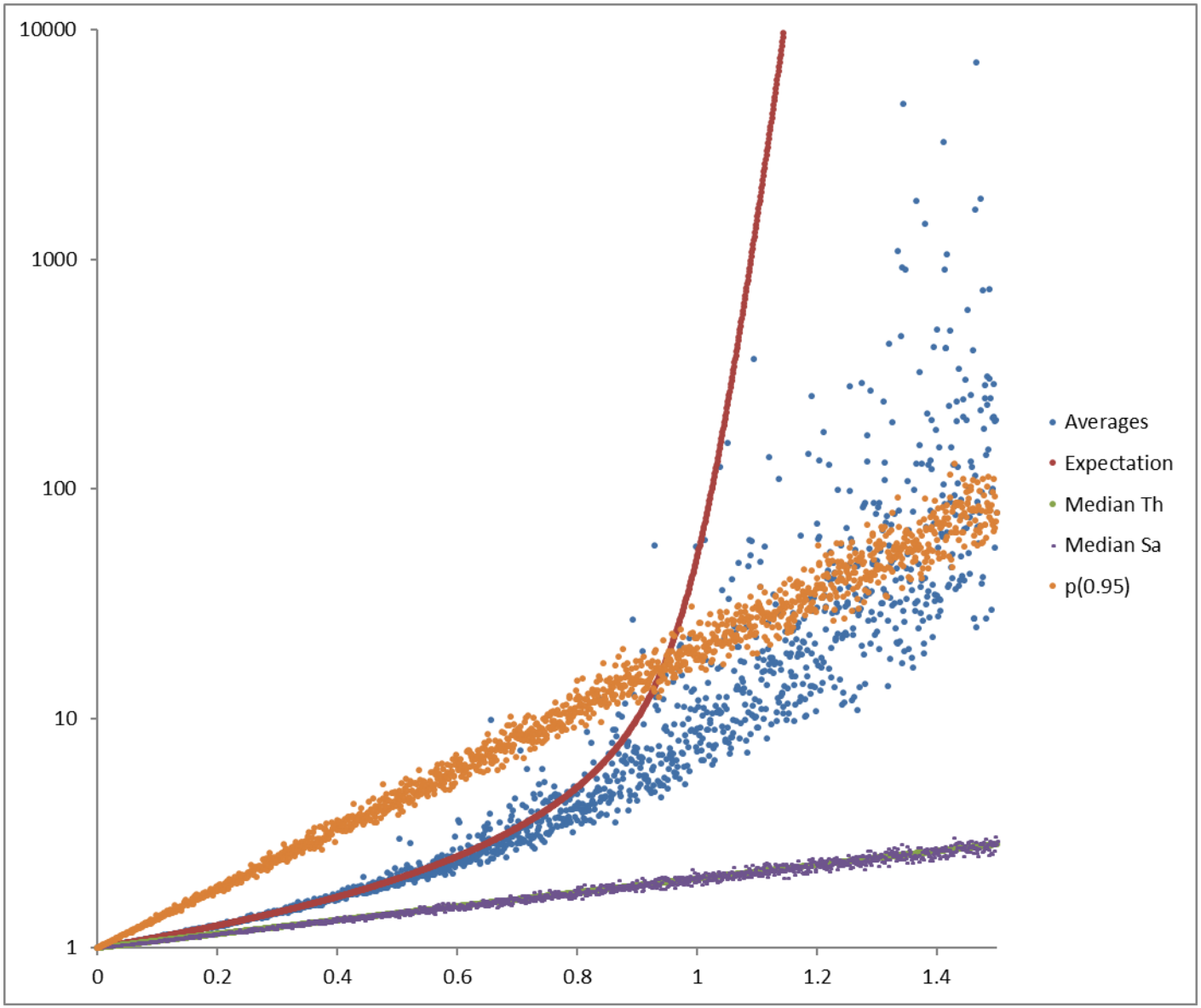
$A = 10$


10000
1000
100
10
1
0
0.2
0.4
0.6
0.8
1
1.2
1.4
Averages
Expectation
Median Th
Median Sa
p(0.95)

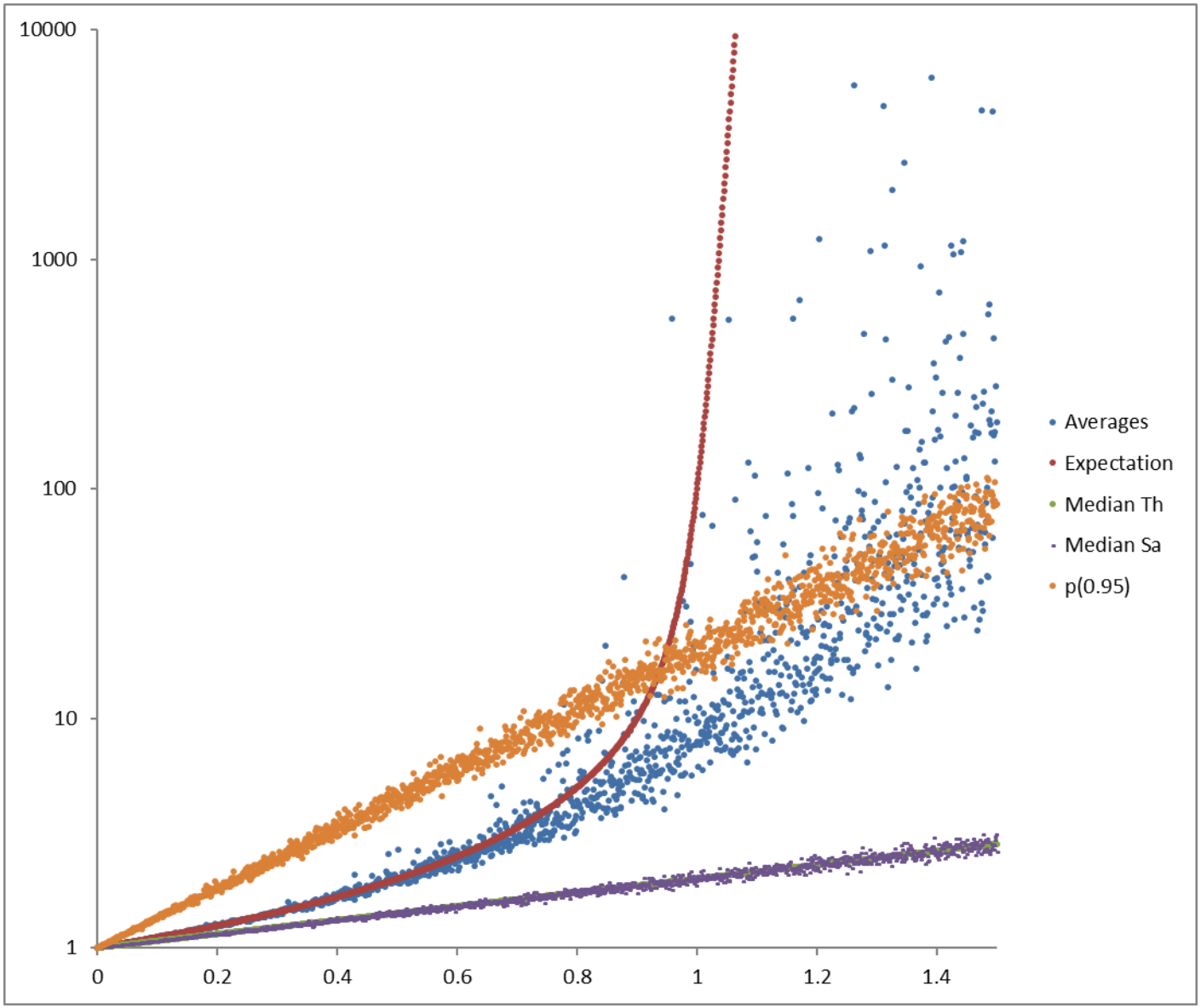
$A = 20$


10000
1000
100
10
1
0
0.2
0.4
0.6
0.8
1
1.2
1.4
Averages
Expectation
Median Th
Median Sa
p(0.95)

$A = 50$


10000
1000
100
10
1
0
0.2
0.4
0.6
0.8
1
1.2
1.4
Averages
Expectation
Median Th
Median Sa
p(0.95)

$A = 100$


10000
1000
100
10
1
0
0.2
0.4
0.6
0.8
1
1.2
1.4
Averages
Expectation
Median Th
Median Sa
p(0.95)