Stochastic modelling in viral and immunological systems

Mario Castro

School of Engineering, Comillas Pontifical University
marioc@comillas.edu

Albuquerque, 16th of July 2015

Outline

- General ideas and methods
  - Why stochastic?
  - When stochastic?
  - Tools (review)
    - Gillespie
    - Chemical Master Equation
    - Van Kampen's approximation
- Case studies
  - Virology: to extinct or not to extinct
  - Immunology: how to count molecular events
- Hands-on lab session:
  - Tinker Cell (14:15-16:00 Computer Lab 8 Student Union Building, Room 3018)

General ideas and methods

Handling ignorance

	Continuum	**Single molecules**
Well mixed	ODEs (rate equations)	Stochastic methods (Gillespie, VanKampen)
Spatial gradients	Reaction diffusion	Random walks

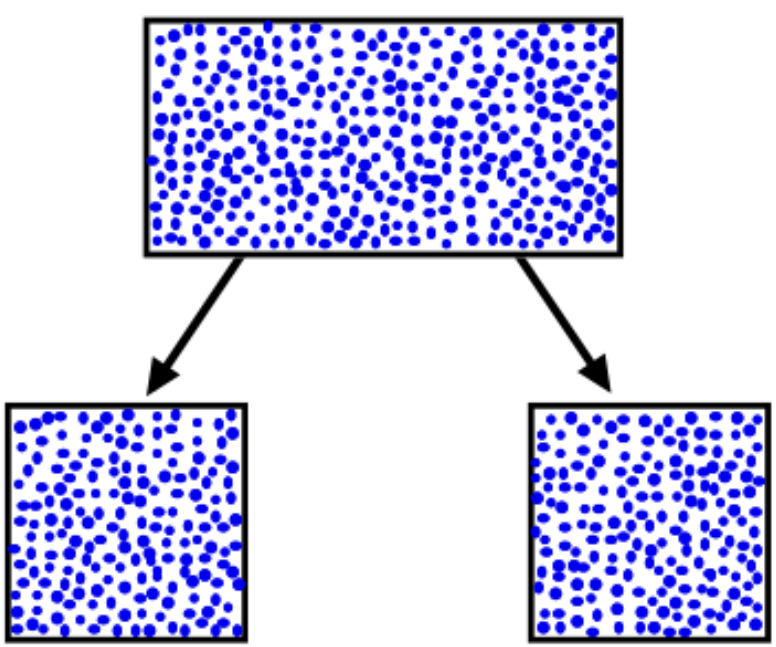
Why: low populations

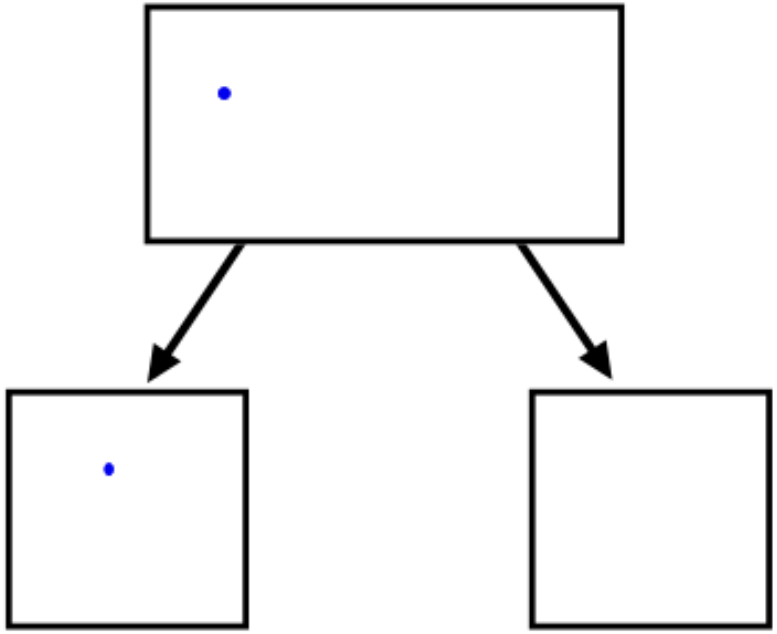
- ODEs are great when we can define a density:

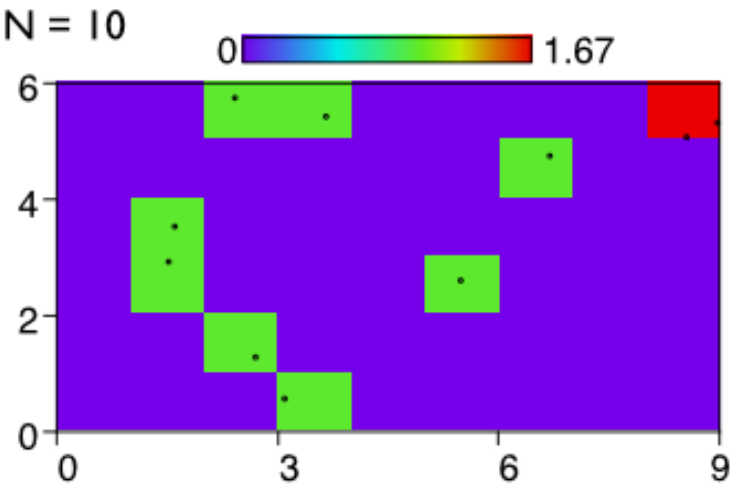
$$\text{density} = \frac{\text{Indistinguishable particles}}{\text{Volume}}$$

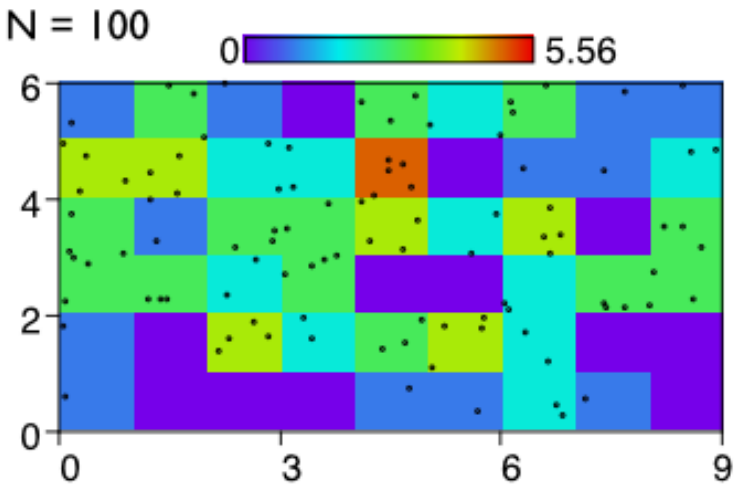
- But, what if half the volume does not contain half the particles?

Why: low populations







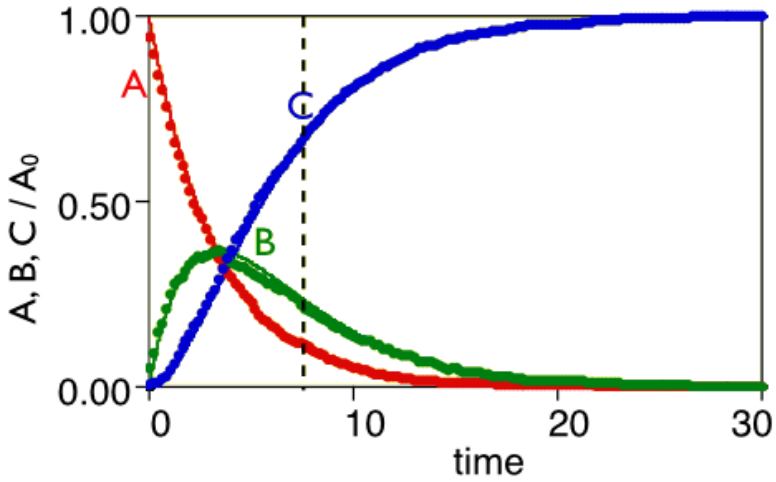


Why: low populations⇒large fluctuations

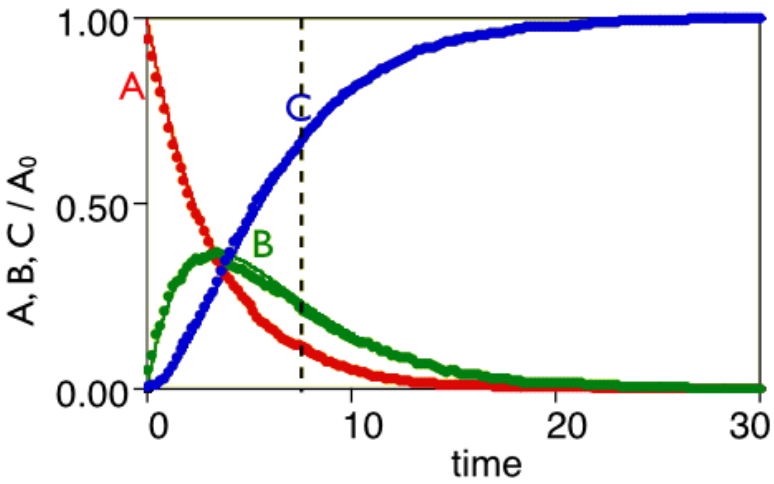
- Example:

$$A \rightarrow B \rightarrow C$$

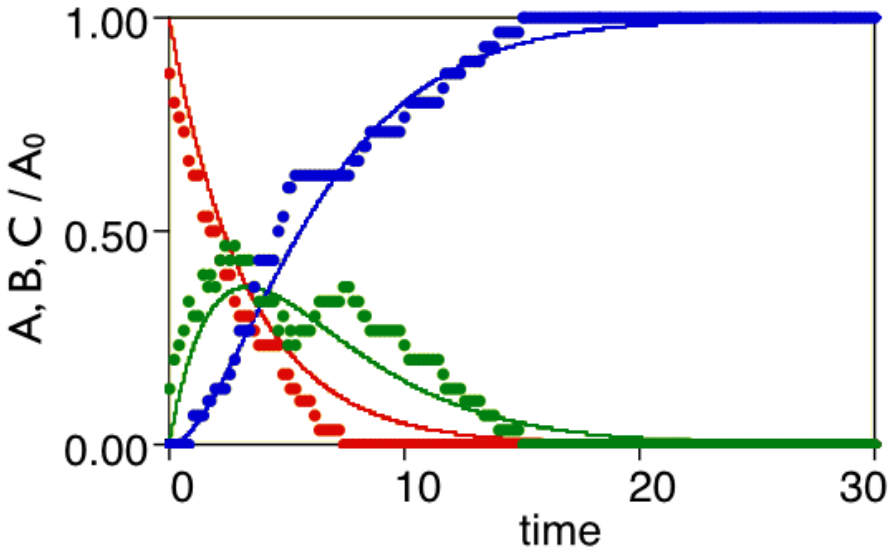
$$\begin{aligned}
\frac{dA}{dt} &= -k_1 A, \\
\frac{dB}{dt} &= k_1 A - k_2 B, \\
\frac{dC}{dt} &= k_2 B,
\end{aligned}$$



Why: low populations⇒large fluctuations

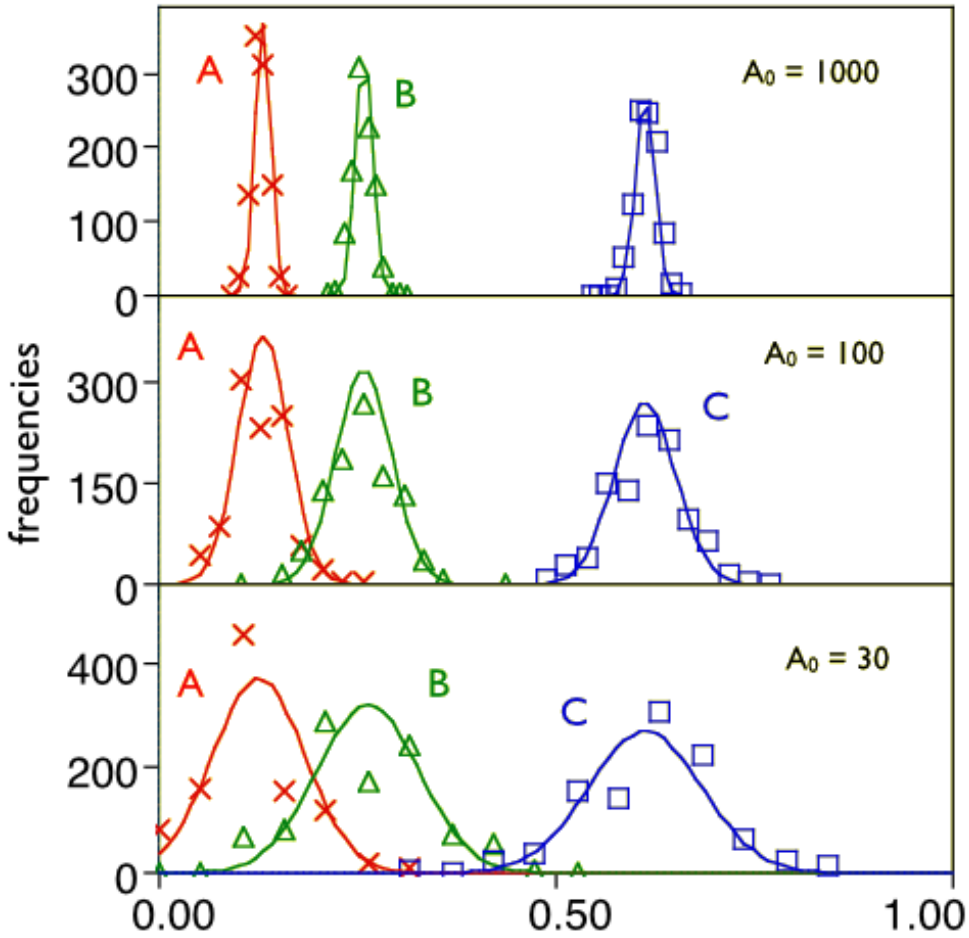


(1000 particles)



(30 particles)

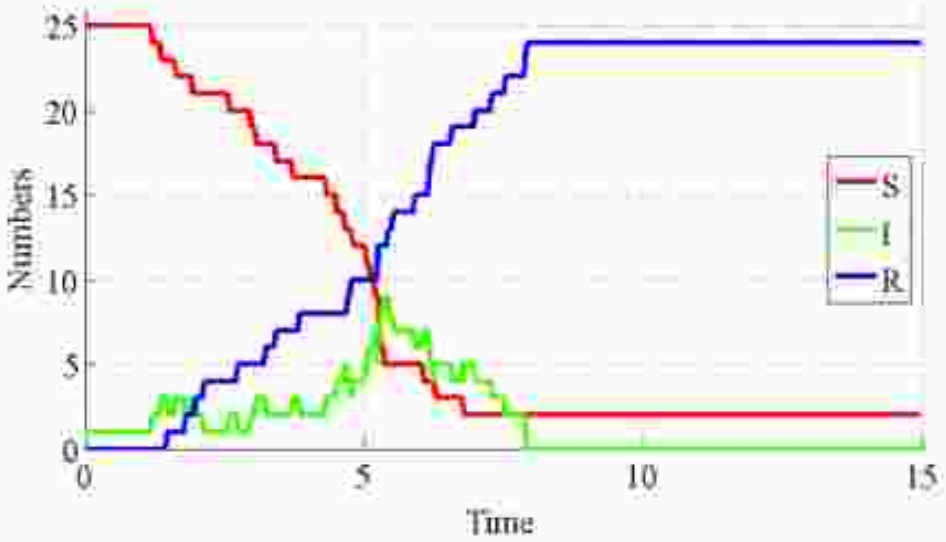
Why: low populations⇒large fluctuations



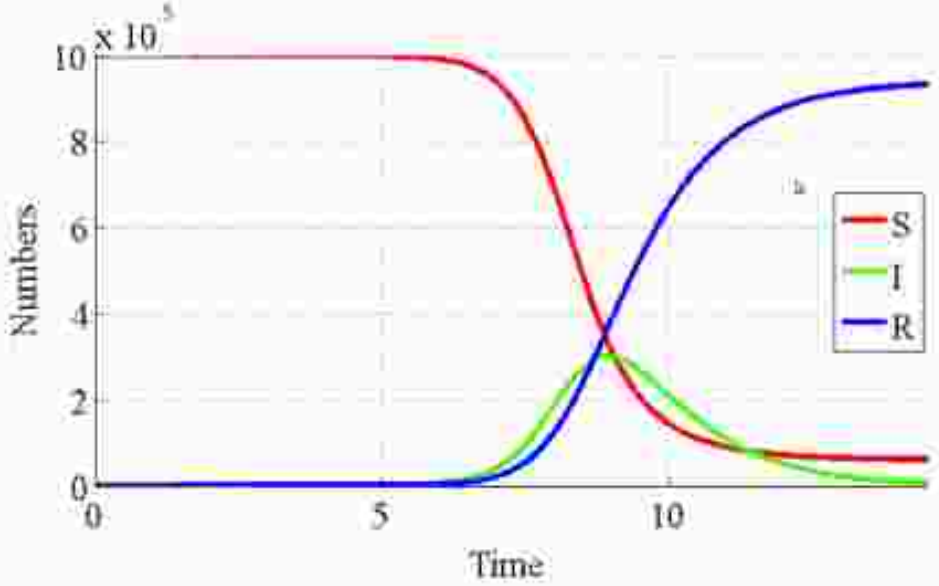
Keep this figure in mind: Almost a gaussian with $\sigma \sim N^{-1/2}$.

More examples: Deterministic or stochastic?

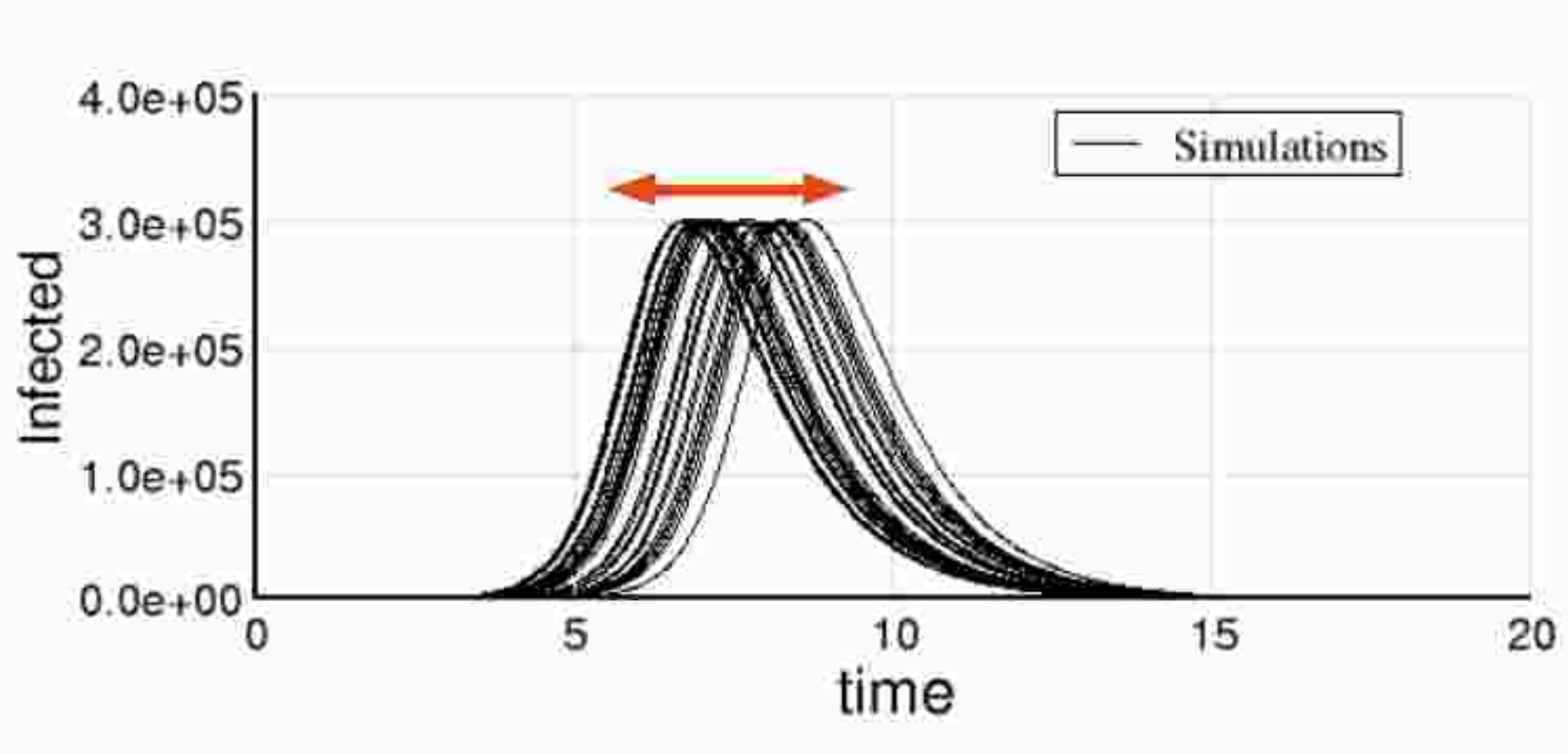
$S(0) = 25, I(0) = 1, R(0) = 0$



$S(0) = 10^6, I(0) = 1, R(0) = 0$

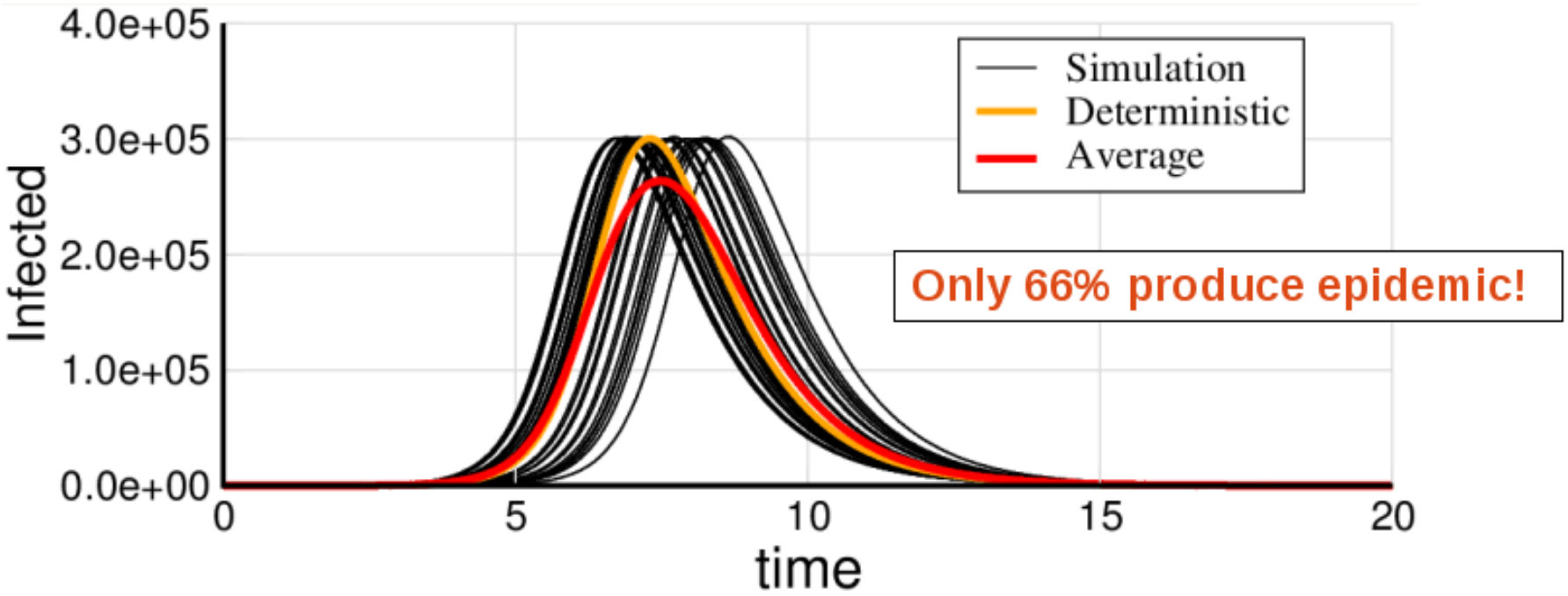


More examples: Deterministic or stochastic?

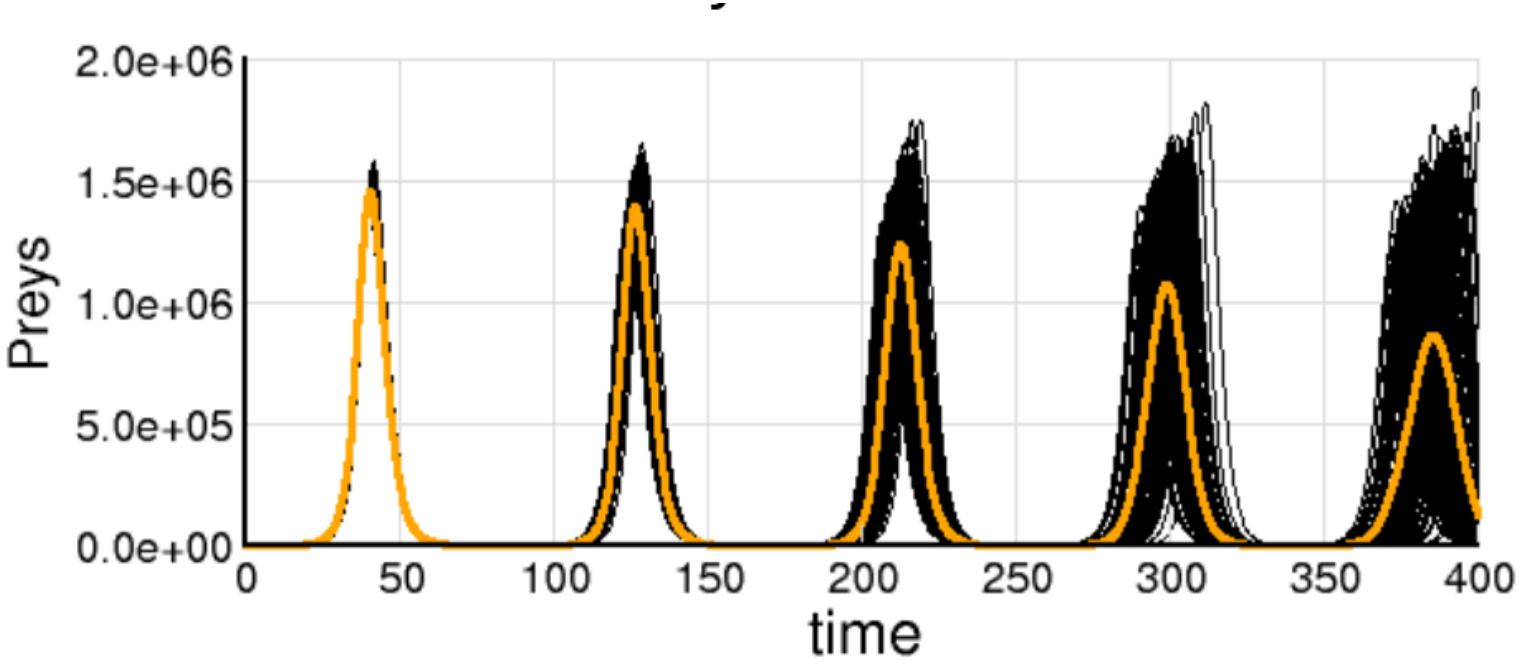


One infected in a population if **a million** susceptibles . . .

Beware simple averaging...



Beware simple averaging...



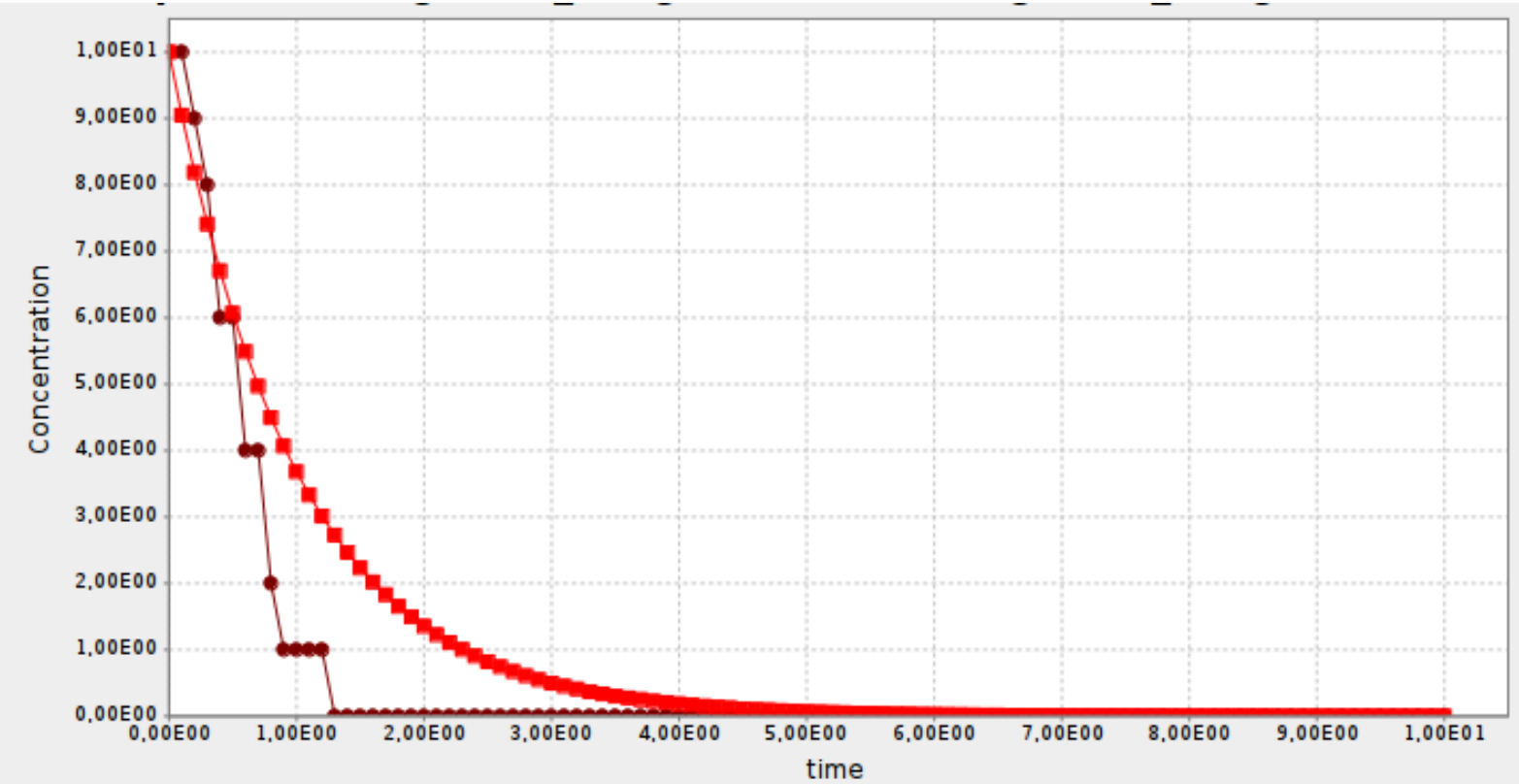
Averaging sucks!

When should we think *stochastically*?

- When can we **safely** use ODEs?
  - If (population) numbers are large
  - Far from extinction
  - Far from a *bifurcation*
- When **MUST** we use stochastic methods
  - Close to extinction or bifurcation points
  - If we are interested on **individuals**
  - Multiple stationary states

Life close to extinction

- A very simple example: $A -> \emptyset$.

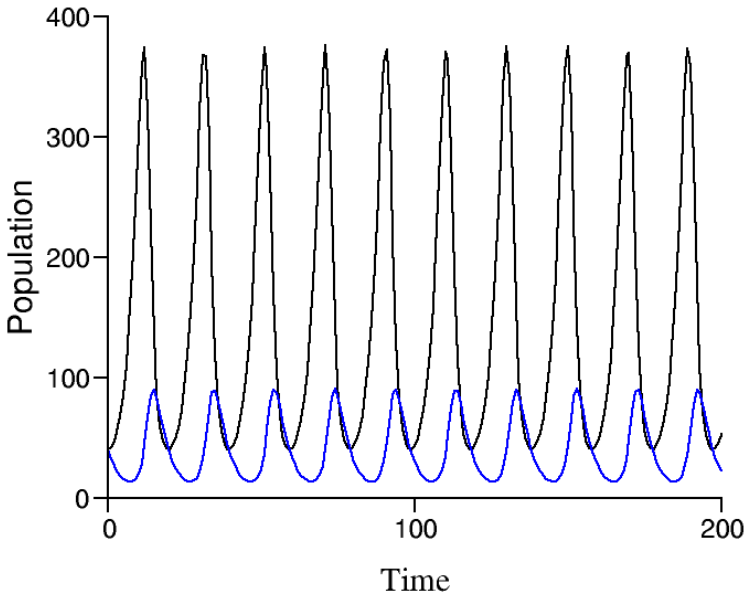


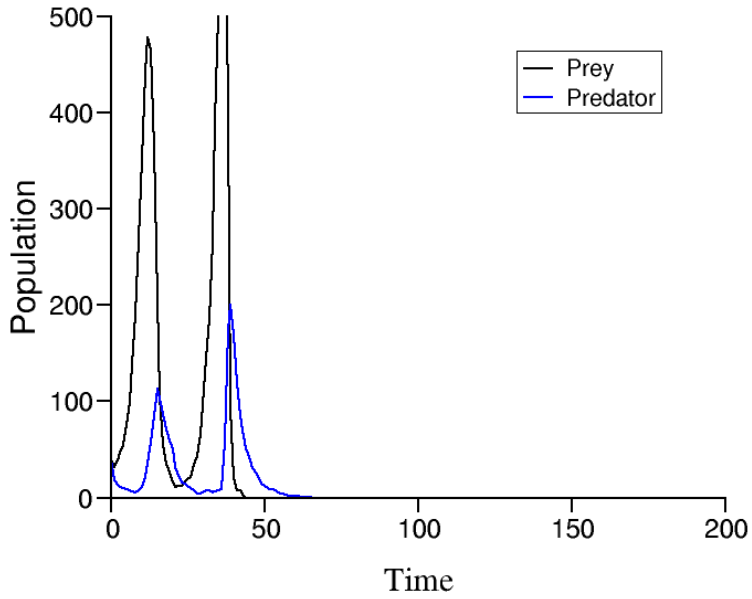
- **Conclusion**: different times to extinction and randomly distributed.

SO WHAT!!!

Life close to extinction

- A more interesting example: Predator-prey dynamics (Lotka-Volterra)

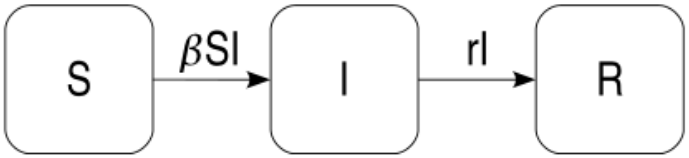


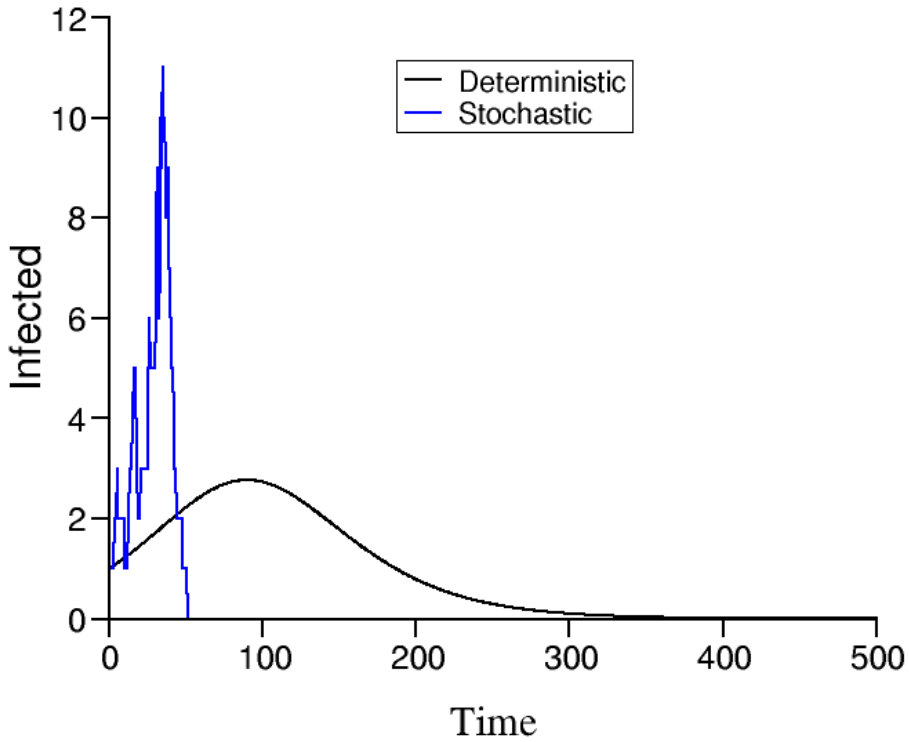


Stochastic life sucks (for these guys)!!!

Life close to a bifurcation point

- The classical SIR model (Susceptible-Infected-Recovered)





$R_0 = 1.2$ (for instance, seasonal influenza)

Life close to a bifurcation point

- The classical SIR model: some conclusions
  - Stochasticity changes our notion of R_0
  - The intensity of the epidemic depends strongly on fluctuations
  - Sometimes there is not epidemic even with $R_0 > 1$!!!

Life close to a bifurcation point

- The classical SIR model: the interest on these problems is relatively new (see this link)

ELSEVIER

Mathematical Biosciences 163 (2000) 1–33

Mathematical Biosciences

www.elsevier.com/locate/mbs

Comparison of deterministic and stochastic SIS and SIR models in discrete time

Linda J.S. Allen *,[1], Amy M. Burgin

Department of Mathematics and Statistics, Texas Tech University, Lubbock, TX 79409-1042, USA

Received 2 June 1998; received in revised form 16 August 1999; accepted 25 August 1999

Life close to a bifurcation point

- The classical SIR model: the interest on these problems is relatively new (or this other link)

ELSEVIER

Available online at www.sciencedirect.com

ScienceDirect

Mathematical Biosciences 208 (2007) 76–97

Mathematical Biosciences

www.elsevier.com/locate/mbs

Some properties of a simple stochastic epidemic model of SIR type

Henry C. Tuckwell [a,*], Ruth J. Williams [b]

[a] *Max Planck Institute for Mathematics in the Sciences Inselstr. 22, Leipzig D-04103, Germany*
[b] *Department of Mathematics, University of California San Diego, La Jolla, CA 92093, USA*

Received 27 May 2005; received in revised form 1 May 2006; accepted 20 September 2006
Available online 11 October 2006

Life close to a bifurcation point

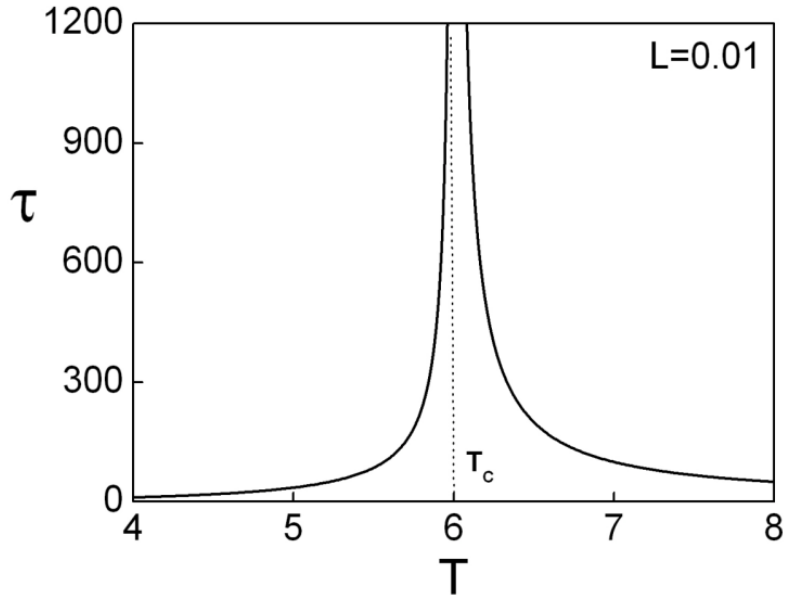
- The classical SIR model: the interest on these problems is relatively new (or even this thesis)

THE STOCHASTIC DYNAMICS OF EPIDEMIC MODELS

A THESIS SUBMITTED TO THE UNIVERSITY OF MANCHESTER
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY
IN THE FACULTY OF ENGINEERING AND PHYSICAL SCIENCES

Life close to a bifurcation point

- An analogy (I couldn't resist!!!)



That could be the susceptibility (magnetic field fluctuations), specific heat (energy fluctuations), ...

Let's recap: when?

- Close to extinction or bifurcations
- When interested on **individuals**
- Multiple stationary states

Individuals

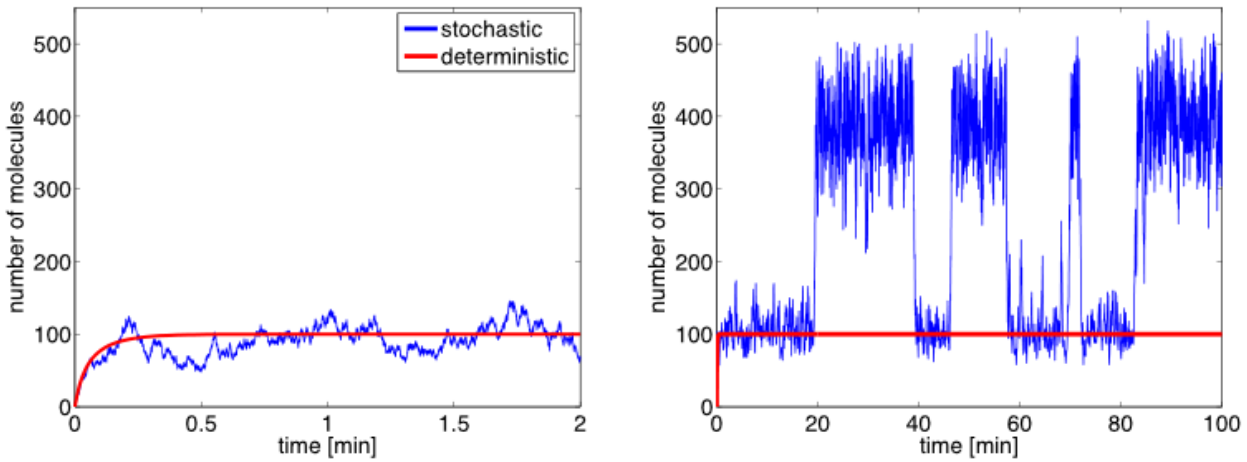
- Information on individuals matter. For instance
  - First passage times (see next slide)
  - Average time in a given state (link with molecular information)
  - Distribution of a given metric (observable)

Multiple stationary states

$$2A \underset{k_2}{\overset{k_1}{\rightleftharpoons}} 3A, \quad \emptyset \underset{k_4}{\overset{k_3}{\rightleftharpoons}} A$$

the deterministic equation is simply

$$\frac{dA}{dt} = k_1 A^2 - k_2 A^3 + k3 - k_4 A$$



The toolbox

- Numerics:
  - Naive integration
  - Gillespie (classical)
  - Tau-leap
  - More sophisticated tools (Munksy)
- Analytical:
  - Exact results: branching processes
  - Exact results: (chemical) master equation
  - Approximations: Van Kampen
    - The most general reference to learn about this method is Van Kampen's book

Naive vs Gillespie integration

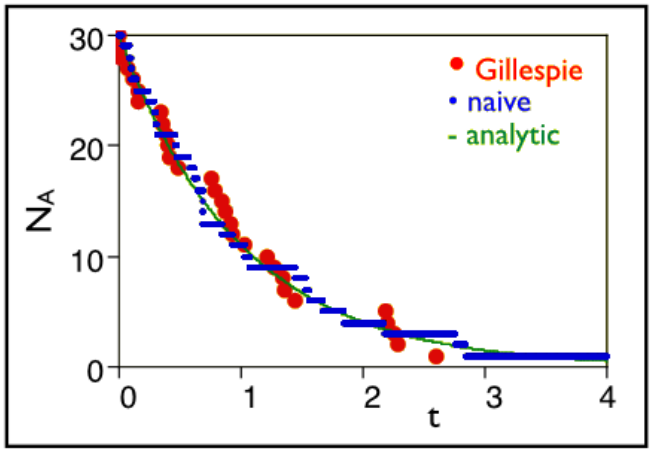
Our old friend: $A \rightarrow \emptyset$

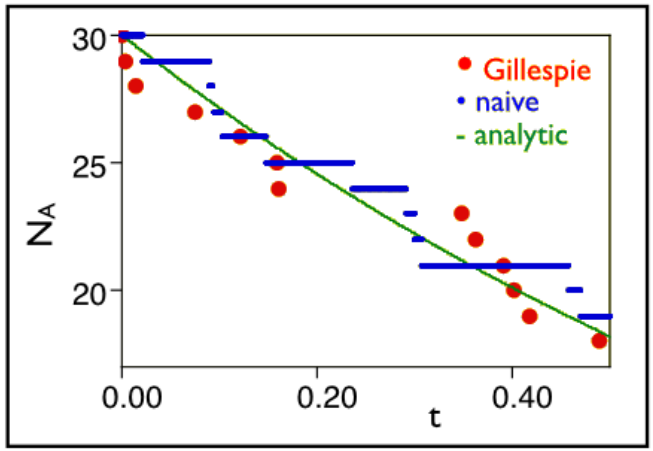
NAIVE (or τ-leap)

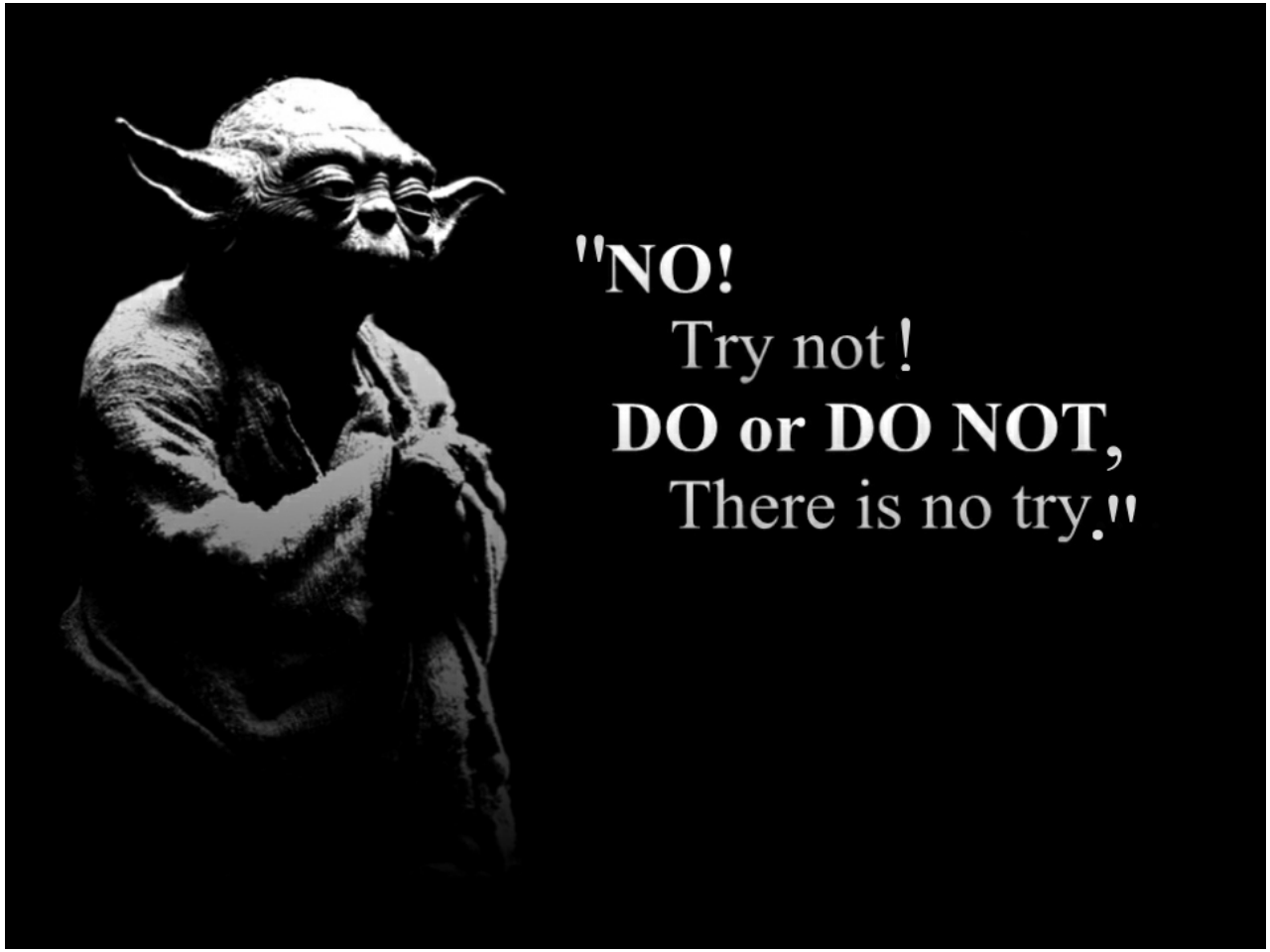
- Probability of occurring in the interval Δt (or τ)
- Fixed timesteps
- 1st order approximation
- *Trial and error*

GILLESPIE

- How long will it take the next event?
- Variable time steps
- Exact!
- Yoda's philosophy






"NO!
Try not!
DO or DO NOT,
There is no try."

Chemical Master equation

Let's go back to the simplest reaction: $A \xrightarrow{k} \emptyset$

- Deterministic approach: $A(t) = A_0 e^{-kt}$
- Stochastic:
  - Discrete number of *molecules* A: N_A
  - Probability that we loose 1 molecule in interval Δt: $N_A(t)k\Delta t$.
  - Compute the probability of having n molecules: $p_n(t)$.
  - Two possibilities for having $N_A(t+\Delta t) = n$:
    - $N_A(t) = n$ (no reaction during interval Δt.
    - Having $N_A(t) = n+1$ and loosing one.
  - Doing the math

$$p_n(t+\Delta t) = \underbrace{p_n(t)(1-nk\Delta t)}_{\text{nothing happened}} + \underbrace{p_{n+1}(t)(n+1)k\Delta t}_{\text{decay}}$$

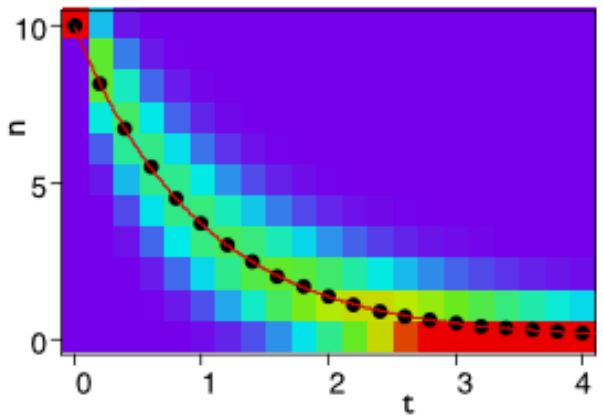
Chemical Master equation

- The master equation is the limit when $\Delta t \to 0$.
- In our simple case (initially we have n_0 particles):

$$\frac{dp_n}{dt} = k(n+1)p_{n+1}(t) - knp_n(t).$$

- Solution:

$$p_n(t) = e^{-nkt}\left(\begin{array}{c} n_0 \\ n \end{array} \right)(1-e^{-kt})^{(n_0-n)}$$



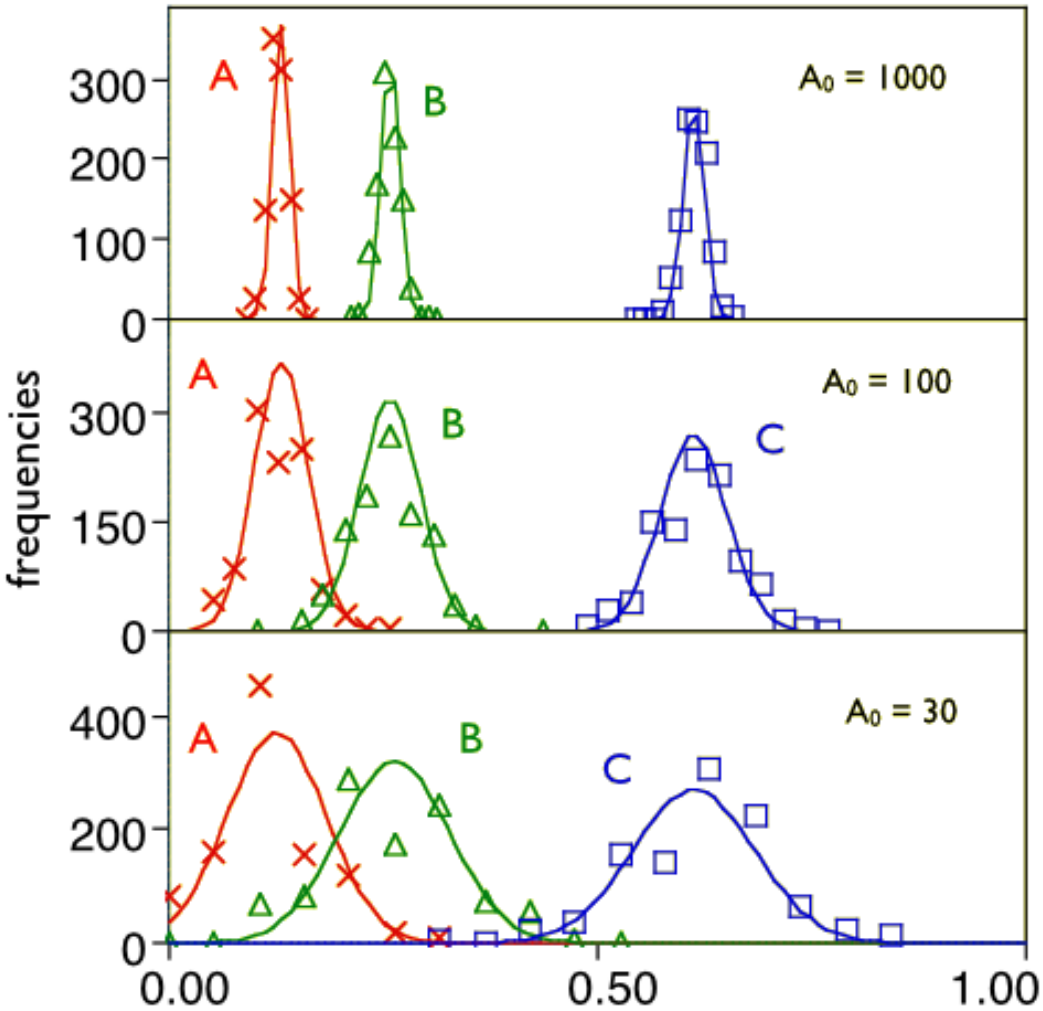
Chemical Master equation

CAVEAT: In this case the stochastic and deterministic provide the same average but remember . . .

- **Deterministic**: First average then integrate
- **Stochastic**: First integrate then average

Van Kampen's approximation

Do you remember that figure?



Van Kampen's approximation

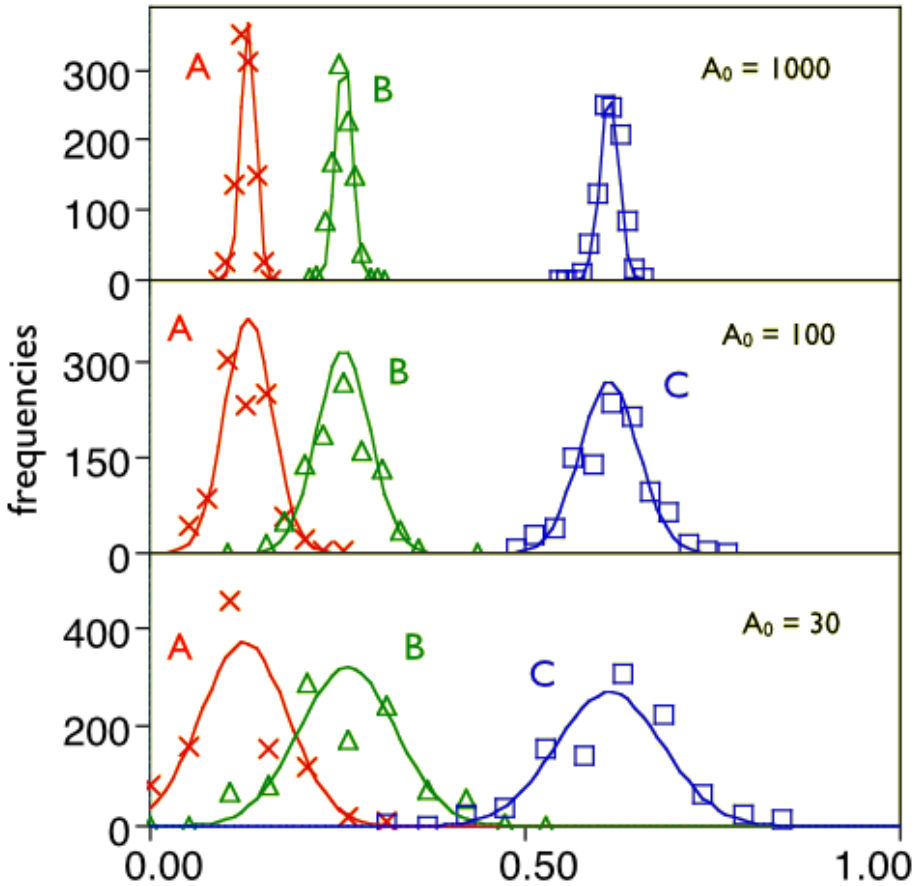
- The idea behind Van Kampen's approximation is to **transform** the master equation into a Langevin equation:

$$\text{Langevin} = \text{Deterministic system} + \text{Gaussian fluctuations (noise)}$$

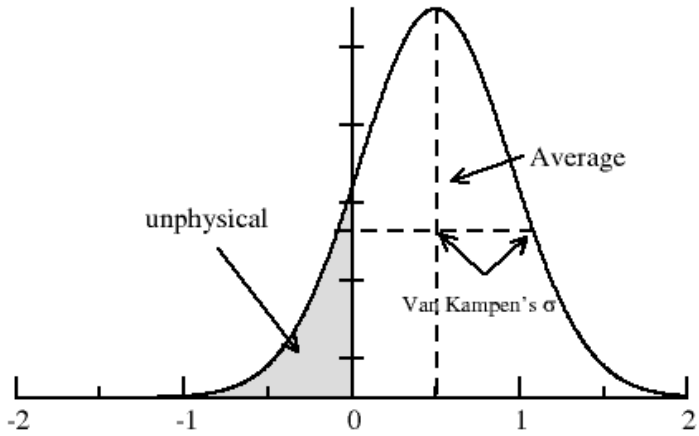
- Also known as Ω-expansion, linear noise or gaussian closure (even has a wikipedia entry).
- In general:
  - The good: It enriches traditional ODE approximation with noise.
  - The bad: Valid far from bifurcation points or extinction values
- Then why is it useful?

Van Kampen's approximation

Then why is it useful?



- It provides a criterion to determine when pure stochastic methods (in practice, Gillespie) are MANDATORY



Van Kampen's approximation

Sketch of the method:

- Write n as the sum of a deterministic and a random variables:

$$n = \Omega x + \sqrt{\Omega}\xi$$

- Rewrite the probabilities: $p_n(t) = p_{\Omega x + \sqrt{\Omega}\xi} \equiv \Pi(\xi, t)$
- Expand every function of $x + \Omega^{-1/2}\xi$ in power series of $\Omega^{-1/2}$
- Use shift operators: $\mathcal{E}f(n) = f(n+1) \Rightarrow \mathcal{E}f(\xi) = f(\xi + \Omega^{-1/2})$
- Arrive at a Fokker-Plank equation of the form, to order $O(\Omega^{-1/2}$:

$$\frac{\partial \Pi}{\partial t} = \sum_{ik} A_{ik} \frac{\partial(\xi_i \Pi)}{\partial \xi_k} + \frac{1}{2} \sum_{ik} [BB^T]_{ik} \frac{\partial^2 \Pi}{\partial \xi_i \partial \xi_k}$$

where A and B are matrices depending on the specific system. The eigenvalues of B give us information about the *variance* of the fluctuations.

Who is using this approach?

- Transcription and translation: Paulsson, Nature (2004)
- Enzyme kinetics: Grima, BMC Sys Biol (2009)
- Enzyme kinetics: Grima, the misterious case of the previously unsolved Michaelis-Menten stochastic dynamics (Phys. Rev. Lett.)
- Autocatalytic reactions: Dauxois et al, Phys. Rev. E (2009)
- General kinetic of reactions: Ben Avraham, 1987

Van Kampen's approximation

- Gooood news: Someone did the work for us

OPEN ACCESS Freely available online

PLoS one

Intrinsic Noise Analyzer: A Software Package for the Exploration of Stochastic Biochemical Kinetics Using the System Size Expansion

Philipp Thomas[1,2,3*], Hannes Matuschek[4*], Ramon Grima[1,2*]

Summary of part one

- When stochastic
  - Close to extinction or bifurcations (**case study 1**)
  - When VK criterion suggests it
  - Multiple stationary states
  - Interested on individuals (**case study 2**)
- Tools
  - Gillespie (or tau-leap)
  - Van Kampen's approximation
  - *Branching processes* (in next life)
  - Averaging after integrating (and no the other way around)

Case studies

- An example from virology: comparison of strategies
- An example from immunology: time needed to activate a T-cell

Case study I: Virology

Two viral strategies

When a cell is infected by a virus, there are two ways in which the virus can use the cell machinery to reproduce

- Bursting (explosive release): The cell dies and K virions are released
- Budding (continuous release): The cell lives infected and serves as a virion *provider*

Two viral strategies

Budding

$$
\begin{aligned}
V + T &\xrightarrow{k} I \\
I &\xrightarrow{N\delta} I + V \\
I &\xrightarrow{\delta} \emptyset \\
V &\xrightarrow{c} \emptyset
\end{aligned}
$$

Bursting

$$
\begin{aligned}
V + T &\xrightarrow{k} I \\
I &\xrightarrow{\delta} NV \\
V &\xrightarrow{c} \emptyset
\end{aligned}
$$

Pearson, Krapivsky and Perelson, PLoS Comp. Biol. **7(2)**, e1001058 (2011).

Bursting vs Budding

Why stochastic matters? Two answers:

- Same deterministic model but very different dynamics
- In early infections the population of virus is small (remember: small $\Rightarrow$ large fluctuations)

Bursting vs Budding: Deterministic equations

Some tips and tricks:

- From A and B means
  - Both A and B lose
  - At a rate proportional to $A \times B$ (*"and"* means *"product"*)
- From either A or B means
  - Both A and B lose
  - In different reactions at rates proportional to either A or B
- Write down *stochiometric* matrix

Budding: Deterministic equations

Budding

$$V \xrightarrow{kT} I \quad (R_1)$$
$$I \xrightarrow{N\delta} I + V \quad (R_2)$$
$$I \xrightarrow{\delta} \emptyset \quad (R_3)$$
$$V \xrightarrow{c} \emptyset \quad (R_4)$$

Stochiometric matrix

	R_1	R_2	R_3	R_4
V	-1	1	0	-1
I	+1	0	-1	0

$$\frac{dV}{dt} = -R_1 + R_2 + 0R_3 - R_4 = N\delta I - (kT + c)V$$

$$\frac{dI}{dt} = R_1 + 0R_2 - R_3 + 0R_4 = kTV - \delta I$$

Bursting: Deterministic equations

Stochiometric matrix

	R_1	R_2	R_3
V	-1	N	-1
I	+1	-1	0

Bursting

$$V + T \xrightarrow{k} I \quad (R_1)$$
$$I \xrightarrow{\delta} NV \quad (R_2)$$
$$V \xrightarrow{c} \emptyset \quad (R_3)$$

$$\frac{dV}{dt} = -R_1 + NR_2 - R_3 = N\delta I - (kT + c)V$$

$$\frac{dI}{dt} = R_1 - R_2 + 0R_3 = kTV - \delta I$$

Bursting vs Budding: Stochastic dynamics

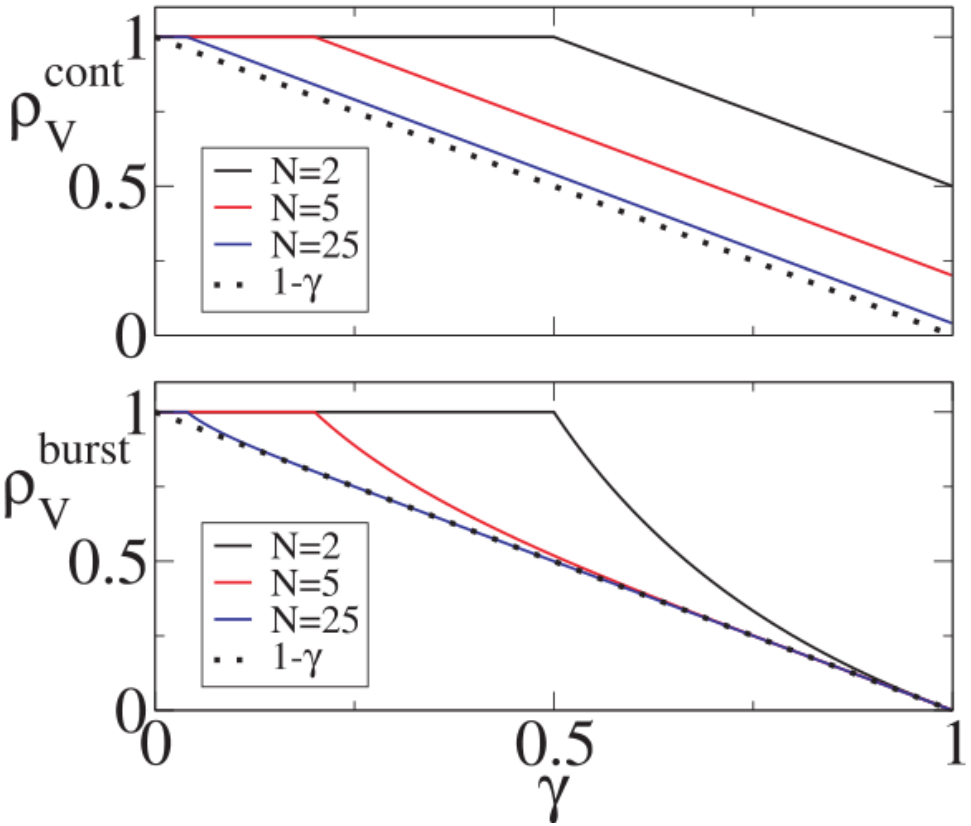
- Same deterministic equations $\Rightarrow$ same $R_0 \equiv \frac{NkT}{kT+c}$
- However, different stochastic probabilities of extinction

$$\begin{aligned} \rho_V^{burst} &= min(\rho^*, 1) \\ \rho_V^{bud} &= min(1-(R_0-1)/N, 1) \end{aligned} \tag{1}$$

where ρ^* is a positive root of

$$\frac{1-\rho_V}{1-(\rho_V)^N} = \frac{R_0}{N} \equiv \gamma$$

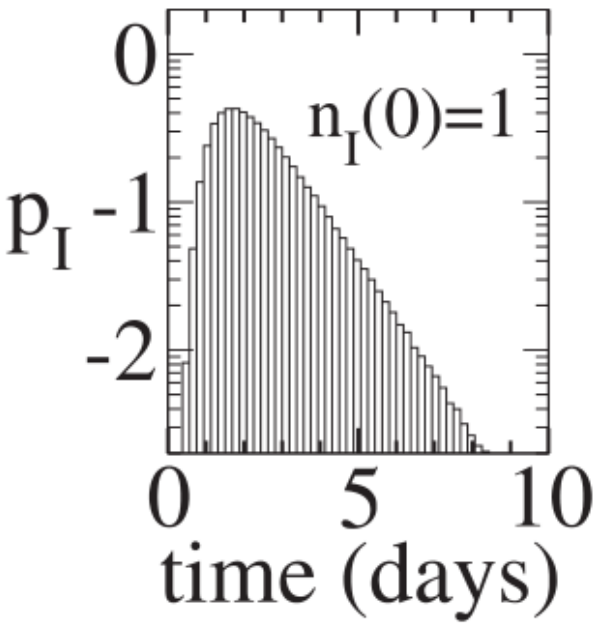
Bursting vs Budding: Stochastic dynamics

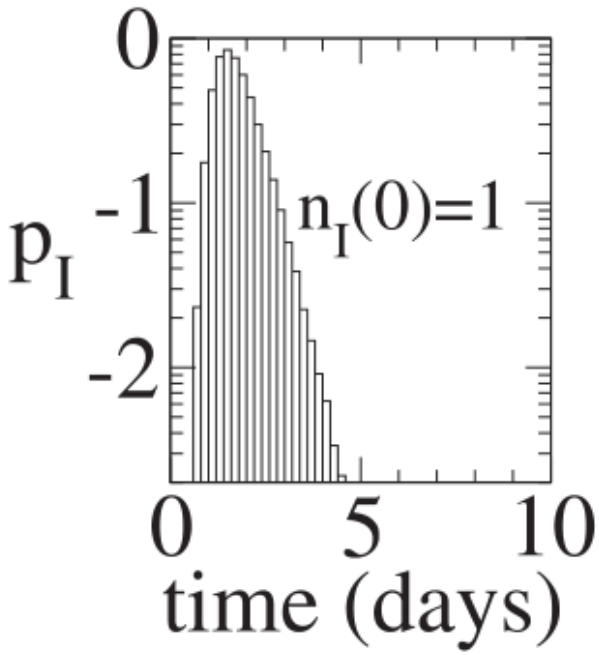


Bursting vs Budding: Stochastic dynamics

Additional benefits of stochastic dynamics

- Frequency distributions
- First passage times

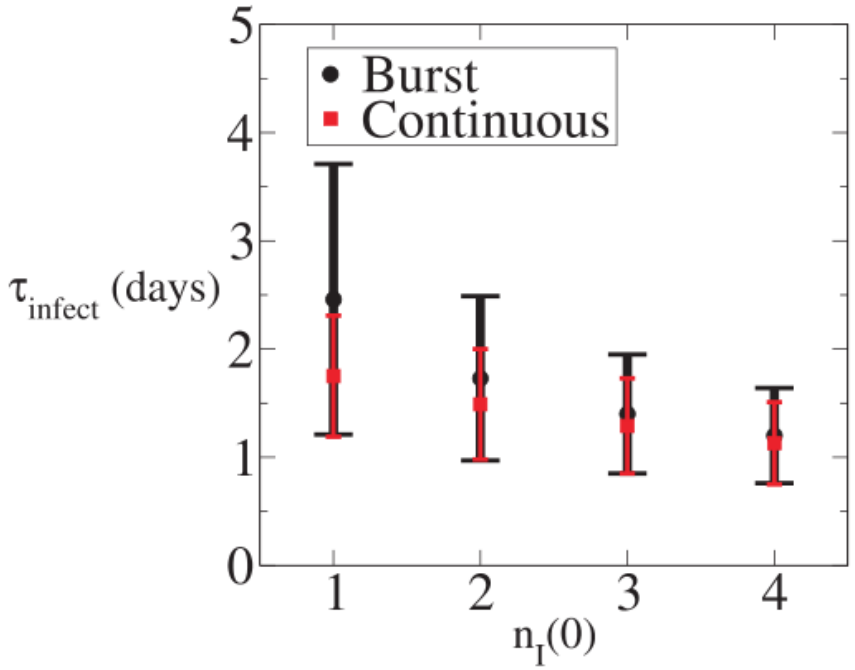




Bursting vs Budding: Stochastic dynamics

Additional benefits of stochastic dynamics

- Frequency distributions
- First passage times



Case study II: Immunology

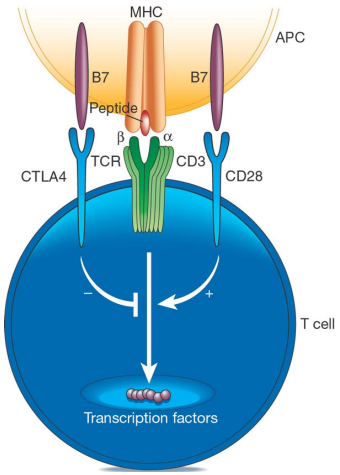
A stochastic T cell response criterion

What is the average time needed to activate a T-cell?

- Here is the reference

A stochastic T cell response criterion

The biological system



- An APC (antigen presenting cell) presents pieces of antigen to a T-cell
- The T-cell has receptors (TCR) able to recognize specific antigens
- The *matching* between the ligand and the receptor elicits a response (T-cell activation)

A stochastic T cell response criterion

Experimental evidence

- A few agonist pMHC ligands can suffice to trigger T cell responses
- Sufficiently long TCRpMHC engagements are required to initiate the signalling cascade, resulting in productive signal transduction
- T-cells can integrate signals; that is, counting devices are at work in T cells to allow signal accumulation, decoding and translation into biological responses

HYPOTHESIS: T cell responses take place once a given number of TCRs (and not necessarily in a simultaneous way), N, have been engaged with ligand for at least a dwell time, τ, each

http://rsif.royalsocietypublishing.org/content/early/2012/06/27/rsif.2012.0205.full

First attempt: deterministic criterion

Remember: Time needed to have N ligand-receptor engagements a time τ.

- The simplest model

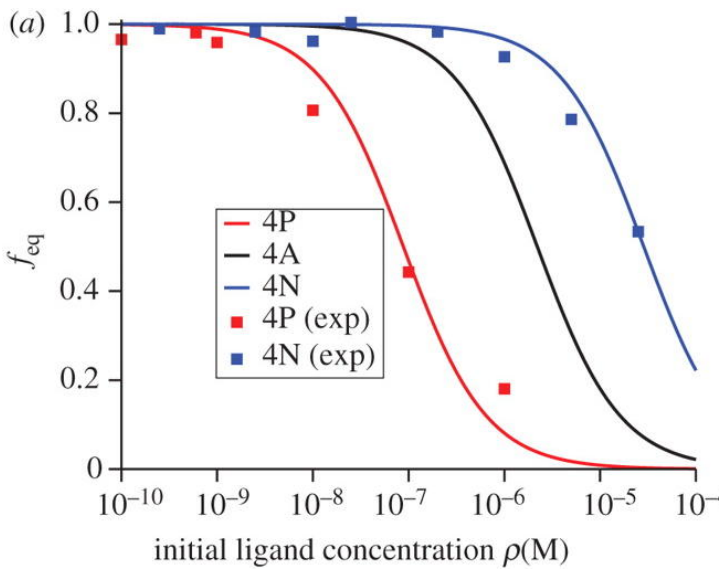
- Deterministically, we find (eliminating the concentration of free ligands and empty receptors)

$$\frac{dz}{dt} = -k_{\text{off}} + k_{\text{on}}(N_R - z)\left(\rho - \frac{N_c z}{V N_A}\right)$$

First attempt: deterministic criterion

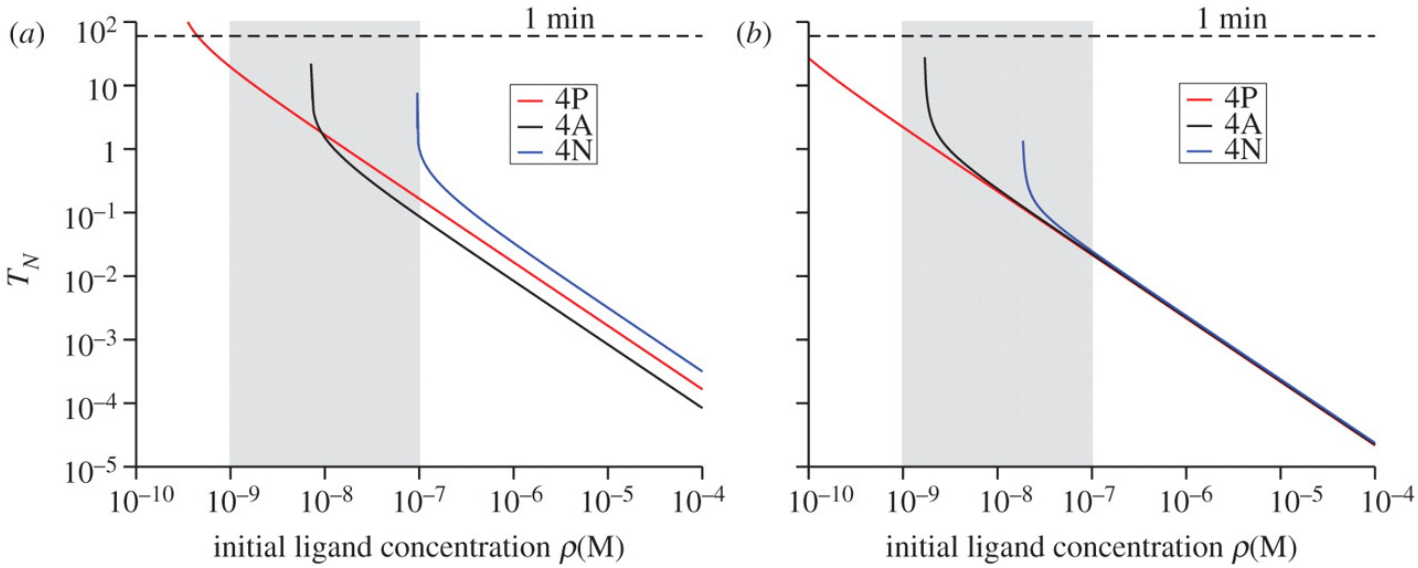
Remember: Time needed to have N ligand-receptor engagements a time τ.

- First: test the model



First attempt: deterministic criterion

- Second: the best we can do is wait until we have N engaged and then wait for another τ seconds

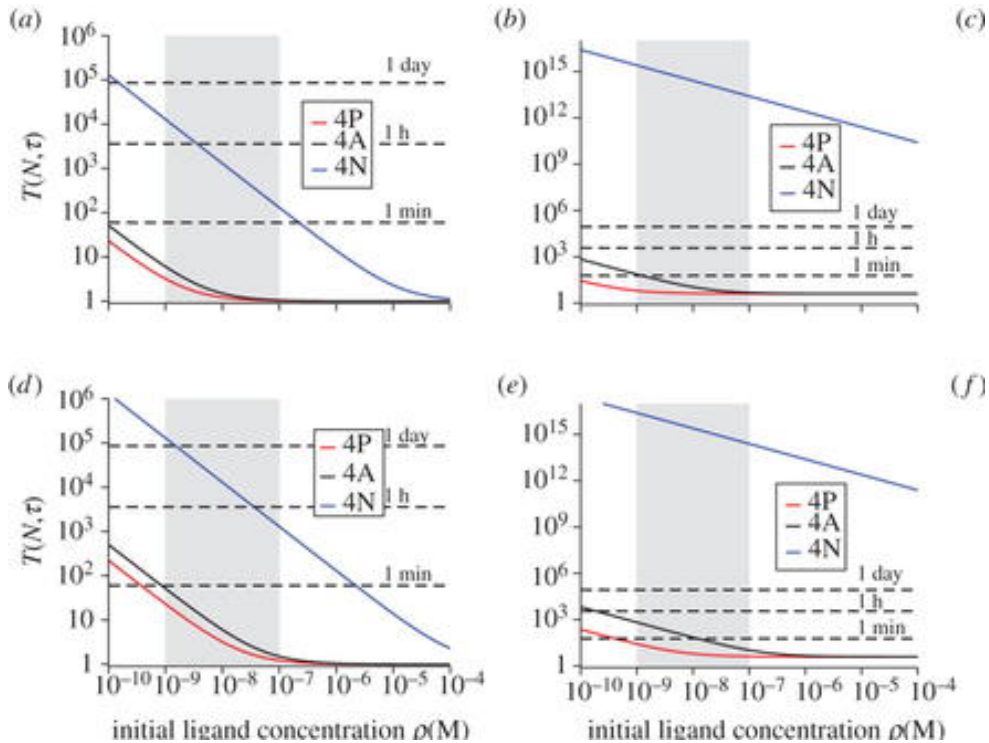


It doesn't predict how things work in nature

So nature works intrinsically in a stochastic way

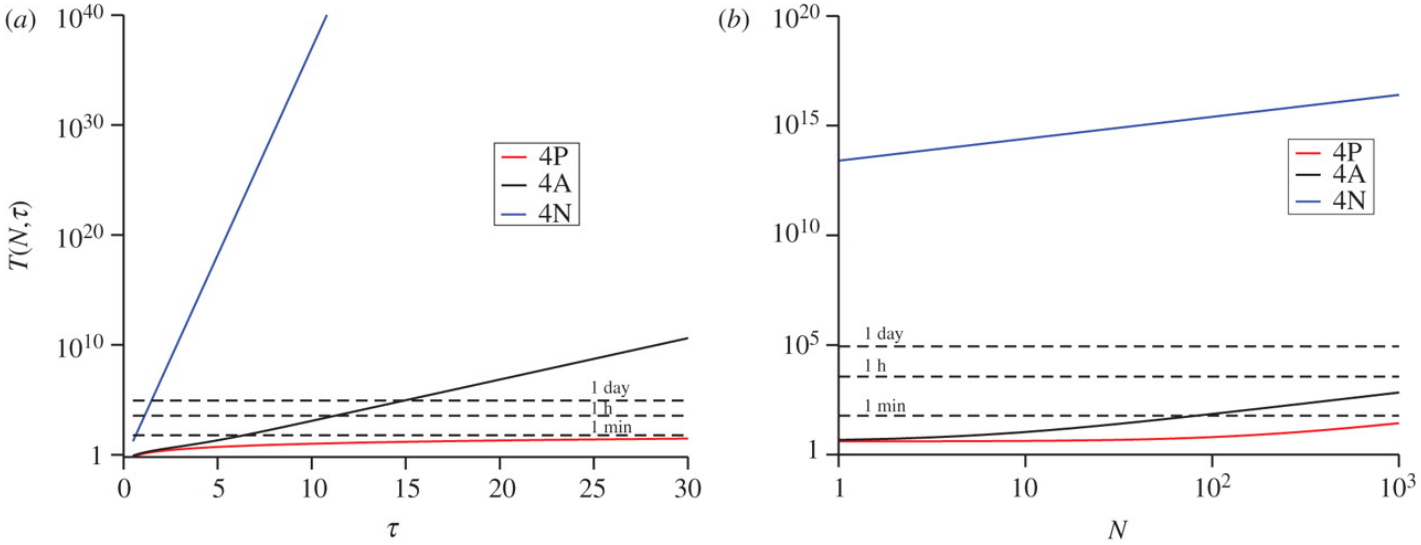
Second attempt: stochastic criterion

- Stochastic dynamics allows us to count engaged receptors individually
- Whenever a receptor is engaged for a time longer than τ, count $+\mathbf{1}$.
- Wait until N have been engaged at least τ.
- Experimental data suggests that $\tau \in [1, 10]$ and $N \in [10, 100]$



Second attempt: stochastic criterion

- This model provides invaluable testable predictions



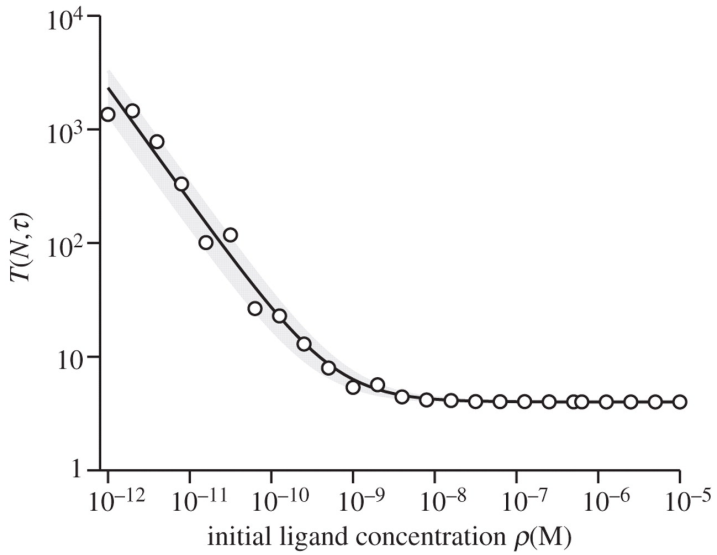
Second attempt: stochastic criterion

- This simple model is amenable to analytical calculations
  - Mean First Passage Time:

$$T(N,\tau) = \tau + \frac{Ne^{k_{\text{off}}\tau}}{k_{\text{on}} N_R \rho}$$

  - Variance

$$Var(T) = \frac{Ne^{k_{\text{off}}\tau}}{(k_{\text{on}} N_R \rho)^2}$$



In a nutshell

We ALWAYS need stochastic methods
if we are concerned with *labeled* individuals

Final remarks . . .

Some computational tools

- Bionetgen (Gillespie, ODE). Simple to codify
- Intrinsic Noise Analyzer (Van Kampen vs Gillespie)
- Mathematica: analytical (*e.g.,* Van Kampen)
- General tools:
  - Matlab
  - Python
  - C++
  - R
  - . . .

Thanks for your attention