Stochastic models in biology

David F. Anderson*

*anderson@math.wisc.edu

Department of Mathematics
University of Wisconsin - Madison

CIBM Seminar

December 7th, 2010

Outline

1. → General discussion of mathematical models and the difference between deterministic and stochastic?

2. What are the basic stochastic models used?

3. How can we understand the behavior of stochastic models both analytically and computationally?

Outline

1. → General discussion of mathematical models and the difference between deterministic and stochastic?

2. What are the basic stochastic models used?

3. How can we understand the behavior of stochastic models both analytically and computationally?

   (not-so-hidden) Ulterior motives:

   i "Sell" the usefulness of probability and mathematics in this arena.

   ii Recruit for my Stochastic Models in Biology course in Fall 2011 (Math 605).

What is a mathematical model?

1. A mathematical description of real world phenomena.
2. Can be used to make predictions of behavior of the system.
3. Experimentally test predictions made.
4. Tweak the mathematical model and repeat.

Stochastic versus deterministic models

A process is *deterministic* if its future is completely determined by its present and past. Examples include

- solutions to differential equations.
- solutions to difference equations.

Stochastic versus deterministic models

A process is *deterministic* if its future is completely determined by its present and past. Examples include

- solutions to differential equations.
- solutions to difference equations.

Example

The initial value problem

$$\frac{d}{dt}x(t) = 3x(t) \qquad x(0) = 2,$$

has the solution $x(t) = 2e^{3t}$. □

Stochastic versus deterministic models

A process is *deterministic* if its future is completely determined by its present and past. Examples include

- solutions to differential equations.
- solutions to difference equations.

Example

The initial value problem

$$\frac{d}{dt}x(t) = 3x(t) \qquad x(0) = 2,$$

has the solution $x(t) = 2e^{3t}$. □

Example

Consider the difference equation

$$F_1 = F_2 = 1$$
$$F_n = F_{n-1} + F_{n-2}, \quad \text{for } n > 2.$$

Then $\{F_n\}_{n=1}^{\infty}$ is the well known Fibonacci sequence: $\{1, 1, 2, 3, 5, 8, \dots\}$. □

Stochastic versus deterministic models

On the other hand, a *stochastic process* is a random process evolving in time.

Informally: even if you have full knowledge of the state of the system (and it's entire past), you can not be sure of it's value at future times.

Stochastic versus deterministic models

On the other hand, a *stochastic process* is a random process evolving in time.

Informally: even if you have full knowledge of the state of the system (and it's entire past), you can not be sure of it's value at future times.

Example

Consider rolling a die multiple times. Let S_n denote the sum of the first n rolls. Then,

$$S_0 = 0$$

and

$$S_1 \in \{1, \ldots, 6\}, \quad S_2 \in \{2, \ldots, 12\}, \text{ etc.}$$

Knowing that $S_2 = 8$ only guarantees that $S_3 \in \{9, \ldots, 14\}$.

Why study stochastic models of intracellular processes?

Stochastic models have a long history in biology (Galton/Watson 1873, Max Delbrück, JCP, 1940); however, over the past 15 years their use has exploded.

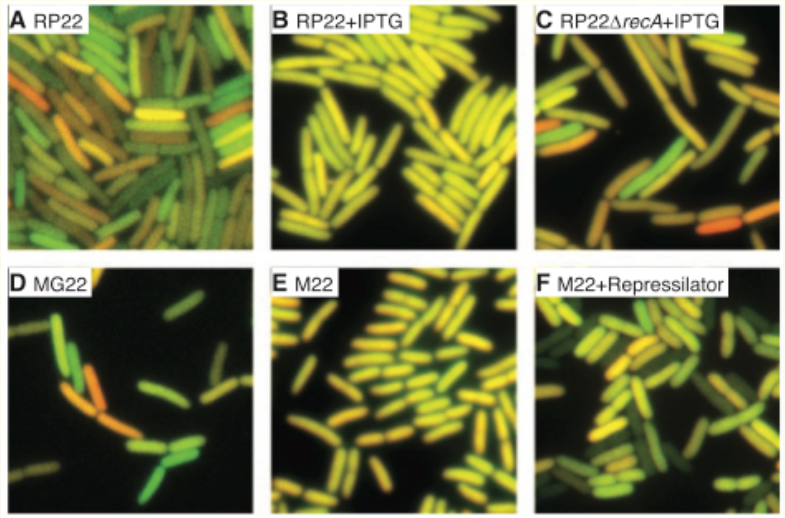
One reason:

- recent advances in experimental methods in biology, such as green fluorescent protein, have enabled quantitative measurements at the single cell, and even single molecule, level.

- Such experiments show time and time again that the dynamics at this level are intrinsically stochastic, or "noisy," and that that noise can have large implications for the qualitative dynamics.

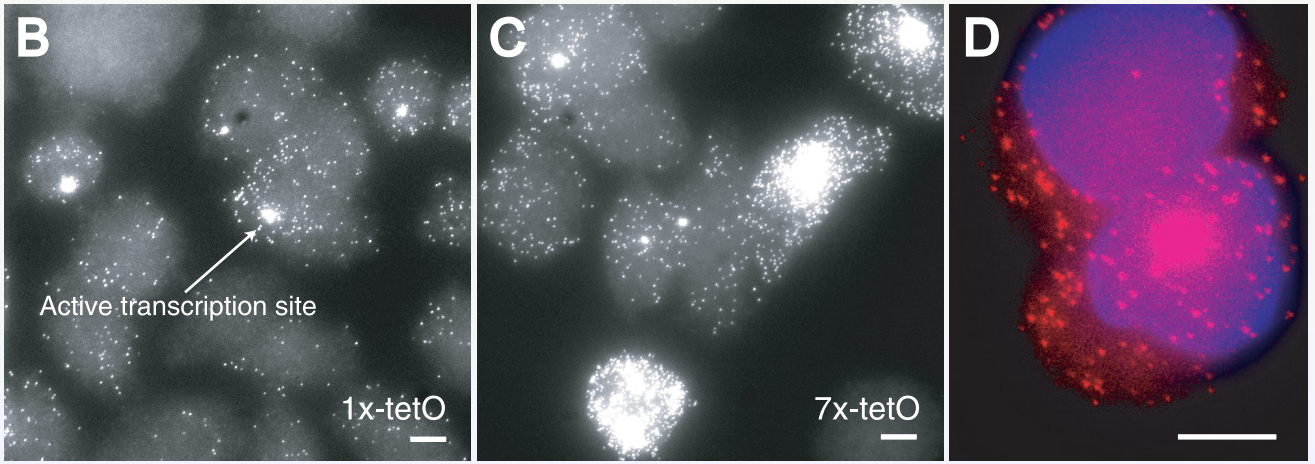
Why study stochastic models of intracellular processes?

Clonal populations of cells exhibit substantial phenotypic variation:



[1]

Different levels of mRNA in genetically homogeneous populations:



[2]

[1] Elowitz et al., Science, **297**, 2002.

[2] Raj et al., PLoS Biology, **4**(10), 2006.

What are the differences? Example: Bacterial Growth

Let's consider two oversimplified models for bacterial growth (by *growth* here, I mean the growth of the size of the colony, not of an individual bacterium):

- one deterministic
- one stochastic.

What are the differences? Example: Bacterial Growth

Let's consider two oversimplified models for bacterial growth (by *growth* here, I mean the growth of the size of the colony, not of an individual bacterium):

- one deterministic
- one stochastic.

We suppose

- there are 10 bacteria at time zero.
- each bacteria divides at an "average" rate of once per three hours.

What are the differences? Example: Bacterial Growth

Let's consider two oversimplified models for bacterial growth (by *growth* here, I mean the growth of the size of the colony, not of an individual bacterium):

- one deterministic
- one stochastic.

We suppose

- there are 10 bacteria at time zero.
- each bacteria divides at an "average" rate of once per three hours.

Deterministic model: a "reasonable" model would be

$$\frac{d}{dt}x(t) = \frac{1}{3}x(t) \qquad x(0) = 10, \tag{1}$$

with solution

$$x(t) = 10e^{t/3},$$

where the units of t are hours.

Example: Bacterial Growth

Stochastic Model: Without going into the finer details yet, assume

1. Each bacteria divides after a random (independent, exponential) amount of time with an average wait of 3 hours.

Example: Bacterial Growth

Stochastic Model: Without going into the finer details yet, assume

1. Each bacteria divides after a random (independent, exponential) amount of time with an average wait of 3 hours.

Similar to equation (1) for the deterministic model, it is possible to write down systems of equations describing the time evolution of model

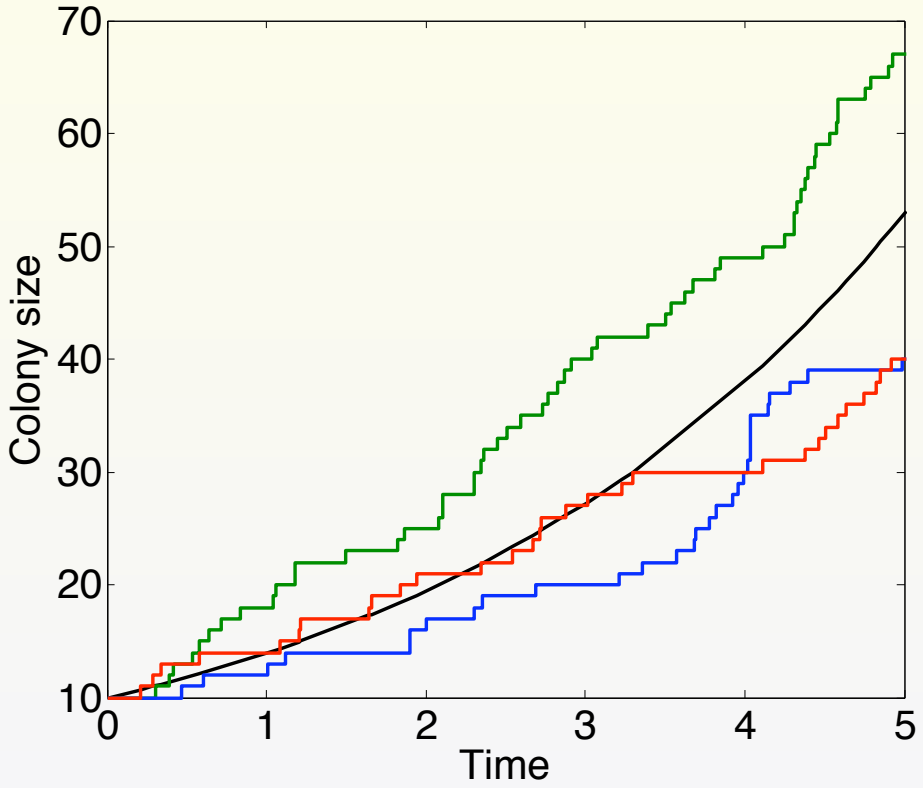
1. Evolution of individual sample paths – instance of experiment (like the ODE model)

2. Evolution of the distribution (probability of being in certain states)

However, I will postpone doing so until later.

Example: Bacterial Growth - evolution of sample paths

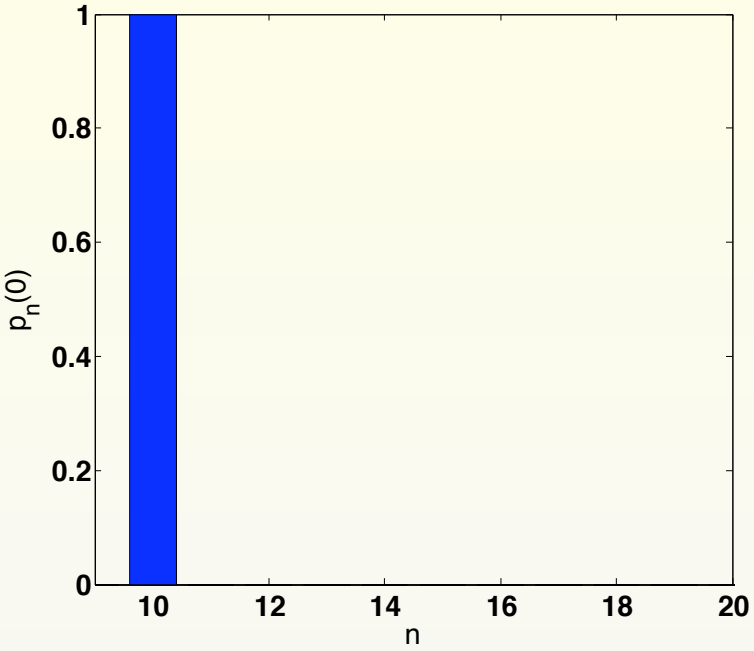
- Below is a plot of the solution of the deterministic system versus three different realizations of the stochastic system.



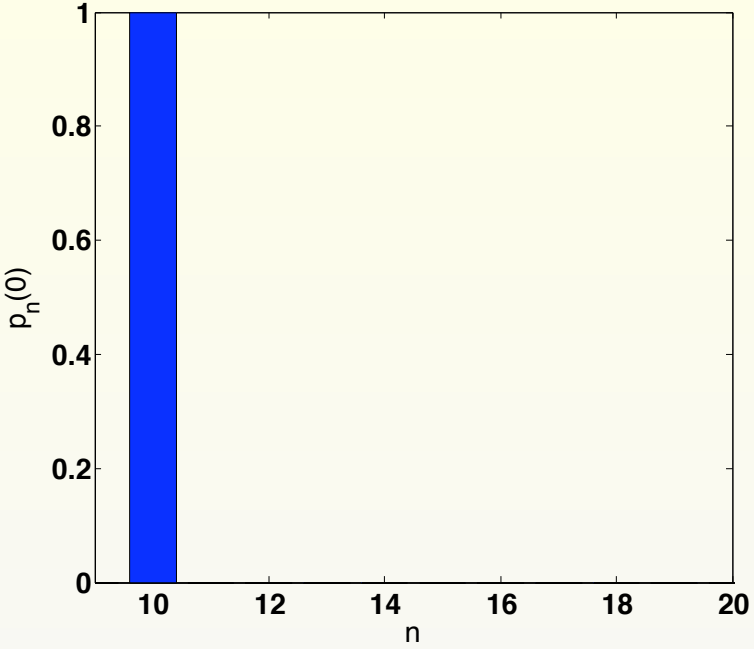
- Stochastic realizations/experiments appear to follow the deterministic system in a “noisy” way.

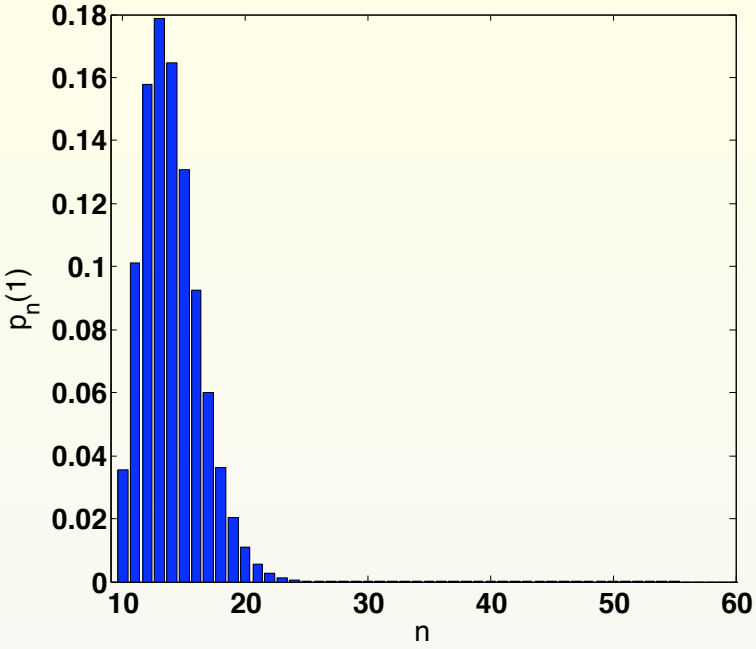
- It is clear that the behavior of a single realization or experiment of the stochastic system can not be predicted with absolute accuracy.

Example: population growth - evolution of distribution

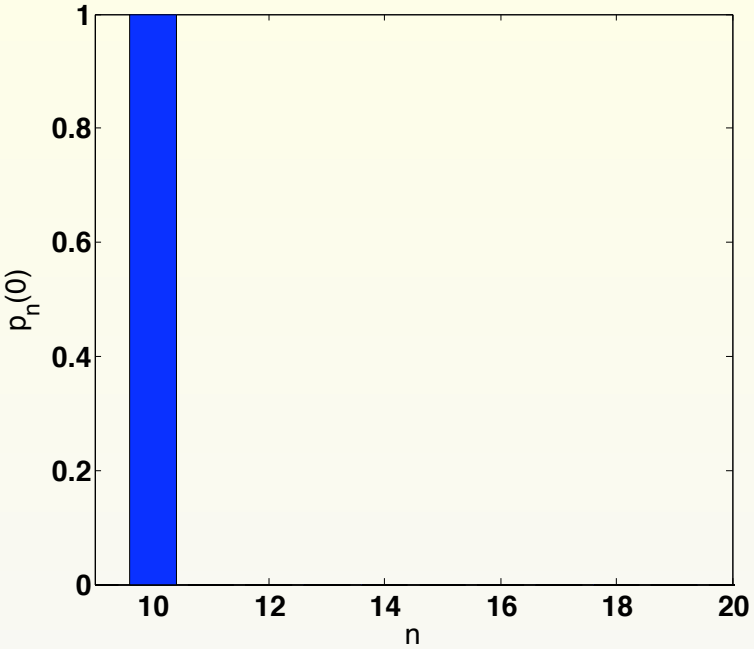


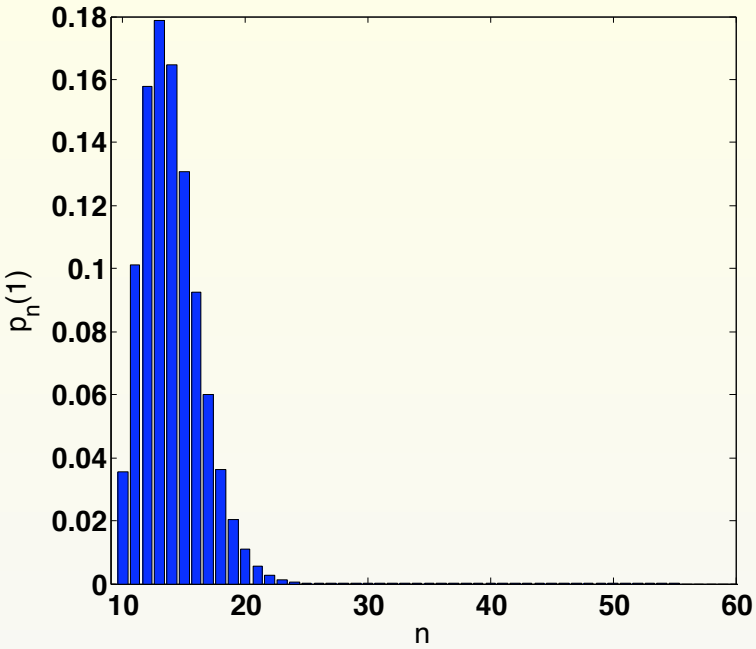
Example: population growth - evolution of distribution

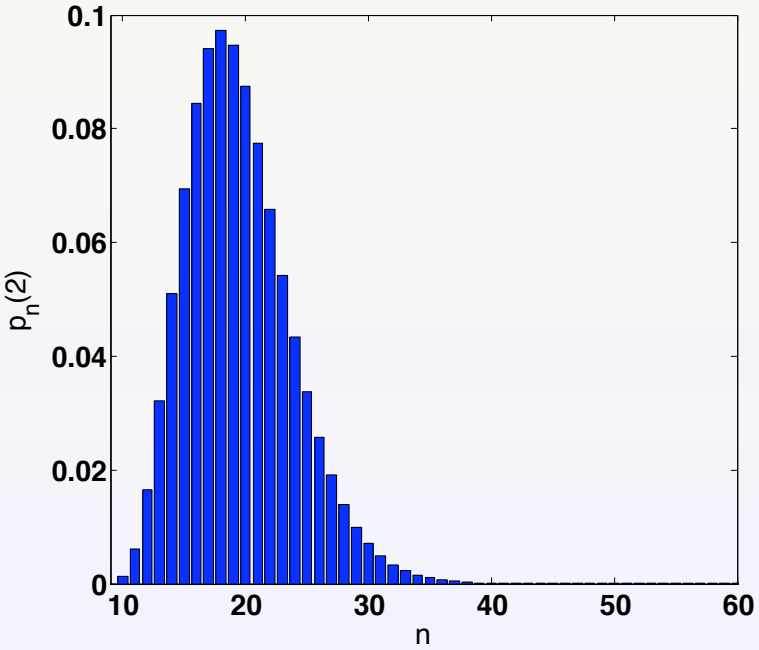




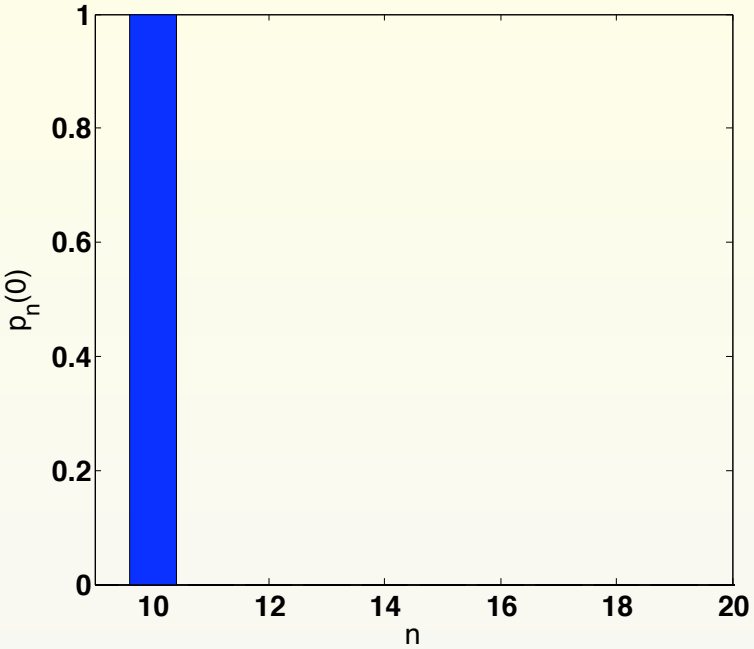
Example: population growth - evolution of distribution

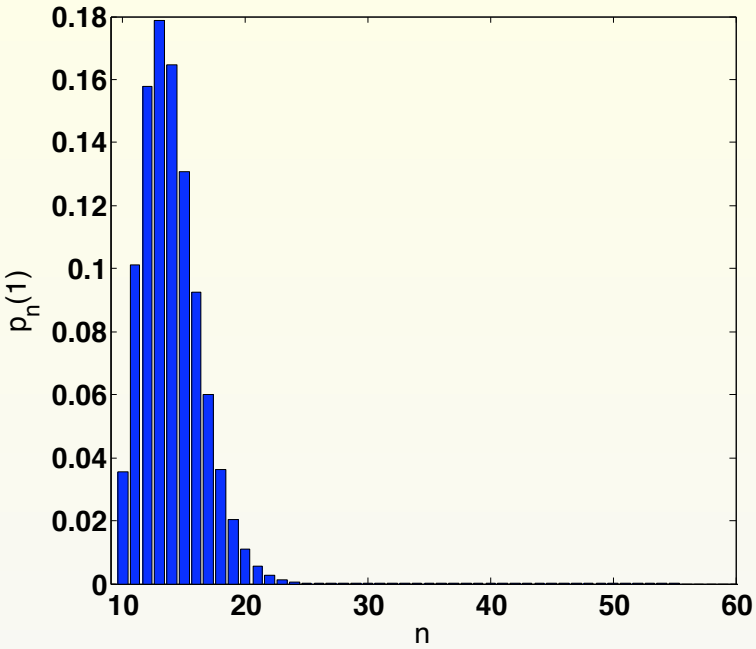


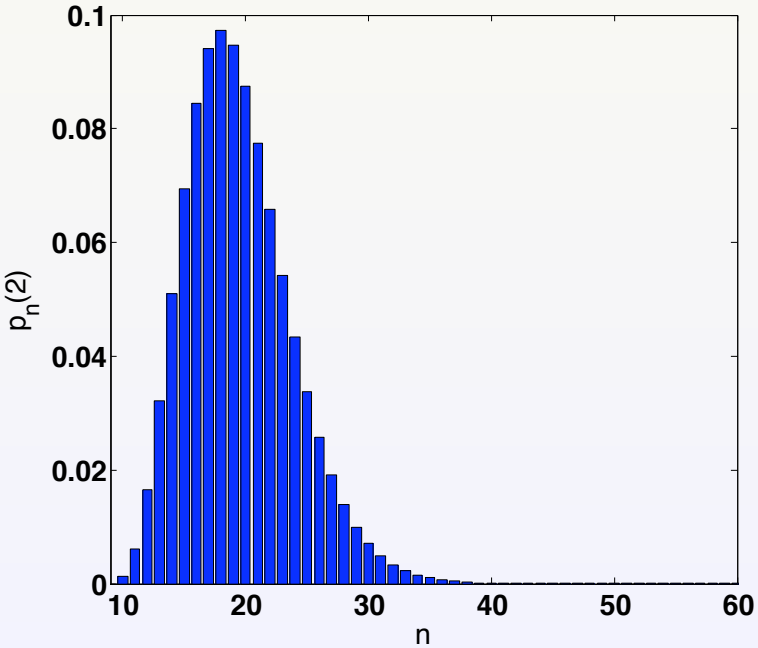


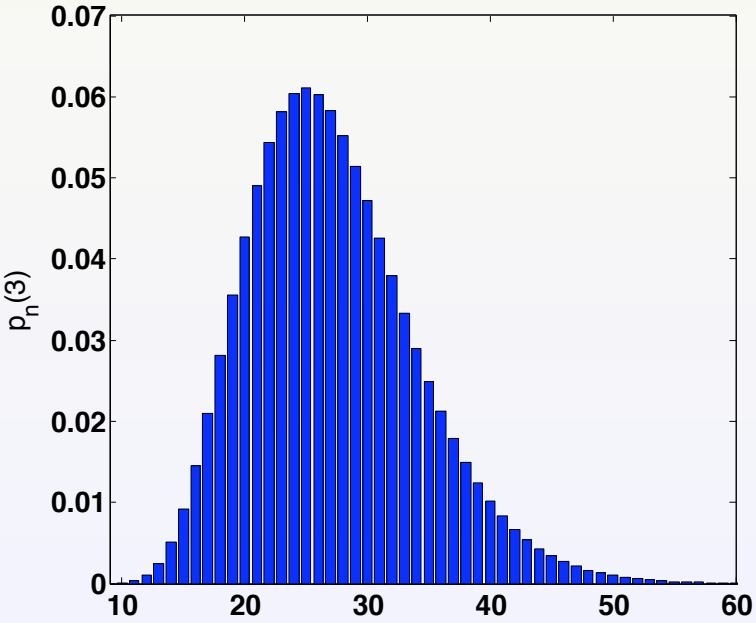


Example: population growth - evolution of distribution









Example: Bacterial Growth and Death

Now suppose that we change the model “slightly” in that:

1. we allow bacteria to die as well as divide.
2. we suppose we begin with only two bacteria.

We suppose that they die after about five hours.

Example: Bacterial Growth and Death

Now suppose that we change the model "slightly" in that:

1. we allow bacteria to die as well as divide.
2. we suppose we begin with only two bacteria.

We suppose that they die after about five hours.

Our new deterministic model could be

$$\dot{x}(t) = \frac{1}{3}x(t) - \frac{1}{5}x(t) = \frac{2}{15}x(t), \qquad x(0) = 2,$$

with solution

$$x(t) = 2e^{2t/15}.$$

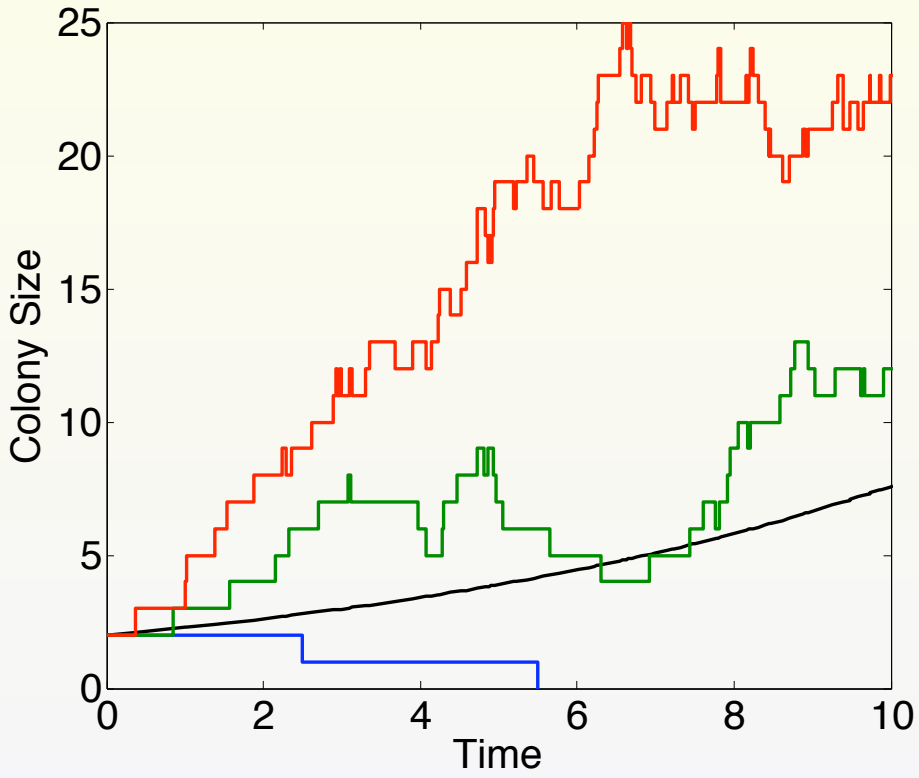
Example: Bacterial Growth and Death

For the stochastic model, we now model the two possible changes to the size of the colony separately. That is, the next event is *either*

1. a growth event (via a division) or
2. a decrease event (via a death).

Example: Bacterial Growth and Death

- Deterministic vs. three realizations/experiments of stochastic system.



- The models now behave *qualitatively* differently:

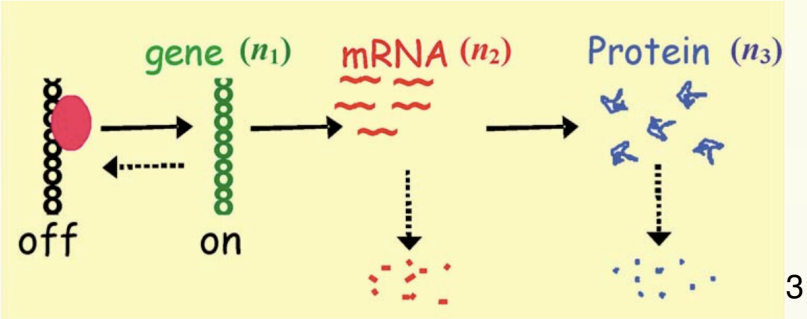
  one of the realizations of the stochastic model (i.e. one of the colonies under observation) has been completely wiped out, *something not possible in the deterministic modeling context*.

Stochastic models for biochemical processes

Gene transcription & translation:

$$
\begin{aligned}
&\vdots \\
G_i &\longrightarrow G_i + M_i \quad \textit{transcription} \\
M_i &\longrightarrow M_i + P_i \quad \textit{translation} \\
M_i &\longrightarrow \emptyset \quad \textit{degradation} \\
P_i &\longrightarrow \emptyset \quad \textit{degradation} \\
G_i + P_j &\longrightarrow \textit{Bound}_i \\
&\vdots
\end{aligned}
$$

Cartoon representation:



[3]

[3] J. Paulsson, Physics of Life Reviews, **2**, 2005 157 – 175.

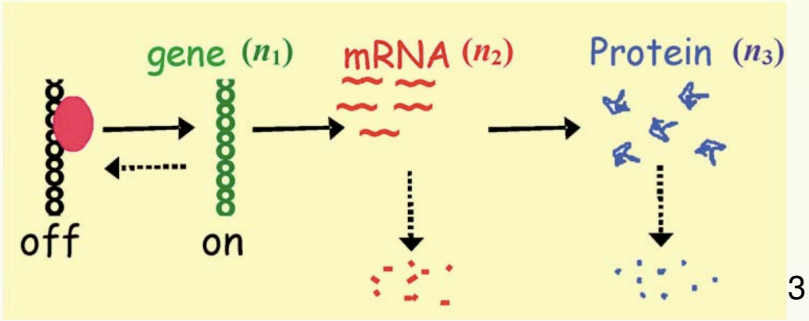
[4] Hye Won Kang, presentation at SPA in 2007.

Stochastic models for biochemical processes

Gene transcription & translation:

$$
\begin{aligned}
&\vdots \\
G_i &\longrightarrow G_i + M_i \quad \textit{transcription} \\
M_i &\longrightarrow M_i + P_i \quad \textit{translation} \\
M_i &\longrightarrow \emptyset \quad \textit{degradation} \\
P_i &\longrightarrow \emptyset \quad \textit{degradation} \\
G_i + P_j &\longrightarrow \textit{Bound}_i \\
&\vdots
\end{aligned}
$$

Cartoon representation:



[3]

E. coli Heat Shock Response Model. 9 species, 18 reactions.

Reaction	Intensity	Reaction	Intensity
$\emptyset \longrightarrow A_8$	4.00×10^{0}	$A_6 + A_8 \longrightarrow A_9$	$3.62 \times 10^{-4} X_{A_6} X_{A_8}$
$A_2 \longrightarrow A_3$	$7.00 \times 10^{-1} X_{A_2}$	$A_8 \longrightarrow \emptyset$	$9.99 \times 10^{-5} X_{A_8}$
$A_3 \longrightarrow A_2$	$1.30 \times 10^{-1} X_{A_3}$	$A_9 \longrightarrow A_6 + A_8$	$4.40 \times 10^{-5} X_{A_9}$
$\emptyset \xrightarrow{A_1} A_2$	$7.00 \times 10^{-3} X_{A_1}$	$\emptyset \longrightarrow A_1$	1.40×10^{-5}
$\text{stuff} + A_3 \longrightarrow A_5 + A_2$	$6.30 \times 10^{-3} X_{A_3}$	$A_1 \longrightarrow \emptyset$	$1.40 \times 10^{-6} X_{A_1}$
$\text{stuff} + A_3 \longrightarrow A_4 + A_2$	$4.88 \times 10^{-3} X_{A_3}$	$A_7 \xrightarrow{A_4} A_6$	$1.42 \times 10^{-6} X_{A_4} X_{A_7}$
$\text{stuff} + A_3 \longrightarrow A_6 + A_2$	$4.88 \times 10^{-3} X_{A_3}$	$A_5 \longrightarrow \emptyset$	$1.80 \times 10^{-8} X_{A_5}$
$A_7 \longrightarrow A_2 + A_6$	$4.40 \times 10^{-4} X_{A_7}$	$A_6 \longrightarrow \emptyset$	$6.40 \times 10^{-10} X_{A_6}$
$A_2 + A_6 \longrightarrow A_7$	$3.62 \times 10^{-4} X_{A_2} X_{A_6}$	$A_4 \longrightarrow \emptyset$	$7.40 \times 10^{-11} X_{A_4}$

[4]

[3] J. Paulsson, Physics of Life Reviews, **2**, 2005 157 – 175.

[4] Hye Won Kang, presentation at SPA in 2007.

Outline

1. Why do people want to model some processes stochastically as opposed to deterministically?

2. → What are the basic models used?

3. How can we understand the behavior of these models both analytically and computationally?

Basic stochastic models of (bio)chemical reaction networks

Consider the simple system

$$A + B \rightarrow C$$

where one molecule each of A and B is being converted to one of C.

Basic stochastic models of (bio)chemical reaction networks

Consider the simple system

$$A + B \to C$$

where one molecule each of A and B is being converted to one of C.

Intuition for standard model is that the probability of reaction occurring in a small time interval $(t, t + \Delta t]$ should be

$$P\{\text{reaction occurs in } (t, t + \Delta t]\} \approx \kappa X_A(t) X_B(t) \Delta t$$

where

- κ is a positive constant, the reaction rate constant.

Models of interest

$$A + B \to C$$

Simple book-keeping: if $X(t) = (X_A(t), X_B(t), X_C(t))$ gives the state at time t then

$$X(t) = X(0) + R(t) \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}, \tag{2}$$

where

- $R(t)$ is the # of times the reaction has occurred by time t and
- $X(0)$ is the initial condition.

Models of interest

$$A + B \to C$$

Simple book-keeping: if $X(t) = (X_A(t), X_B(t), X_C(t))$ gives the state at time t then

$$X(t) = X(0) + R(t) \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}, \tag{2}$$

where

- $R(t)$ is the # of times the reaction has occurred by time t and
- $X(0)$ is the initial condition.

Note:

- $R(0) = 0$ and
- R is constant except for jumps of plus one.

Goal: represent R in terms of Poisson process.

The Poisson process

A Poisson process, Y, is a model for a series of random observations occurring in time.

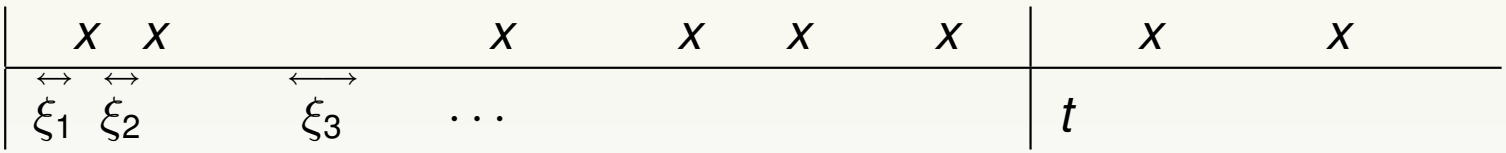
(a) Let $\{\xi_i\}$ be i.i.d. exponential random variables with parameter one.

The Poisson process

A Poisson process, Y, is a model for a series of random observations occurring in time.

(a) Let $\{\xi_i\}$ be i.i.d. exponential random variables with parameter one.

(b) Now, put points down on line with spacing equal to the ξ_i.



- Let $Y(t)$ denote the number of points hit by time t.
- In the figure above, $Y(t) = 6$.

Intuition: The unit rate Poisson process is simply the number of points hit when we run along the time frame at rate one.

The Poisson process

Let

- Y be a unit rate Poisson process.
- $Y_\lambda(t) \equiv Y(\lambda t)$,

Then Y_λ is a Poisson process with parameter λ.

Intuition: The Poisson process with rate λ is simply the number of points hit (of the unit-rate point process) when we run along the time frame at rate λ.

Thus, we have "changed time" to convert a unit-rate Poisson process to one which has rate λ.

The Poisson process

Let

- Y be a unit rate Poisson process.
- $Y_\lambda(t) \equiv Y(\lambda t)$,

Then Y_λ is a Poisson process with parameter λ.

Intuition: The Poisson process with rate λ is simply the number of points hit (of the unit-rate point process) when we run along the time frame at rate λ.

Thus, we have "changed time" to convert a unit-rate Poisson process to one which has rate λ.

There is no reason λ needs to be constant in time, in which case

$$Y_\lambda(t) \equiv Y\left(\int_0^t \lambda(s)ds\right).$$

Putting it all together

It turns out that

$$P\{Y_\lambda(t+\Delta t) - Y_\lambda(t) > 0\} \approx 1 - e^{-\lambda(t)\Delta t} \approx \lambda(t)\Delta t.$$

Putting it all together

It turns out that

$$P\{Y_\lambda(t+\Delta t) - Y_\lambda(t) > 0\} \approx 1 - e^{-\lambda(t)\Delta t} \approx \lambda(t)\Delta t.$$

Recall that for $A + B \longrightarrow C$ we wanted to model

$$P\{\text{reaction occurs in } (t, t+\Delta t]\} \approx \kappa X_A(t) X_B(t) \Delta t.$$

Putting it all together

It turns out that

$$P\{Y_\lambda(t+\Delta t) - Y_\lambda(t) > 0\} \approx 1 - e^{-\lambda(t)\Delta t} \approx \lambda(t)\Delta t.$$

Recall that for $A + B \longrightarrow C$ we wanted to model

$$P\{\text{reaction occurs in } (t, t+\Delta t]\} \approx \kappa X_A(t) X_B(t) \Delta t.$$

This suggests we can model

$$R(t) = Y\left(\int_0^t \lambda(s)ds\right) = Y\left(\int_0^t \kappa X_A(s) X_B(s) ds\right)$$

where Y is unit-rate Poisson process. This is similar to deterministic model!

Putting it all together

It turns out that

$$P\{Y_\lambda(t+\Delta t) - Y_\lambda(t) > 0\} \approx 1 - e^{-\lambda(t)\Delta t} \approx \lambda(t)\Delta t.$$

Recall that for $A + B \longrightarrow C$ we wanted to model

$$P\{\text{reaction occurs in } (t, t+\Delta t]\} \approx \kappa X_A(t) X_B(t) \Delta t.$$

This suggests we can model

$$R(t) = Y\left(\int_0^t \lambda(s)ds\right) = Y\left(\int_0^t \kappa X_A(s) X_B(s) ds\right)$$

where Y is unit-rate Poisson process. This is similar to deterministic model! Hence

$$\begin{pmatrix} X_A(t) \\ X_B(t) \\ X_C(t) \end{pmatrix} \equiv X(t) = X(0) + \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix} Y\left(\int_0^t \kappa X_A(s) X_B(s) ds\right).$$

This equation uniquely determines X for all $t \geq 0$.

Stochastic models of (bio)chemical reactions

- We consider a network of reactions involving d chemical species, $S_1, \ldots, S_d$:

$$\sum_{i=1}^{d} \nu_{ik} S_i \longrightarrow \sum_{i=1}^{d} \nu'_{ik} S_i$$

Stochastic models of (bio)chemical reactions

- We consider a network of reactions involving d chemical species, $S_1, \ldots, S_d$:

$$\sum_{i=1}^{d} \nu_{ik} S_i \longrightarrow \sum_{i=1}^{d} \nu'_{ik} S_i$$

- The state of the system, $X(t) \in \mathbb{Z}^d_{\geq 0}$, gives the number of molecules of each species in the system at time t.

- ν_k: vector giving number of molecules of each chemical species consumed in the kth reaction.

- ν'_k: vector giving number of molecules of each chemical species created in the kth reaction.

Stochastic models of (bio)chemical reactions

- If kth reaction occurs at time t, the new state becomes

$$X(t) = X(t-) + \nu'_k - \nu_k.$$

- The rate of kth reaction is $\lambda_k : \mathbb{Z}^d_{\geq 0} \to \mathbb{R}$.

- By analogy with before

$$X(t) = X(0) + \sum_k Y_k \left(\int_0^t \lambda_k(X(s))ds \right) (\nu'_k - \nu_k).$$

Mass-action kinetics

The standard intensity function chosen is mass-action kinetics:

$$\lambda_k(x) = \kappa_k(\prod_i \nu_{ik}!)\binom{x}{\nu_k} = \kappa_k \prod_i \frac{x_i!}{(x_i - \nu_{ik})!}.$$

Rate is proportional to the number of distinct subsets of the molecules present that can form inputs for the reaction. (this assumes vessel is "well-stirred".)

Example: If $S_1 \to$ anything, then $\lambda_k(x) = \kappa_k x_1$.

Example: If $S_1 + S_2 \to$ anything, then $\lambda_k(x) = \kappa_k x_1 x_2$.

Example: If $S_1 + 2S_2 \to$ anything, then $\lambda_k(x) = \kappa_k x_1 x_2(x_2 - 1)$.

Population Example: Lotka-Volterra predator-prey model

Think of A as a prey and B as a predator.

$$A \xrightarrow{\kappa_1} 2A, \qquad A + B \xrightarrow{\kappa_2} 2B, \qquad B \xrightarrow{\kappa_3} \emptyset,$$

with $A(0) = B(0) = 1000$ and $\kappa_1 = 2$, $\kappa_2 = .002$, $\kappa_3 = 2$.

Population Example: Lotka-Volterra predator-prey model

Think of A as a prey and B as a predator.

$$A \xrightarrow{\kappa_1} 2A, \qquad A + B \xrightarrow{\kappa_2} 2B, \qquad B \xrightarrow{\kappa_3} \emptyset,$$

with $A(0) = B(0) = 1000$ and $\kappa_1 = 2$, $\kappa_2 = .002$, $\kappa_3 = 2$.

Deterministic model. Let $x(t) = [A(t), B(t)]^T$.

$$x(t) = x(0) + \kappa_1 \int_0^t x_1(s)ds \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \kappa_2 \int_0^t x_1(s)x_2(s)ds \begin{bmatrix} -1 \\ 1 \end{bmatrix} + \kappa_3 \int_0^t x_2(s)ds \begin{bmatrix} 0 \\ -1 \end{bmatrix}$$

Population Example: Lotka-Volterra predator-prey model

Think of A as a prey and B as a predator.

$$A \xrightarrow{\kappa_1} 2A, \qquad A + B \xrightarrow{\kappa_2} 2B, \qquad B \xrightarrow{\kappa_3} \emptyset,$$

with $A(0) = B(0) = 1000$ and $\kappa_1 = 2$, $\kappa_2 = .002$, $\kappa_3 = 2$.

Deterministic model. Let $x(t) = [A(t), B(t)]^T$.

$$x(t) = x(0) + \kappa_1 \int_0^t x_1(s)ds \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \kappa_2 \int_0^t x_1(s)x_2(s)ds \begin{bmatrix} -1 \\ 1 \end{bmatrix} + \kappa_3 \int_0^t x_2(s)ds \begin{bmatrix} 0 \\ -1 \end{bmatrix}$$

Stochastic model. Let $X(t) = [A(t), B(t)]^T$.

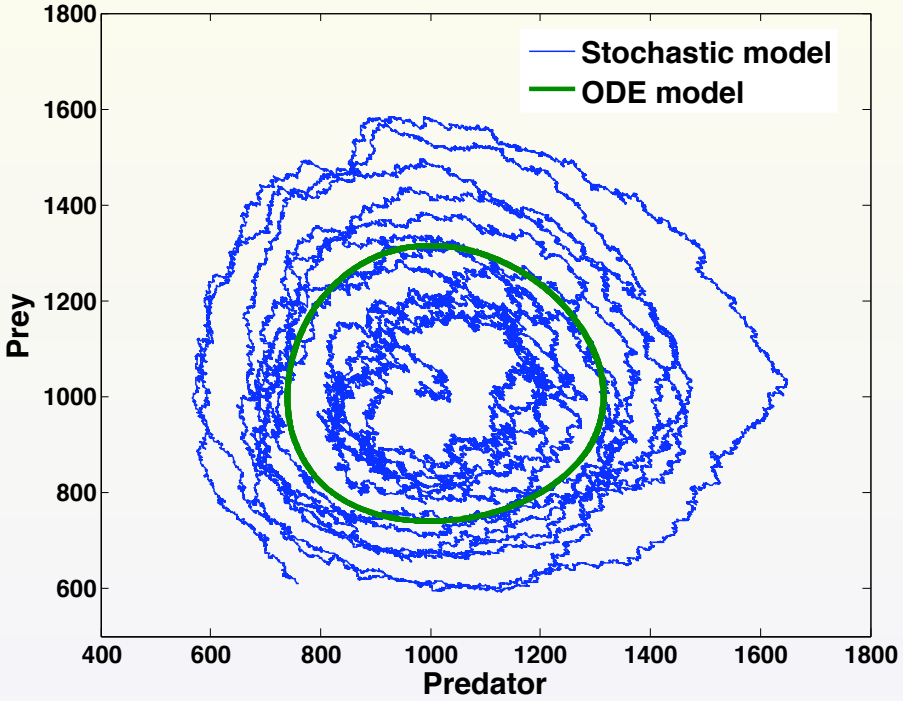
$$\begin{aligned} X(t) = X(0) &+ Y_1\left(\kappa_1 \int_0^t X_1(s)ds\right) \begin{bmatrix} 1 \\ 0 \end{bmatrix} + Y_2\left(\kappa_2 \int_0^t X_1(s)X_2(s)ds\right) \begin{bmatrix} -1 \\ 1 \end{bmatrix} \\ &+ Y_3\left(\kappa_3 \int_0^t X_2(s)ds\right) \begin{bmatrix} 0 \\ -1 \end{bmatrix} \end{aligned}$$

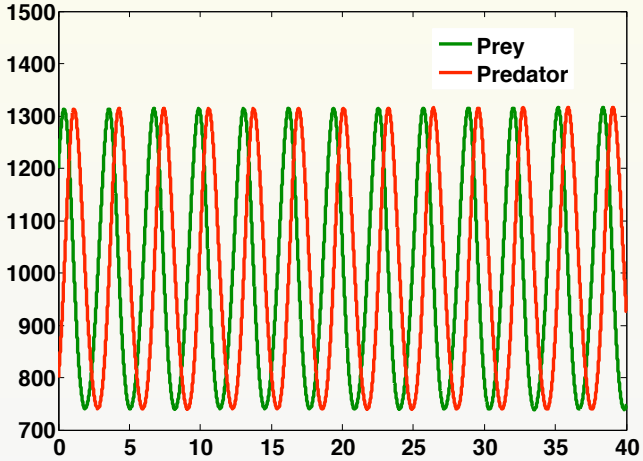
Lotka-Volterra

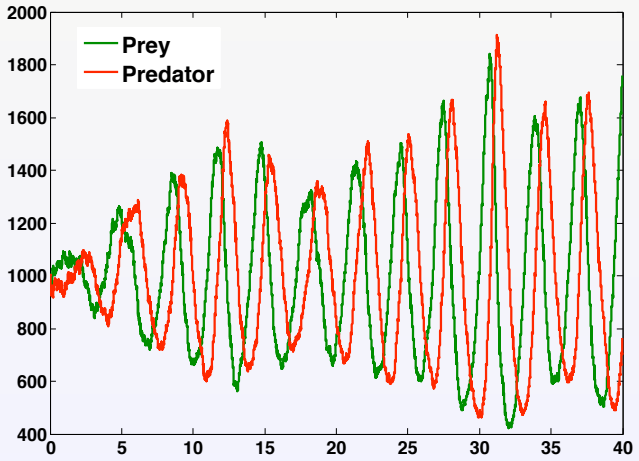
Think of A as a prey and B as a predator.

$$A \xrightarrow{\kappa_1} 2A, \qquad A + B \xrightarrow{\kappa_2} 2B, \qquad B \xrightarrow{\kappa_3} \emptyset,$$

with $A(0) = B(0) = 1000$ and $\kappa_1 = 2$, $\kappa_2 = .002$, $\kappa_3 = 2$.







Outline

1. Why do people want to model some processes stochastically as opposed to deterministically?

2. What are the basic models used?

3. → How can we understand the behavior of these models both analytically and computationally?

How can the models be understood?

1. Understand how the distribution of the process behaves.
2. Understand how paths behave.

How can the models be understood?

1. Understand how the distribution of the process behaves.
2. Understand how paths behave.

Kolmogorov's forward equation ("Chemical Master Equation") describes the evolution of the distribution of the state of the system

$$\frac{d}{dt}P(x,t) = \sum_k \lambda_k(x - \nu'_k + \nu_k)P(x - \nu'_k + \nu_k, t) - \sum_k \lambda_k(x)P(x,t),$$

where $P(x,t)$ is probability $X(t) = x$.

How can the models be understood?

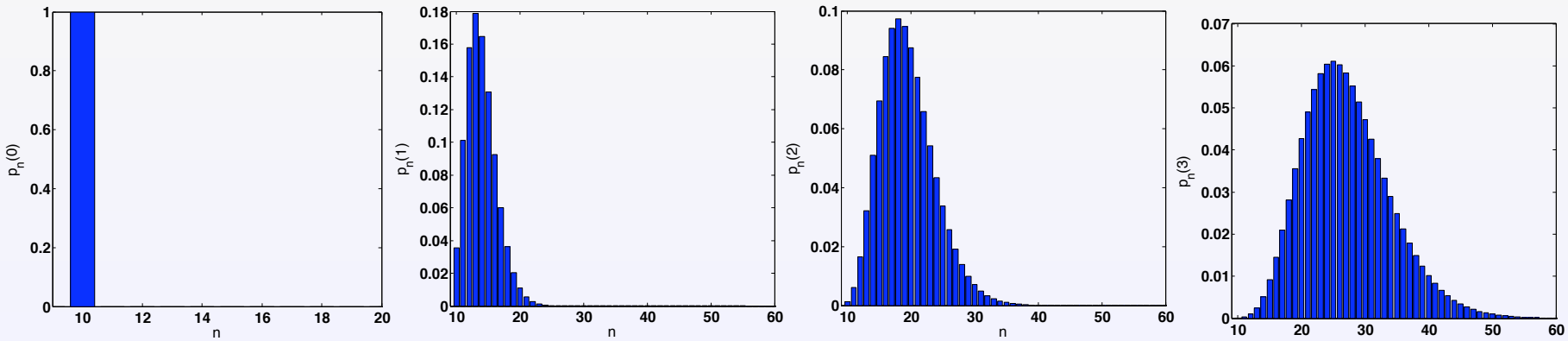
1. Understand how the distribution of the process behaves.
2. Understand how paths behave.

Kolmogorov's forward equation ("Chemical Master Equation") describes the evolution of the distribution of the state of the system

$$\frac{d}{dt}P(x,t) = \sum_k \lambda_k(x - \nu_k' + \nu_k)P(x - \nu_k' + \nu_k, t) - \sum_k \lambda_k(x)P(x,t),$$

where $P(x,t)$ is probability $X(t) = x$.

$$\frac{dP_n(t)}{dt} = (1/3)(n-1)P_{n-1}(t) - (1/3)nP_n(t),$$



Example: Stationary distributions

$$\frac{d}{dt}P(x,t) = \sum_k \lambda_k(x - \nu_k' + \nu_k)P(x - \nu_k' + \nu_k, t) - \sum_k \lambda_k(x)P(x,t),$$

Large (oftentimes infinite), linear system: very difficult to solve.

Example: Stationary distributions

$$\frac{d}{dt}P(x,t) = \sum_k \lambda_k(x - \nu_k' + \nu_k)P(x - \nu_k' + \nu_k, t) - \sum_k \lambda_k(x)P(x,t),$$

Large (oftentimes infinite), linear system: very difficult to solve.

- Shows stationary distribution (analog of a fixed point) satisfies

$$0 = \sum_k \lambda_k(x - \nu_k' + \nu_k)\pi(x - \nu_k' + \nu_k) - \sum_k \lambda_k(x)\pi(x).$$

1. These can sometimes be found and can determine “large time behavior.”

Example: Stationary distributions

$$\frac{d}{dt}P(x,t) = \sum_k \lambda_k(x - \nu_k' + \nu_k)P(x - \nu_k' + \nu_k, t) - \sum_k \lambda_k(x)P(x,t),$$

Large (oftentimes infinite), linear system: very difficult to solve.

- Shows stationary distribution (analog of a fixed point) satisfies

$$0 = \sum_k \lambda_k(x - \nu_k' + \nu_k)\pi(x - \nu_k' + \nu_k) - \sum_k \lambda_k(x)\pi(x).$$

1. These can sometimes be found and can determine "large time behavior."
2. Mathematicians spend time thinking of ways to connect network structure with existence/form of stationary distribution.
3. Example (A., Craciun, Kurtz, Bull. Math. Biol. 2010):
   weakly reversible + deficiency zero $\implies$ $\pi(x)$ is of very special form:
   - Product form.
   - Product of Poisson's.

$$\pi(x) = \prod_{i=1}^{d} \frac{c_i^{x_i}}{x_i!} e^{-c_i}, \qquad x \in \mathbb{Z}_{\geq 0}^d.$$

Point of analysis: result will be useful in myriad applications.

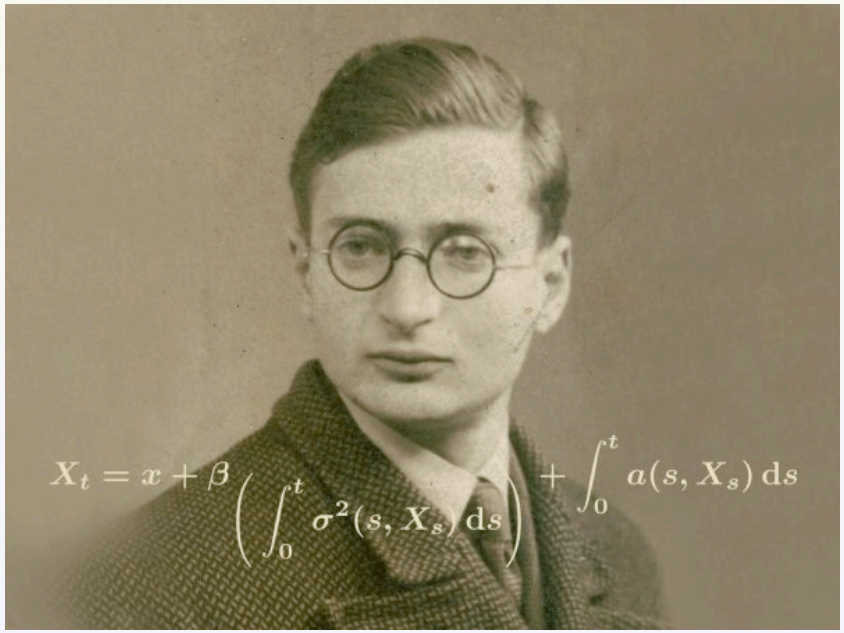
Pathwise Representations – Random time changes

A representation for path-wise solutions of our model is given by random time-changes of Poisson processes

$$X(t) = X(0) + \sum_k Y_k \left(\int_0^t \lambda_k(X(s))ds \right) (\nu'_k - \nu_k),$$

where the Y_k are independent, unit-rate Poisson processes.

Random time changes have interesting history:



(Wolfgang Döblin)

Methods of investigation: numerical simulation

$$X(t) = X(0) + \sum_k Y_k \left(\int_0^t \lambda_k(X(s))ds \right) (\nu_k' - \nu_k),$$

(GOOD NEWS) There are a number of numerical methods that produce statistically exact sample paths:

1. Gillespie's algorithm.
2. The first reaction method.
3. The next reaction method.

For each step of these methods one must find :

(i) the amount of time that passes until the next reaction takes place:

$$\Delta_n \sim exp\left(\sum_k \lambda_k(X(t)) \right)$$

(the minimum of exponential RVs)

(ii) which reaction takes place at that time.

Methods of investigation: numerical simulation

$$X(t) = X(0) + \sum_k Y_k \left(\int_0^t \lambda_k(X(s))ds \right) (\nu_k' - \nu_k),$$

(GOOD NEWS) There are a number of numerical methods that produce statistically exact sample paths:

1. Gillespie's algorithm.
2. The first reaction method.
3. The next reaction method.

For each step of these methods one must find :

(i) the amount of time that passes until the next reaction takes place:

$$\Delta_n \sim exp\left(\sum_k \lambda_k(X(t)) \right)$$

(the minimum of exponential RVs)

(ii) which reaction takes place at that time.

(BAD NEWS) If $\sum_k \lambda_k(X(t)) \gg 1$, then $\Delta_n \approx \frac{1}{\sum_k \lambda_k(X(t))} \ll 1$

$*$ time to produce a single path over an interval $[0, T]$ can be prohibitive.

Tau-leaping

Explicit "τ-leaping" [5] was developed by Dan Gillespie in an effort to overcome the problem that Δ_n may be prohibitively small.

Tau-leaping is essentially an Euler approximation of $\int_0^t \lambda_k(X(s))ds$:

[5] D. T. Gillespie, J. Chem. Phys., **115**, 1716 – 1733.

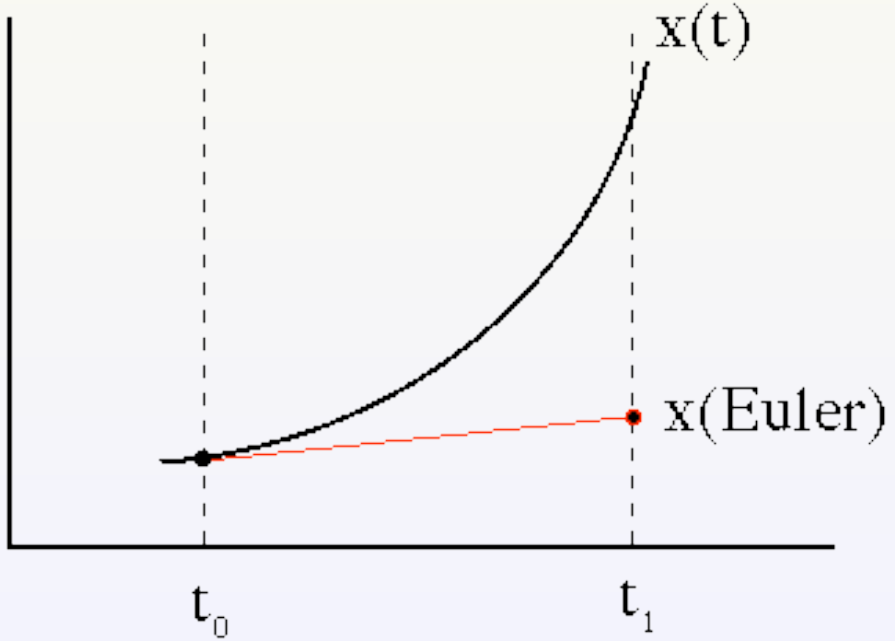
Tau-leaping

Explicit "τ-leaping" [5] was developed by Dan Gillespie in an effort to overcome the problem that Δ_n may be prohibitively small.

Tau-leaping is essentially an Euler approximation of $\int_0^t \lambda_k(X(s))ds$:

Say: $x(t) = x(t_0) + \int_{t_0}^t x(s)ds$ (i.e. $\frac{d}{dt}x(t) = x(t)$)

Use approximation: $\tilde{x}(t) = x(t_0) + (t - t_0)x(t_0)$.



[5] D. T. Gillespie, J. Chem. Phys., **115**, 1716 – 1733.

Tau-leaping

Explicit "τ-leaping" [6] was developed by Dan Gillespie in an effort to overcome the problem that Δ_n may be prohibitively small.

Tau-leaping is essentially an Euler approximation of $\int_0^t \lambda_k(X(s))ds$:

$$
\begin{aligned}
Z(\tau) &= Z(0) + \sum_k Y_k \left(\int_0^\tau \lambda_k(Z(s))\, ds \right) (\nu_k' - \nu_k) \\
&\approx Z(0) + \sum_k Y_k \left(\int_0^\tau \lambda_k(Z(0))\, ds \right) (\nu_k' - \nu_k) \\
&= Z(0) + \sum_k Y_k \Big(\lambda_k(Z(0))\, \tau \Big) (\nu_k' - \nu_k) \\
&\stackrel{d}{=} Z(0) + \sum_k \text{Poisson} \Big(\lambda_k(Z(0))\, \tau \Big) (\nu_k' - \nu_k).
\end{aligned}
$$

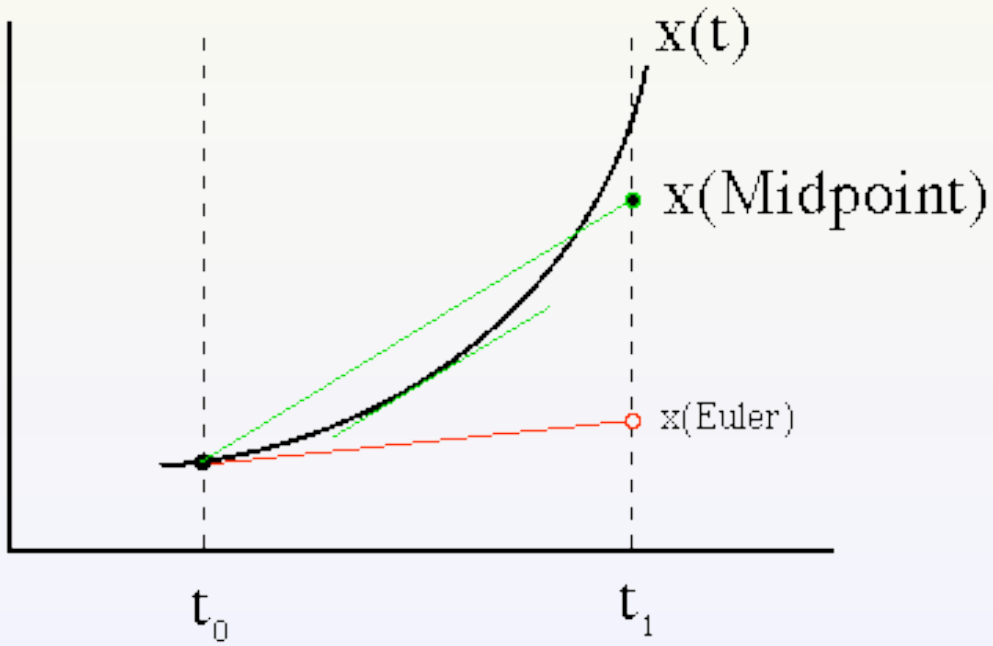
[6] D. T. Gillespie, J. Chem. Phys., **115**, 1716 – 1733.

Another algorithm: A midpoint method

For a time discretization $0 = t_0 < t_1 < \cdots < t_N = T$, with $\tau = t_n - t_{n-1}$, let

$$\rho(z) = z + \frac{1}{2}\tau \sum_k \lambda_k(z)(\nu_k' - \nu_k),$$

be a "deterministic" midpoint approximation



Another algorithm: A midpoint method

For a time discretization $0 = t_0 < t_1 < \cdots < t_N = T$, with $\tau = t_n - t_{n-1}$, let

$$\rho(z) = z + \frac{1}{2}\tau \sum_k \lambda_k(z)(\nu_k' - \nu_k),$$

be a "deterministic" midpoint approximation

and let $\mathcal{Z}(t)$ solve:

$$\begin{aligned}
\mathcal{Z}(\tau) &= \mathcal{Z}(0) + \sum_k Y_k \left(\int_0^\tau \lambda_k(\mathcal{Z}(s))\, ds \right) (\nu_k' - \nu_k) \\
&\approx \mathcal{Z}(0) + \sum_k Y_k \left(\int_0^\tau \lambda_k(\rho(\mathcal{Z}(0)))\, ds \right) (\nu_k' - \nu_k) \\
&\stackrel{d}{=} \mathcal{Z}(0) + \sum_k \text{Poisson}\Big(\lambda_k(\rho(\mathcal{Z}(0)))\, \tau \Big) (\nu_k' - \nu_k).
\end{aligned}$$

Error analysis

Under the scaling $\tau \longrightarrow 0$:

1. Li[7] and also Rathinam, Petzold, Cao, and Gillespie[8] showed Euler tau-leaping is a first order method.

2. The midpoint method has the same order of accuracy as explicit Euler tau-leaping as $\tau \longrightarrow 0$.

[7] T. Li, SIAM Multi. Model. Simul., **6**, 2007, 417 – 436.

[8] M. Rathinam et al., SIAM Multi. Model. Simul., **4**, 2005, 867 – 895.

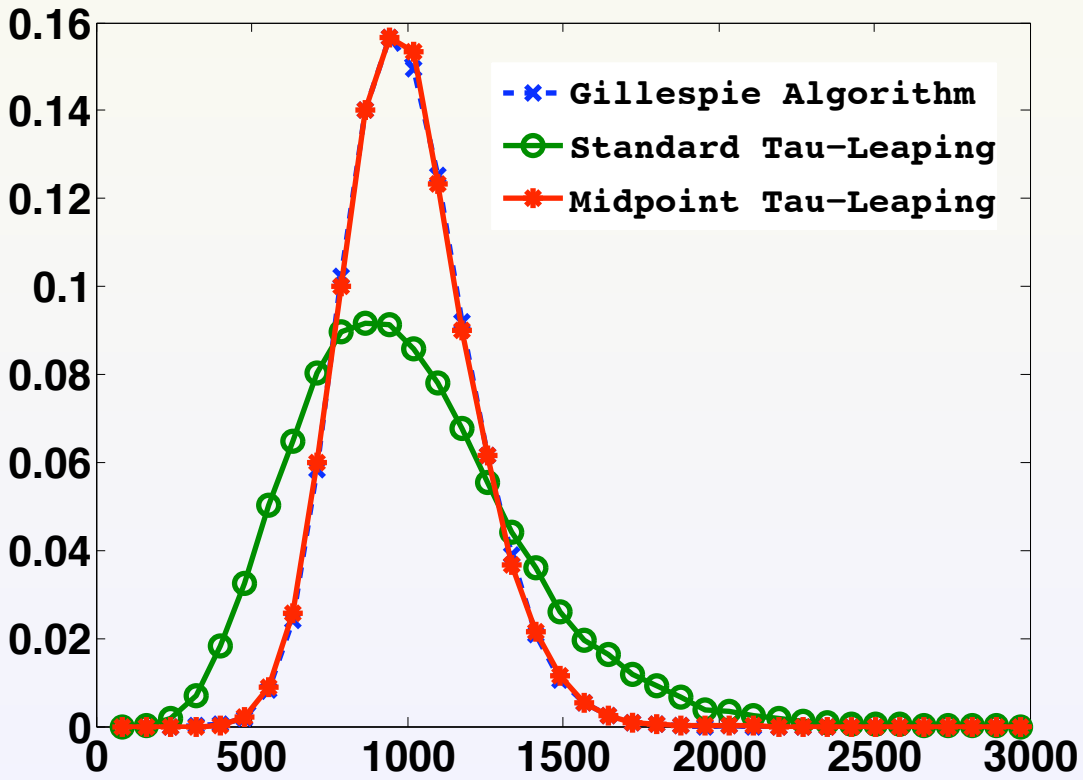
Example

Again think of A as a prey and B as a predator.

$$A \xrightarrow{\kappa_1} 2A, \qquad A + B \xrightarrow{\kappa_2} 2B, \qquad B \xrightarrow{\kappa_3} \emptyset,$$

with $A(0) = B(0) = 1000$ and $\kappa_1 = 2$, $\kappa_2 = .002$, $\kappa_3 = 2$.

Letting $\tau = 1/20$ and simulating 30,000 sample paths with each method yields the following approximate distributions for $B(10)$:



Another perspective

Recall, tau-leaping methods are used when $\tau \gg \Delta_n$, for otherwise an exact method would be performed. Therefore, we should require that

$$\tau \gg \frac{1}{\sum_k \lambda_k(X(t))} \approx \Delta_n \qquad \text{while} \qquad \sum_k \lambda_k(X(t)) \gg 1.$$

So $\tau \longrightarrow 0$ does not tell the whole story.

Another perspective

Recall, tau-leaping methods are used when $\tau \gg \Delta_n$, for otherwise an exact method would be performed. Therefore, we should require that

$$\tau \gg \frac{1}{\sum_k \lambda_k(X(t))} \approx \Delta_n \qquad \text{while} \qquad \sum_k \lambda_k(X(t)) \gg 1.$$

So $\tau \longrightarrow 0$ does not tell the whole story.

Begs the question: how else can one perform a useful error analyses?

Another perspective

Recall, tau-leaping methods are used when $\tau \gg \Delta_n$, for otherwise an exact method would be performed. Therefore, we should require that

$$\tau \gg \frac{1}{\sum_k \lambda_k(X(t))} \approx \Delta_n \qquad \text{while} \qquad \sum_k \lambda_k(X(t)) \gg 1.$$

So $\tau \longrightarrow 0$ does not tell the whole story.

Begs the question: how else can one perform a useful error analyses?

- Perform non-standard error analysis. Take natural scales into account.

Another perspective

Take natural scales into account. Suppose that

(*i*) Numbers of molecules $X_i^V = \mathcal{O}(V)$ for some V large (100's, 1000's, ...).

(*ii*) $\lambda_k(X^V(t)) = \mathcal{O}(V)$.

(*iii*) Timestep not too small.

Another perspective

Take natural scales into account. Suppose that

(*i*) Numbers of molecules $X_i^V = \mathcal{O}(V)$ for some V large (100's, 1000's, ...).

(*ii*) $\lambda_k(X^V(t)) = \mathcal{O}(V)$.

(*iii*) Timestep not too small.

Can prove that (A., Ganguly, Kurtz, 2011, to appear in Ann. of Appl. Prob.):

1. Euler $\tau-$leaping is a first order method.
2. Midpoint $\tau-$leaping is a second order method.

Another perspective

Take natural scales into account. Suppose that

(*i*) Numbers of molecules $X_i^V = \mathcal{O}(V)$ for some V large (100's, 1000's, ...).

(*ii*) $\lambda_k(X^V(t)) = \mathcal{O}(V)$.

(*iii*) Timestep not too small.

Can prove that (A., Ganguly, Kurtz, 2011, to appear in Ann. of Appl. Prob.):

1. Euler $\tau-$leaping is a first order method.
2. Midpoint $\tau-$leaping is a second order method.

Proof makes explicit use of representation and uses facts of Poisson process:

$$\overline{X}^V(t) = \overline{X}^V(0) + \sum_k \frac{1}{V} Y_k \left(V \int_0^t \overline{\lambda}_k(\overline{X}^V(s))ds \right) (\nu_k' - \nu_k)$$

Note: this informs what algorithm should be used in what situation.

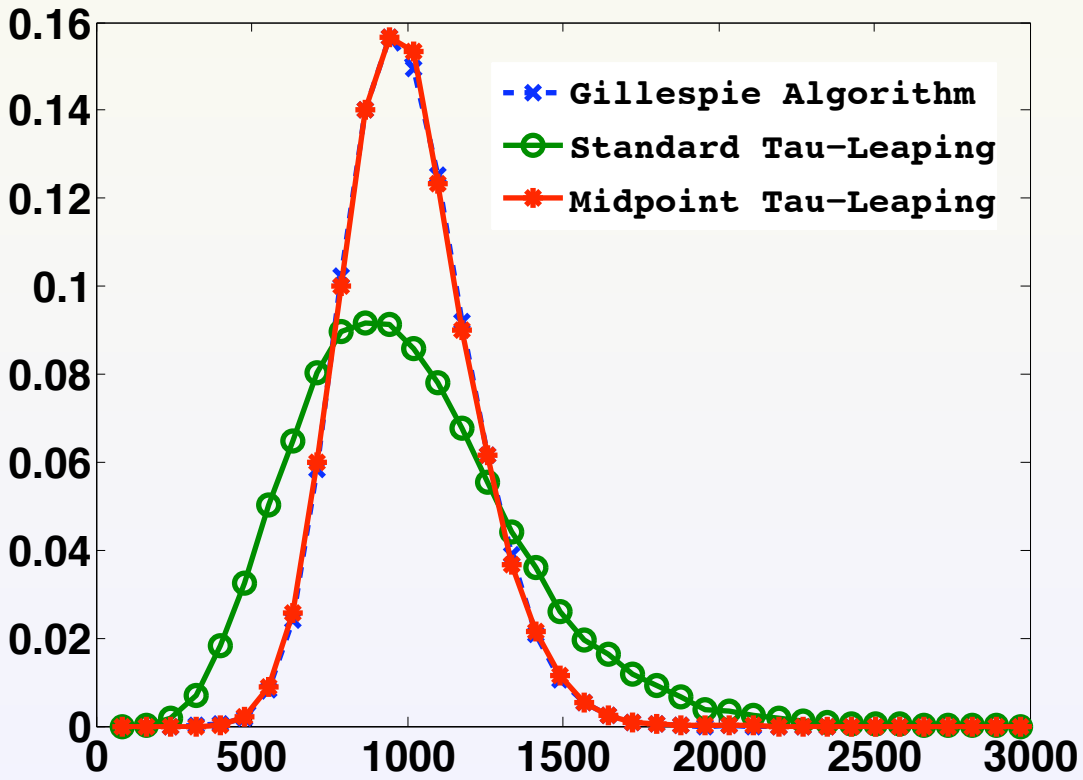
Example

Consider A as a prey and B as a predator.

$$A \xrightarrow{\kappa_1} 2A, \qquad A + B \xrightarrow{\kappa_2} 2B, \qquad B \xrightarrow{\kappa_3} \emptyset,$$

with $A(0) = B(0) = 1000$ and $\kappa_1 = 2$, $\kappa_2 = .002$, $\kappa_3 = 2$.

$V = 1{,}000$ and $\tau = 1/20 = 1/V^{.434}$. Simulating 30,000 sample paths yields the following approximate distributions for $B(10)$:



Post-leap check: Poisson Bridge

What else can be done with good representation?

Problem: $\tau-$leaping can lead to negative molecular counts.

[8] David F. Anderson, Incorporating postleap checks in tau-leaping, J. Chem. Phys., 2008

Post-leap check: Poisson Bridge

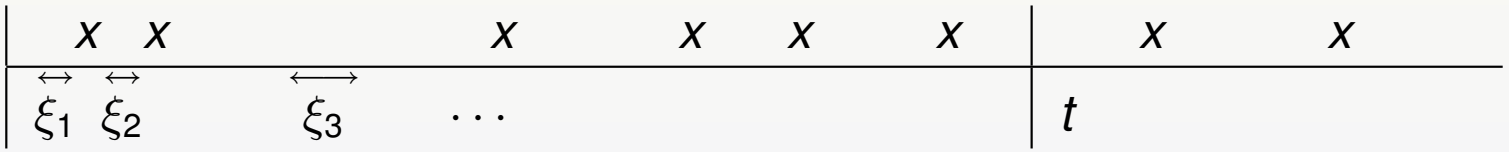
What else can be done with good representation?

Problem: $\tau-$leaping can lead to negative molecular counts.

Solution: Postleap check or Poisson bridge. Main ideas:

- Given $Y_k(t)$ and $Y_k(s)$ with $s < t$, for $r \in (s, t)$

$$Y_k(r) = \text{Binomial}\left(Y_k(t) - Y_k(s), \frac{r-s}{t-s}\right).$$



[8] David F. Anderson, Incorporating postleap checks in tau-leaping, J. Chem. Phys., 2008

Post-leap check: Poisson Bridge

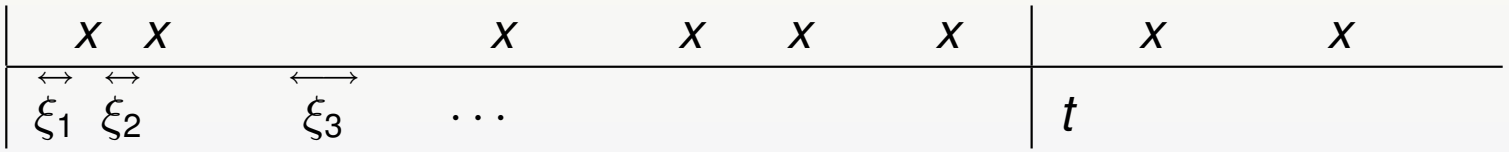
What else can be done with good representation?

Problem: $\tau-$leaping can lead to negative molecular counts.

Solution: Postleap check or Poisson bridge. Main ideas:

- Given $Y_k(t)$ and $Y_k(s)$ with $s < t$, for $r \in (s, t)$

$$Y_k(r) = \text{Binomial}\left(Y_k(t) - Y_k(s), \frac{r-s}{t-s}\right).$$



Som benefits:

1. Guaranteed that any "leap condition" can be enforced with prob. 1.
2. Easy to implement.
3. Naturally avoids negative population numbers without biasing Y_k.

[8] David F. Anderson, Incorporating postleap checks in tau-leaping, J. Chem. Phys., 2008

Some conclusions

1. Simulation of continuous time Markov chains is easy.

Some conclusions

1. Simulation of continuous time Markov chains is easy.....unless it's not.

2. Mathematics has a role to play:

   2.1 Error analysis informs choice of algorithm in different setting.

   2.2 Post-leap checking serves as core of error reduction strategy in multiple algorithms.

---

3. Stochastic models do have a place in biology.

4. Mathematical tools just as sophisticated for stochastic as opposed to deterministic models.

Some conclusions

1. Simulation of continuous time Markov chains is easy.....unless it's not.

2. Mathematics has a role to play:

   2.1 Error analysis informs choice of algorithm in different setting.

   2.2 Post-leap checking serves as core of error reduction strategy in multiple algorithms.

---

3. Stochastic models do have a place in biology.

4. Mathematical tools just as sophisticated for stochastic as opposed to deterministic models.

5. What can be said about different type of modes? Langevin?

Diffusion/Langevin Approximation

Classical result:

$$\frac{1}{V} Y(Vu) \approx u + \frac{1}{\sqrt{V}} W(u) \quad \text{for large } V,$$

where W is a Brownian motion (Functional LLN and CLT).

Diffusion/Langevin Approximation

Classical result:

$$\frac{1}{V} Y(Vu) \approx u + \frac{1}{\sqrt{V}} W(u) \quad \text{for large } V,$$

where W is a Brownian motion (Functional LLN and CLT).

Under the classical scaling, can use diffusion approximation

$$\begin{aligned} X(t) &= X(0) + \frac{1}{V} \sum_k Y_k \left(V \int_0^t \lambda(X(s)) ds \right) (\nu_k' - \nu_k) \\ &\approx X(0) + \sum_k \int_0^t \lambda(X(s)) ds \, (\nu_k' - \nu_k) + \frac{1}{\sqrt{V}} \sum_k W_k \left(\int_0^t \lambda_k(X(s)) ds \right) (\nu_k' - \nu_k). \end{aligned}$$

Diffusion/Langevin Approximation

Classical result:

$$\frac{1}{V} Y(Vu) \approx u + \frac{1}{\sqrt{V}} W(u) \quad \text{for large } V,$$

where W is a Brownian motion (Functional LLN and CLT).

Under the classical scaling, can use diffusion approximation

$$\begin{aligned} X(t) &= X(0) + \frac{1}{V} \sum_k Y_k \left(V \int_0^t \lambda(X(s)) ds \right) (\nu_k' - \nu_k) \\ &\approx X(0) + \sum_k \int_0^t \lambda(X(s)) ds \, (\nu_k' - \nu_k) + \frac{1}{\sqrt{V}} \sum_k W_k \left(\int_0^t \lambda_k(X(s)) ds \right) (\nu_k' - \nu_k). \end{aligned}$$

Hence can approximate X with solution to

$$dZ(t) = \sum_k \lambda_k(Z(t)) \, (\nu_k' - \nu_k) \, dt + \frac{1}{\sqrt{V}} \sum_k (\nu_k' - \nu_k) \sqrt{\lambda(Z(t))} \, dW_k(t).$$

Diffusion/Langevin Approximation

Classical result:

$$\frac{1}{V} Y(Vu) \approx u + \frac{1}{\sqrt{V}} W(u) \quad \text{for large } V,$$

where W is a Brownian motion (Functional LLN and CLT).

Under the classical scaling, can use diffusion approximation

$$\begin{aligned} X(t) &= X(0) + \frac{1}{V} \sum_k Y_k \left(V \int_0^t \lambda(X(s)) ds \right) (\nu_k' - \nu_k) \\ &\approx X(0) + \sum_k \int_0^t \lambda(X(s)) ds \, (\nu_k' - \nu_k) + \frac{1}{\sqrt{V}} \sum_k W_k \left(\int_0^t \lambda_k(X(s)) ds \right) (\nu_k' - \nu_k). \end{aligned}$$

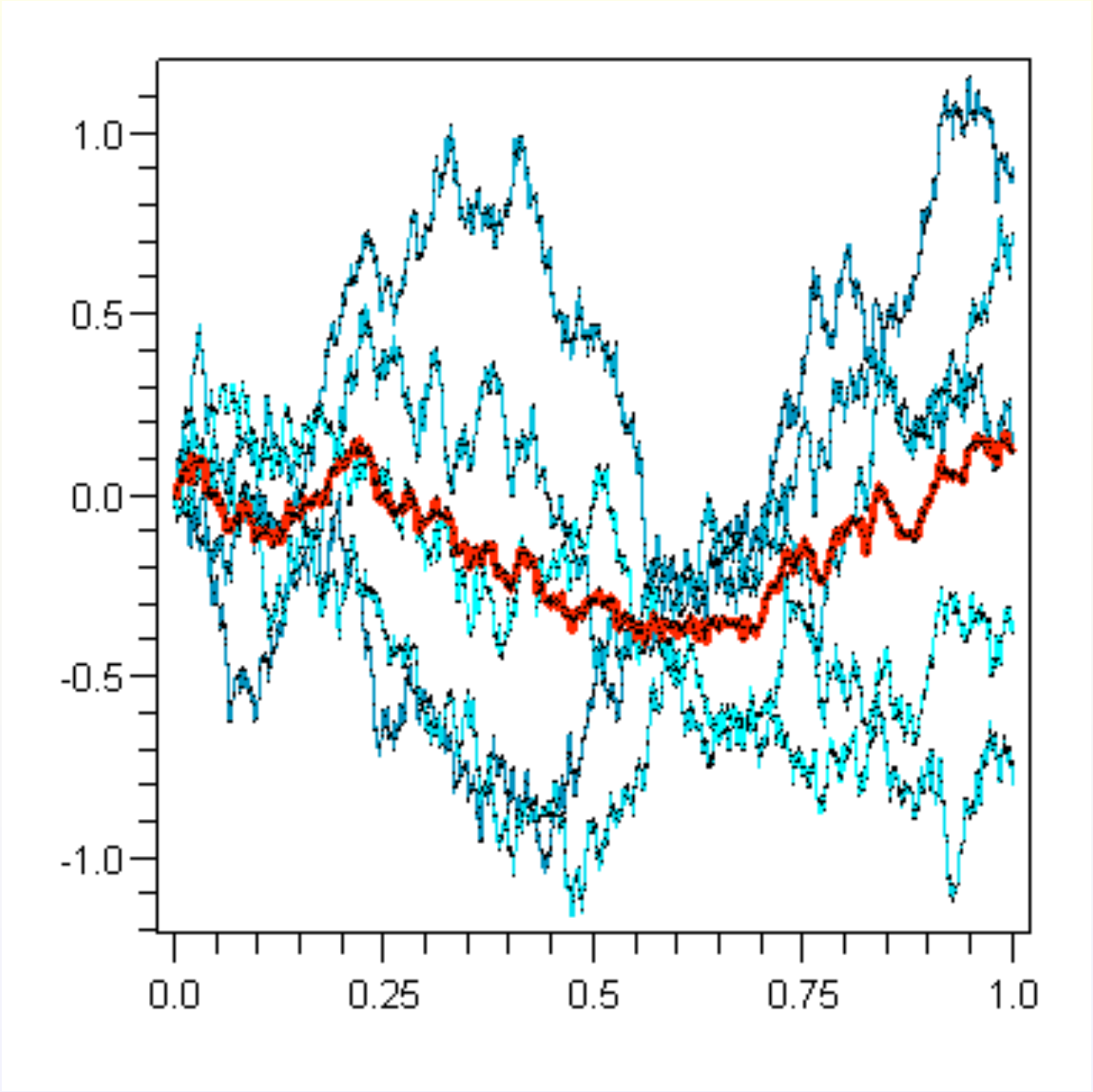
Hence can approximate X with solution to

$$dZ(t) = \sum_k \lambda_k(Z(t)) \, (\nu_k' - \nu_k) \, dt + \frac{1}{\sqrt{V}} \sum_k (\nu_k' - \nu_k) \sqrt{\lambda(Z(t))} \, dW_k(t).$$

This is:

- Continuous.
- Used *very often*.

Brownian motions



Want a good algorithm: change representation!

System

$$dX(t) = \int_0^t b(X(s))ds + \sum_k \int_0^t \sigma_k(X(s))\ \nu_k\ dW_k(s),$$

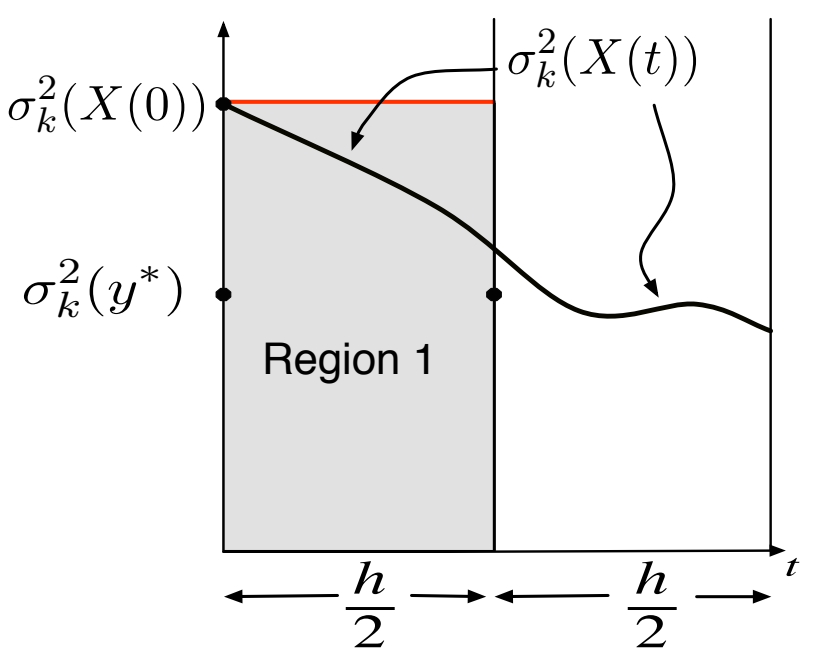
is equivalent to

$$X(t) = X(0) + \int_0^t b(X(s))ds + \sum_k \nu_k \int_0^\infty \int_0^t 1_{[0,\sigma_k^2(X(s)))}(u)\tilde{W}_k(ds \times du),$$

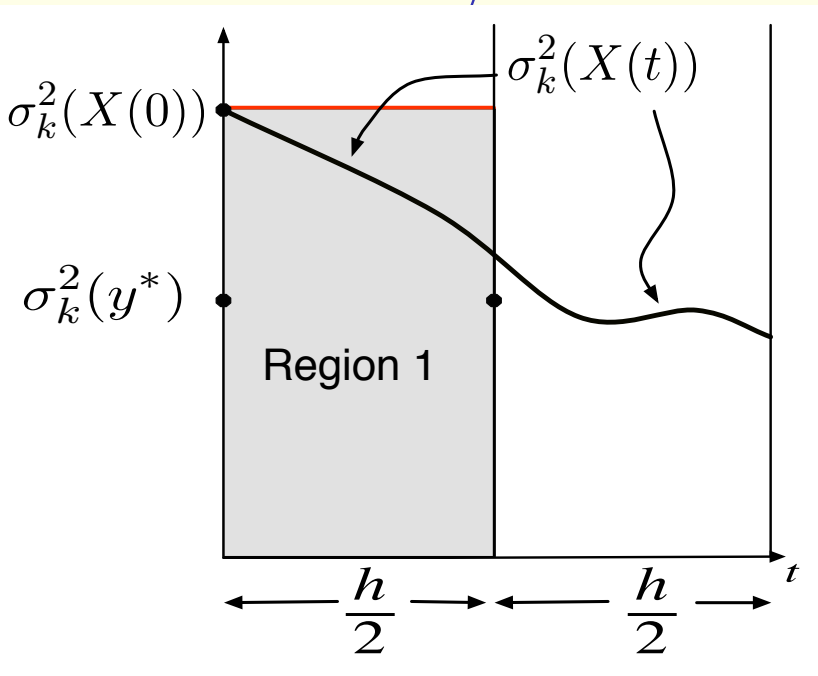
where the $\tilde{W}_k$ are independent space-time white noise processes.

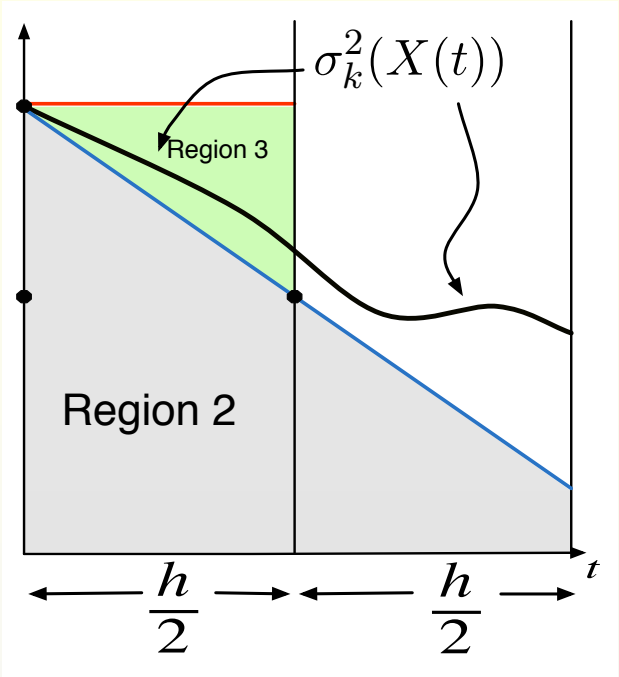
Challenge is in approximating diffusion term.

Why this works: $\theta = 1/2$

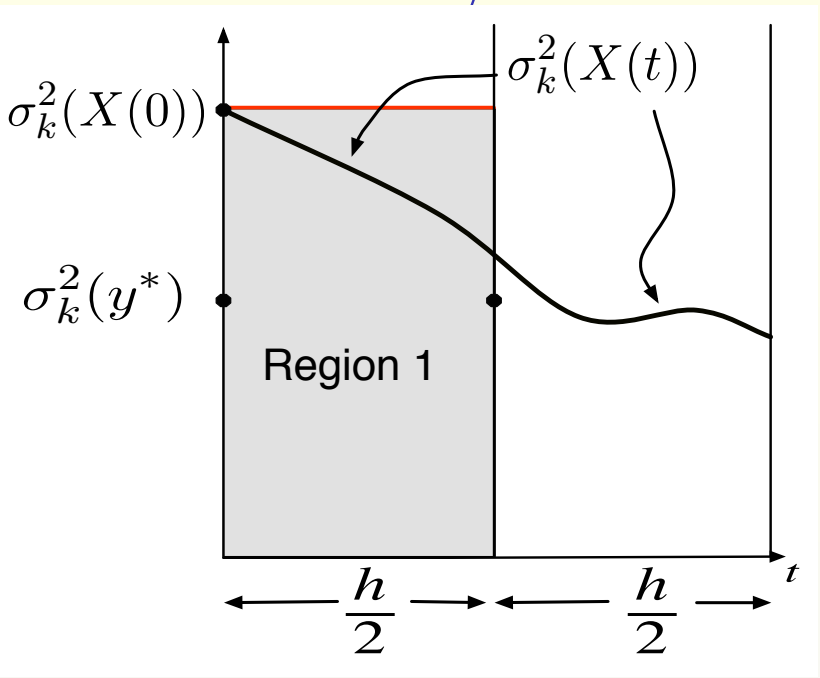


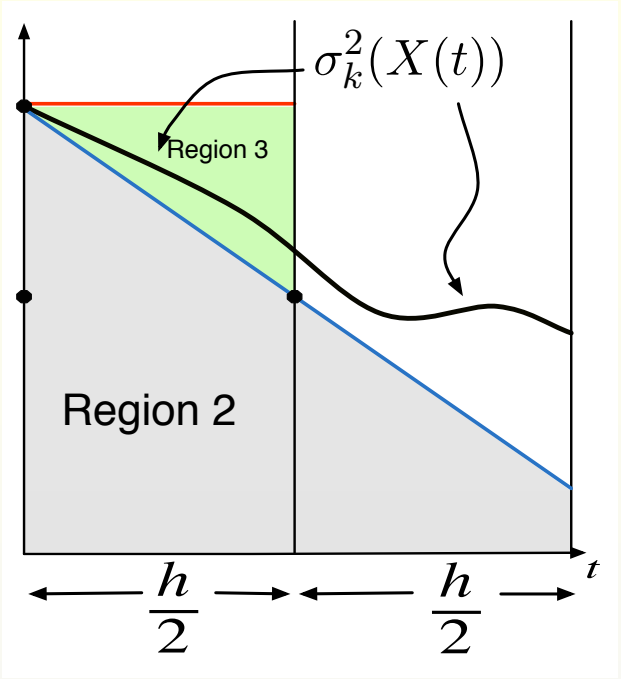
Why this works: $\theta = 1/2$

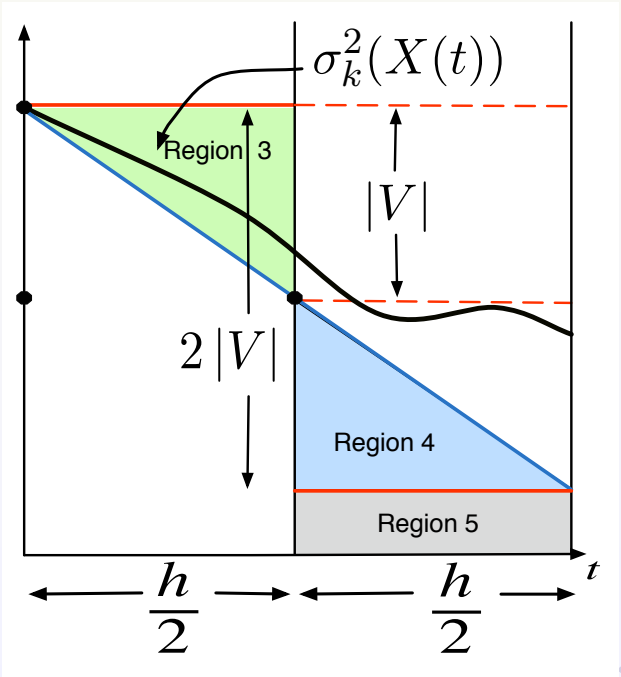




Why this works: $\theta = 1/2$







Approximating the diffusion approximation

Let approximate path be Y_i. The following is 2nd order accurate in weak sense.

ALGORITHM [1] (D. Anderson and J. Mattingly, 2011). *Fixing a* $\theta \in (0,1)$*, we define*

$$\alpha_1 \stackrel{def}{=} \frac{1}{2}\frac{1}{\theta(1-\theta)} \qquad \text{and} \quad \alpha_2 \stackrel{def}{=} \frac{1}{2}\frac{(1-\theta)^2+\theta^2}{\theta(1-\theta)}\,. \tag{3}$$

Next fixing a discretization step h, for each $i \in \{1,2,3,\dots\}$ *we repeat the following steps in which we first compute a* θ*-midpoint* y^* *and then the new value* Y_i*:*

Step 1. *Set*

$$y^* = Y_{i-1} + b(Y_{i-1})\theta h + \sum_k \sigma_k(Y_{i-1})\,\nu_k\,\eta_{1k}^{(i)}\sqrt{\theta h}$$

Step 2. *Set*

$$Y_i = y^* + (\alpha_1 b(y^*) - \alpha_2 b(Y_{i-1}))(1-\theta)h + \sum_k \sqrt{\left[\alpha_1\sigma_k^2(y^*) - \alpha_2\sigma_k^2(Y_{i-1})\right]^+}\,\nu_k\,\eta_{2k}^{(i)}\sqrt{(1-\theta)h}.$$

[1] D. F. Anderson and J. C. Mattingly, Comm. Math. Sci., 2011