

Bioinformatics 3

# V23 – Dynamic Modelling: Examples

Tue, Jan. 24, 2012

# Mass Action Kinetics

Again:  $A + B \rightleftharpoons AB$

Objective: mathematical description of (= formulas for)  
the changes of  $[A]$ ,  $[B]$ , and  $[AB]$

Consider  $[A]$ :

Gain from dissociation  $AB \Rightarrow A + B$

Loss from association  $A + B \Rightarrow AB$

$$\frac{d}{dt}[A] = G_A - L_A$$

$AB$  falls apart

$\Rightarrow G_A$  depends only on  $[AB]$

$A$  has to find  $B$

$\Rightarrow L_A$  depends on  $[A]$  and  $[B]$

$$G_A = k_r [AB]$$

phenomenological  
proportionality constant

$$L_A = k_f [A] [B]$$

$$\frac{d}{dt}[A] = k_r [AB] - k_f [A] [B]$$

# Dynamics Are Different

## **Continuous/deterministic:**

- integration: doing better than Euler
- effective vs. explicit models (Michaelis-Menten)
- pathway: static vs. dynamic behavior, length

## **Stochastic:**

- the know-all ansatz: Master equation
- fluctuations in Lotka-Volterra
- graded response via fluctuations

# Forward-Euler Integration Scheme

Linear approximation to (true)  $A(t)$ :

$$\begin{aligned} A(t) &\approx A(0) + t \cdot \frac{dA}{dt}(0) + O(t^2) \\ &\approx A(0) + t \cdot f(A(0), B(0)) + O(t^2) \end{aligned}$$



initial condition      increment      error

For  $t \rightarrow 0$ :

$$t \cdot \frac{dA}{dt}(0) \gg \frac{t^2}{2} \cdot \frac{d^2A}{dt^2}(0) \gg \dots$$

Use linear approximation for small time step  $\Delta t$ :

$$A(t + \Delta t) = A(t) + \Delta t \cdot \frac{dA}{dt}(t)$$

"forward Euler" algorithm

# Euler-Problems: Lotka–Volterra

Reactions: (1)  $A + X \rightarrow 2X$       rate constants:  $k_1$

(2)  $X + Y \rightarrow 2Y$        $k_2$

(3)  $Y \rightarrow B$        $k_3$

predator–prey–system

X: rabbits

Y: foxes

A: grass

B: dead foxes

possible modifications:

(0)  $B \rightarrow A$

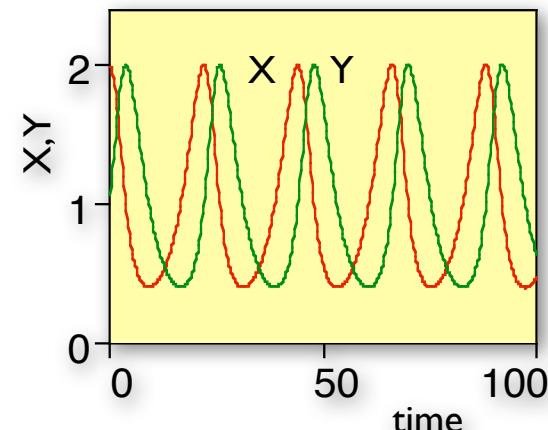
(1a)  $A + 2X \rightarrow 10X$

Rate equations:

$$\frac{dX}{dt} = k_1 AX - k_2 XY$$

$$\frac{dY}{dt} = k_2 XY - k_3 Y$$

$$A = \text{const}$$



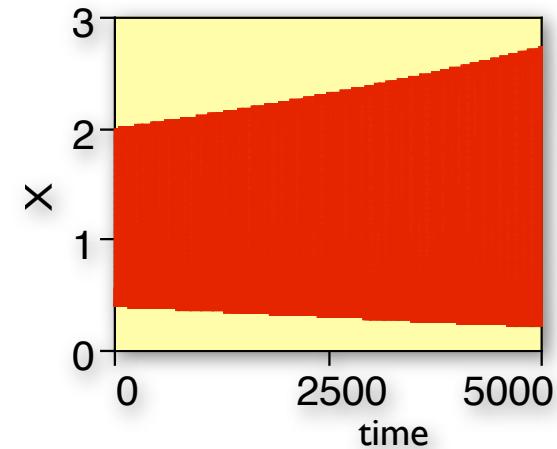
$$k_1 = k_2 = k_3 = 0.3 \quad \Delta t = 0.002$$

# Lotka–Volterra II

Long time deviations: error accumulates

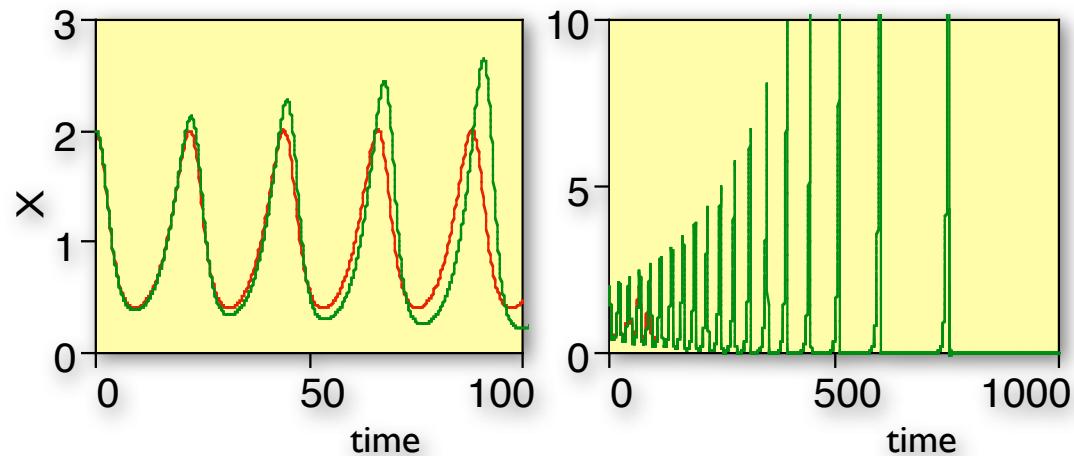
=> total amount of X,Y changes

$$\frac{T}{\Delta t} = 10^4 \Leftrightarrow \text{many small steps per oscillation}$$



$\Delta t = 0.002$  vs.  $\Delta t = 0.1$

=> period of the oscillation changes



=> reliability ?????

# Problems with “forward–Euler”

linear extrapolation from the beginning of the interval

second order of  
Taylor expansion  
neglected

asymmetry: only  
from one side of  
the interval

Pro:      => extremely simple implementation

Con:      => only very small  $\Delta t$   
              => only for well behaved problems (damping!)  
              => don't use it for serious work!!!

# Midpoint algorithm

Taylor expansion from beginning of the interval (size h):

$$f(t) = f(0) + t f'(0) + \frac{t^2}{2!} f''(0) + \frac{t^3}{3!} f'''(0) + \dots$$

Symmetrized: Taylor expansion around midpoint  $t' = h/2$  of the interval

$$\tilde{f}(t) = \tilde{f}(h/2) + (t - h/2) \tilde{f}'(h/2) + \frac{(t - h/2)^2}{2!} \tilde{f}''(h/2) + \frac{(t - h/2)^3}{3!} \tilde{f}'''(h/2) + \dots$$

Determine  $f(h)$  at the end of the interval  $h$ :

$$\begin{aligned} f(h) &= f(0) + [\tilde{f}(h) - \tilde{f}(0)] \\ &= f(0) + \left[ \tilde{f} + \frac{h}{2} \tilde{f}' + \frac{h^2}{12} \tilde{f}'' + \dots - \tilde{f} + \frac{h}{2} \tilde{f}' - \frac{(-h)^2}{12} \tilde{f}'' - \dots \right] \\ &= f(0) + h \tilde{f}'(h/2) + O(h^3) \end{aligned}$$

# Midpoint algorithm II

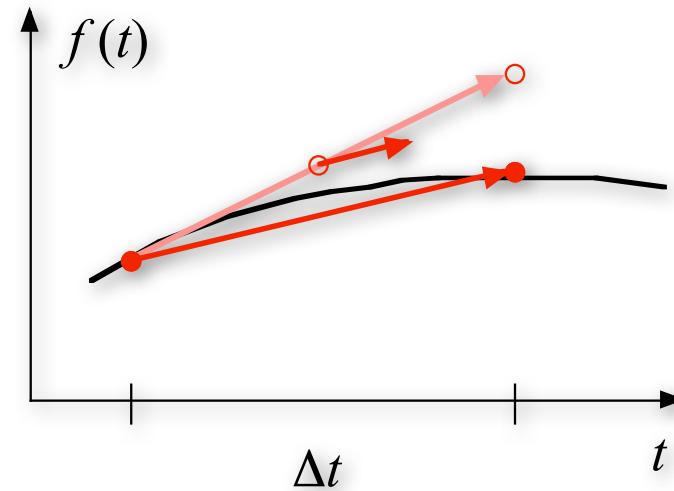
Estimate  $f'(h/2)$

$$\Delta = h/2 \cdot f'(0, f(0))$$

$$f(h) = f(0) + h \cdot f'(h/2, \Delta) + O(h^3)$$

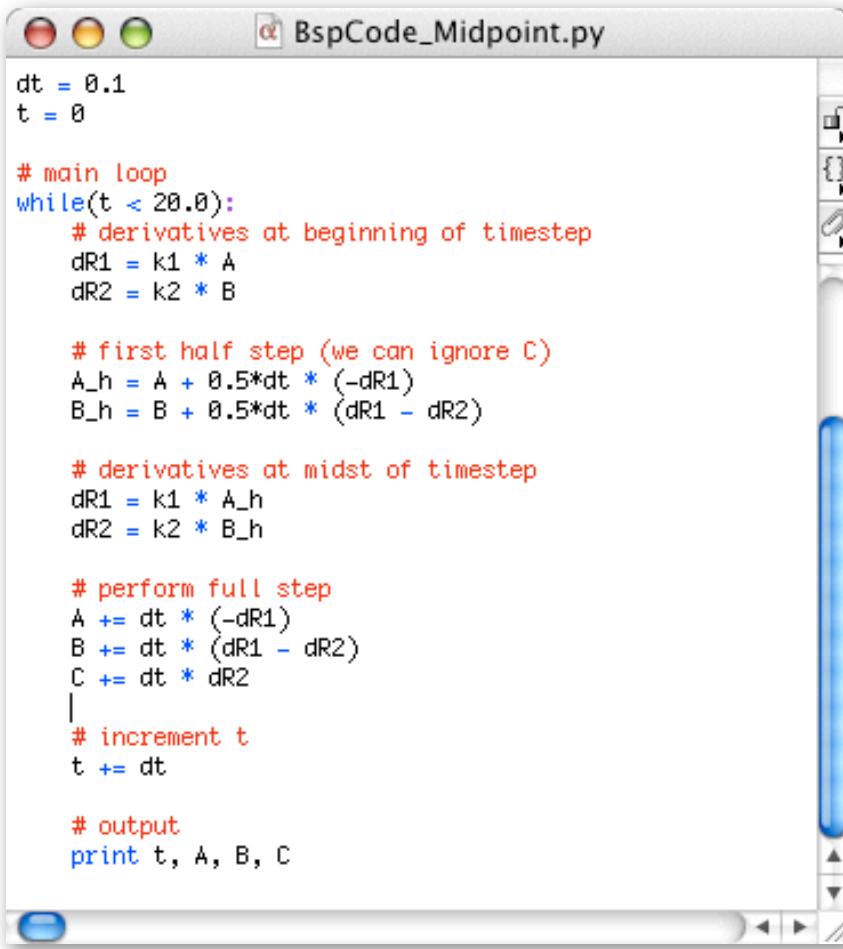
=> 2nd order algorithm

“2nd order Runge–Kutta”



- higher order algorithm
  - nearly symmetric
- } => larger time steps possible for same error  
=> more efficient than forward Euler  
if  $\Delta t$  at least 2x larger

# Example Code: Midpoint



```
dt = 0.1
t = 0

# main loop
while(t < 20.0):
    # derivatives at beginning of timestep
    dR1 = k1 * A
    dR2 = k2 * B

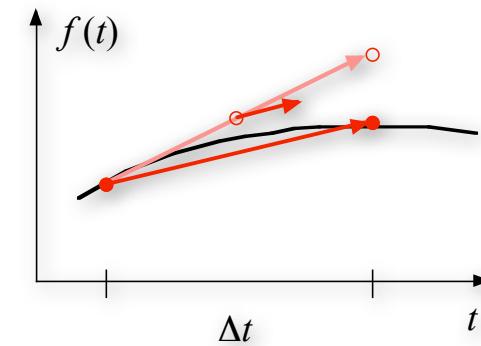
    # first half step (we can ignore C)
    A_h = A + 0.5*dt * (-dR1)
    B_h = B + 0.5*dt * (dR1 - dR2)

    # derivatives at midst of timestep
    dR1 = k1 * A_h
    dR2 = k2 * B_h

    # perform full step
    A += dt * (-dR1)
    B += dt * (dR1 - dR2)
    C += dt * dR2
    |
    # increment t
    t += dt

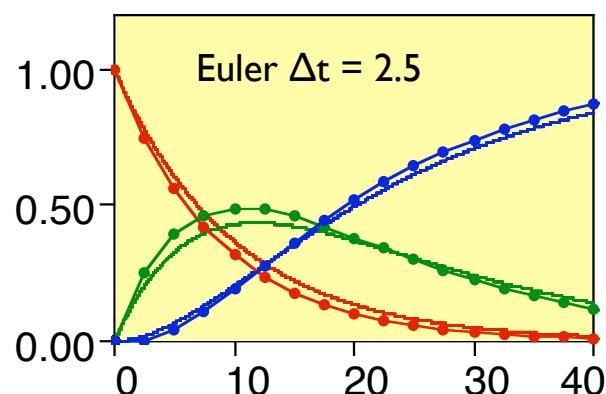
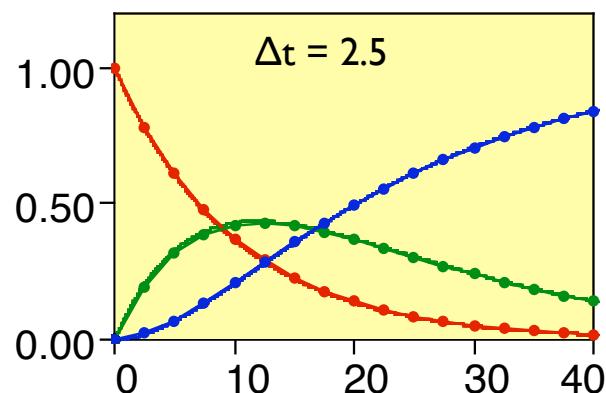
    # output
    print t, A, B, C
```

A => B => C

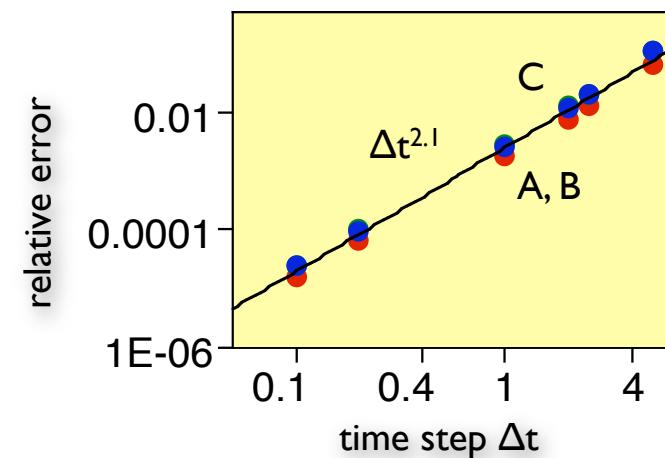


# Example I: chained reactions

Midpoint



Relative error vs.  $\Delta t$   
at  $t = 10$ :



=> larger time steps than forward Euler

## Example 2: Lotka–Volterra

For any  $\Delta t < 2$

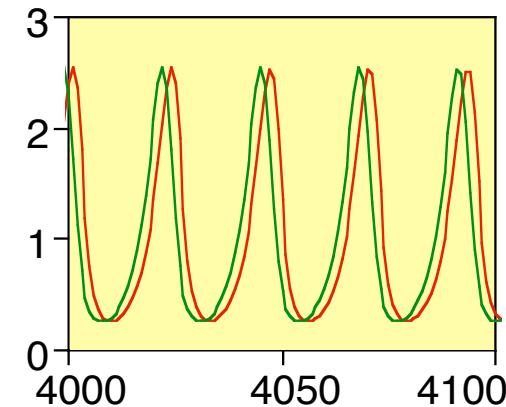
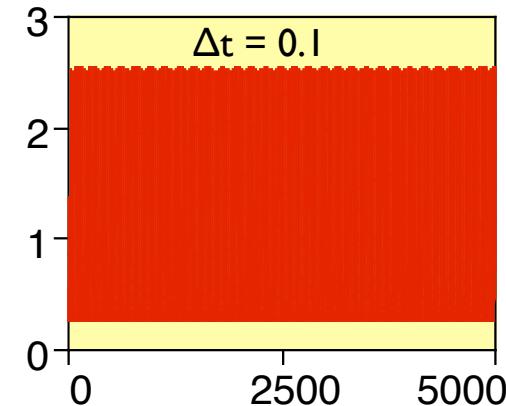
=> Constant amplitude

midpoint methods tend to “undershoot”

=> inherent stability

$\Delta t = 0.01$  vs.  $0.2$

=> oscillation frequency changes



# Runge–Kutta 4th order

Extend ideas of midpoint method:  
cancel higher order errors

$$\Delta_1 = h f'(0, f(0))$$

$$\Delta_2 = h f'\left(\frac{h}{2}, f(0) + \frac{\Delta_1}{2}\right)$$

$$\Delta_3 = h f'\left(\frac{h}{2}, f(0) + \frac{\Delta_2}{2}\right)$$

$$\Delta_4 = h f'(h, f(0) + \Delta_3)$$

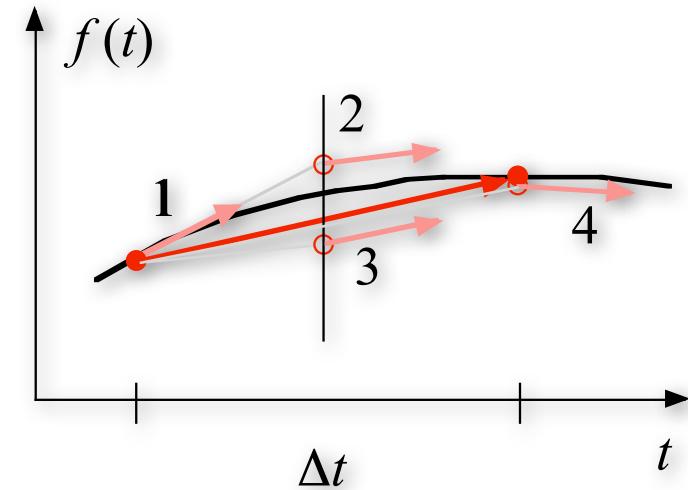
$$f(h) = f(0) + \frac{\Delta_1 + \Delta_4}{6} + \frac{\Delta_2 + \Delta_3}{3} + O(h^5)$$

midpoint:

2 evaluations of the derivatives

Runge–Kutta 4th:

4 evaluations of the derivatives



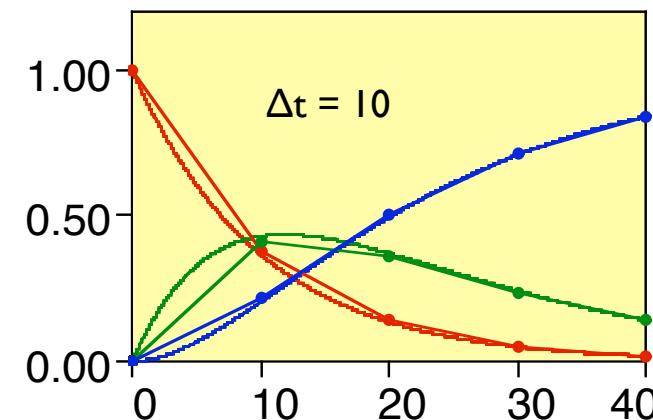
RK4 more  
efficient, if  $\Delta t$  at  
least twice as large.

# Example – chained reactions

Error  $< 10^{-7}$  for  $\Delta t \leq 0.5$

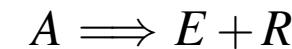
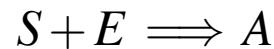
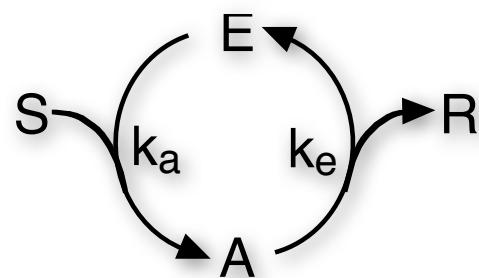
Error at  $t = 10$

	RK4	midpoint	forward Euler
<rel. error>	0.04	0.02	0.04
time step	10	2.5	1
#derivatives	4	8	10



# Michaelis–Menten Kinetics

enzymatically catalyzed reaction



$$M = A + E$$

$$\begin{aligned}\frac{dA}{dt} &= k_a SE - k_e A \\ &= k_a S(M - A) - k_e A \\ &= k_a SM - (k_a S + k_e) A\end{aligned}$$

$$\frac{dR}{dt} = k_e A$$

**Stationary state:**  $\frac{dR}{dt} = k_e A = \frac{k_e M S}{S + k_e/k_a} = \frac{v_M S}{S + S_M} \Rightarrow \text{flux}$

maximal speed  $v_M = k_e M$

substrate amount for half speed  $S_M = k_e/k_a$

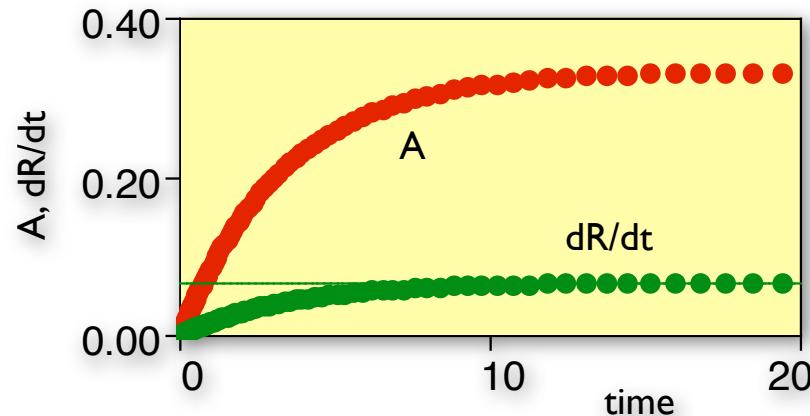
# Michaelis–Menten II

Start simulation with **arbitrary initial** values

=> system will **equilibrate** to steady state

=> scan S to get response curve

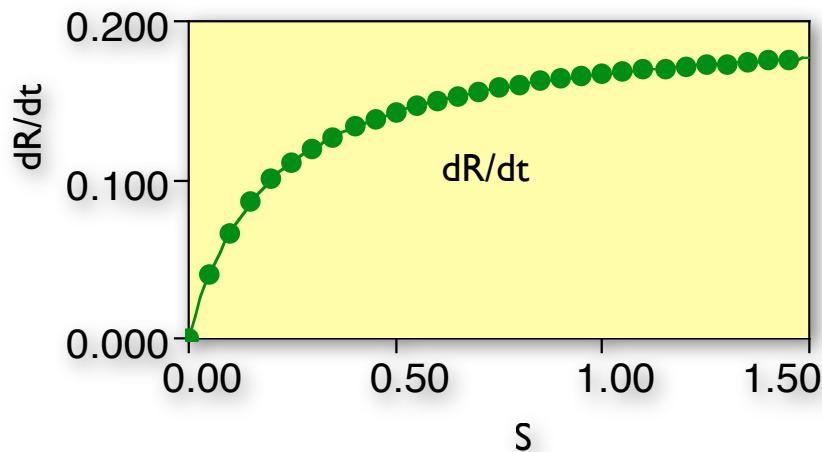
Here:  $k_a = 1$ ,  $k_e = 0.2$ ,  $M = 1$ ,  $S = 0.1$



=> use time dependent method to determine steady state properties  
with constant boundary conditions  
=> numerical search for  $dA/dt = 0$

# Michaelis–Menten III

Run simulations for a range of values for S, tabulate  $dR/dt$  after convergence



Numerical results reproduce analytic formula

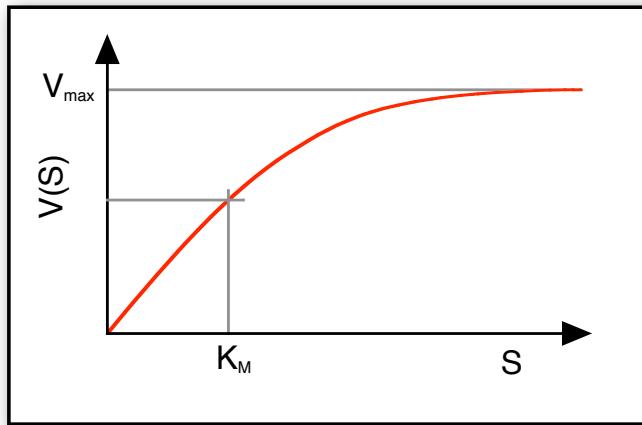
$$\frac{dR}{dt} = \frac{k_e M S}{S + k_a/k_e}$$

# The Equation

Effective throughput for MM:  $V = V_{max} \frac{S}{S + K_M}$

$$V_{max} = k_{off} E_T$$

$$K_M = \frac{k_{off}}{k_{on}}$$



Pro:

- **analytical** formula for the turnover
- **interpretation** of the characteristic with  $V_{max}$ ,  $K_M$
- enzyme can be **ignored** in the modelling

Con:

fewer kinetic parameters => information is **ignored**  
 $k_{on}, k_{off}, E_T \Rightarrow V_{max}, K_M$

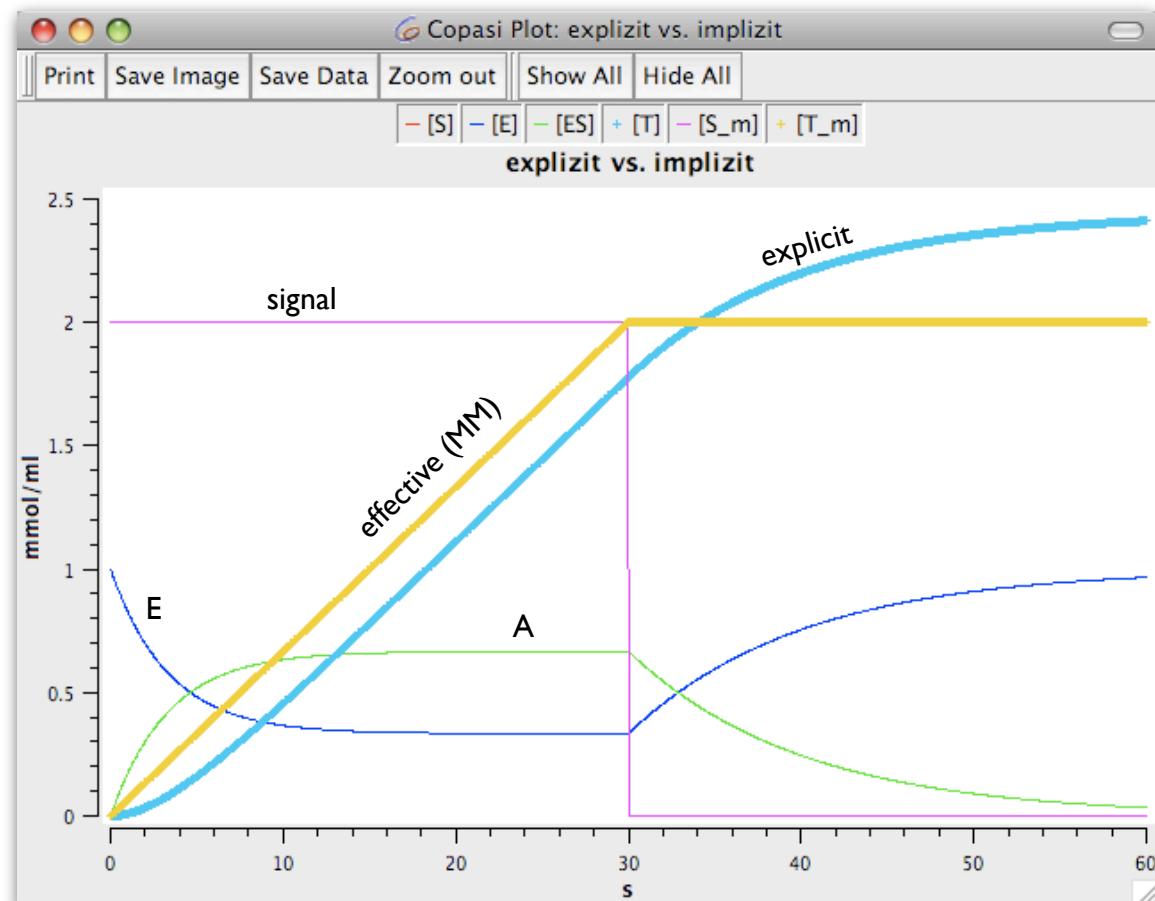
# MM vs. explicit Modelling

If E catalyzes **multiple substrates** => **MM not** (directly) applicable

Time evolutions:  
MM-kinetics vs.  
explicit modelling

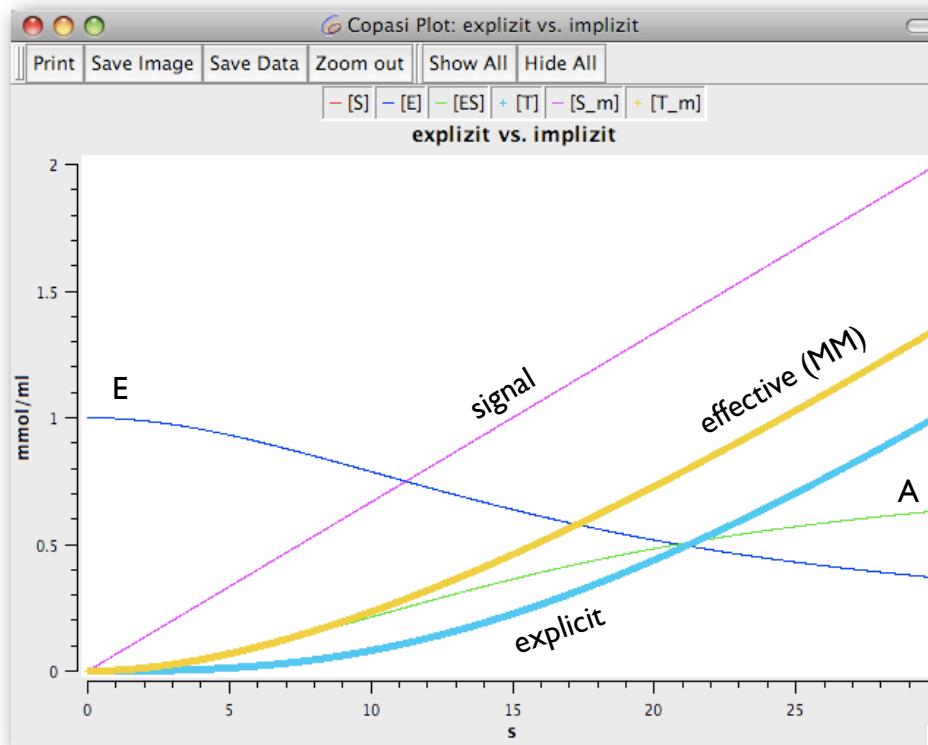
=> transients

=> different total  
turnover



# Again: explicit vs. MM

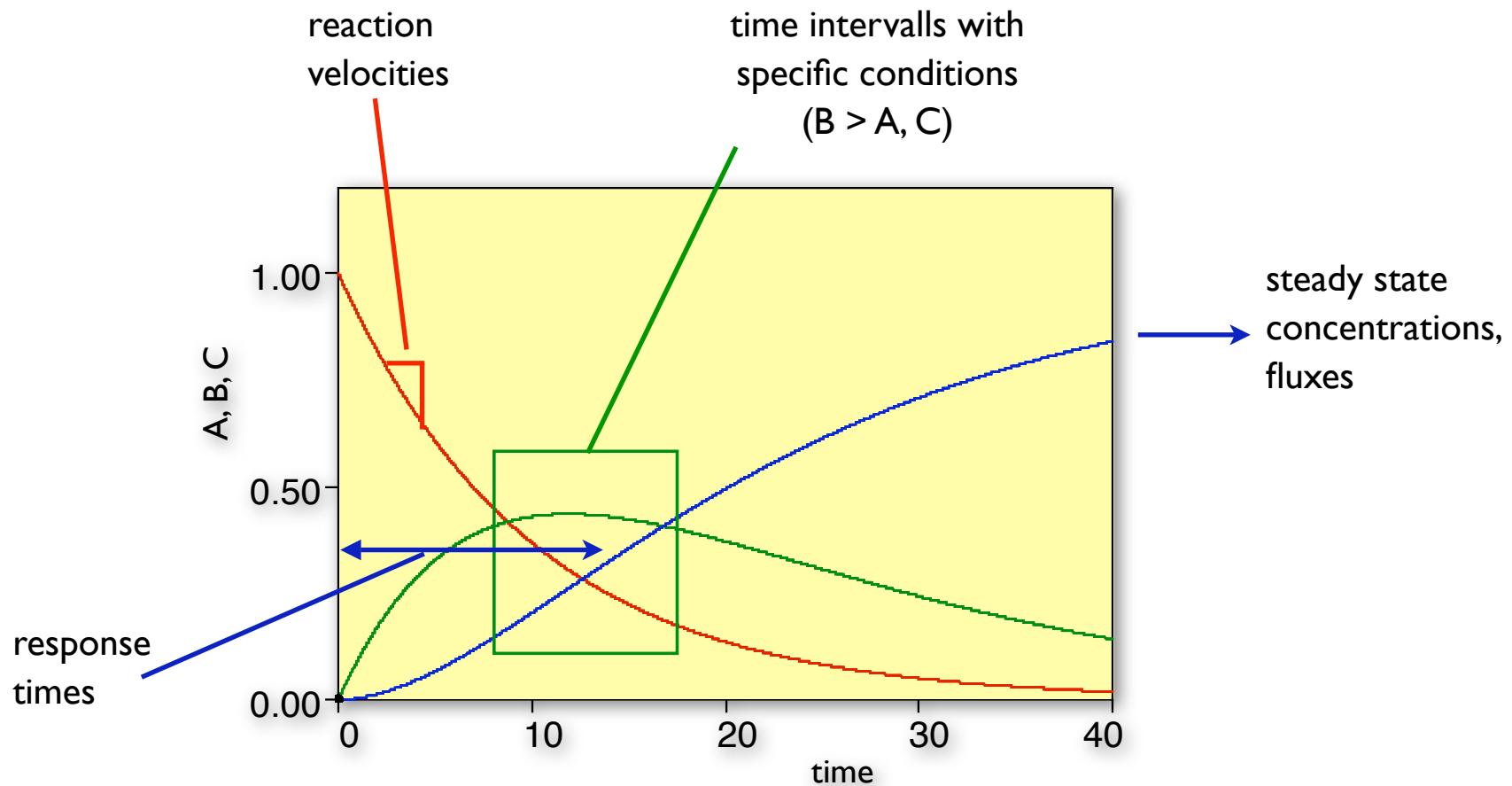
linear rise of S: => output from MM starts instantaneously  
=> delay with multiple-step explicit model



=> Effective kinetics usable when **changes** are much **slower** than  $k_i^{-1}$

# Information from a simulation

Available: any information in the model



# Signal Propagation in a Chain of Reactions

"Pathway":  $A_1 \xrightarrow{k_1} A_2 \xrightarrow{k_2} A_3 \xrightarrow{k_3} A_4 \xrightarrow{k_4} \dots \xrightarrow{k_{N-2}} A_{N-1} \xrightarrow{k_{N-1}} A_N$

rates:  $\frac{dR_i}{dt} = k_i A_i$        $\frac{dA_i}{dt} = k_{i-1} A_{i-1} - k_i A_i$

No forks, intersections => **simplified** version with the same ends:



Steady state:

Concentrations along the chain:

$$\frac{dA_i}{dt} = 0 \Rightarrow A_i = \frac{k_{i-1}}{k_i} A_{i-1}$$

Fluxes are identical! (no sinks or sources)

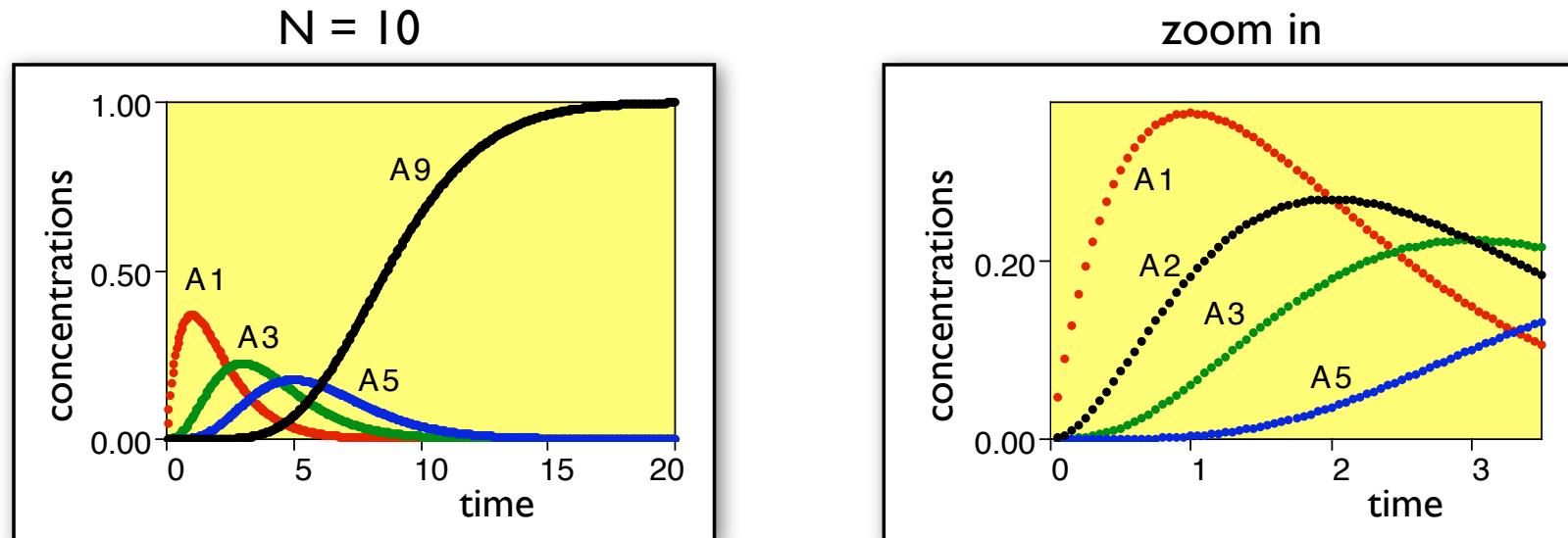
$$\frac{dR_i}{dt} = k_1 A_1$$

=> Same steady state throughput in long and short version

# Signal in a Chain

Signal propagation for:  $A_0(0) = 1$ ,  $A_i(0) = 0$ ,  $k_i = 1$

$$A_0 \Rightarrow A_1 \Rightarrow A_2 \Rightarrow A_3 \Rightarrow \dots \Rightarrow A_{N-1}$$

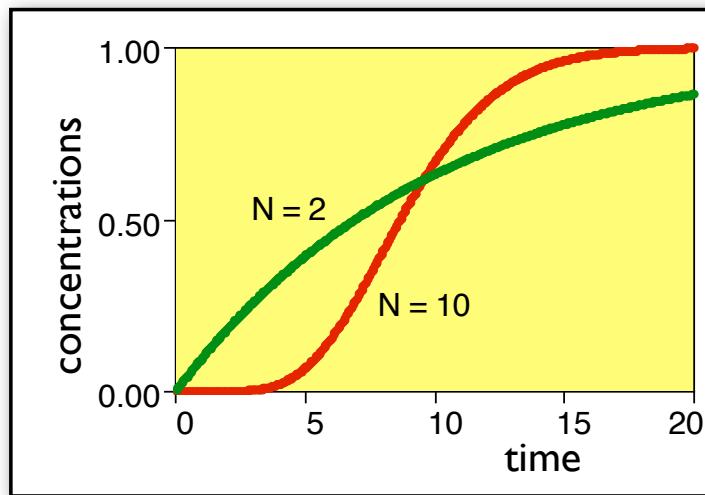


=> Signal delayed and broadened (washed out) along the chain

# Response of a Chain

Compare  $A_N$ :

$$N = 10, k_i = 1 \quad \Leftrightarrow \quad N = 2, k_i = 0.1$$



=> **different** response curve for **simplified** pathway  
(even with adapted  $k_i$ )

Be careful with simplifications when time courses matter!!!

# Different $k_i$ along the Chain

$N = 10$

$$k_i = 1$$

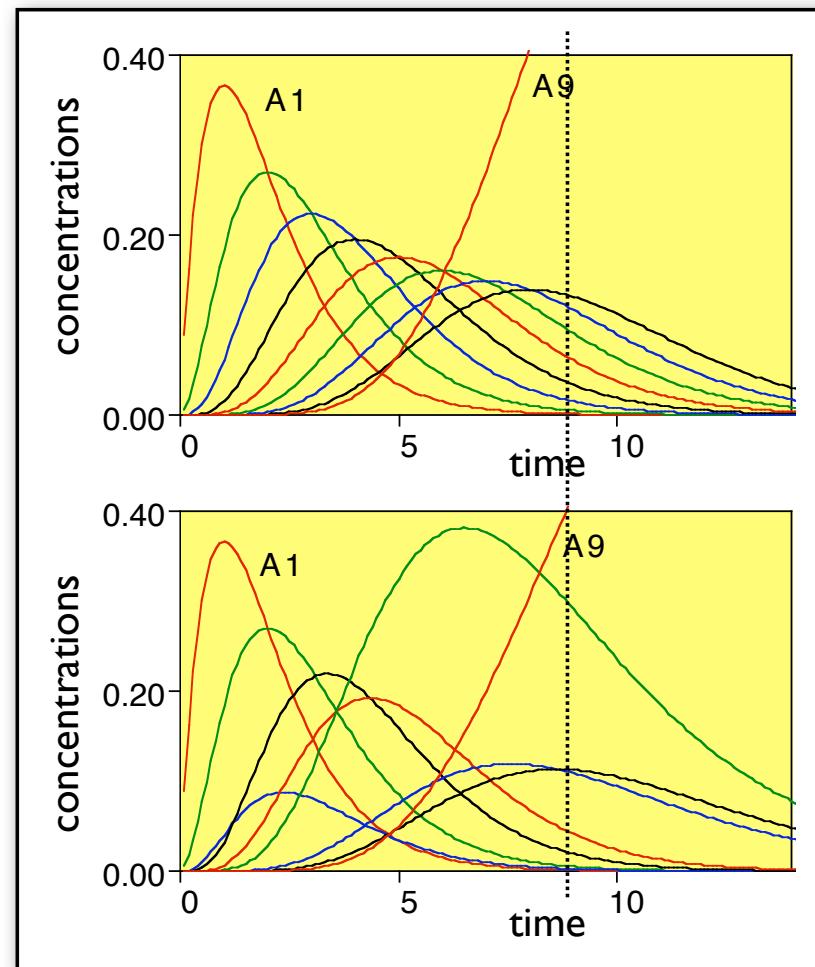


$$k_i = 1$$

$$k_3 = 3$$

$$k_6 = 0.33$$

**slowest** reaction  
determines **response** time  
 $\Rightarrow$  fast intermediate steps  
can (often) be neglected



# Probability Distribution

Consider decay reaction:



Continuous solution:

$$\frac{dA}{dt} = -kA \Rightarrow A(t) = A_0 e^{-kt}$$

Stochastic description:

- discrete particle number  $N_A$
- reaction probability per  $\Delta t$ :  $P_A = k \Delta t = c_A \Delta t$

=> probability that at time  $t$  one  $A$  decays during  $\Delta t$  =  $N_A(t) c_A \Delta t$

Probability that the system contains  $n$  particles at  $t$  =  $p_n(t)$

Dynamics governed by **probabilities** => yields only probabilities!!!

=> state described by probability distribution  $P_n(t) = \{p_0(t), p_1(t), p_2(t), \dots\}$

# Evolution of $P_n(t)$

For  $A \Rightarrow \emptyset$  assume that  $N_A(t+dt) = n$       ( $dt$  is short!!!)

Can result from:

- (i)  $N_A(t) = n$       and      no reaction during  $dt$
- (ii)  $N_A(t) = n+1$       and      one decay event during  $dt$

Propagate  $p_n(t)$ :

$$p_n(t+dt) = \underbrace{p_n(t)(1 - nc_A dt)}_{\begin{array}{l} \text{probability that the system was in state } n \\ \times \\ \text{probability for no decay during } dt \end{array}} + \underbrace{p_{n+1}(t)(n+1)c_A dt}_{\begin{array}{l} \text{decay from } (n+1) \\ \times \\ \text{probability that the system was in state } (n+1) \\ \times \\ \text{probability for one decay in } n+1 \text{ particles during } dt \end{array}}$$

# Master Equation

New  $p_n(t+dt)$ :

$$\begin{aligned} p_n(t + dt) &= p_n(t) (1 - n c_A dt) + p_{n+1}(t) (n + 1) c_A dt \\ &= p_n(t) - n c_A p_n(t) dt + p_{n+1}(t) (n + 1) c_A dt \end{aligned}$$

Sort for  $p_n$ :

$$\frac{p_n(t + dt) - p_n(t)}{dt} = c_A (n + 1) p_{n+1}(t) - c_A n p_n(t)$$

Limit of  $dt \Rightarrow 0$ :

$$\frac{dp_n}{dt} = c_A (n + 1) p_{n+1} - c_A n p_n$$

"Chemical Master equation" for the system  $A \Rightarrow \emptyset$

Master equation describes the evolution of the probability distribution  $P_n(t)$   
 $\Rightarrow$  (Infinite) set of coupled differential equations

# Analytic Solution

Usually: Master equation too complex for analytic solution  
(especially with multiple species)

For  $A \Rightarrow \emptyset$  with  $N_A(0) = n_0$ :

There will never be more than  $n_0$  particles  $\Rightarrow p_n = 0$  for  $n > n_0$

Evolution of  $p_{n0}$ :

$$\frac{dp_{n_0}}{dt} = -c_A n_0 p_{n_0} \quad \Rightarrow \quad p_{n_0}(t) = e^{-c_A n_0 t}$$

Notes:

- (i) this does not describe the evolution of the particle number
- (ii)  $p_{n_0}(t)$  decays faster for larger  $n_0$

# Analytic Solution II

Plug in  $p_{n_0}(t)$ :

$$\begin{aligned}\frac{dp_{n_0-1}}{dt} &= c_A n_0 p_{n_0} - c_A (n_0 - 1) p_{n_0-1} \\ &= c_A n_0 e^{-c_A n_0 t} - c_A (n_0 - 1) p_{n_0-1}\end{aligned}$$

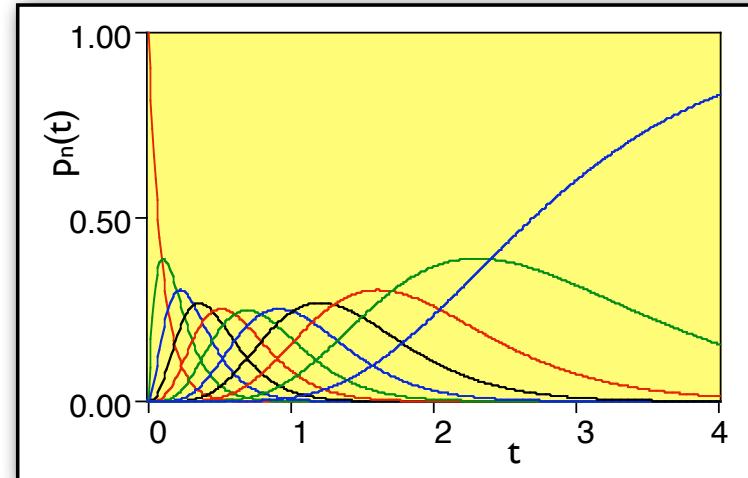
With  $p_{n_0-1}(0) = 0$ :

$$p_{n_0-1}(t) = e^{-c_A(n_0-1)t} n_0 (1 - e^{-c_A t})$$

By induction:

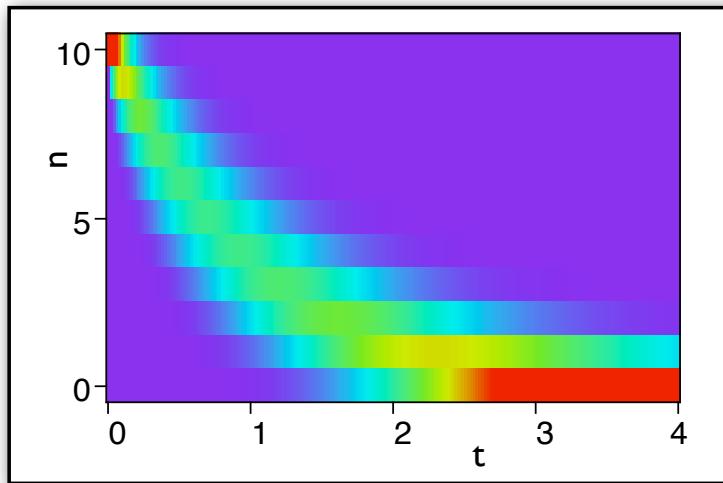
$$p_n(t) = e^{-c_A n t} \binom{n_0}{n} (1 - e^{-c_A t})^{n_0-n}$$

At a given  $t$ : system can  
be in any of the states  $0 \leq n \leq n_0$   
(with corresponding  $p_n(t)$ )

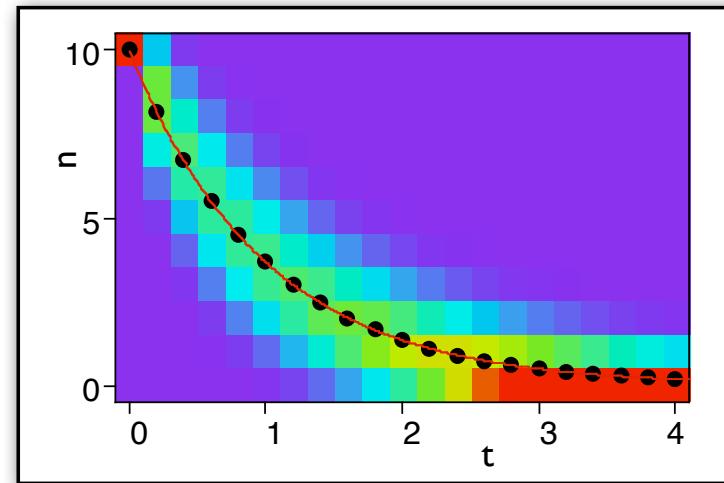


# Time Evolution

$$P_n(t) \text{ for } n_0 = 10, c_A = 1$$



$$dt = 0.02$$

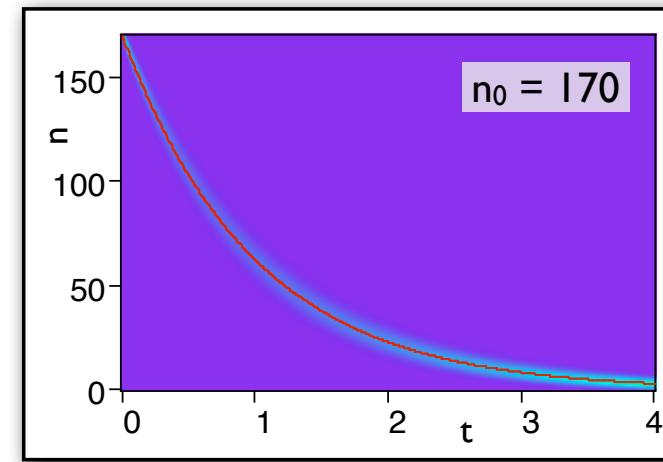
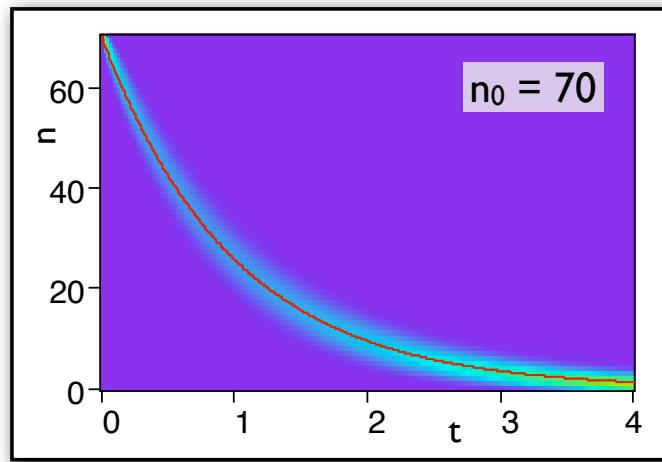
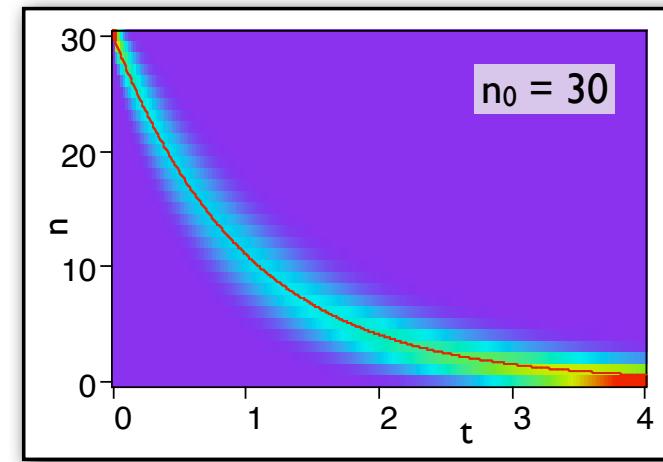
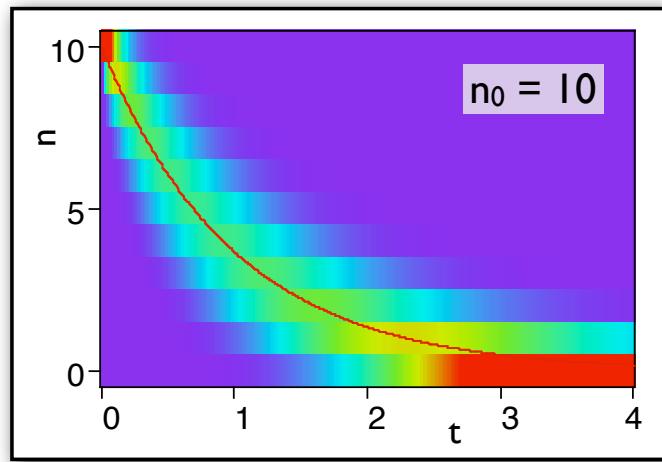


$$dt = 0.2$$

Average  $n(t)$ : 
$$\langle n \rangle(t) = \sum_{n=0}^{n_0} n p_n(t) = \dots = n_0 e^{-c_a t}$$

Beware: this is **one** of the few **examples** where stochastic and continuous description give the **same average number!!!**

# Larger Numbers => Less Fluctuations



# Another Example



Chemical master equation ( $N_A = n$ ,  $N_B = m$ ):

$$\begin{aligned}\frac{dp_{n,m}}{dt} = & k_1 p_{n-1,m} - k_1 p_{n,m} + \\ & k_2 p_{n,m-1} - k_2 p_{n,m} + \\ & k_3 (n+1)(m+1) p_{n+1,m+1} - k_3 n m p_{n,m}\end{aligned}$$

$$P(n_1, n_2, t) = P_{n_1, n_2}(t) = \{p_{0,0}(t), p_{1,0}(t), \dots, p_{0,1}(t), \dots\}$$

Observations:

Two species  $\Rightarrow$  two indices,  $N_{\max}(A) \times N_{\max}(B)$  terms in  $P_{n_1, n_2}(t)$

Gain and loss terms from the "creation" reactions R1 and R2  
 $\Rightarrow$  these change the particle numbers, too.

# One Equation to Rule Them All

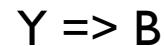
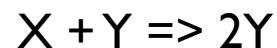
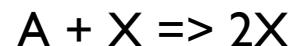
**Master equation** = propagation equation of the density distribution  $P(n_1, n_2, \dots, n_k, t)$  for the particle numbers  $n_1, n_2, \dots$

- built from simple **gain** and **loss terms**
- $P(n_1, n_2, \dots, n_k, t)$  **contains all** statistical measures for the system (averages, fluctuations, cross-correlations, ...)
- very few analytically solvable cases
- for simple systems **direct integration** possible (though many dimensions)
- numerical hybrid methods: Verena Wolf
- further reading: Erban, Chapman, Maini, arXiv:0704.1908v2 [q-bio.SC]

=> more simple and practical description of stochastic processes?

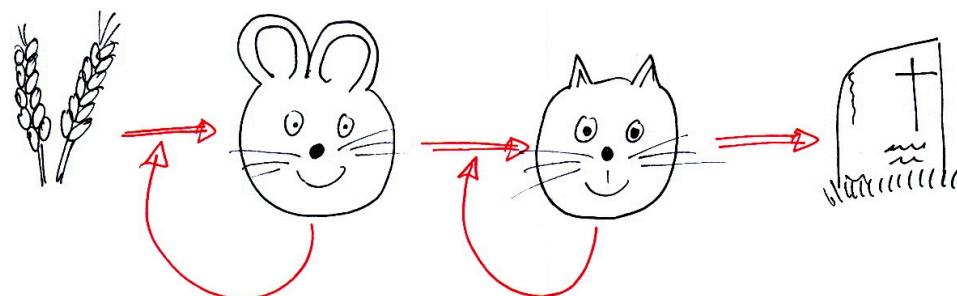
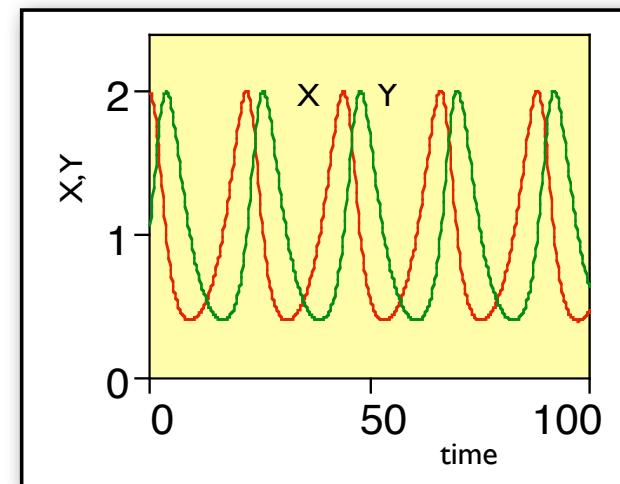
# Lotka-Volterra with Small Populations

Simplified population dynamics:



Autokatalysis  $\Rightarrow$  stable oscillations

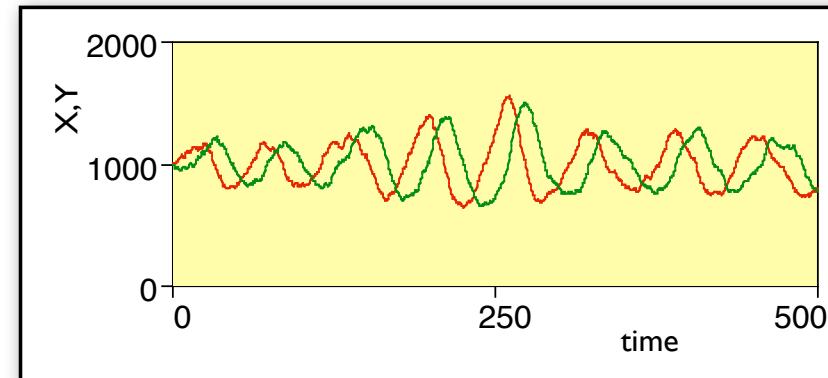
Steady state:  $Y = \frac{k_1}{k_2} A$      $X = \frac{k_3}{k_2}$



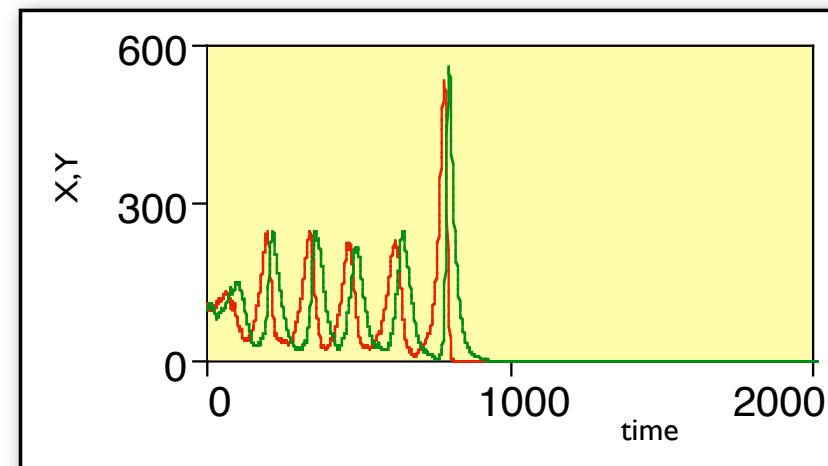
# Lotka–Volterra with Fluctuations

With stochastic fluctuations:

=> stationary state is not  
stationary any more  
("noise")



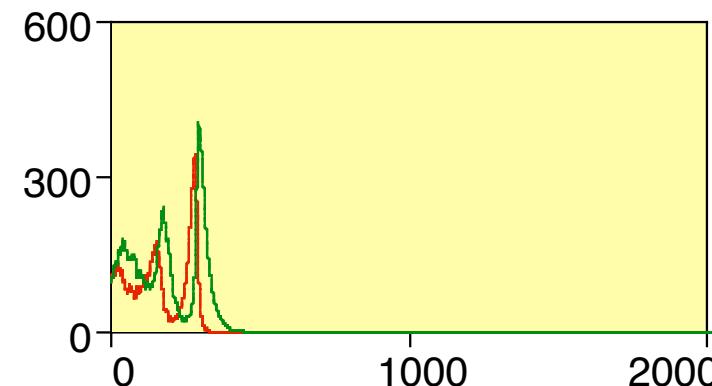
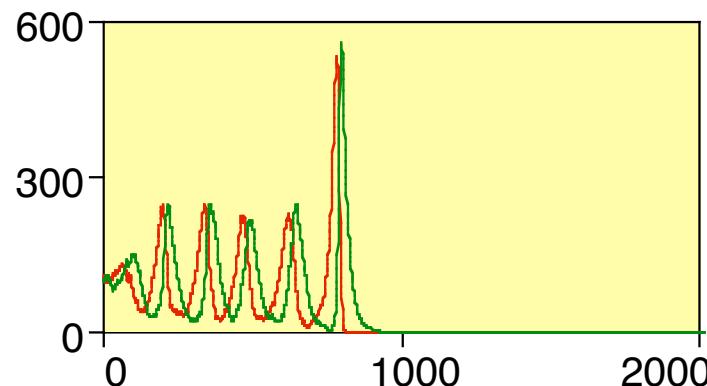
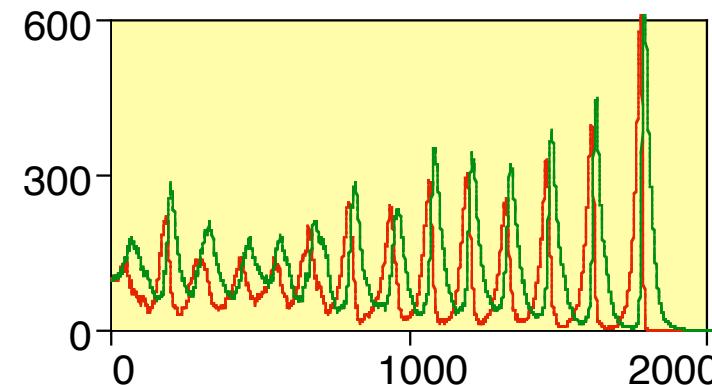
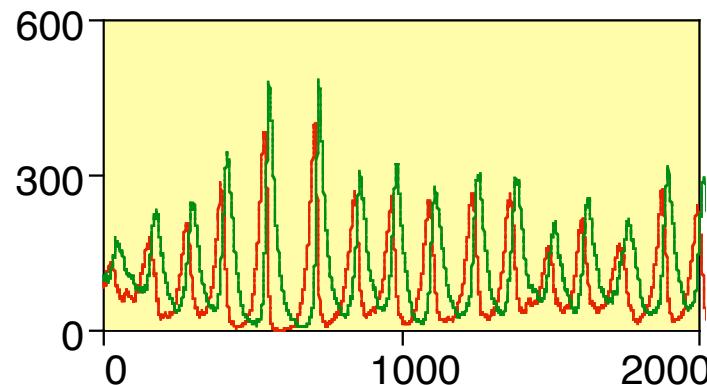
Discrete species sizes  
=> extinction of small populations  
=> different behavior  
  
(survival time vs.  
population size?)



# Smaller Populations

Stochastic simulation of the stationary configuration with fewer particles:

$$A = X = Y = 100 \quad k_1 = k_2 = 0.0005 \quad k_3 = 0.05$$



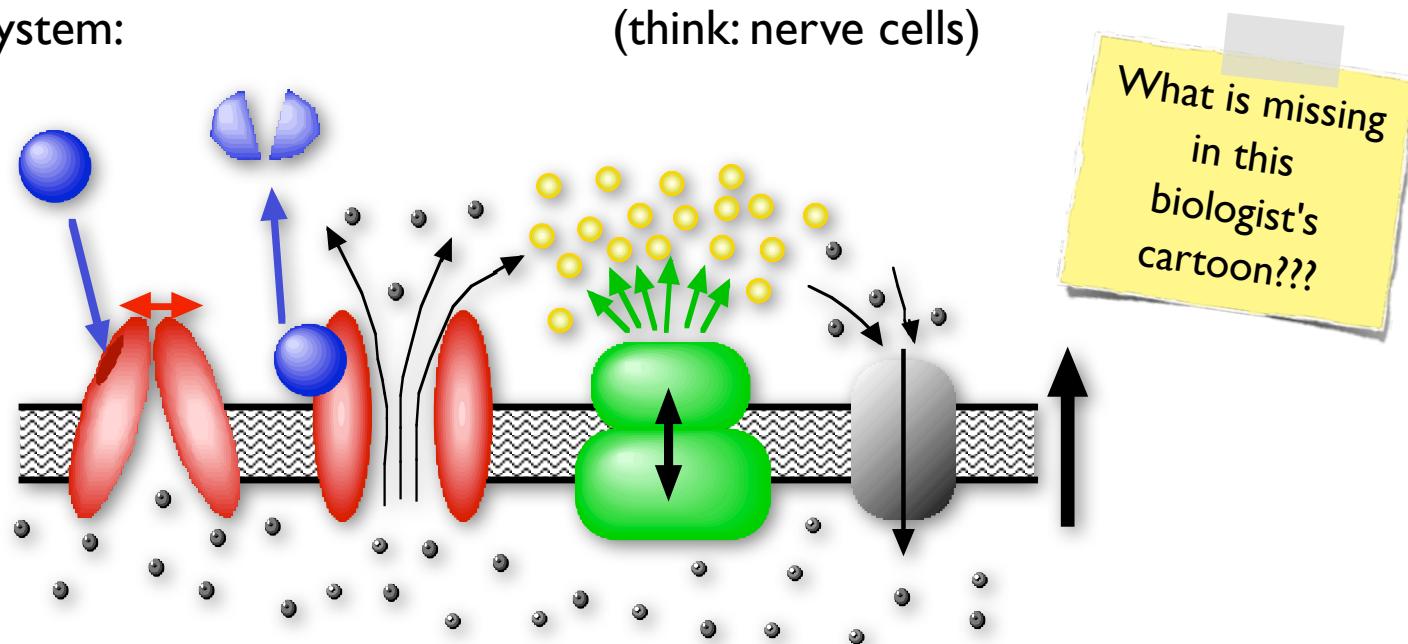
if too few => fluctuations lead to extinction

# (Stochastic) Response to a Signal

How can a few single molecules induce a macroscopic response?

Example system:

(think: nerve cells)



(i) Signalling molecule M binds to ion channel

(ii) Ions are released until M is deactivated, ion gradient decreases

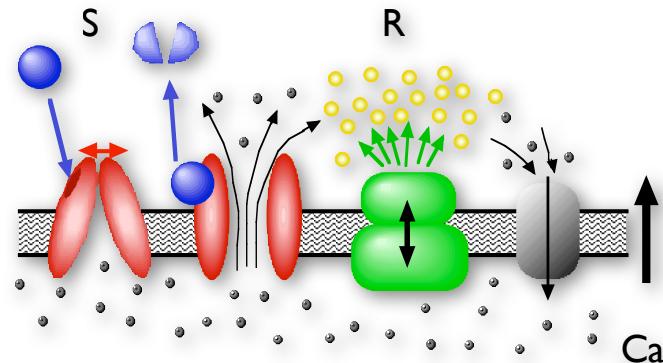
(iii) Response molecules R are released if ion gradient is small enough (vesicle breaks open)

(0) Ion pump restores rest state and compensates for leakage

# Signalling example: the reactions

Species: ions “inside”  
signalling molecules  
response molecules

Ca  
S  
R



Ions: constant pumping, leakage and release induced by S

$$\frac{dCa}{dt} = P_0 - k_L Ca - k_s Ca S$$

Signalling molecules: release between  $T_{on}$  and  $T_{off}$  and decay

$$\frac{dS}{dt} = w(t) S_0 - k_d S \quad w(t) = 1 \text{ for } T_{on} < t < T_{off}$$

Response molecules: release when Ca concentration is low

$$\frac{dR}{dt} = R_0 \quad \text{if } Ca < Ca_{min}$$

# Parameter set

Initial conditions:

$$C_a(0) = 100$$

$$S(0) = 0$$

$$R(0) = 0$$

Reaction constants (reasonable units??):

Pump rate:  $P_0 = 10$

Leakage:  $k_L = 0.1$

Release:  $k_S = 1.0$

Signal strength:  $S_0 \geq 0$

Signal interval:  $T_{on} = 40$     $T_{off} = 140$

Deactivation rate:  $k_d = 1.0$

Response release:  $R_0 = 1$

# Stochastic implementation

```
my $Ca = 100.0;
my $M = 0;
my $R = 0;

for(my $t=0; $t<= $tEnd; $t += $dt)
{
    my $dM = 0;
    my $dCa = 0;
    my $dR = 0;

    # my $dCa = $dt * ($D0 - $kL * $Ca - $ks * $M);
    if(rand(1) <= ($dt * $D0)) {
        $dCa++;
    }

    for(my $i=0; $i<$Ca; ++$i) {
        if(rand(1) <= $PL) {
            $dCa--;
        }

        for(my $i=0; $i<$M; ++$i) {
            if(rand(1) <= $Ps) {
                $dCa--;
            }
        }
    }

    # my $dM = $dt * ($M0 - $kM * $M);
    if (($t >= $tOn) && ($t <= $tOff)) {
        if(rand(1) <= ($dt * $M0)) {
            $dM++;
        }
    }

    for(my $i=0; $i<$M; ++$i) {
        if(rand(1) <= $PM) {
            $dM--;
        }
    }

    if($Ca <= $CaMin) {
        $dR += $dt;
    }

    $Ca += $dCa;
    $M += $dM;
    $R += $dR;
}
```

Gated release of S

=> Gillespie misses increase in S

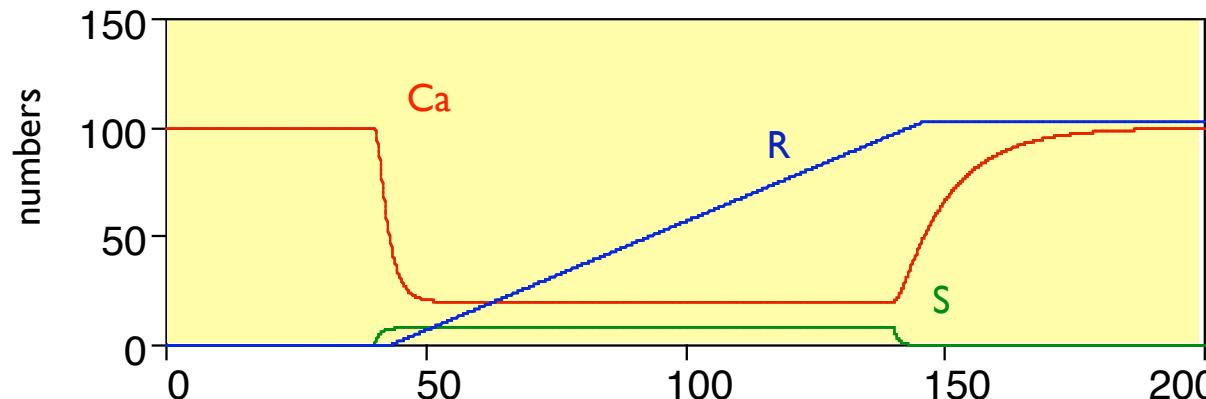
$$\alpha_0(T=0) = 0 \Rightarrow s = \infty$$

=> naive implementation used here

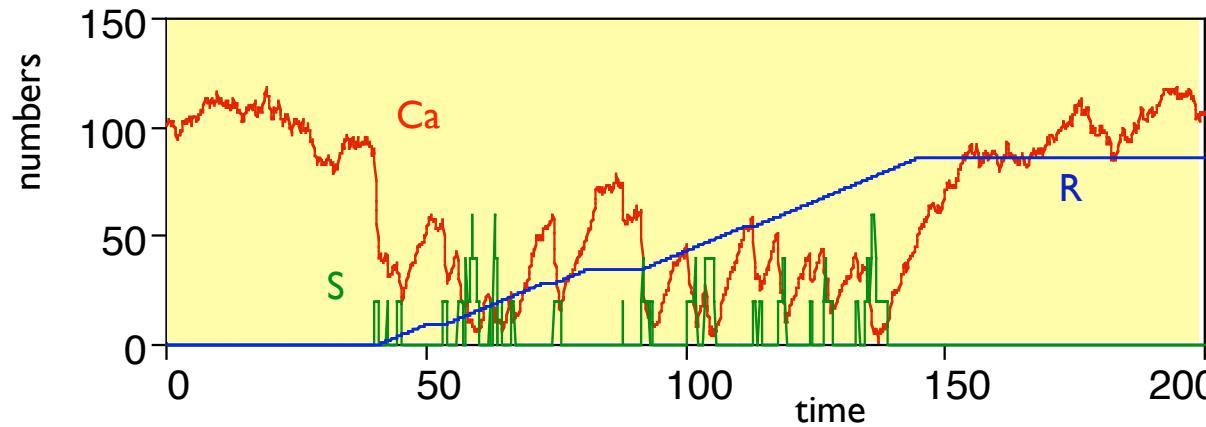
# Example trajectories

$S_0 = 0.4 \Rightarrow 40$  molecules of S on average  
 $\Rightarrow$  similar behavior from both approaches

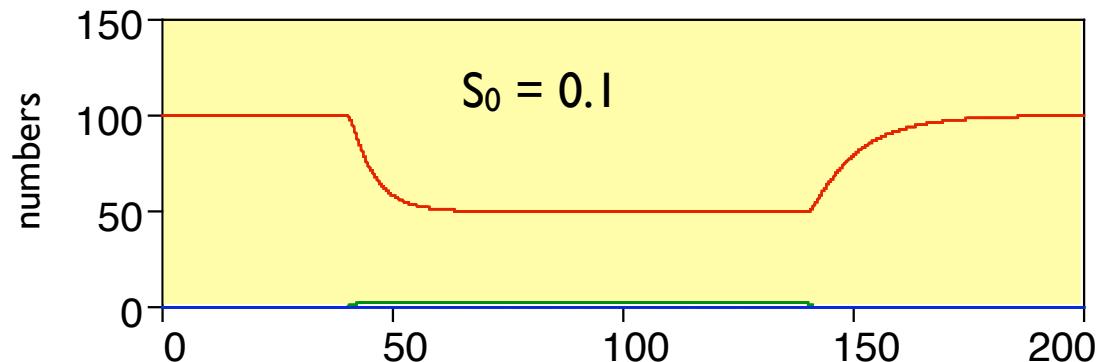
Rate  
equations:



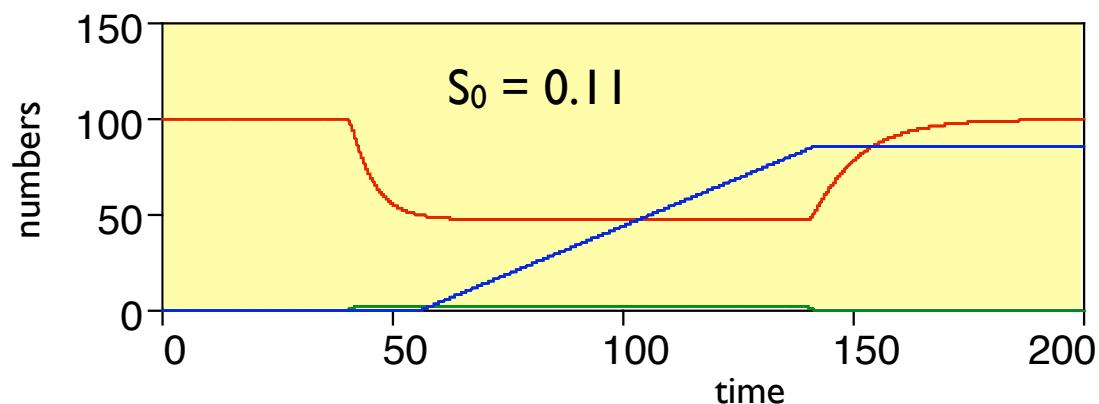
Stochastic:



# Threshold in the continuous treatment



total output  
 $R(200) = 0$



$R(200) = 85$

Threshold behavior of the response:

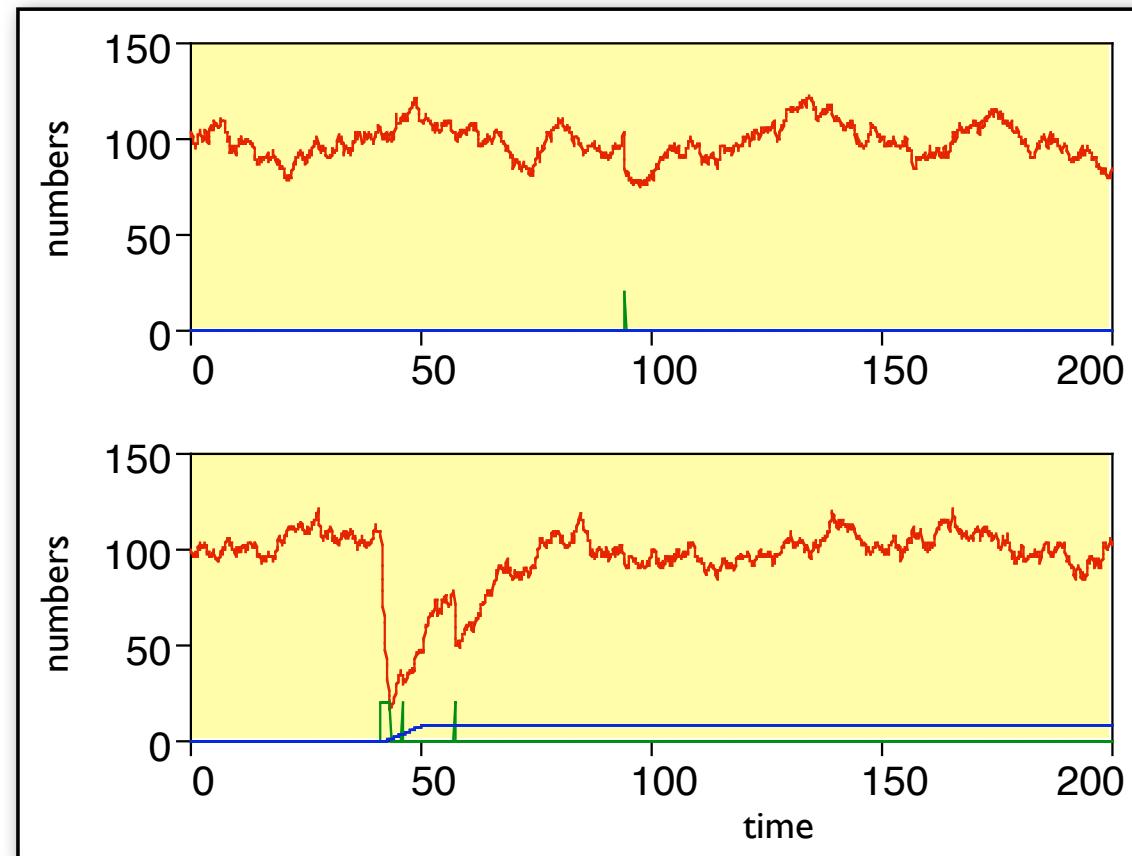
$\begin{cases} S_0 \leq 0.1 & \Leftrightarrow R = 0 \\ S_0 > 0.1 & \Leftrightarrow R \approx 100 \end{cases}$

# Threshold and fluctuations

Stochastic approach with  $S_0 = 0.05$  — below the threshold

most of the time:  
=> nothing happens

sometimes:  
=> one S survives  
long enough  
=> short burst of R  
=> (weak) signal  
below threshold

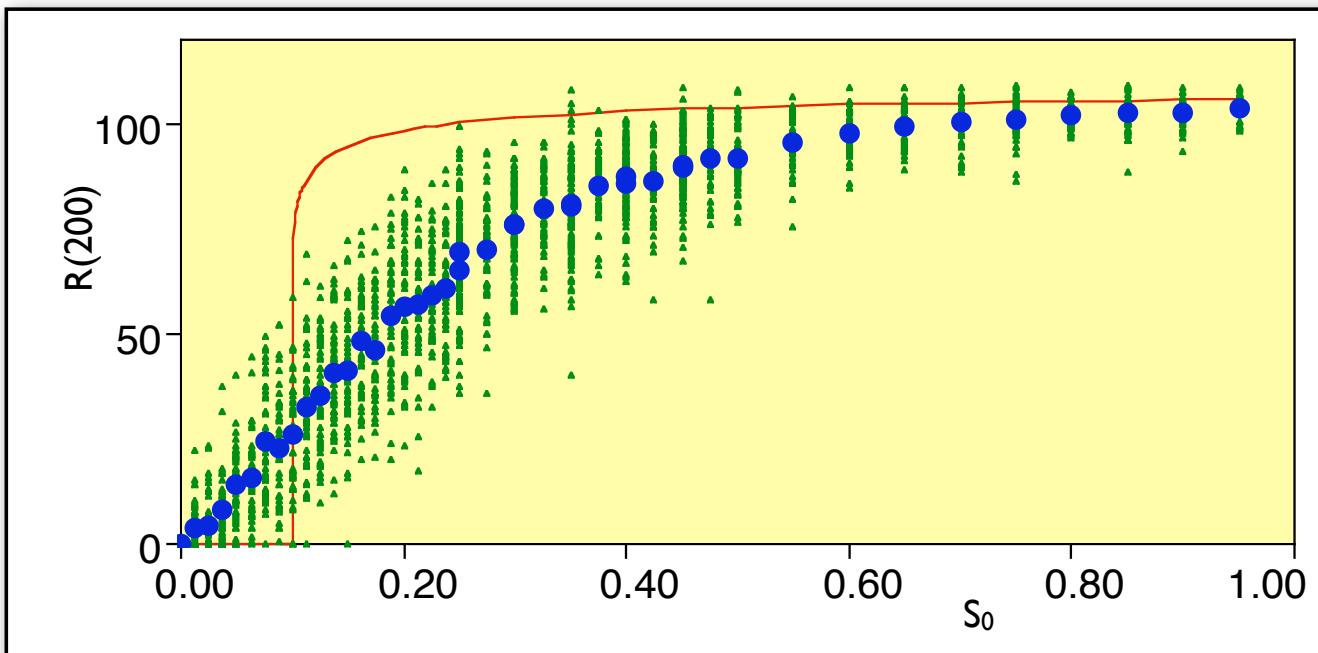


=> A single molecule can trigger a response (in one ion channel)

# Average response of many channels

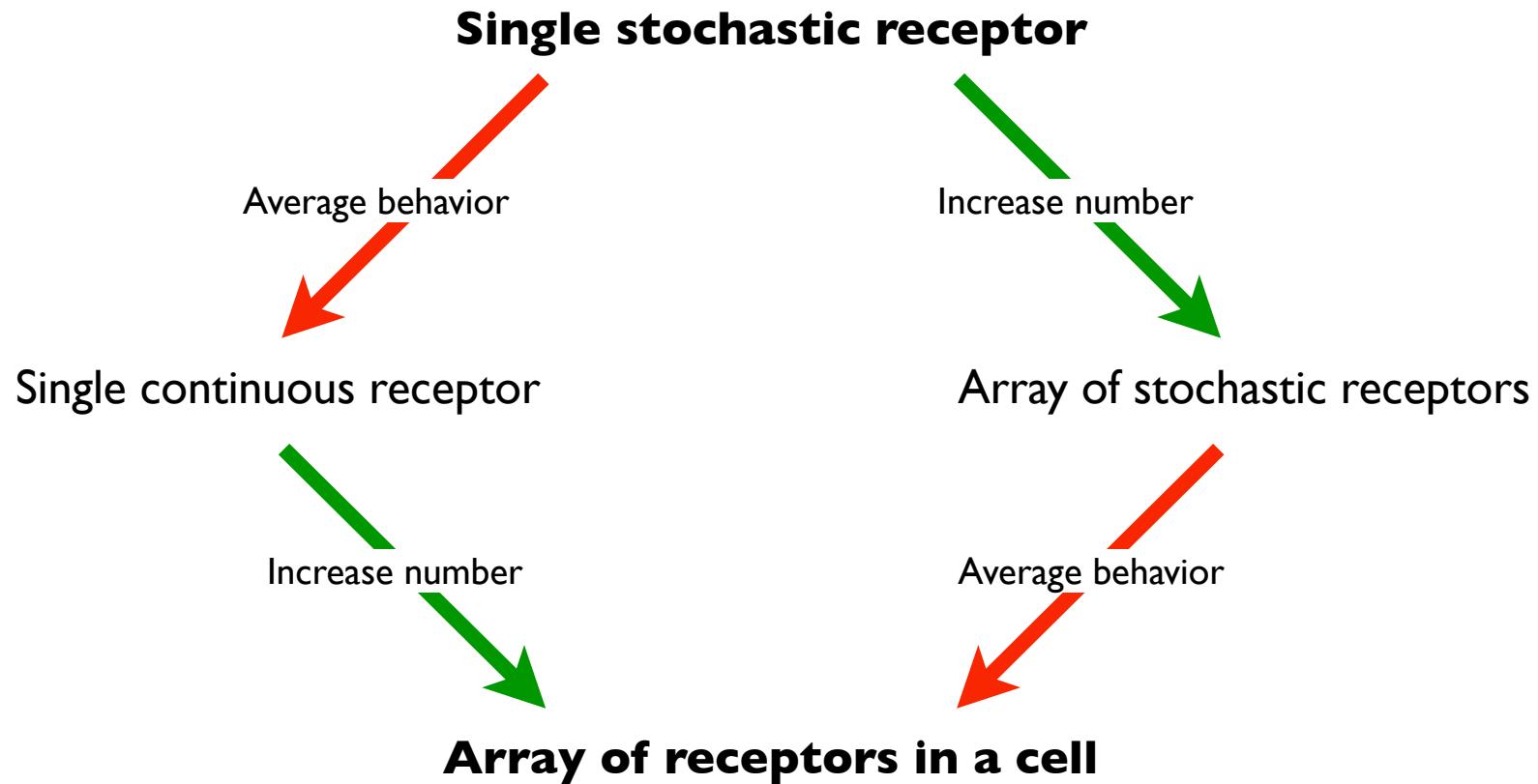
Averaged behavior vs. averaged response

Continuous vs. stochastic



Ensemble-averaged response from stochastic model:  
=> sensitive to single molecules  
=> smoothed response curve despite threshold behavior

# Different averaging strategies



# Summary

## Today:

- Better integration schemes: Midpoint, Runge-Kutta
- Effective models
  - => delayed response of pathways and enzymes
- Stochastic integration: Master equation
- Different behavior with fluctuations
  - => extinction in MM
  - => graded response in signalling

## Next lecture:

- Pools-and-proteins ansatz
- some Systems Biology