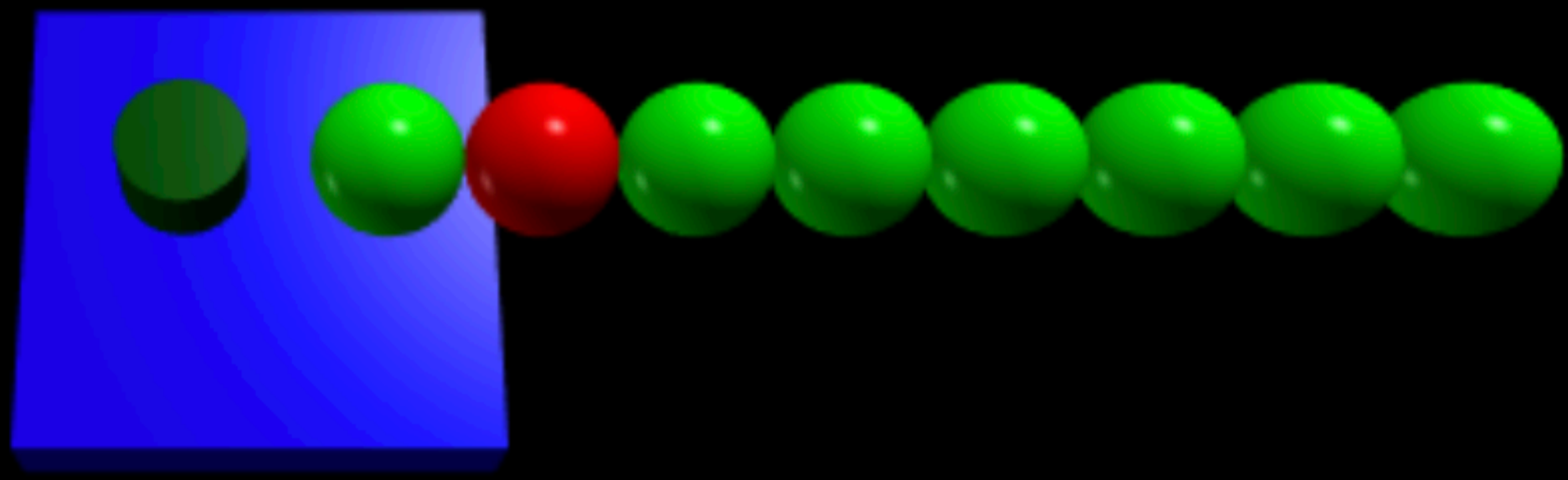
Phil Nelson
University of
Pennsylvania

Stochastic simulation



For these slides see:
www.physics.upenn.edu/~pcn

This talk

Ion channels

Gillespie algorithm

Bursting in transcription

Kinetic proofreading

Adaptation in chemotaxis

Birth of single-molecule biophysics

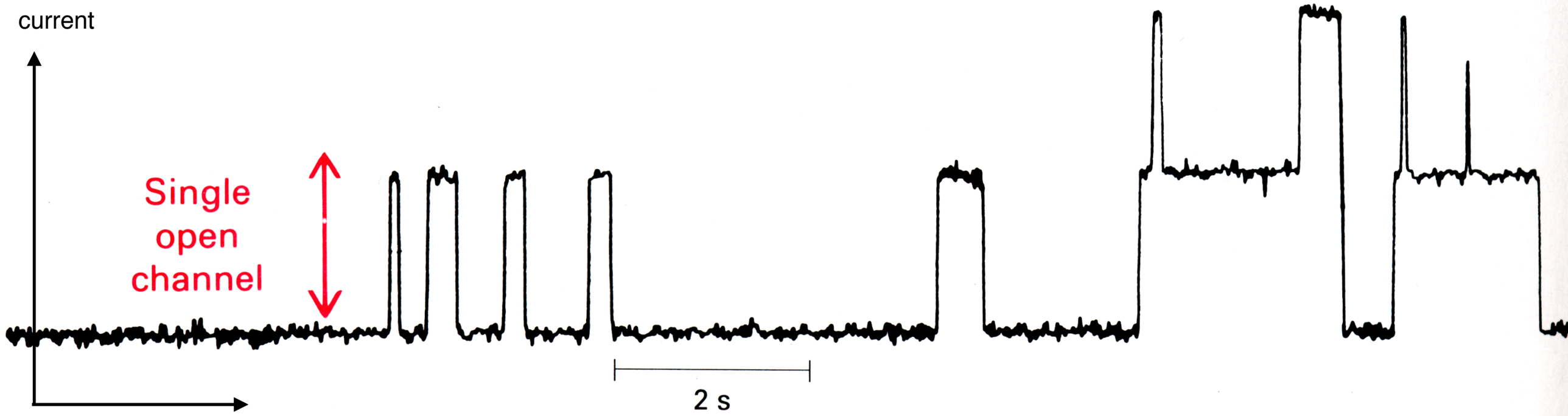
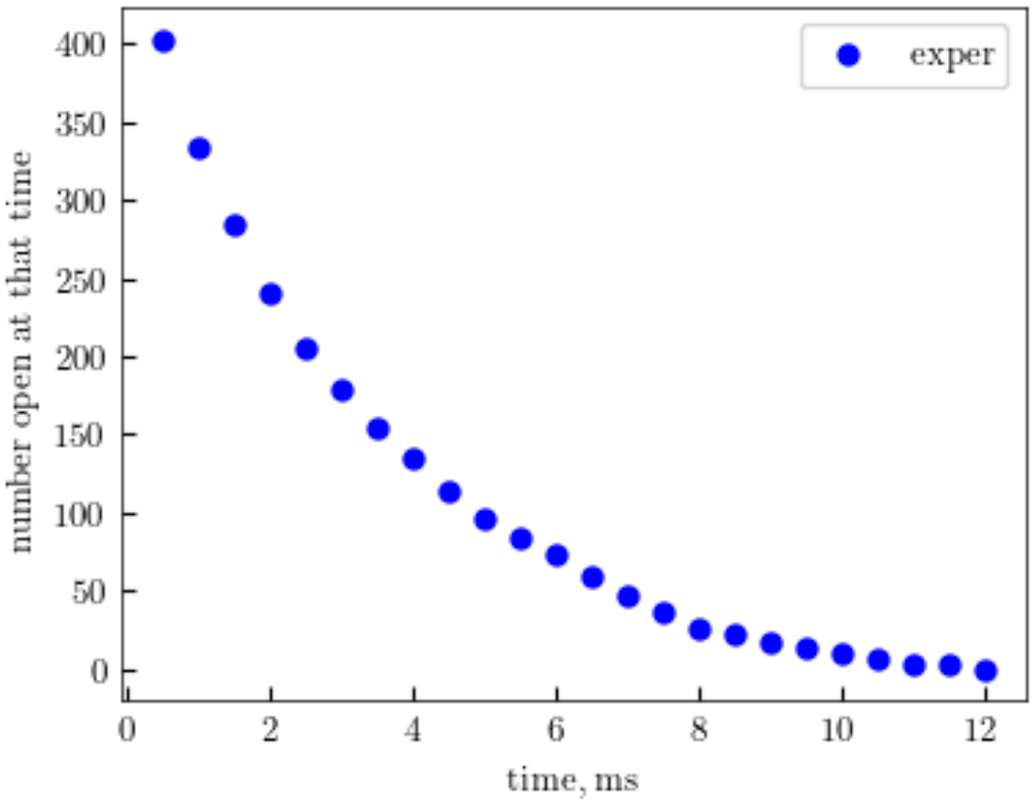


Figure 11-23
The conductance of a lipid bilayer membrane containing a few molecules of gramicidin A fluctuates in a step-by-step manner. The smallest step in conductance arises from a single open channel.



A patch of muscle membrane (postsynaptic) was exposed to the neurotransmitter acetylcholine (Ach) until all of its ion channels were open. Then the Ach was rapidly washed out. Each channel was then stuck in the "wrong" state for the new conditions. But each waited for a random time before making the one-way transition to the state that is "right." So the reaction scheme is simply

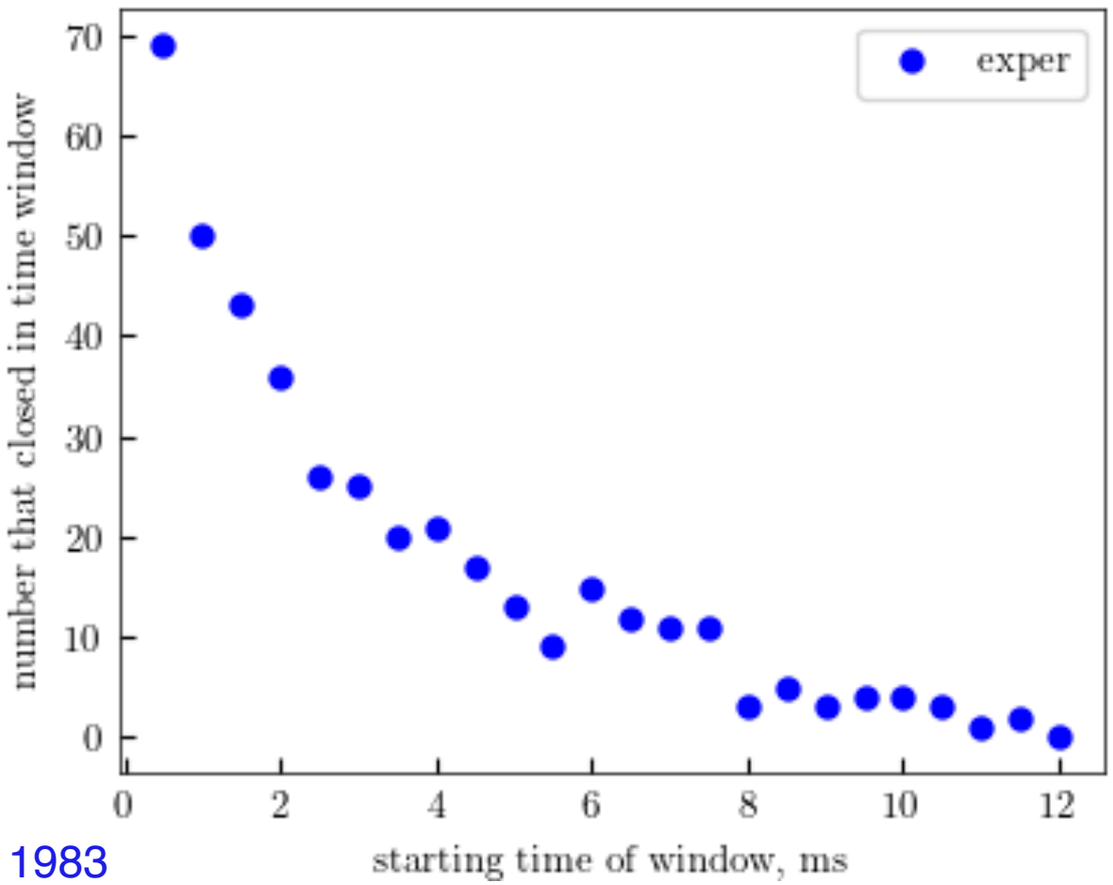
open --> closed.

Thus, at first the total electrical conductance remained high, but then it began to drop.

The conductance dropped in discrete steps of equal size, which were well enough defined to count how many happened in each time bin. There were a total of 430 such steps. Here are the numbers that closed in each of 24 time windows.

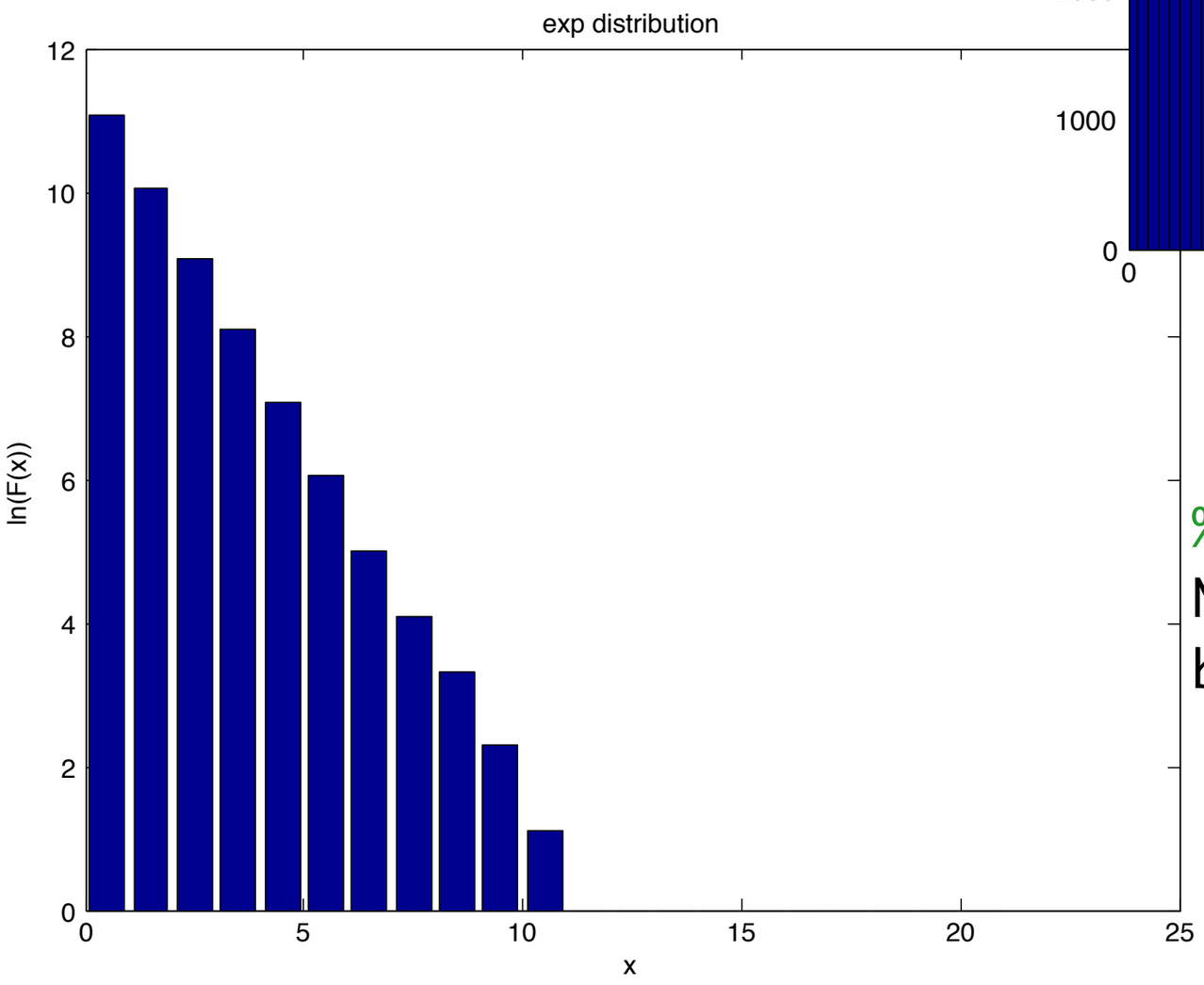
A simple physical model for these (partly random) data:

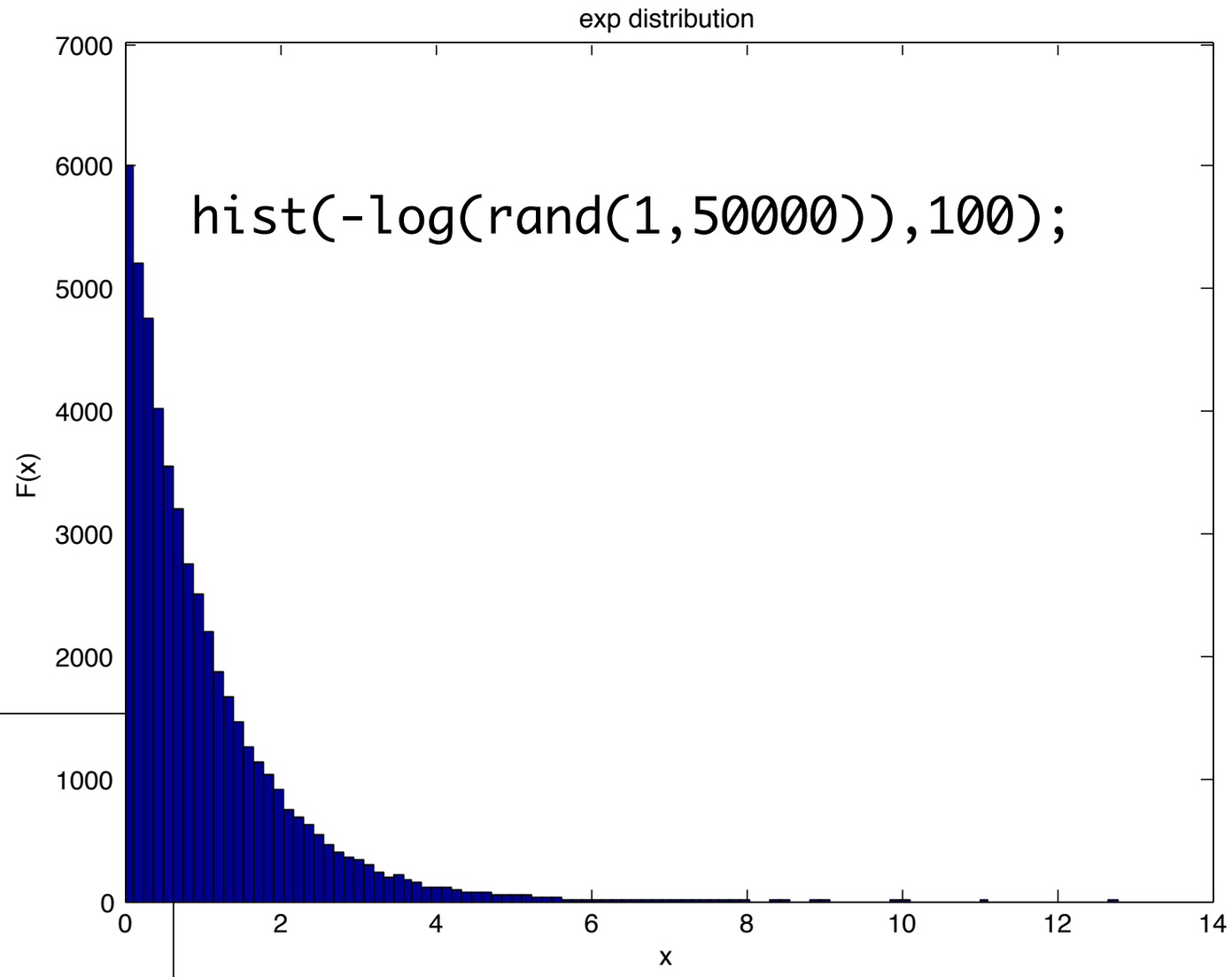
- channels make independent transitions...
- with exponentially-distributed wait times.



Data from Colquhoun and Hawkes in "Single channel recording" 1983

Can we get a computer to draw from an exponential distribution?





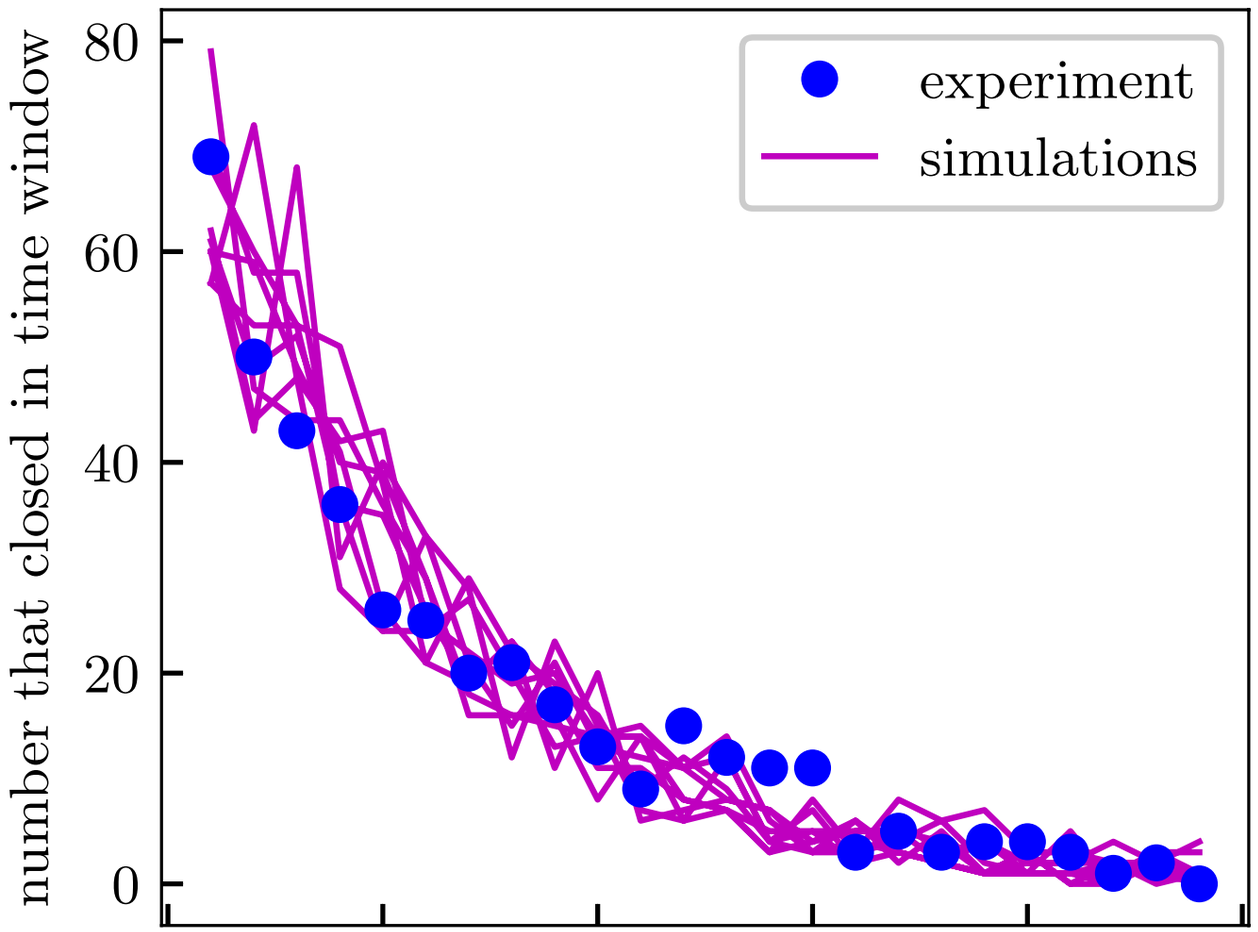

```
hist(-log(rand(1,50000)),100);
```

```
% make it nicer by plotting log freqs:
N=histc(-log(rand(1,100000)),0:20);
bar(.5+(0:20),log(N));
```

PN, *Physical models of living systems*

Phil Nelson

A single fit parameter (rate constant) successfully reproduces both the mean behavior and also the observed noise.



Thermally-activated hopping

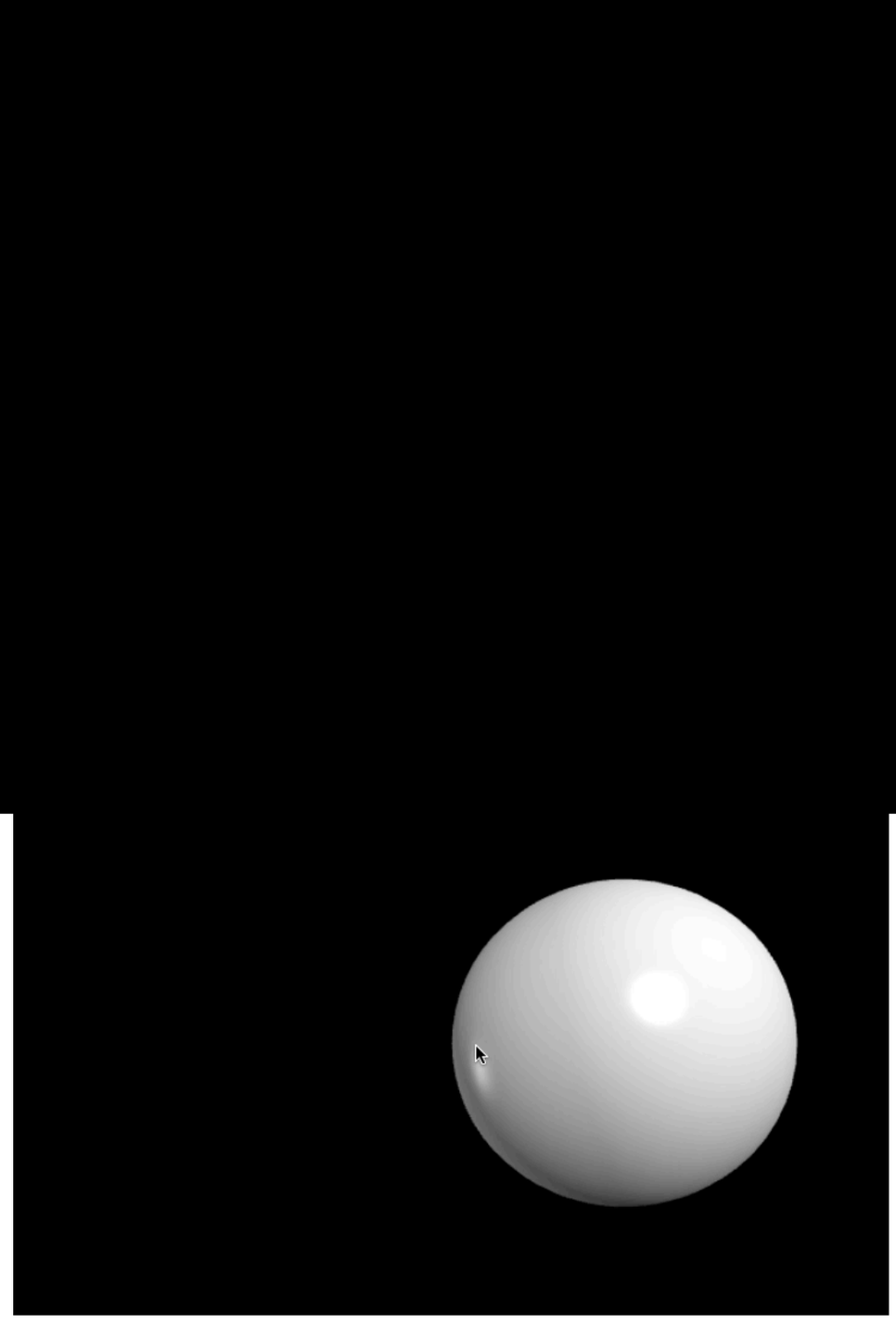
Now let's move from the very simplest possible scheme:

A --> B

to the next simplest: a single instance of:

A <--> B

Just *a few lines* of code succeed in reproducing realistic behavior:



Video micrograph courtesy Adam J Simon; see Simon and Libchaber. Phys. Rev. Lett. 68, 3375 (1992).
http://www.physics.upenn.edu/biophys/PMLS/Media/brownian/BeadJump.mov

```
from numpy.random import random as rng
```

```
framerate = 10  # frame rate in Hz, so 10 means 100ms per frame
```

```
dts = -.5*np.log(rng(100)) # waiting times [s] with mean (0.5)s
```

```
ts = np.cumsum(dts)        # transition times [s]
lasttime = ts[-1]
print(lasttime)
```

```
59.983106299387686
```

```
scene = canvas()
myobject = sphere(pos=vector(1,0,0))
```

```
for i in range(15*framerate): rate(framerate)# pause 15 seconds
j = 0                        # which transition we're on now
for timenow in np.arange(0, lasttime-1, 1/framerate):  # which video fram
    while ts[j] < timenow:   # decide whether a new event has happened
        j+=1                 # since last frame
    myobject.pos = vector((-1)**j,0,0)
    rate(framerate)          # pause till it's time to show the frame
```

This talk

Ion channels

Gillespie algorithm

Bursting in transcription

Kinetic proofreading

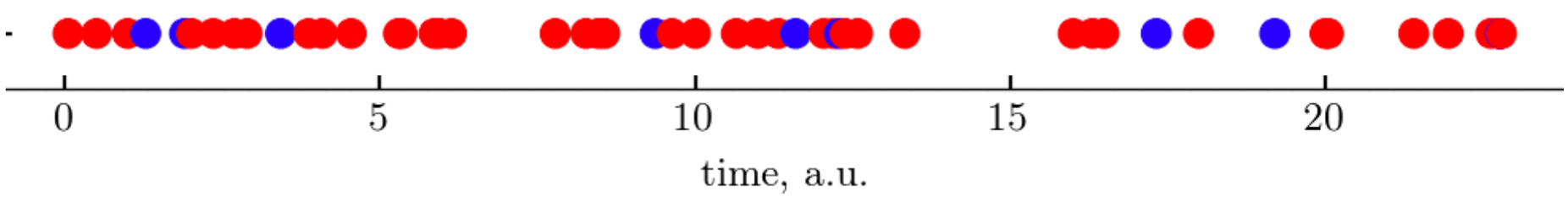
Adaptation in chemotaxis

Phil Nelson

More complicated processes

Perhaps I want two *independent* processes (not required to alternate), with *different* rate constants.

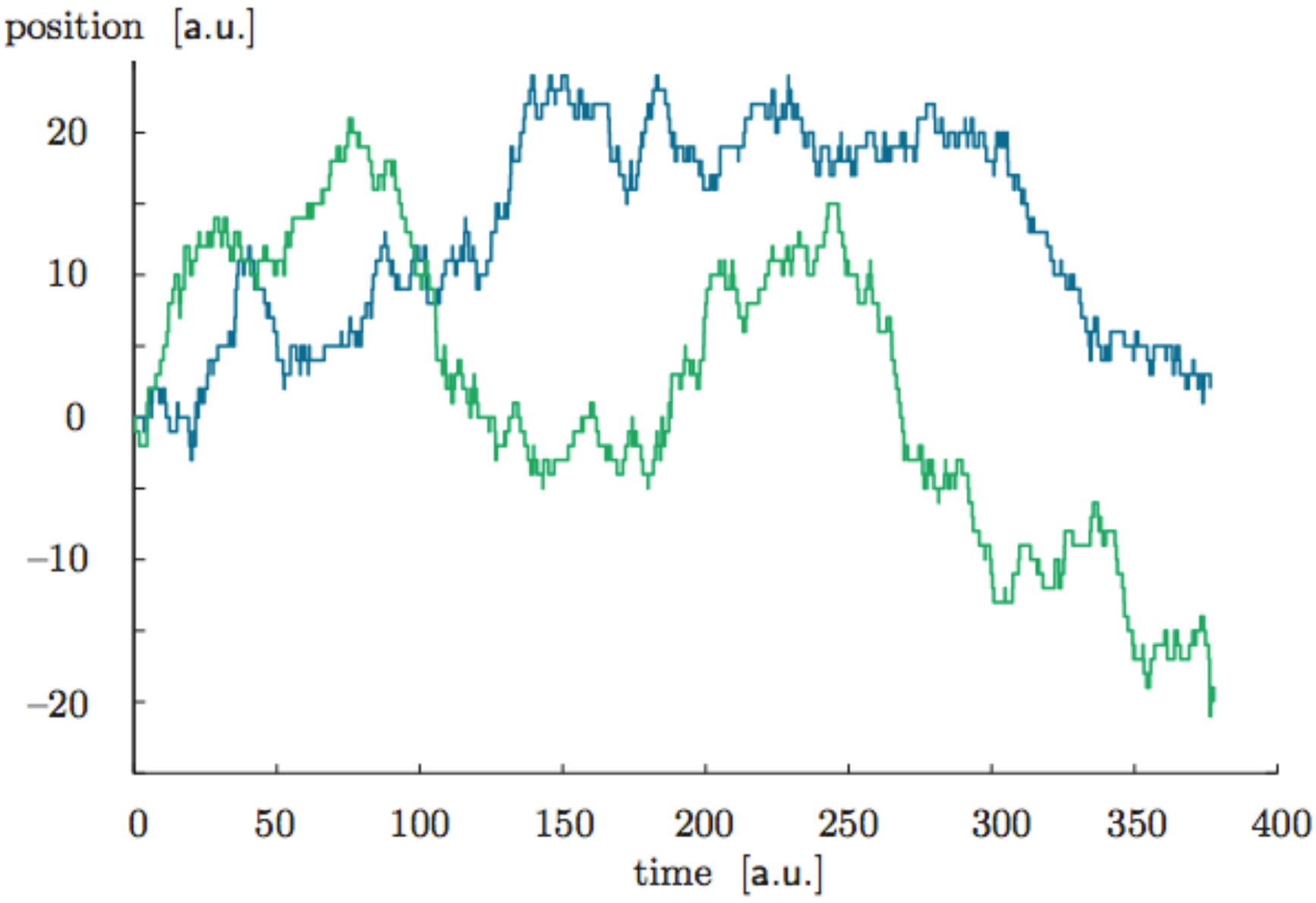

```
betaA = .5
betaB = 2   # type B comes faster than A
Nsteps = 50 # how many steps

betatot = betaA + betaB
deltats = -np.log(np.random.random(Nsteps))/betatot; # waiting times
ts = np.cumsum(deltats);                             # event times
types = [] # event types
for i in range(len(ts)):
    if np.random.random() < (betaA/betatot):
        types = types + ['b']
    else:
        types = types + ['r']
plt.scatter(ts, np.zeros(Nsteps),c=types)
```



We just invented the "Gillespie algorithm" for "stochastic simulation."

["Created by Joseph L. Doob and others (circa 1945) and popularized by Dan Gillespie in 1977." – `http://en.wikipedia.org/wiki/Gillespie algorithm`]

From there, it's a short "step" to random walks (the two processes are left and right stepping):



Phil Nelson

This talk

Ion channels

Gillespie algorithm

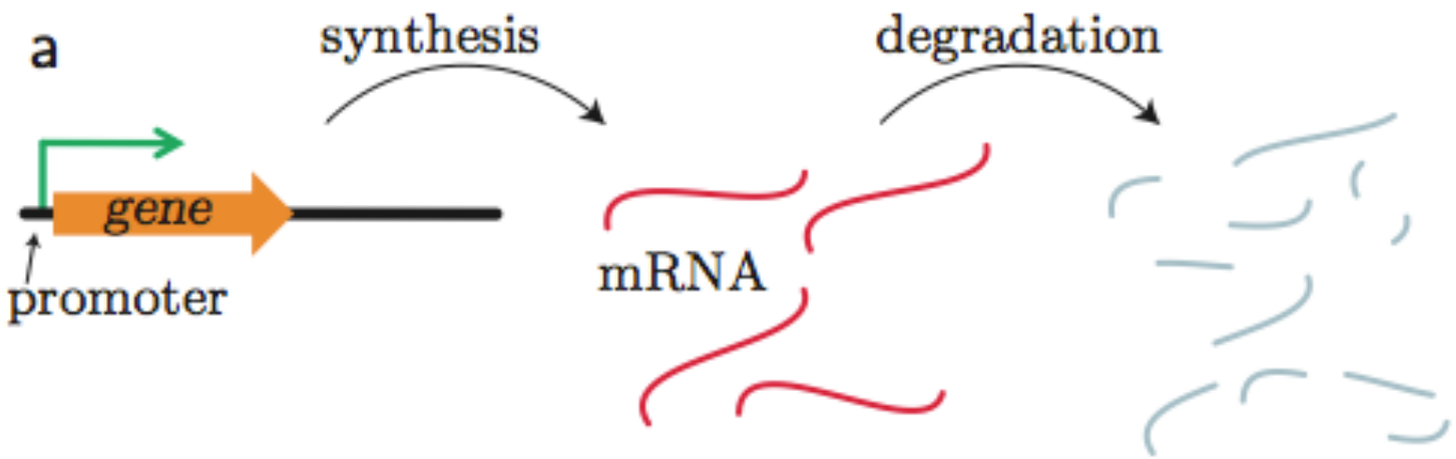
Bursting in transcription

Kinetic proofreading

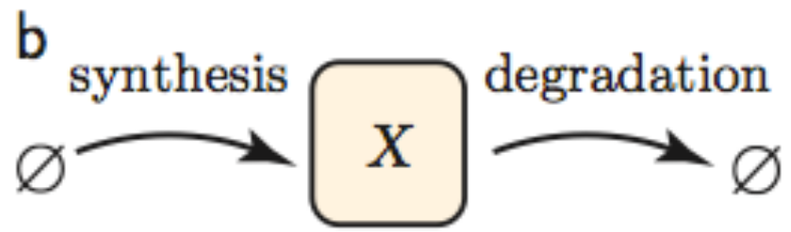
Adaptation in chemotaxis

Gene expression: Birth-death process

One abstract representation:

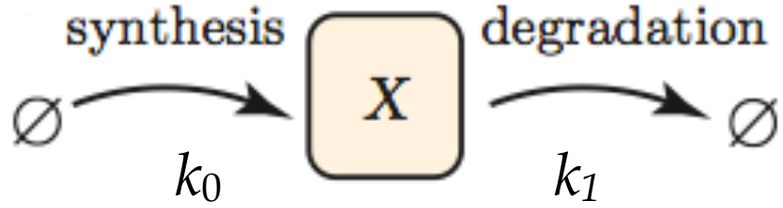


Even more abstract representation:



What's new here is that the propensities are not constants. Doesn't matter!

Upgrade code to handle the birth/death process:

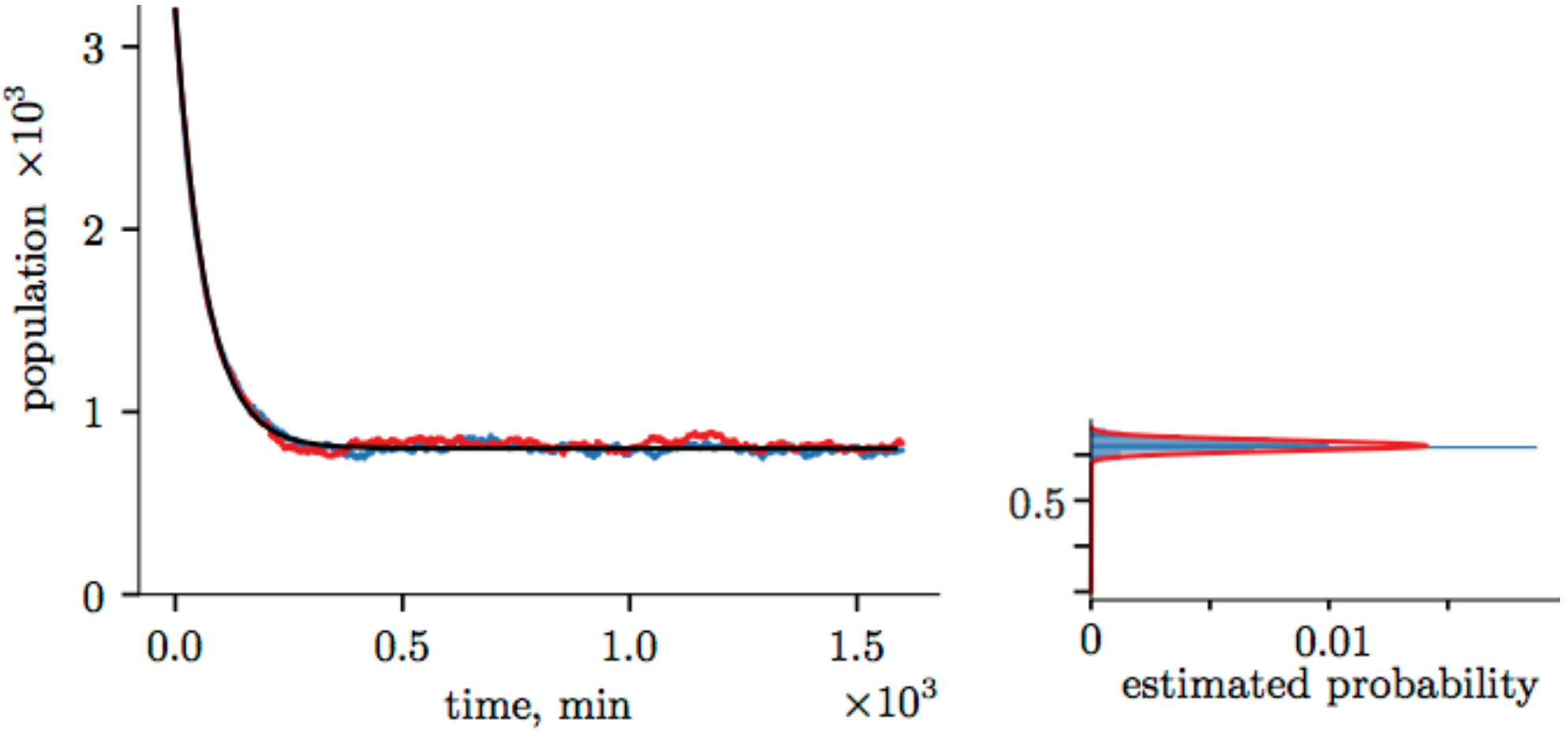



```
def transcrip2rxn(lini, T):
    ''' inputs:
    lini = initial number of mRNA
    T = total time to run
    outputs:
    ts = times at which x changed
    ls = running values of x just after those times'''
#%% Parameters:
    ks =  np.array([0.15, 0.014]); # rate constants in 1/minute
    stoich = np.array([0, 1]);     # reaction orders
#%% initialize
    t = 0;    # current time
    x = lini; # current mRNA population
    ts = [t]; # histories
    ls = [x];
    while t<T:
        a = (x**stoich) * ks                    # propensities
        atot = np.sum(a);                       # total rate for anything to happen
        t = t - np.log(np.random.random())/atot
# make birth/death decision based on the relative propensity:
        mu = 1 - 2*(a[0]/atot < np.random.random())
        x = x + mu
        ts = ts + [t]
        ls = ls + [x]
    return (ts, ls)
```

Phil Nelson

Birth-death process

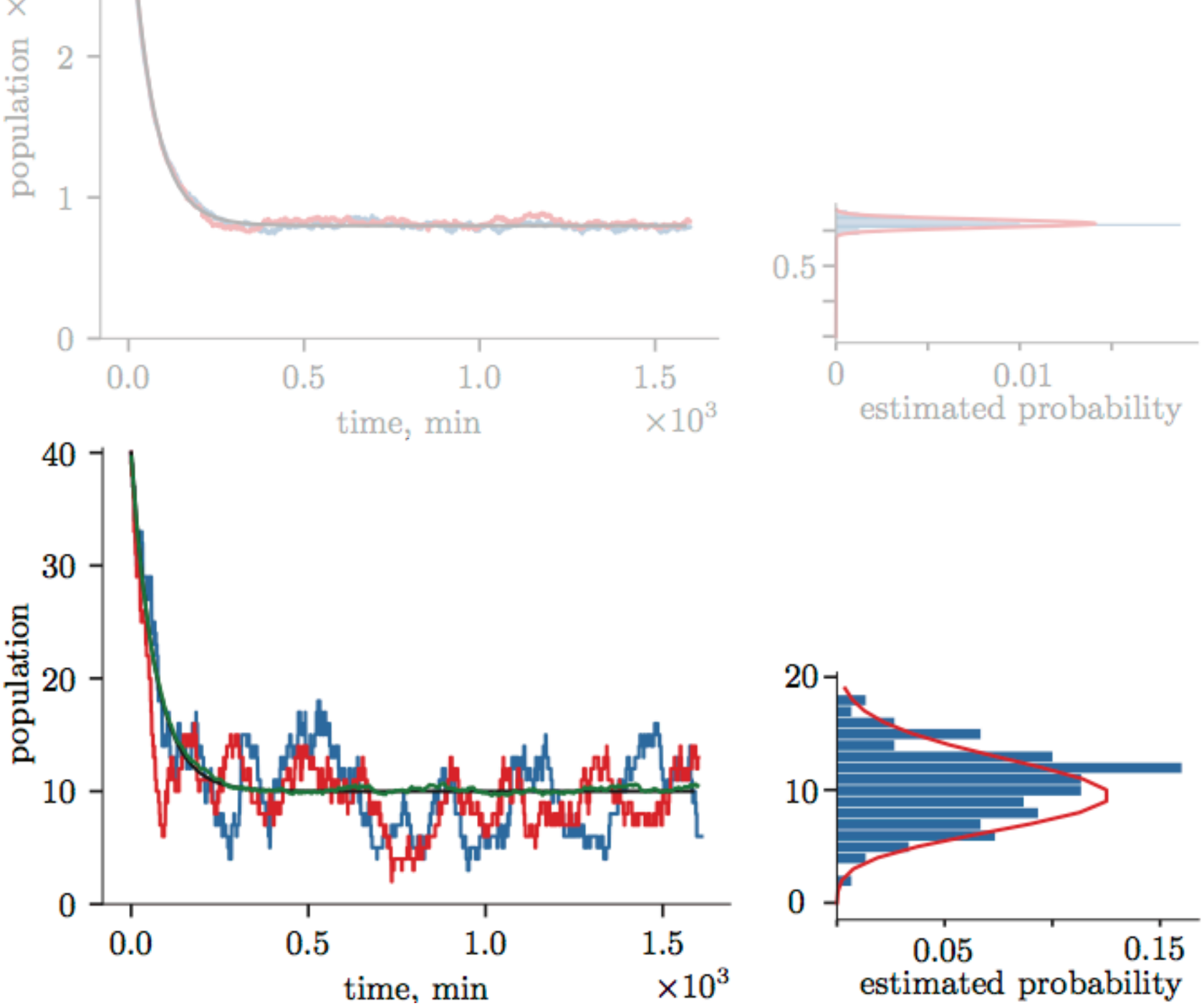
Here's an example of two time series (*red, blue*), for the case in which the molecule count starts out higher than its steady value.
Black trace is the continuous deterministic approximation.



Hmm, seems like a lot of work just to recover exponential decay! Was it worth the effort?

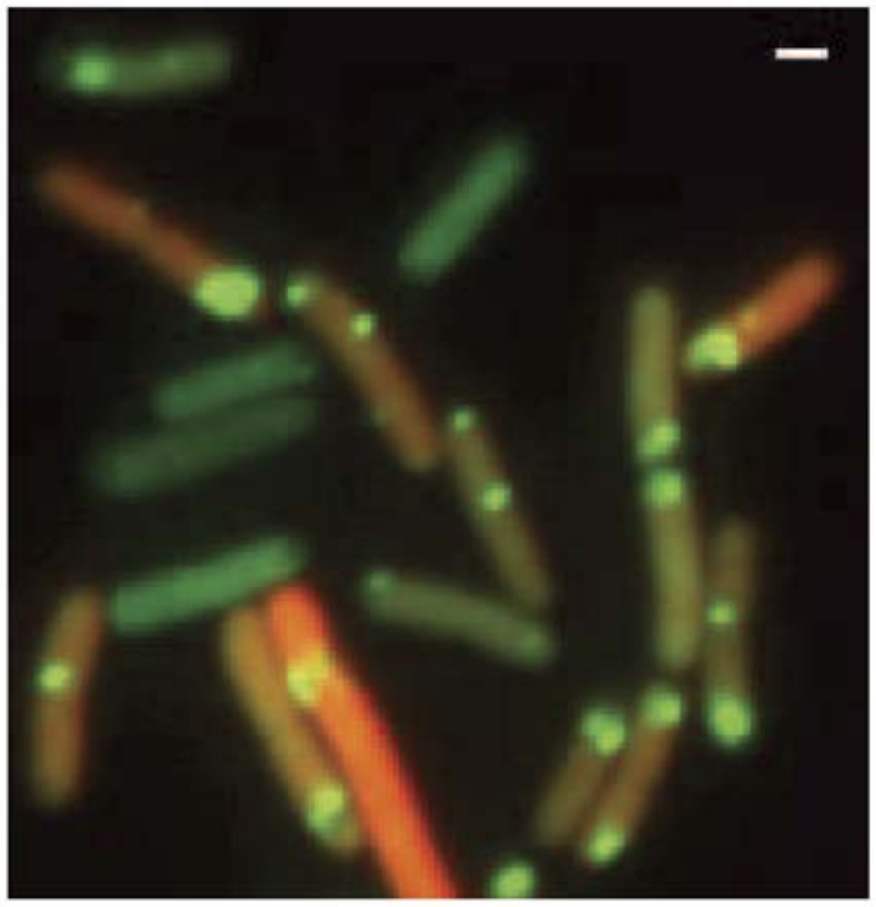
PN, *Physical models of living systems*.

Something much more interesting happens when the total numbers are not large. We see that the "steady" state can actually be pretty lively (big fluctuations). And interesting–those fluctuations follow a very famous distribution.

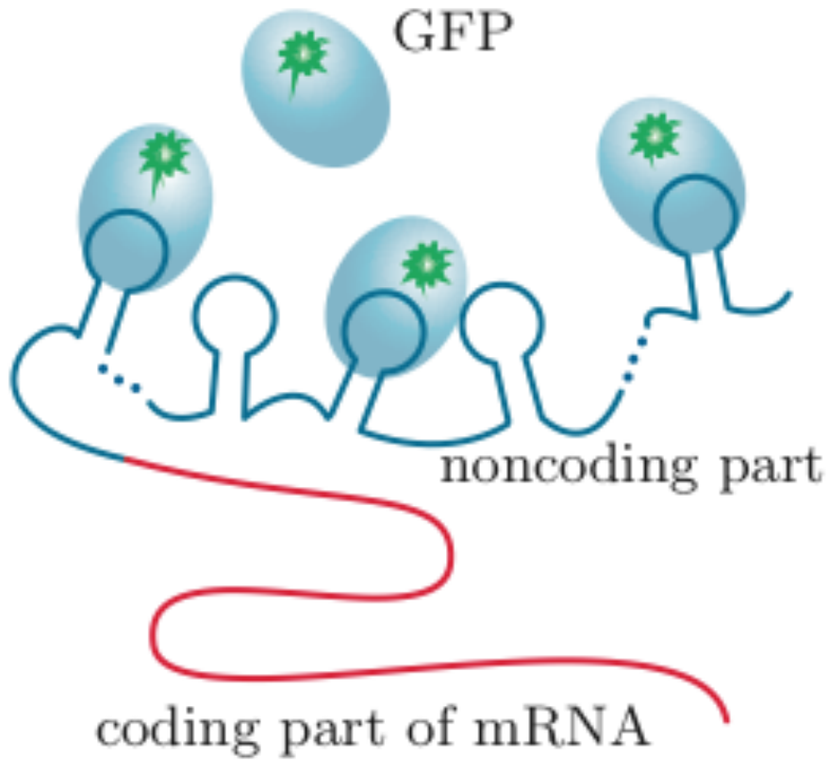


Um… Why not just see this via the Master equation? Well, for one thing there's no substitute for watching individual time courses, which after all is what single-molecule experiments see. Also, often the moment you add one little extra bit of realism, then the analytic solutions are lost. Let's look at some of that realism.

Counting single molecules in live cells



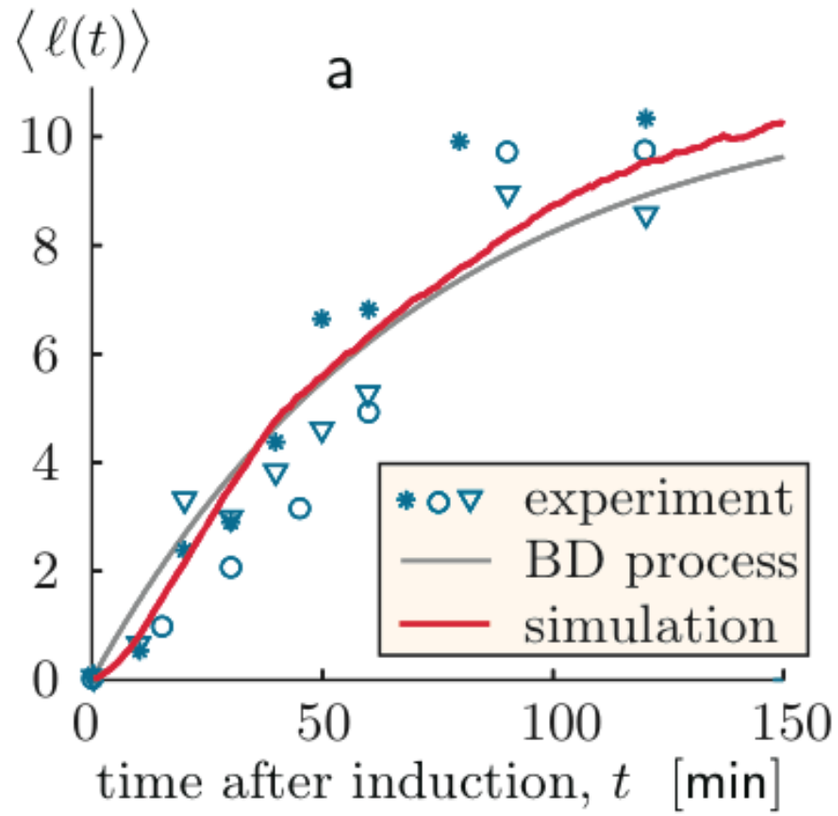
Here, Red color indicates gene product (RFP1). (Green foci are each one or m mRNA.)

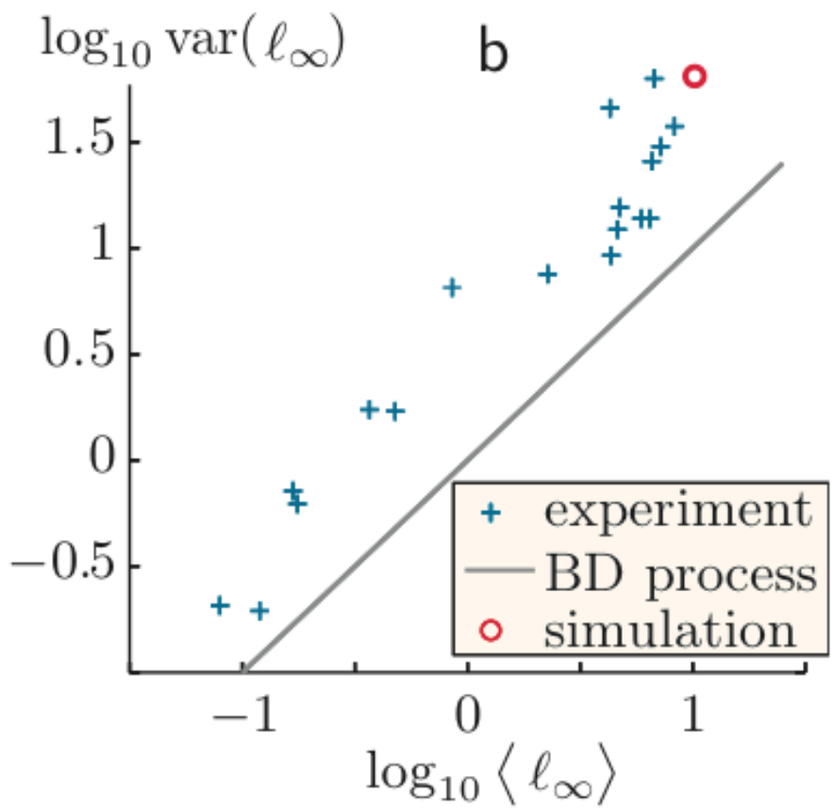


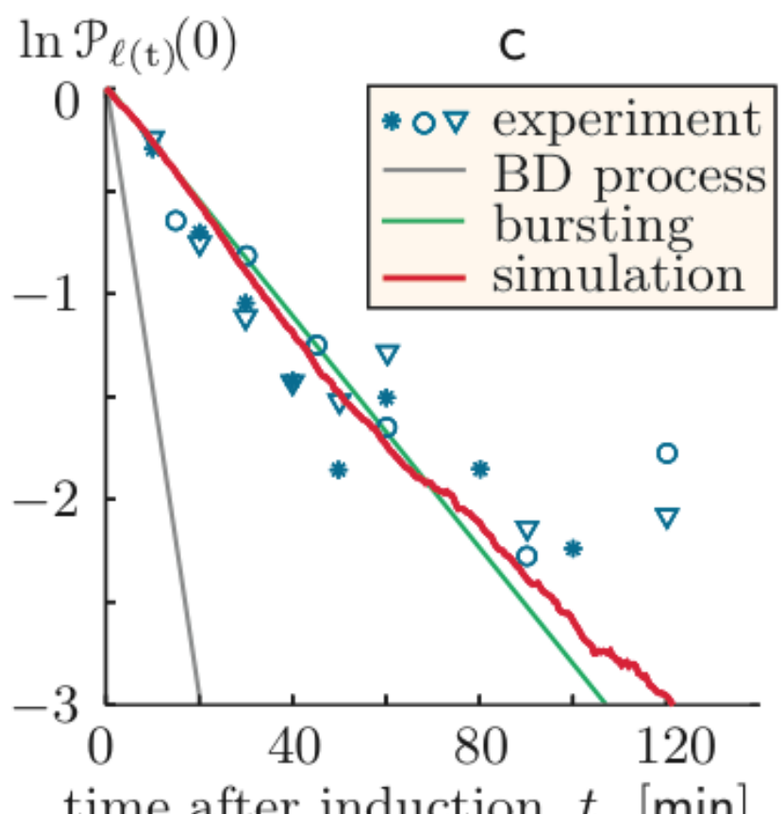
Detection of mRNA and protein in living cells. The picture is a false-colored overlay of the green and red channels. Scale bar, 1 μm.

Golding, I., Paulsson, J., Zawilski, S. M., & Cox, E. C. (2005). *Cell*, *123*(6), 1025. 10.1016/j.cell.2005.09.031

Failure of birth-death model







mRNA dynamics appears at first to be Poisson, rising and saturating as usual.

But the distribution in steady state is not Poisson!

Nor does the probability to have zero transcripts fall in the way expected.

These quantitative failures of the birth-death model led to a *bursting hypothesis*:
* If each burst makes k copies, but the burst-initiation rate is k times slower than in the B-D model, then the first time course will look about the same and we retain that good agreement.
* But then the variance increases by a factor of k^2, so we fix the second graph.
* And the initial slope of the last graph also decreases by a factor of k, fixing it too.

PN, *Physical models of living systems*. Data from Golding, I., Paulsson, J., Zawilski, S. M., & Cox, E. C. (2005). Cell, 123(6), 1025. 10.1016/j.cell.2005.09.031

Go long

Simulation sometimes gets a bad rap because admittedly it is often unilluminating, doesn't give general results, we can't evaluate its validity, etc. Generally speaking, that happens when we have to introduce extra ad hoc assumptions to get the system into simulatable form.

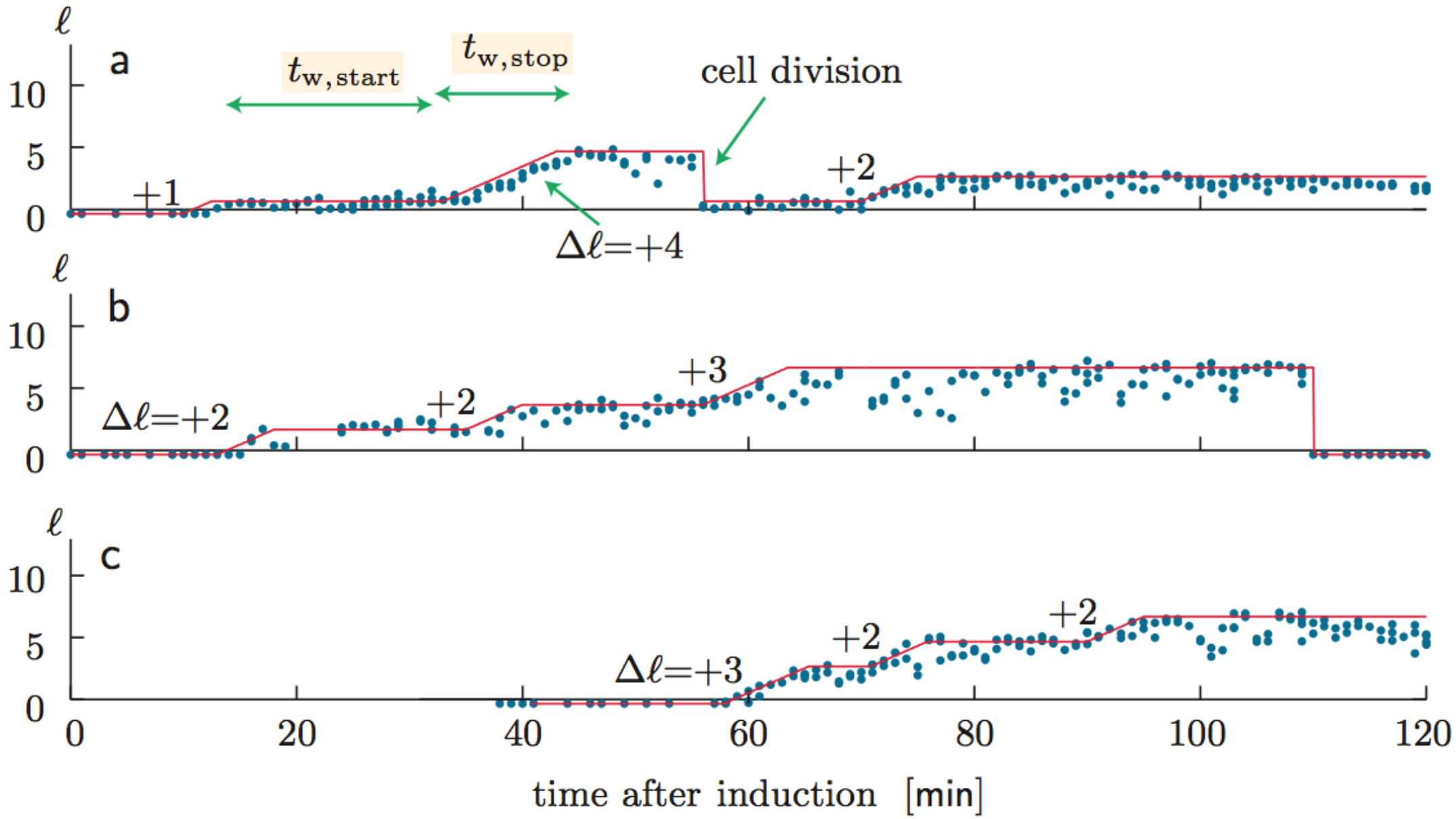
However, if your model's *primary definition* is in terms of transitions among discrete states, then simulation is its *most direct* implementation; here it can be the case that *analytic* results require suspect approximations, not the other way round.

What we just saw was a reasonable-looking model getting falsified because its predictions, which required simulation, could not fit data.

> Often we don’t remember the crucial role of such indirect inference because it only served to motivate a more direct experiment. The new experiment gets all the glory, but often it *would never have been done* (or not till years later) without the kick in the butt from the indirect argument.

Phil Nelson

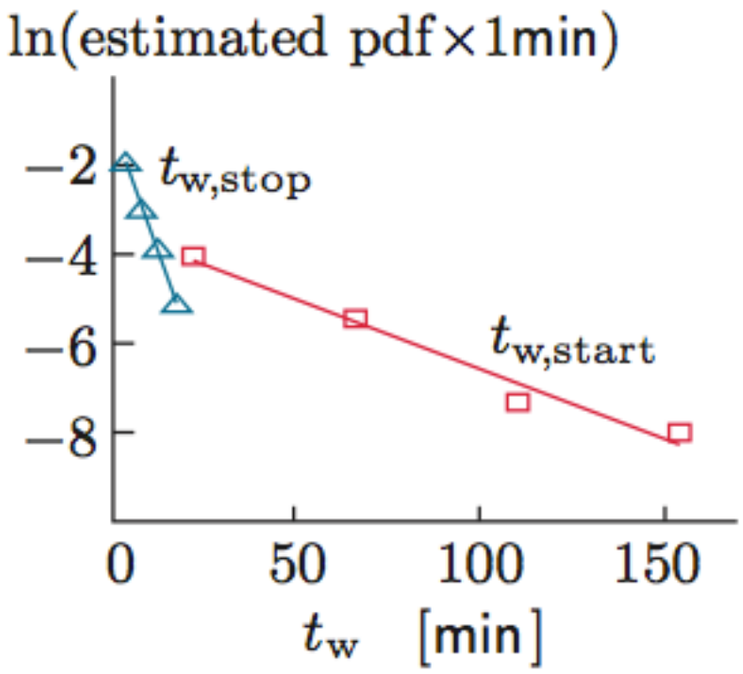
Indeed, Golding et al. were *then* able to observe "bursts" of mRNA synthesis directly:



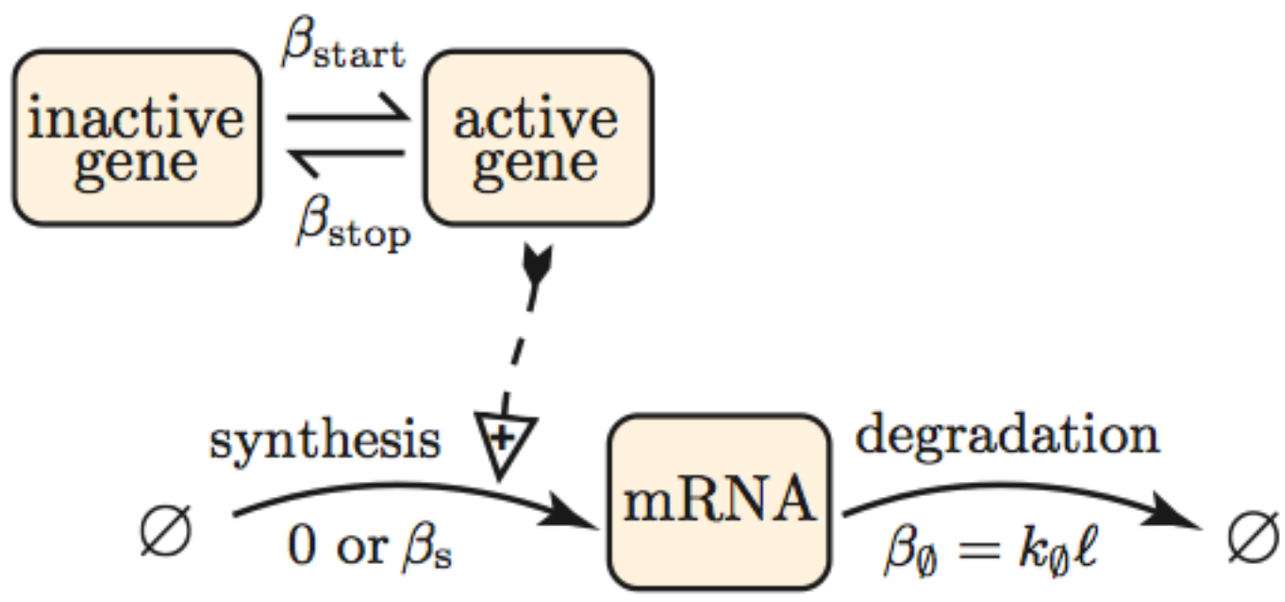
Golding et al 2005

Gene expression: Bursting

The starts and ends of bursts were found to be exponentially distributed (Golding et al 2005):



And that led the authors to propose an almost-simple model:



… a stochastic simulation was needed to confront the new model with experiment, and succeeded…

though its underlying molecular mechanism was only found several years later!

Golding et al 2005

This talk

Ion channels

Gillespie algorithm

Bursting in transcription

Kinetic proofreading

Adaptation in chemotaxis

Everybody knows "the secret of Life is DNA," right?

But it is less well appreciated that the stability of a molecule of DNA does not guarantee the accuracy of its replication and transcription. There is *another big secret* here, just as essential to Life as the well known one.

Can we simulate a model of this process, e.g. the famous one due to John Hopfield? Doing so would help us to understand the claim that the model captures important aspects of Nature's solution.

> Imagine that you run an art museum and wish to find a mechanism that picks out Picasso lovers from among all your museum's visitors. You could open a door from the main hallway into a room with a Picasso painting. Visitors would wander in at random, but those who do not love Picasso would not remain as long as those who do. Thus, the concentration of Picasso lovers in the room would arrive at a steady value (with fluctuations, of course) that is enriched for the desired subpopulation.
>
> To improve the enrichment factor further, you could hire an employee who occasionally closes the door to the main hallway, stopping the dilution of your enriched group by random visitors. Then open a new exit doorway onto an empty corridor. Some of the trapped visitors will gratefully escape, but die-hard Picasso lovers will still remain, leading to a second level of enrichment.
>
> After an appropriate time has elapsed, you can then reward everyone still in the room with, say, tickets to visit the Picasso museum in Paris.

— Paraphrased from *An introduction to systems biology: Design principles of biological circuits* by Uri Alon

Let's try to apply this metaphor and see via simulation whether it really can explain (some of) the high fidelity of the ribosome.

a

0

T

1/2

T

the "second exit"!

+ D

P_i hydrolysis

D

★

3/4

closing the "entrance doorway"

: ribosome

: mRNA

: codon

: tRNA

●,● : amino acids

T : GTP

D : GDP

D

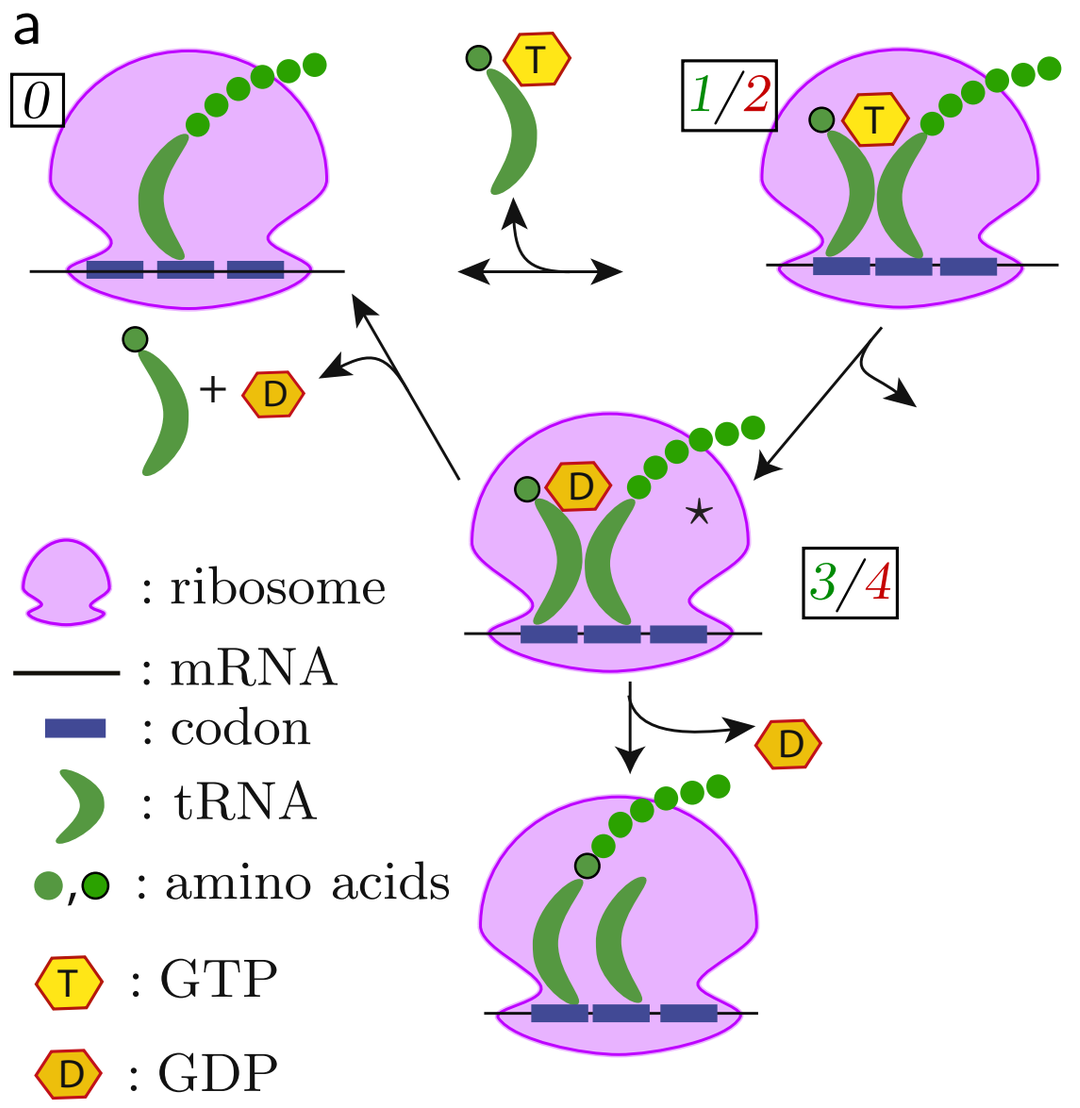
Then shift right and go back to *0*.

Phil Nelson

Fine print

- Each step shown probably represents multiple substeps. For example, GTP hydrolysis is subdivided into GTPase activation followed by actual hydrolysis.

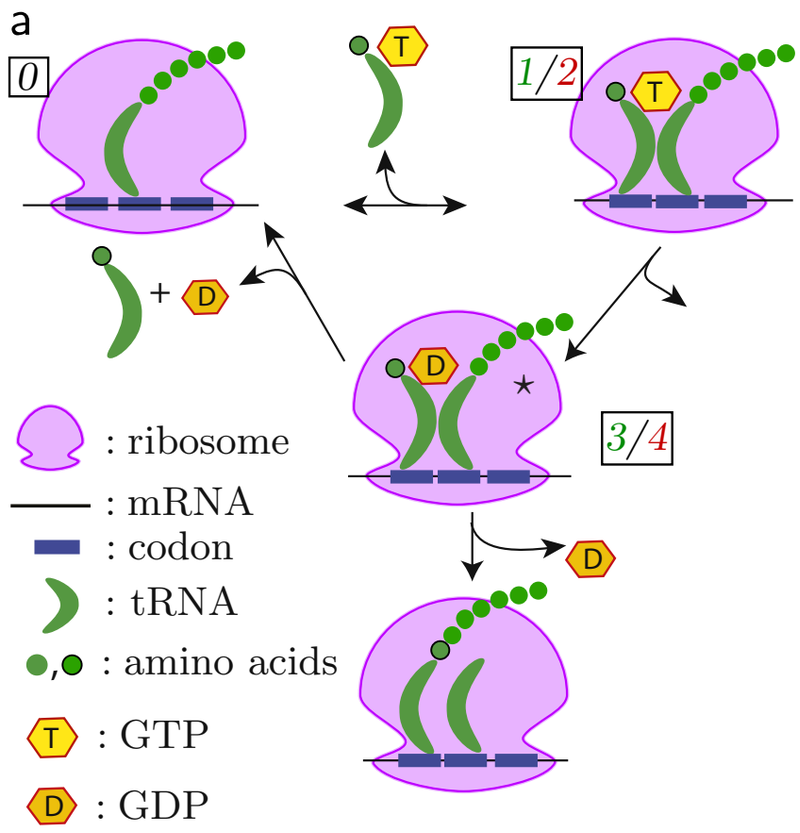
- The mechanism assumes, and there is evidence that,
  - No GTPase activity until the complex is bound to a ribosome (the "clock doesn't start" till then).
  - After hydrolysis the complex can't get back to *0* by swapping the GDP for an ambient GTP from solution.
  - The last "commitment" step cannot happen until after GTP hydrolysis.



- Correct binding can also speed up some forward steps, leading to further reduction in error rate. We will consider only Hopfield's original idea, where the only selection was that *incorrect* binding speeds up the *rejection* steps. This has been confirmed in single-molecule experiments.

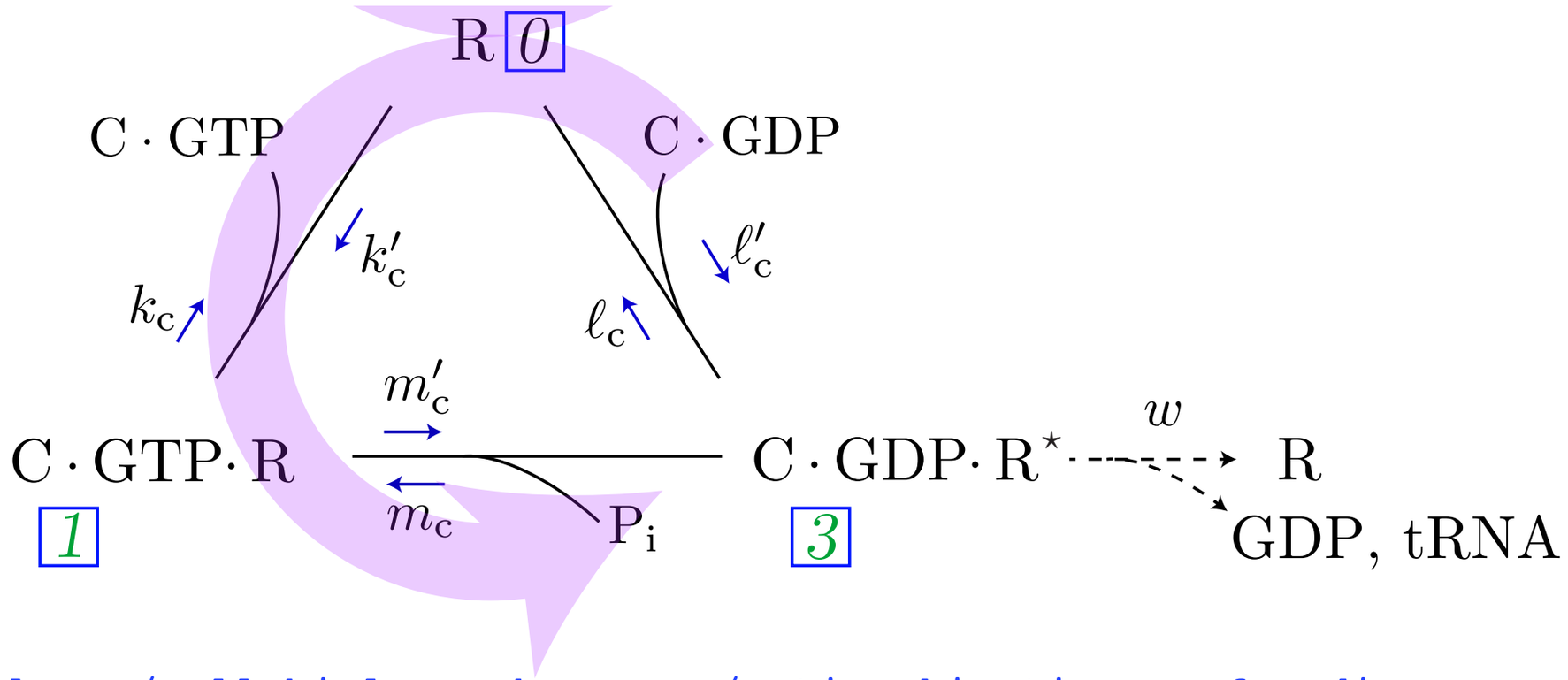
- Recently some evidence of *two* sequential proofreading steps, but we will illustrate with just one.

- We completely neglect the role of EF-G, not implicated in the high fidelity puzzle.



Isn't that terribly inefficient?

We want a stochastic simulation of these five states representing a *single* ribosome in a bath of W's and C's.



Zuckerman,
http://physicallensonthecell.org/cell-biology-phenomena/active-kinetic-proofreading

Phil Nelson

Yes, you can show data in tables and graphs. Sure you can apply lots of sophisticated statistical tools to model it. But it can be important to also present model results in a way that *looks just like your experiment*—to generate a time series and present it to your neural wetware *as a time series*.

That approach can be a great source of intuition.

Phil Nelson

Simulate Hopfield model

KY Chen, DM Zuckerman, PCN (`http://biorxiv.org/cgi/content/short/418772v1`).

Wrap–Proofreading

The ribosome is pretty complex, pretty evolved. I am not claiming that kinetic proofreading explains all (or even most of) its impressive accuracy. I am pointing out that:

- There are known steps corresponding to the ones I have imagined (as well as other steps I didn't show).
- Those steps include initial selection, irreversible GTP hydrolysis, and secondary selection.
- They have now been seen in single-molecule studies. Their sequence is established in part by blocking one step and seeing that the later steps won't happen either.
- All by themselves, without postulating anything more fancy ("induced fit" etc.), these simple steps yield a big enhancement in accuracy…
- As one can confirm by a simple simulation.
- The slowness of the final step (incorporation), crucial for this enhancement, has been observed in single-molecule studies.

This talk

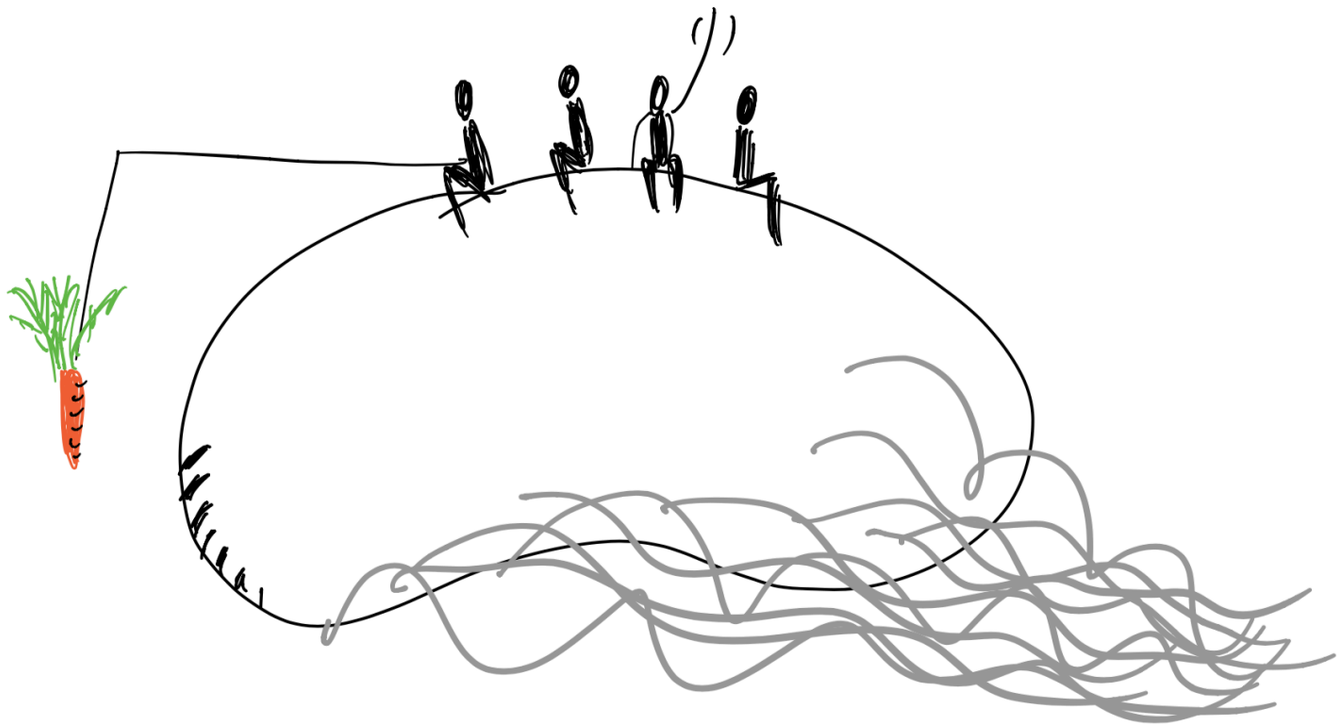
Ion channels

Gillespie algorithm

Bursting in transcription

Kinetic proofreading

Adaptation in chemotaxis

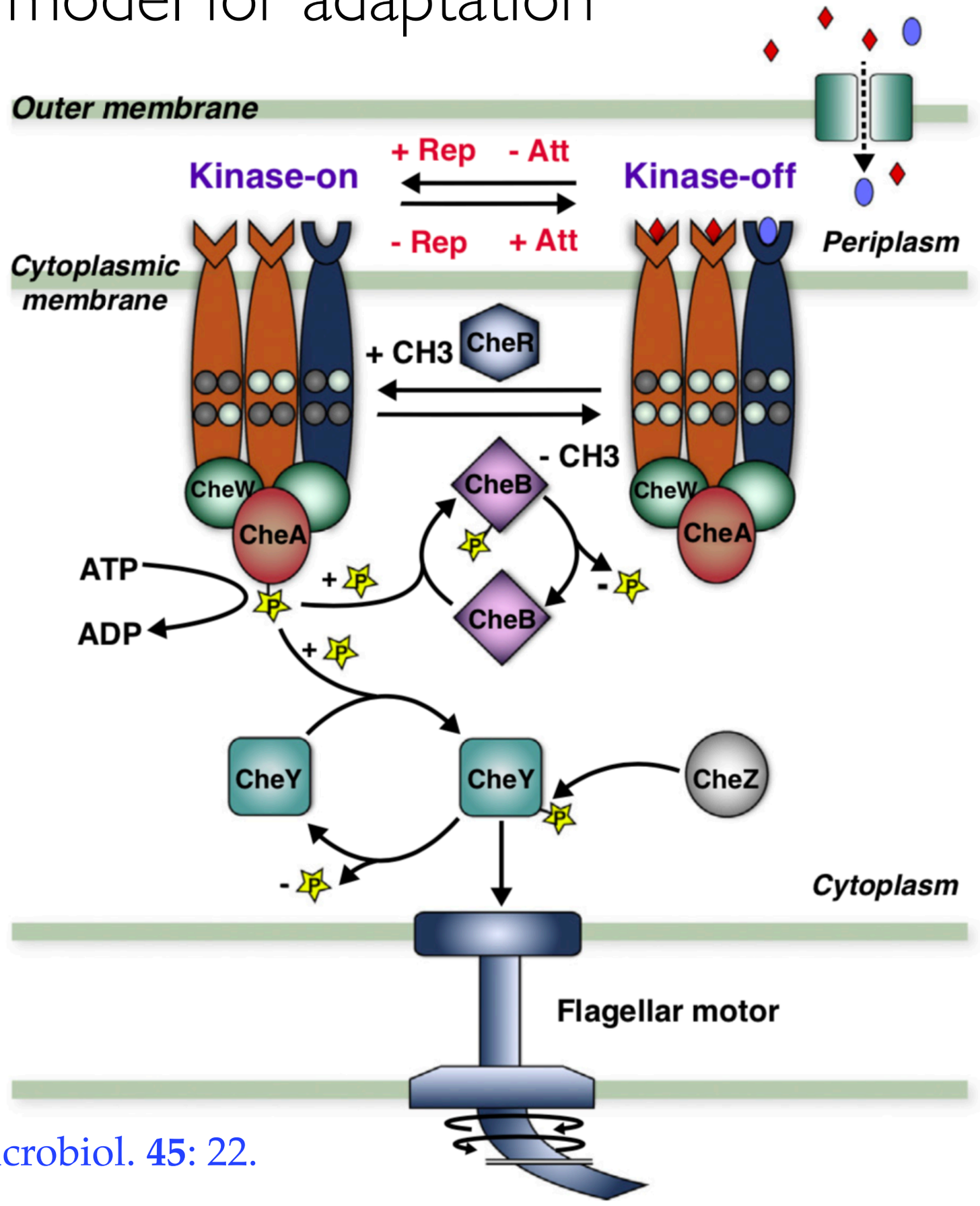


Cartoon: Tina Subic

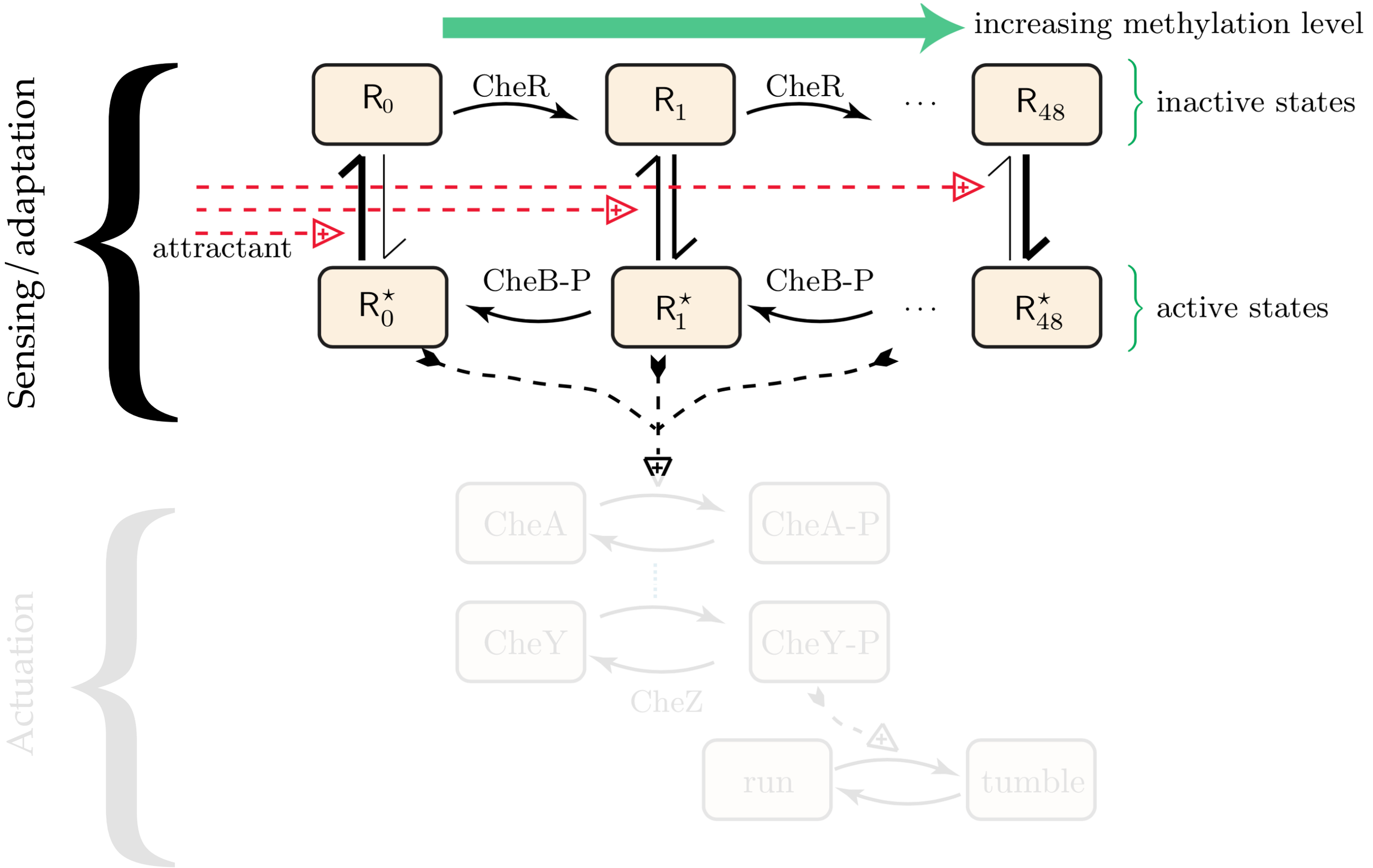
Cartoon model for adaptation

Sensing / adaptation

Actuation



Bi and Sourjik. 2018. Curr. Opin. Microbiol. **45**: 22.

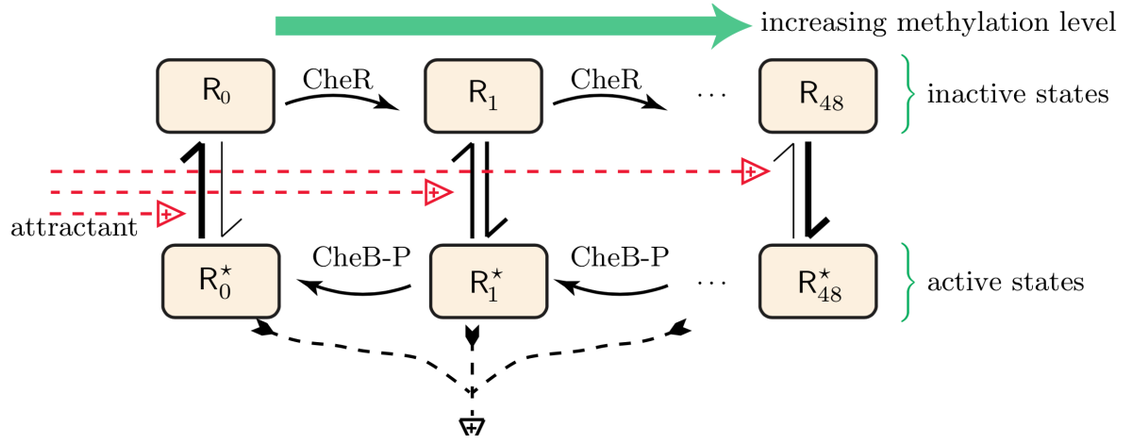


PN, ***From photon to neuron: Light, imaging, vision***. The model was developed by Tom Shimizu, Yuhai Tu, Howard Berg, and coauthors.

Phil Nelson

I asked my students to:

A. See if the model really behaves like the experiment (below).
B. Make a more illuminating graphical representation of *how* the model achieves adaptation.



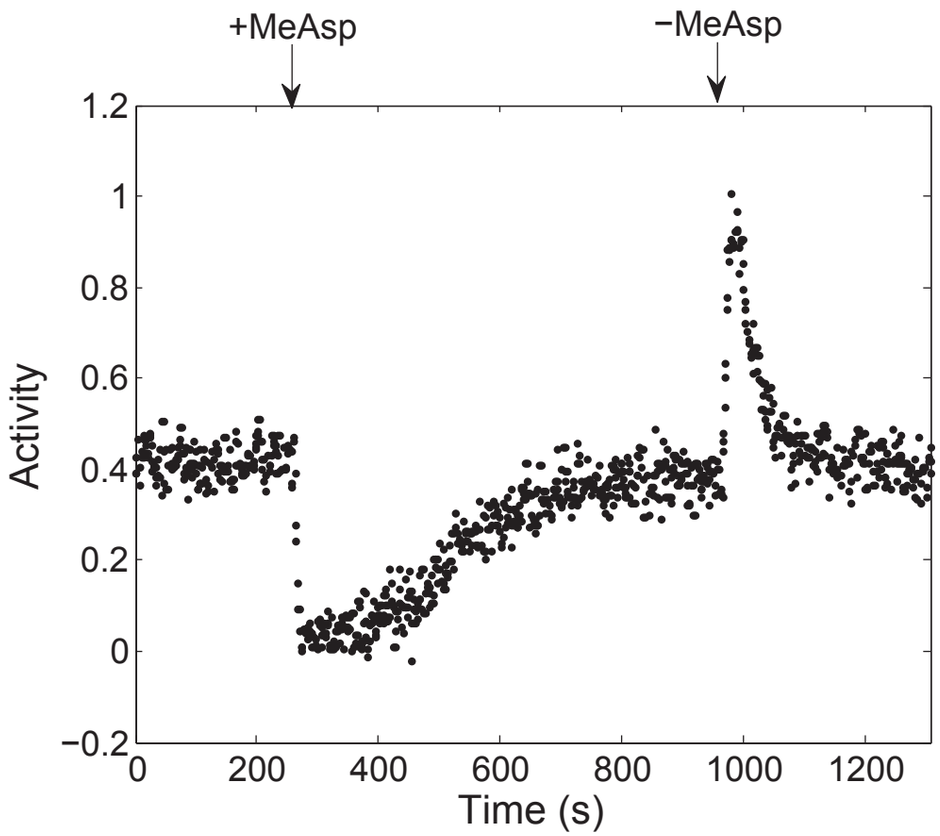
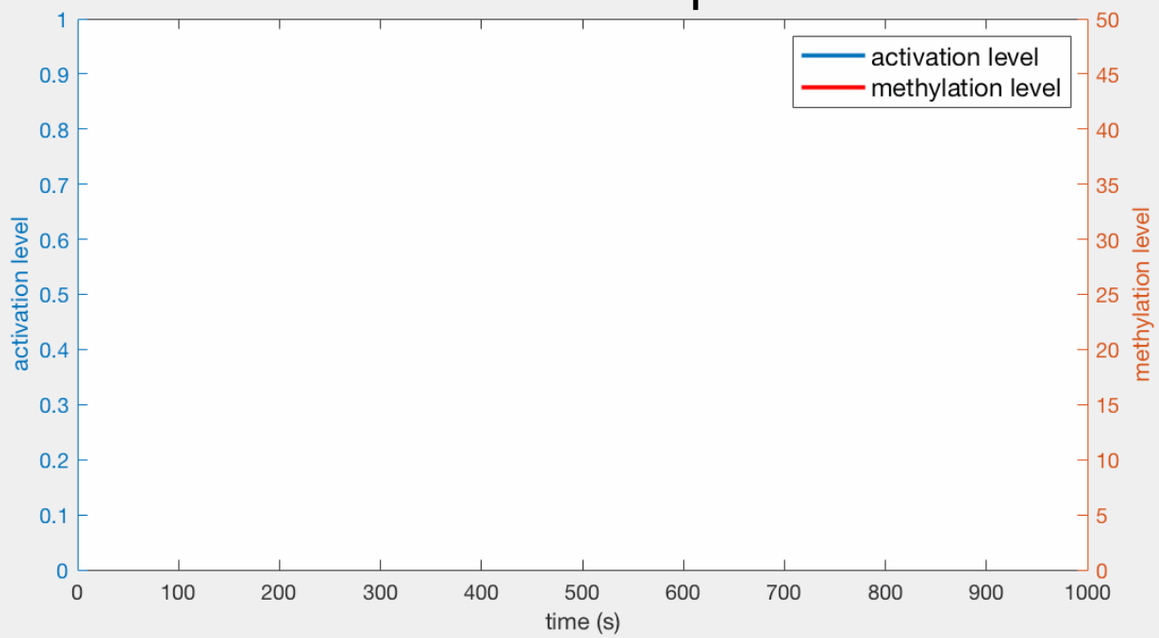


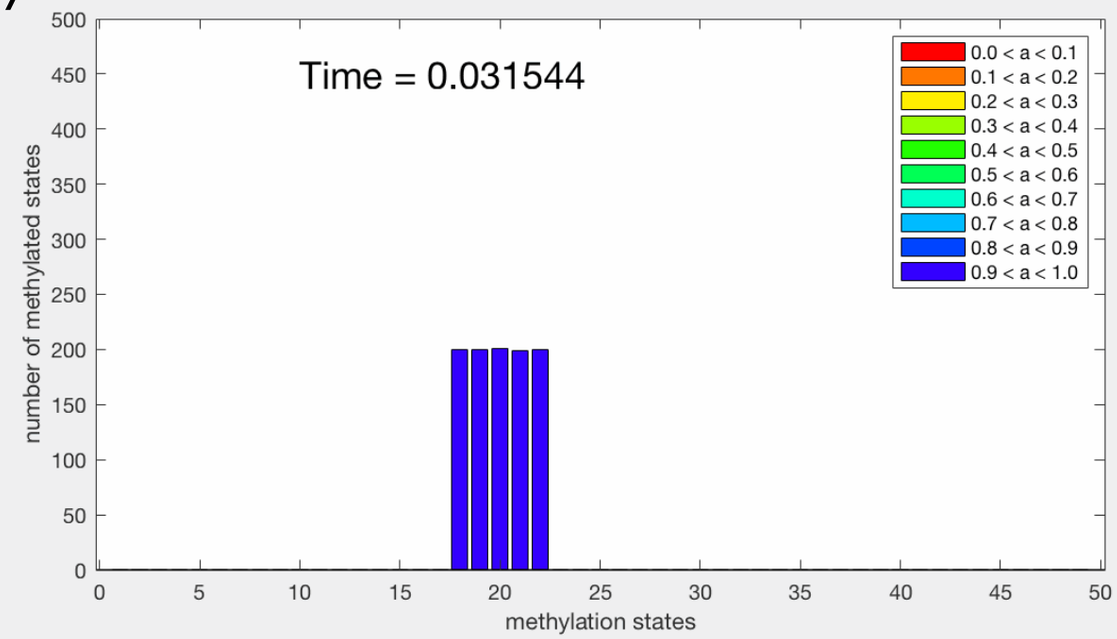
FIGURE 1 Responses of cells of E. coli wild-type strain RP437 to step-addition and removal of 0.05 mM MeAsp, showing the receptor-kinase activity as a function of time. (Arrows) Times of addition and removal of attractant.

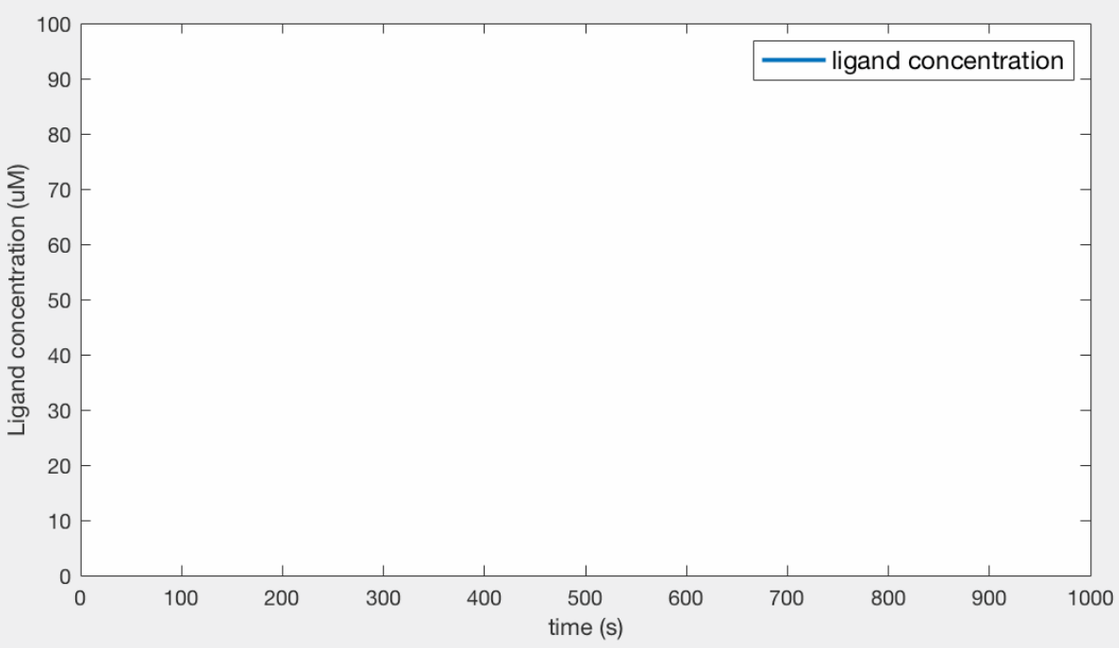
Again we should ask about *noise:*
Again, it's not independently adjustable.
Does the model predict it correctly?
Such questions have been addressed in analytic approaches, but if your model's *primary definition* is in terms of transitions among discrete states, why not address them *directly* in that framework?

Yuan, J., & Berg, H. C. (2012). Biophysical Journal, 103(6), L31–L33.

Adaptation to steady attractant level



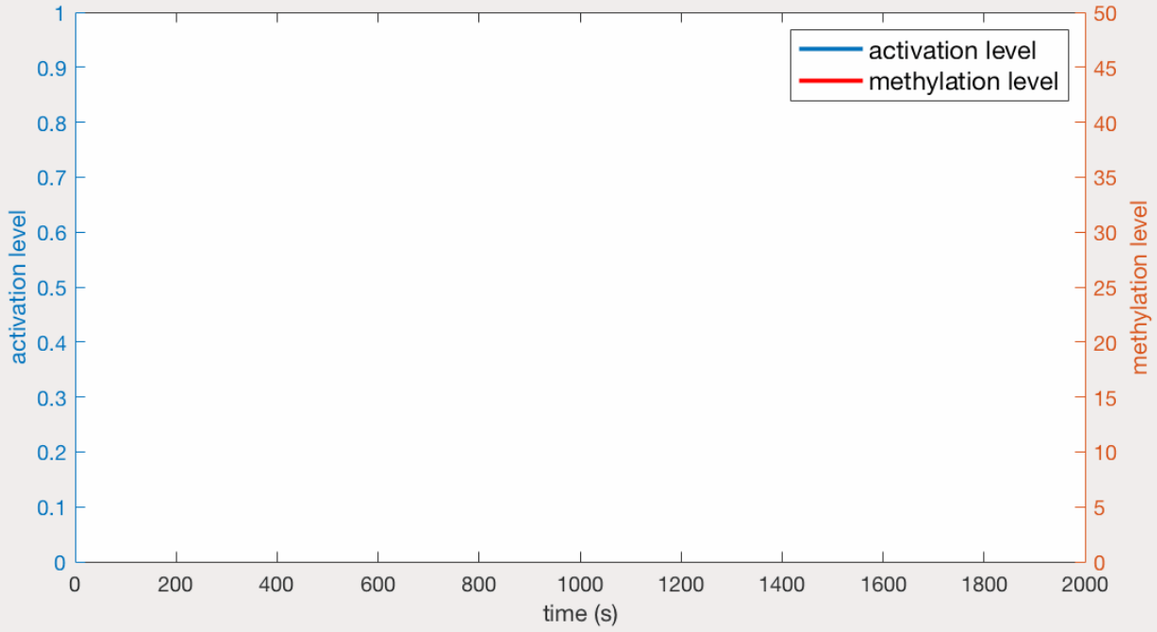


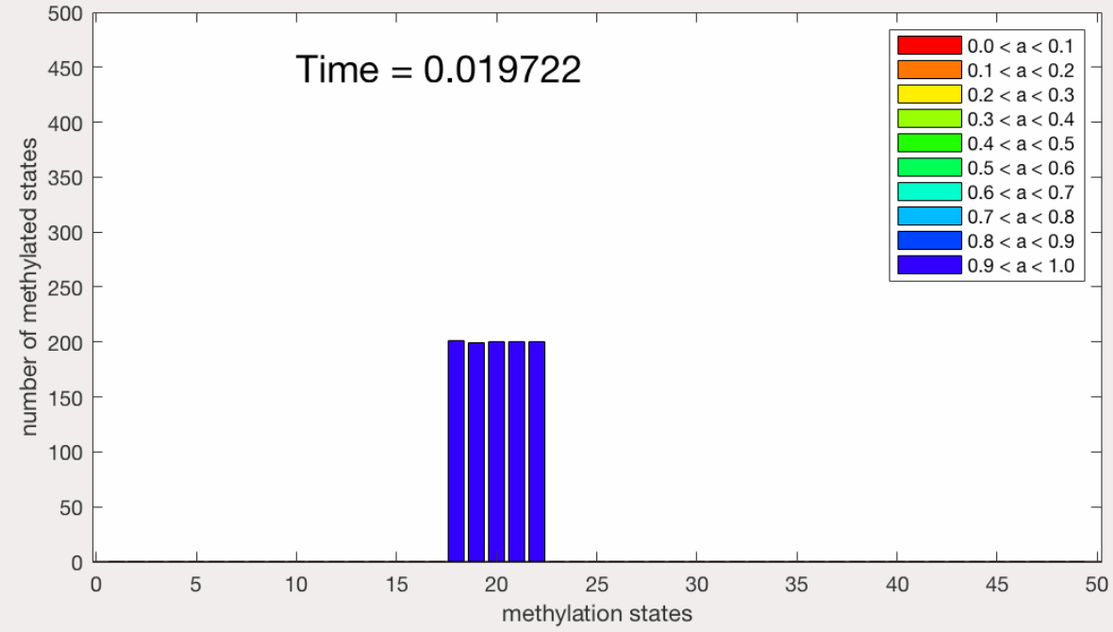


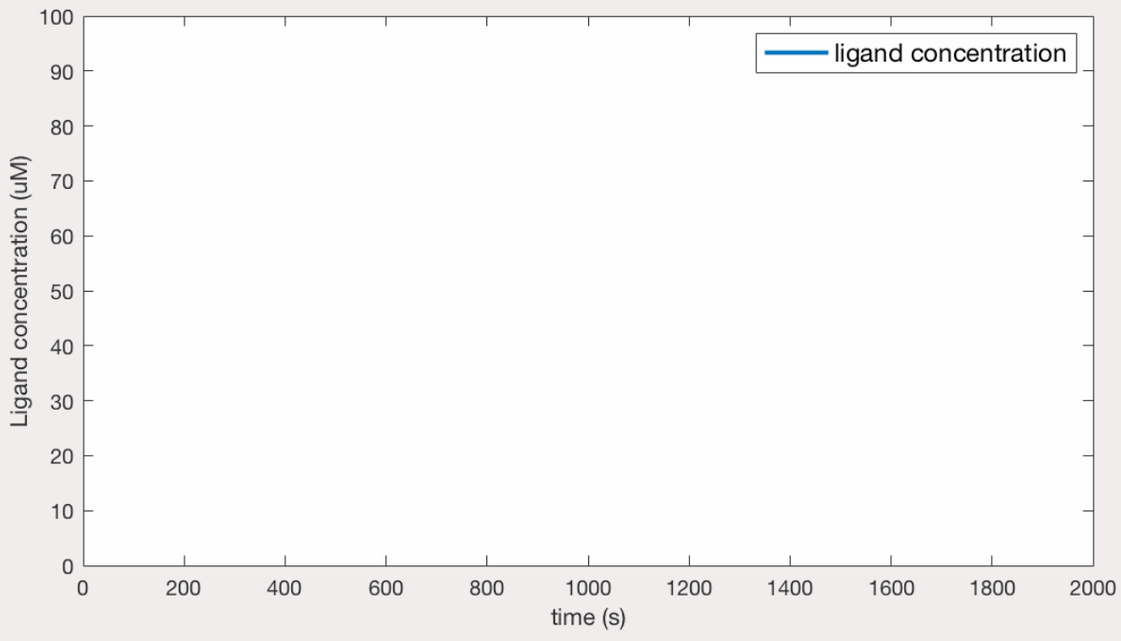
Tina Subic, Edita Bulovaite, and PN, unpublished.

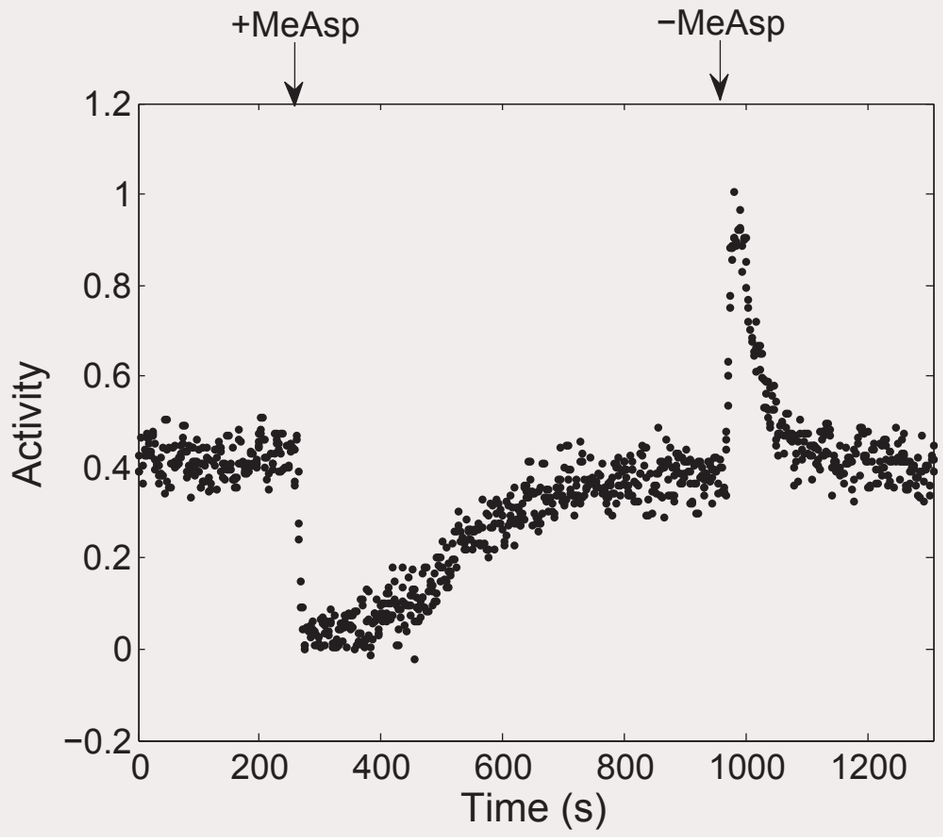
Phil Nelson

Adaptation to a jump of attractant level









Tina Subic, Edita Bulovaite, and PN, unpublished.

Phil Nelson

Wrap

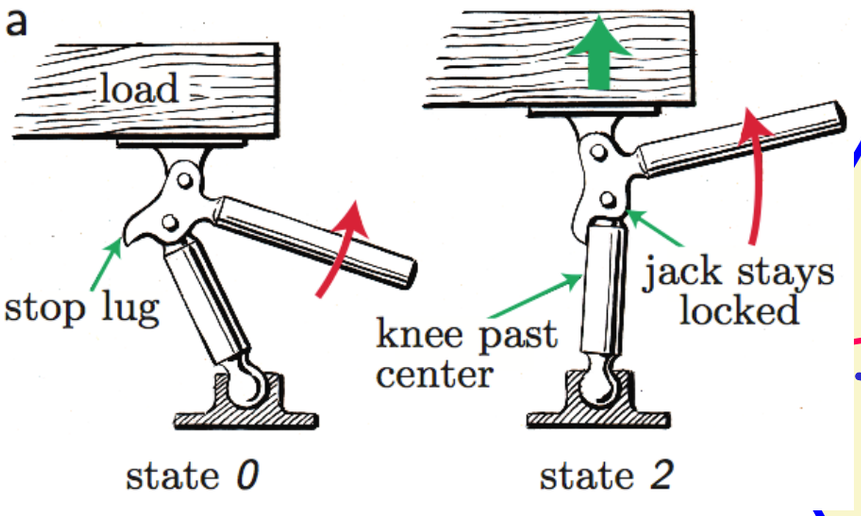
This is a whole *new kind* of modeling, more informative than solving ODEs (and easier than solving master equations).
It can go to places where analytic angels fear to tread.
And it's not (necessarily) hard.

But… um… what *is* physical modeling, anyway?

Don’t want to get all philosophical on you. I say,

It’s a Tetrahedron:

```
Nf1=figure(1);
quiver(dm,theta,dmdot.*scaling,thetadot.*scaling,1);
% nullclines
hold on;
smalltg = [-1.2:.02:1.2];
```



data

$$\frac{dc_1}{dt} = -\frac{c_1}{\tau_1} + \frac{\Gamma_1/V}{1+(c_2/K_{d,2})^{n_1}}$$
$$\frac{dc_2}{dt} = -\frac{c_2}{\tau_2} + \frac{\Gamma_2/V}{1+(c_1/K_{d,1})^{n_2}}$$

“Yadda, yadda... feedback, yadda... bistability, hysteresis, yadda,... bifurcation...

From PN, *Physical models of living systems.*

Phil Nelson

Thanks

Some of this material was taken from a recent textbook:

Physical models of living systems
(www.physics.upenn.edu/biophys/PMLS).

Other bits, including explicit code: KY Chen, DM Zuckerman, PCN (http://biorxiv.org/cgi/content/short/418772v1).

Also see:

DM Zuckerman, http://physicallensonthecell.org/cell-biology-phenomena/active-kinetic-proofreading

A student's guide to Python for physical modeling by Jesse Kinder and PN (Princeton University Press, 2018).

From photon to neuron: Light, imaging, vision (Princeton, 2017; www.physics.upenn.edu/biophys/PtN).

Daniel Thomas Gillespie
(1938 – 2017)

University of Pennsylvania

NSF

For these slides see:
www.physics.upenn.edu/~pcn