

(Tutorial on / brief overview of)
**Methods from statistical physics
to model biological networks**

Eric Fanchon (TIMC-IMAG)

Toward Systems Biology

May 2011

Outline

- Stochastic processes, description levels
- (Chemical) **master equation** (CME)
- Gillespie algorithm

- Link with Mass Action law (macroscopic laws)
- Deviant effects
- Large fluctuations
- Gene expression : multistep reactions

Orders of magnitude

- (Avogadro number : $N_A = 6,02 \cdot 10^{23}$)
- A typical *E coli* cell contains about 10^5 copies of abundant proteins, and on the order of 10 copies of scarce ones
- Number of water molecules in a small cell $\sim 10^8$ - 10^9
- **Molecular dynamics** (with explicit solvent) : position and velocity of each atom
Ex : assembly of viral capsid. 1 million atoms, time
Simulated: $50 \cdot 10^{-9}$ s

Probabilities and stochastic processes in physics

- "Experience has taught us that in spite of our ignorance of most of the microscopic variables it is still possible to detect regularities in the macroscopic behavior and formulate them in general laws" van Kampen
- Several possible description levels
- Not a property of the observer but of the system itself

Sources of stochasticity :

- Uncertainty of initial state, deterministic chaos
- Elimination of variables ('noise')

From mechanics to statistical physics

- Dynamical quantities \longrightarrow Statistical quantities,
Deterministic evolution non-deterministic evolution,
Mesoscopic or macroscopic state
- Reversibility (t/-t) \longrightarrow Irreversibility
- Dynamical instability \longrightarrow Statistical stability

Stochastic processes

- Given a probability space (Ω, E, P) , and a state space X , a **stochastic process** φ is a function $\Omega \times \mathbb{R} \rightarrow X$ such that:

For each $t \in \mathbb{R}$, $\varphi(t): \Omega \rightarrow X$ is a random variable

- To each event $\omega \in \Omega$ is associated a trajectory $\varphi_\omega(t)$
- Hierarchy of distribution functions P_n :
 $P_1(x, t)$ = probability density that φ has the value x at time t .
 $P_n(x_1, t_1; x_2, t_2; \dots; x_n, t_n)$ = joint probability density that $\varphi(t)$ has the value x_1 at time t_1 *and* x_2 at time t_2 *and* ...

Conditional probability

- $P_{1|1}(x_2, t_2 | x_1, t_1)$ is the probability density for φ to take value x_2 at t_2 , given that its value at t_1 is x_1 .
- $P_2(x_1, t_1; x_2, t_2) = P_{1|1}(x_2, t_2 | x_1, t_1) \cdot P_1(x_1, t_1)$
- ...

Markov process

Stochastic process with the property that for any set of successive times $t_1 < t_2 < \dots < t_n$:

$$P_{1|n-1}(x_n, t_n | x_1, t_1; \dots; x_{n-1}, t_{n-1}) = P_{1|1}(x_n, t_n | x_{n-1}, t_{n-1})$$

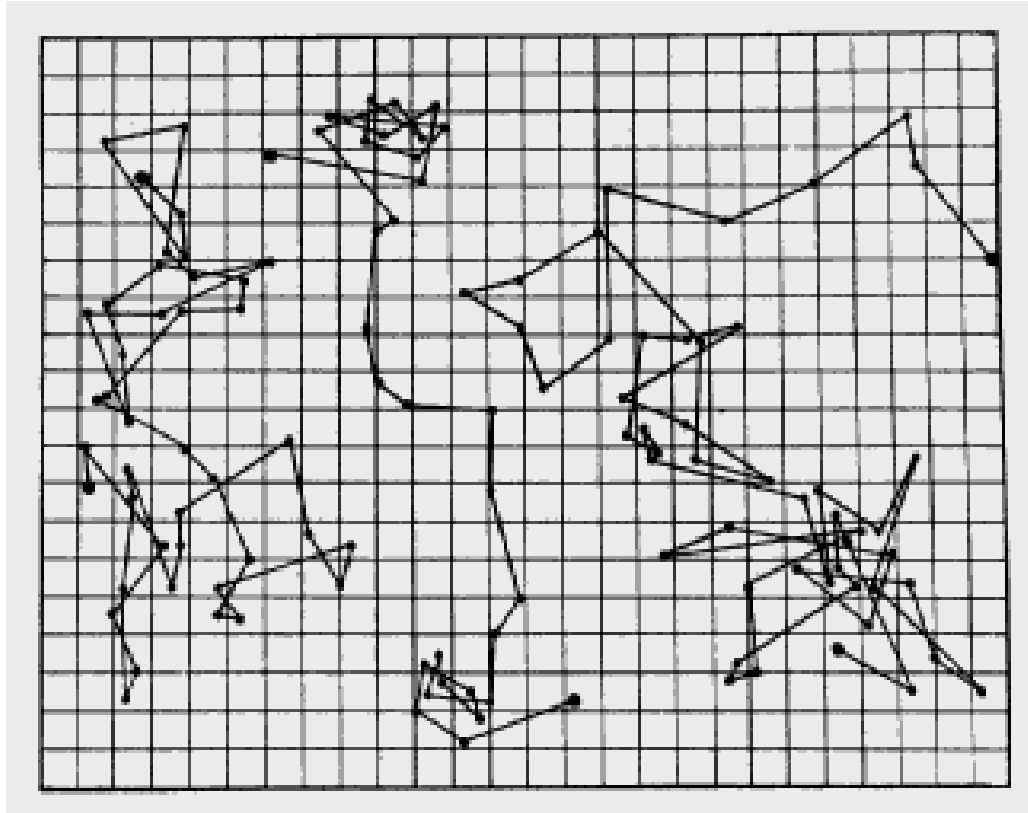
- $P_{1|1}(x_2, t_2 | x_1, t_1)$: **transition probability**

We deduce :

- $P_2(x_1, t_1; x_2, t_2) = P_1(x_1, t_1) \cdot P_{1|1}(x_2, t_2 | x_1, t_1)$
- $P_3(x_1, t_1; x_2, t_2; x_3, t_3) = P_1(x_1, t_1) \cdot P_{1|1}(x_2, t_2 | x_1, t_1) \cdot P_{1|1}(x_3, t_3 | x_2, t_2)$
- ...

A Markov process is fully determined by P_1 and $P_{1|1}$.

Brownian motion (Brown, 1827)



Jean Perrin's observations

Brownian motion

- What is observed is the net displacement resulting after many variations of the velocity
- The position and the velocity are Markov processes

- The Markov property holds approximately :

$x_k - x_{k-1}$ large \rightarrow large velocity at $t_k \rightarrow x_{k+1} - x_k$ large

Auto-correlation time of velocity $> 0 \Rightarrow$ correlation between successive displacements

Ordered flow around the particle \rightarrow memory

- The same physical system can be associated with several Markov processes.

$AB \rightarrow A + B$: concentration, vibrations of AB

Markov processes

- With respect to which set of variables is the system Markovian ?
- The microscopic motion in phase space is deterministic and therefore Markovian.
- Integrating out variables → sub-process not Markovian in general
- Problem : find a much smaller set of variables (defining a 'meso- or macro-scopic state') whose behavior in time can be described as a Markov process.

Reactions (kinetic theory)

- Thermal motion ($k_B T$)
- **Collision** between 2 molecules :
 - 'Efficient' collision → **reaction**
 - 'Inefficient' collision → no reaction (elastic collision)
- The reactional event is very short (10^{-10} s)

Master equation

- For homogeneous Markov processes : $P_{1|1}$ depends only on $\tau = t_2 - t_1$: $P_{1|1}(x_2, t_2 | x_1, t_1) = T_\tau(x_2 | x_1)$
- More manageable than Chapman-Kolmogorov eq
- For a discrete state space :

$$dp_i/dt = \sum_k [W_{ki} \cdot p_k(t) - W_{ik} \cdot p_i(t)]$$

$p_i(t)$ is the probability of state i (given an initial state),

W_{ik} is the **transition probability** $i \rightarrow k$ per unit time.

Master equation (2)

- The ME is a huge system of differential equations !
- It is a linear system
- It is generally stiff (several time scales)
- The matrix is sparse

Two options :

- Solve the ME.
 - Model reduction. Slow manifold (Roussel & Zhu)
- Stochastic simulations

Chemical master equation

- M reaction channels R_j
- Well-stirred medium
- Discrete state space : numbers X_i of molecules (state $\mathbf{X}(t)$)

Each reaction channel R_j is characterized by 2 quantities :

- State-change vector $\mathbf{v}_j = (v_{1j}, \dots, v_{Nj})$. One R_j reaction causes a change of state \mathbf{x} to $\mathbf{x} + \mathbf{v}_j$.
- **Propensity function** $a_j(\mathbf{x})$:
 $a_j(\mathbf{x})dt \equiv$ probability, given $\mathbf{X}(t) = \mathbf{x}$, that one R_j reaction will occur in the next time interval $[t, t+dt[$.

Chemical master equation

$$P(\mathbf{x}, t \mid \mathbf{x}_0, t_0) \equiv \text{Prob}\{\mathbf{X}(t) = \mathbf{x}, \text{ given } \mathbf{X}(t_0) = \mathbf{x}_0\}$$

$$\partial P(\mathbf{x}, t \mid \mathbf{x}_0, t_0) / \partial t =$$

$$\sum_{j=1}^M [a_j(\mathbf{x} - \mathbf{v}_j) \cdot P(\mathbf{x} - \mathbf{v}_j, t \mid \mathbf{x}_0, t_0) - a_j(\mathbf{x}) \cdot P(\mathbf{x}, t \mid \mathbf{x}_0, t_0)]$$

Equation for average state :

$$d\langle \mathbf{X}(t) \rangle / dt = \sum v_j \langle a_j(\mathbf{x}) \rangle_t$$

Simple example

Degradation :

- $S \rightarrow \emptyset$ (c)

$$a(x) = cx$$

$$\partial P(x,t | x_0,0)/\partial t = a(x+1) P(x+1,t | x_0,0) - a(x) P(x,t | x_0,0)$$

Can be solved exactly (because $P(x_0+1,t | x_0,0) = 0$)

Gillespie algorithm

- *Simulate* the stochastic time evolution
- Most collisions are elastic \rightarrow at each t the medium is well-stirred
- $p(\tau, j | \mathbf{x}, t) d\tau \equiv$ probability, given $\mathbf{X}(t) = \mathbf{x}$, that the next reaction in the system will occur in the infinitesimal interval $[t+\tau, t+\tau+d\tau[$ and will be an R_j reaction.
- $p(\tau, j | \mathbf{x}, t) = a_j(\mathbf{x}) \exp(-a_0(\mathbf{x})\tau)$

$$\text{with } a_0(\mathbf{x}) \equiv \sum_k^M a_k(\mathbf{x})$$

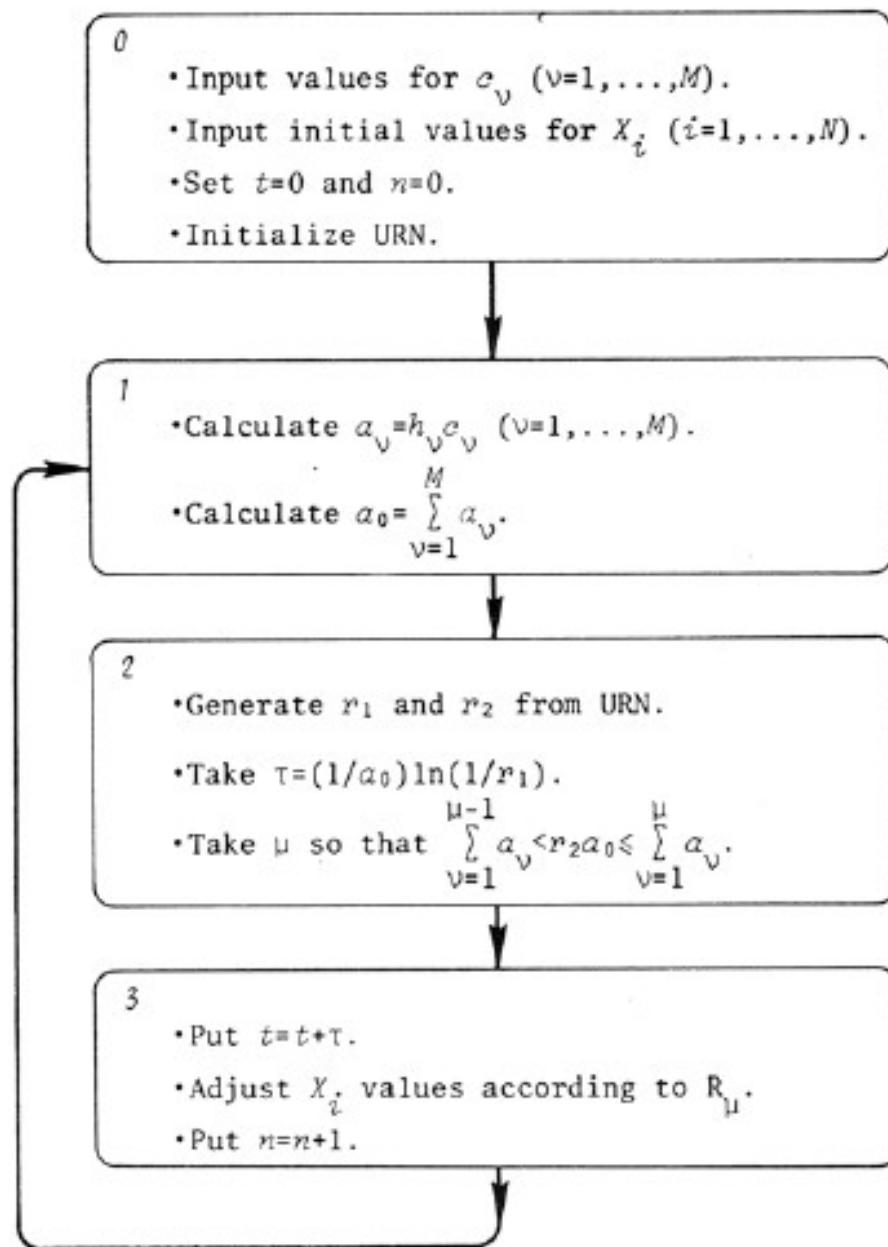
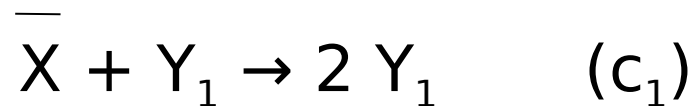


Figure 2. Schematic of the stochastic simulation algorithm.

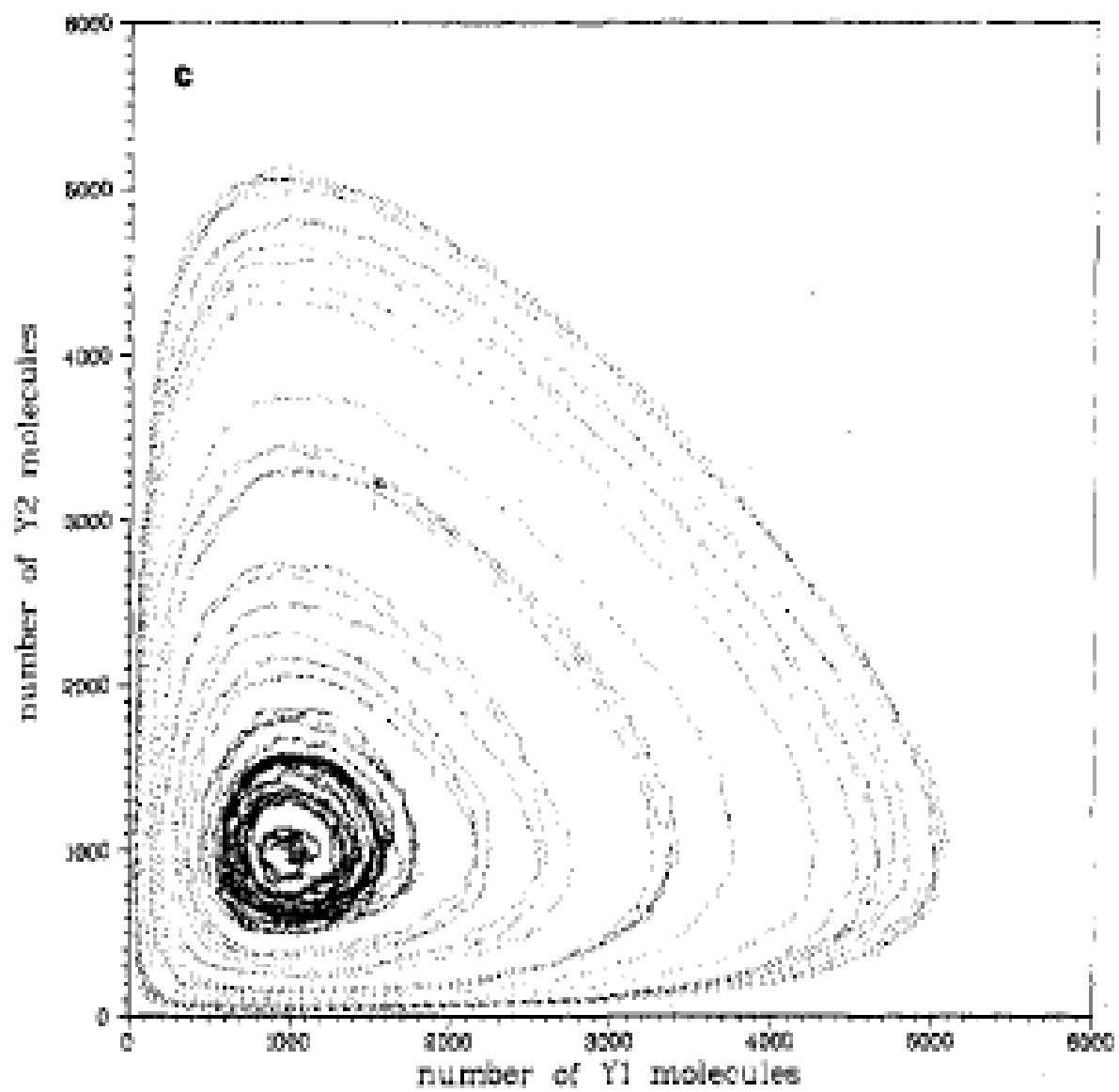
Example : Lotka-Volterra

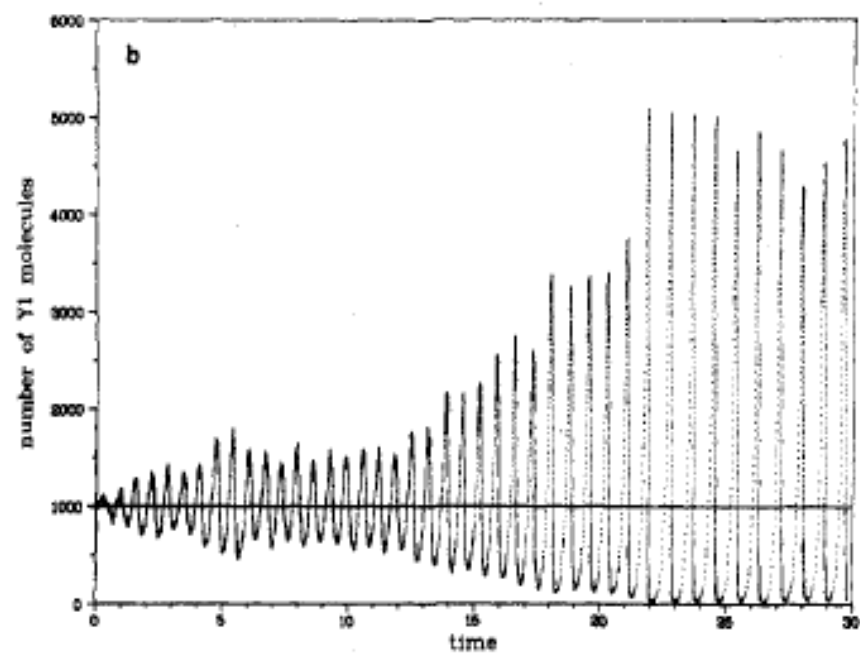
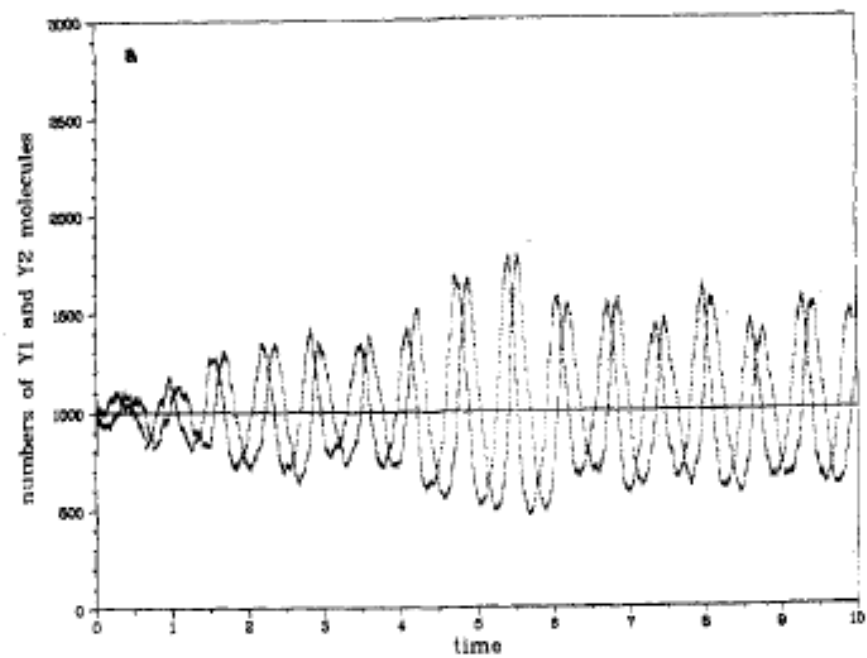
Lotka, 1920 : autocatalytic reaction system



$$dY_1/dt = c_1 X Y_1 - c_2 Y_1 Y_2$$

$$dY_2/dt = c_2 Y_1 Y_2 - c_3 Y_2$$





Approximation schemes

Tau-leaping : fixed time τ , several reactive events

Stiff systems :

- Slow-scale SSA
- Hybrid simulations
- ...

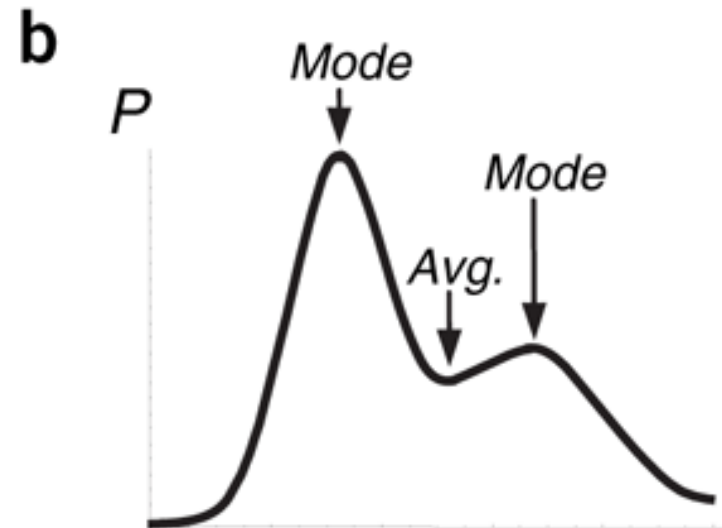
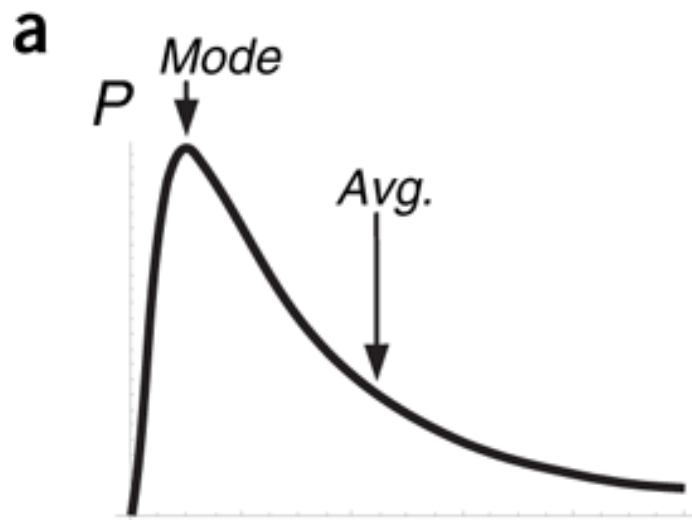
From stochastic to deterministic (macroscopic/dissipative)

- Kubo et al, J Stat Phys, 1973 ; Kitahara, 1973
- Eikonal approx of CME (expansion in $1/\Omega$).

$$P(\mathbf{x},t) = C \exp(-\Omega (s(\mathbf{x},t) + O(1/\Omega)))$$

- Replace into Chemical ME; continuous \mathbf{x}
→ Hamilton-Jacobi equation for the action $s(\mathbf{x},t)$
- The trajectory of the **mode** ($\nabla s(\mathbf{x},t)=0$) is solution of an ODE corresponding to mass-action law
- Linear systems : mode \sim average

Average of distribution and mode(s)



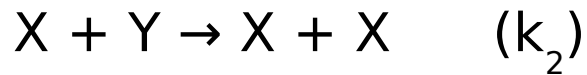
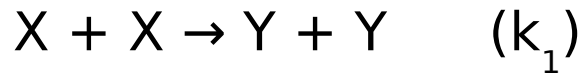
Classification of deviant effects

(Samoilov & Arkin, nature biotech, 2006)

Approximations :

- Type I : Full state probability distribution is approximated by its leading exponent (eikonal approx)
- Type II : Characteristic system features occur on scales much greater than a molecule: Discrete molecular states are approximated by continuous concentrations
- Type III : The evolution of the state distribution is replaced by that of its mode

Type I



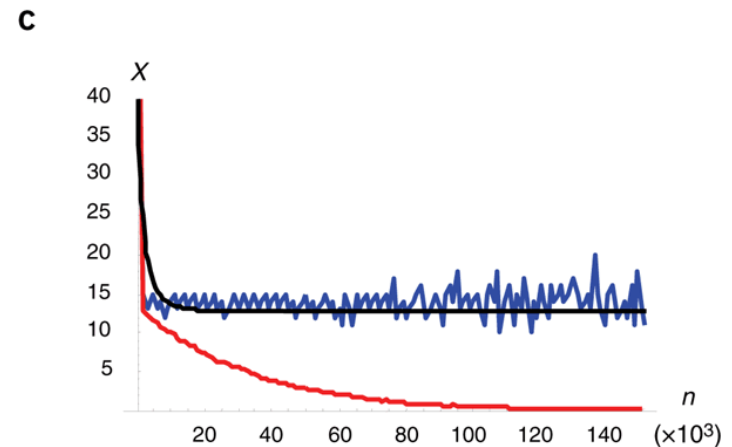
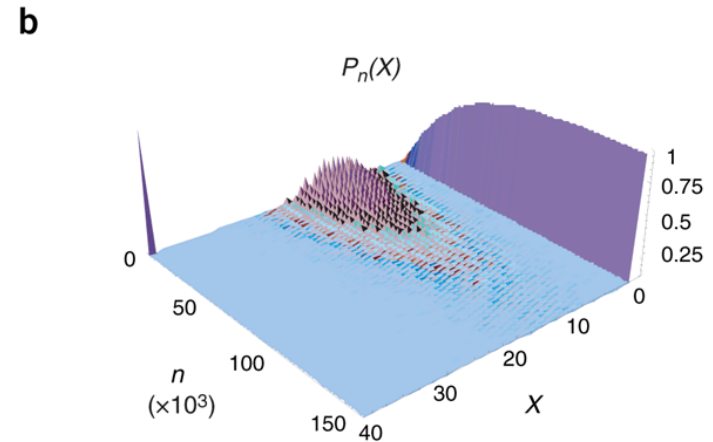
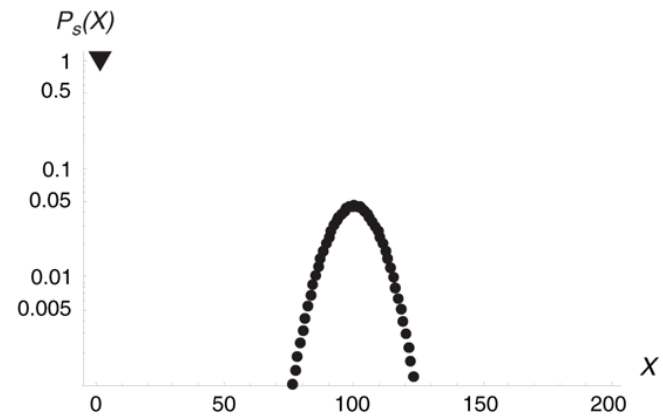
Mass-action law :

- Two stationary states :

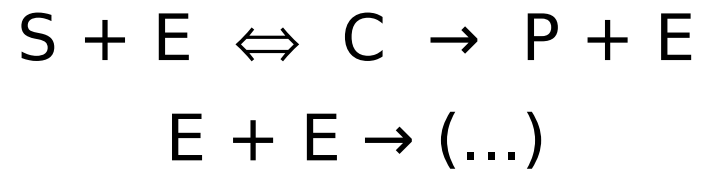
$$x_{ss1} = 0 \quad (\mathbf{unstable})$$

$$x_{ss2} = x_T k_2 / (k_2 + 2k_1) \quad (\mathbf{stable})$$

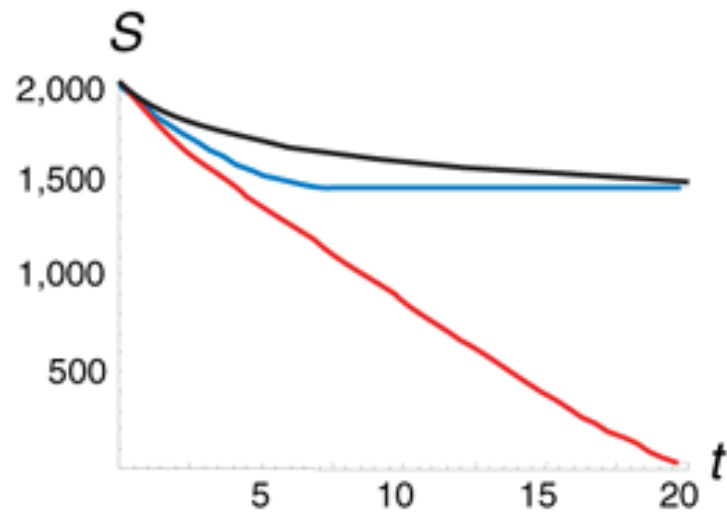
- CME analysis \longrightarrow



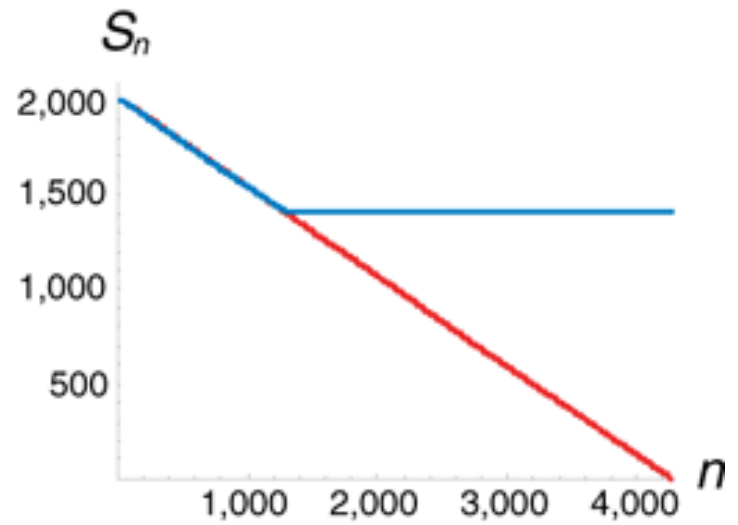
Type II



a

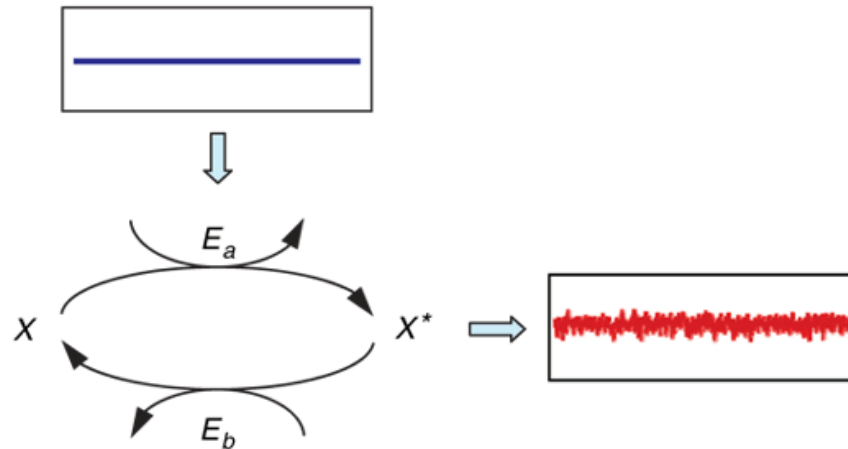


b

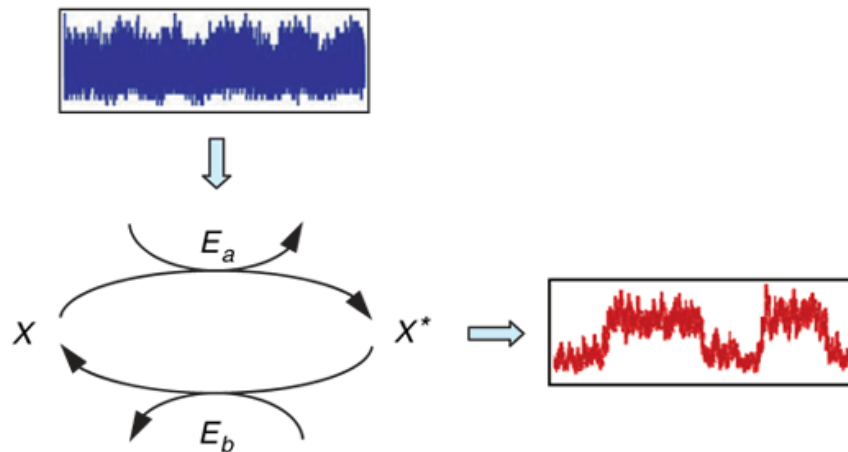


Type III

a



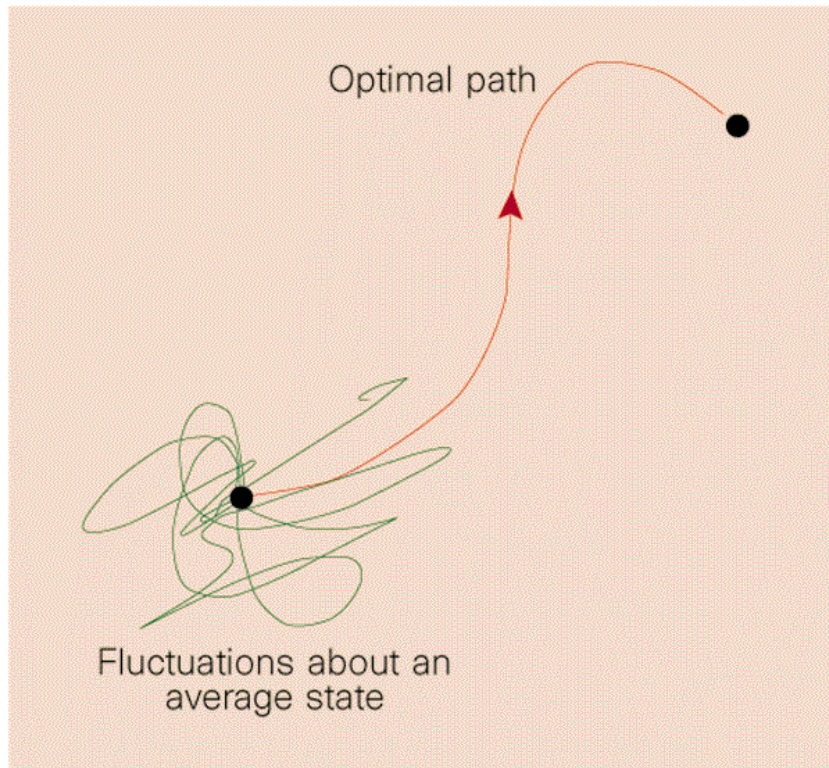
b



Large fluctuations out of equilibrium

- Large and rare, but effect may be important ; trigger change of attractor
- Luchinsky & McClintock, nature, 1997
- Large fluctuations
 - Tail of distribution (exponentially decreasing)
- Probability of transition to other attractor is not zero

Large fluctuations out of equilibrium



Two kinds of random motions :

- Fluctuational (away from SS)
- Relaxational (back to SS)

Optimal path to point far from average state

Heterogeneity of cell populations (genetically identical cells)

Other approaches

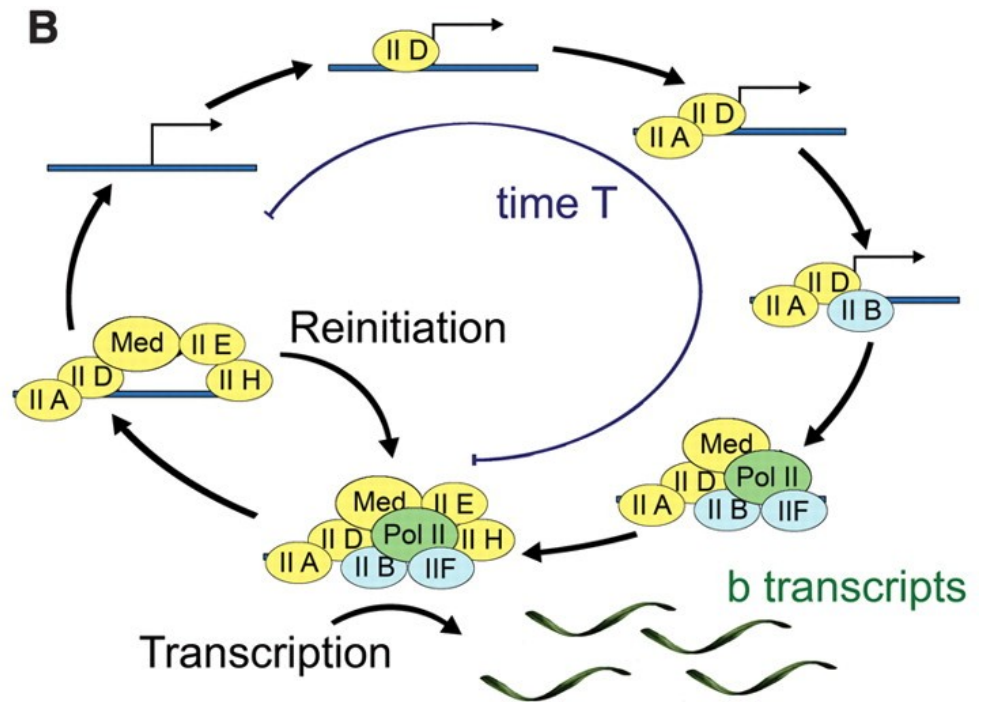
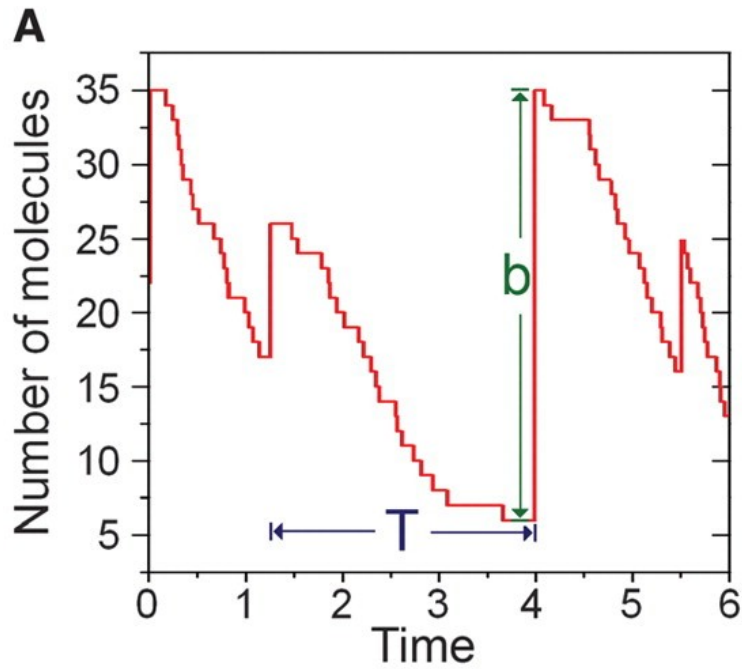
- (chemical and reaction-diffusion) Master eqs → second quantization, operator algebra techniques
- Renormalization group
Fluctuations on several scales (multiscale analysis)

Multistep reactions

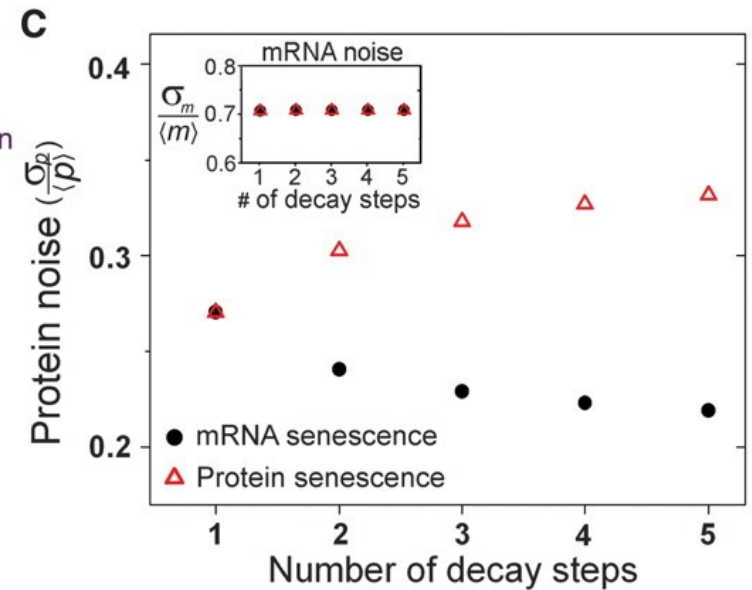
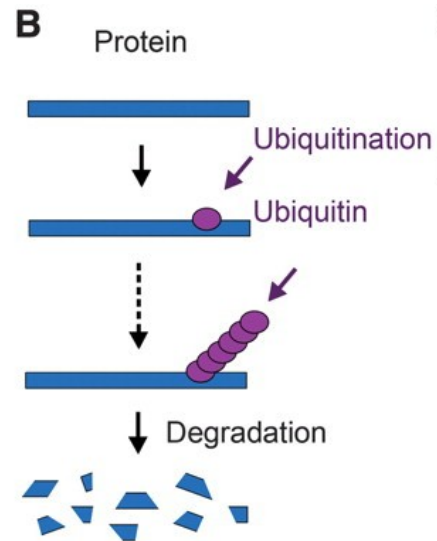
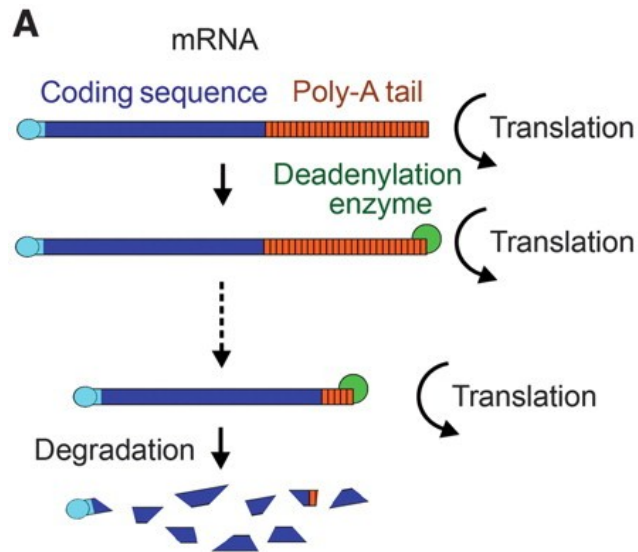
(Pedraza & Paulsson, Science, 2008)

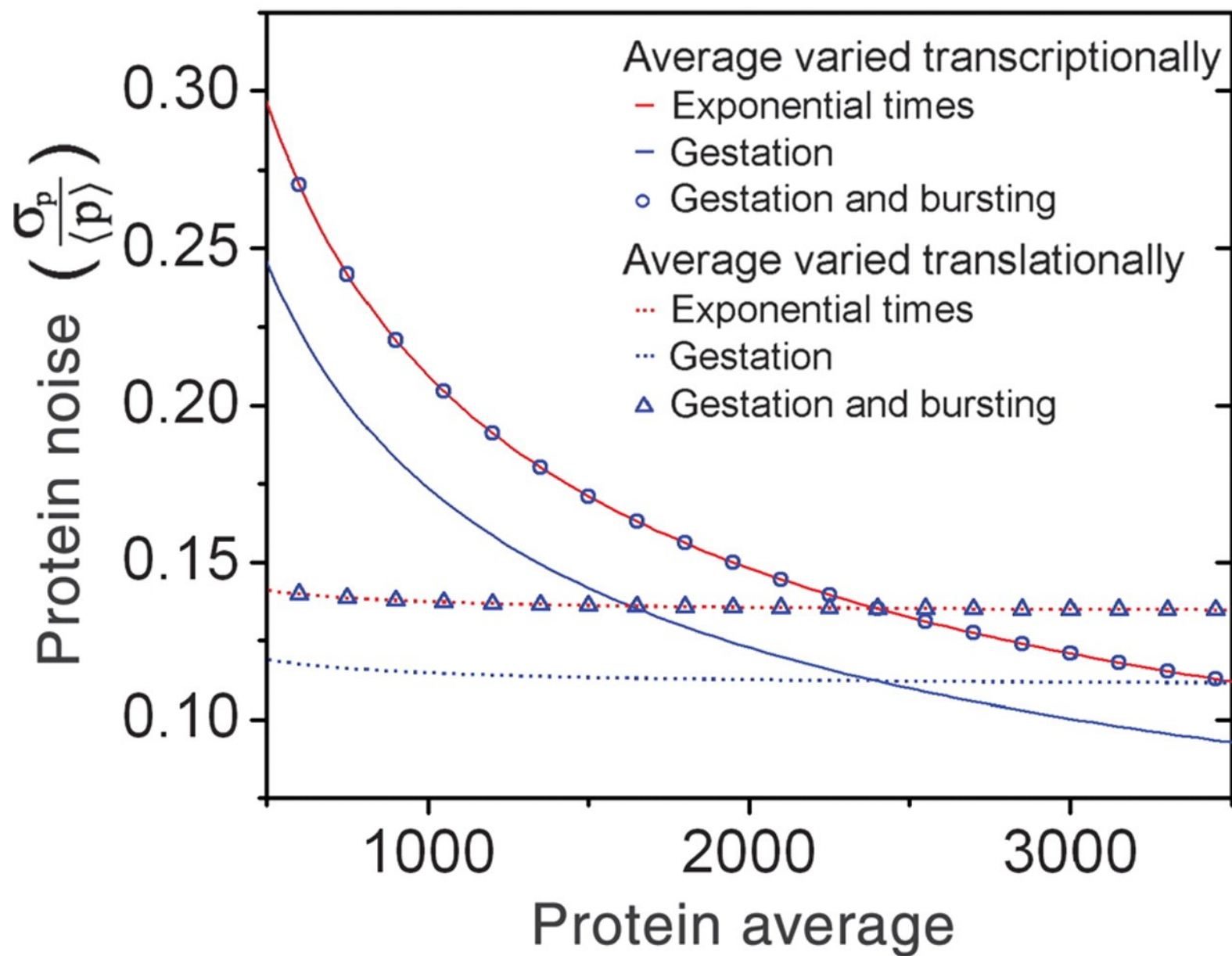
- Complex control of gene expression : several repressors, transcr factors and mediators, chromatin remodelling → non-exponential time intervals between transcription windows
- RNA gestation and senescence

Bursting



Molecular senescence





Summary

- Stochasticity results from a *choice* of description level. It's a property of the system, not of the observer.
- The Markov property is always an approximation for a physical system.
- The macroscopic MA law describe the evolution of the *mode(s)* rather than the *average*.
- Deviant effects can occur at large numbers of molecules
- Proba of large fluctuations; auxiliary hamiltonian system
- Multistep reactions (transcription, ...) → reduction of noise