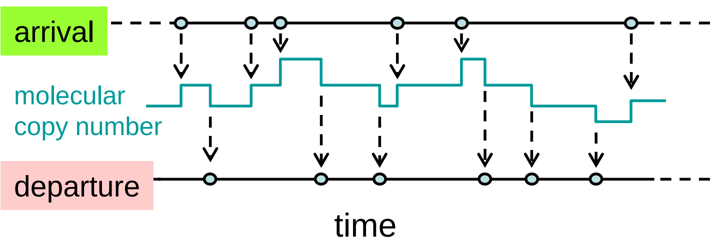
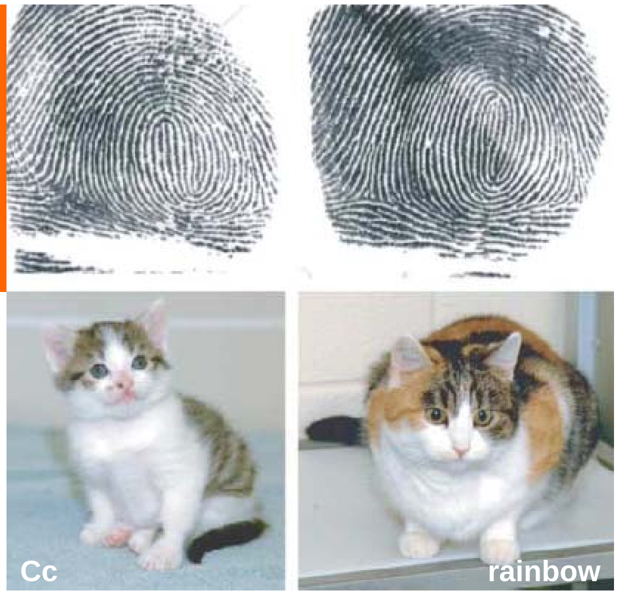
Noise propagation in *biochemical networks*





J Raser and E O'Shea, Science **309**: 2010 (2005)

Li-ping Xiong, Yu-qiang Ma, Lei-Han Tang
Department of Physics, Hong Kong Baptist University
Department of Physics, Nanjing University

The 4th KIAS Conference on Statistical Physics
***Nonequilibrium Statistical Physics of Complex Systems*, 26-29 July, 2010**

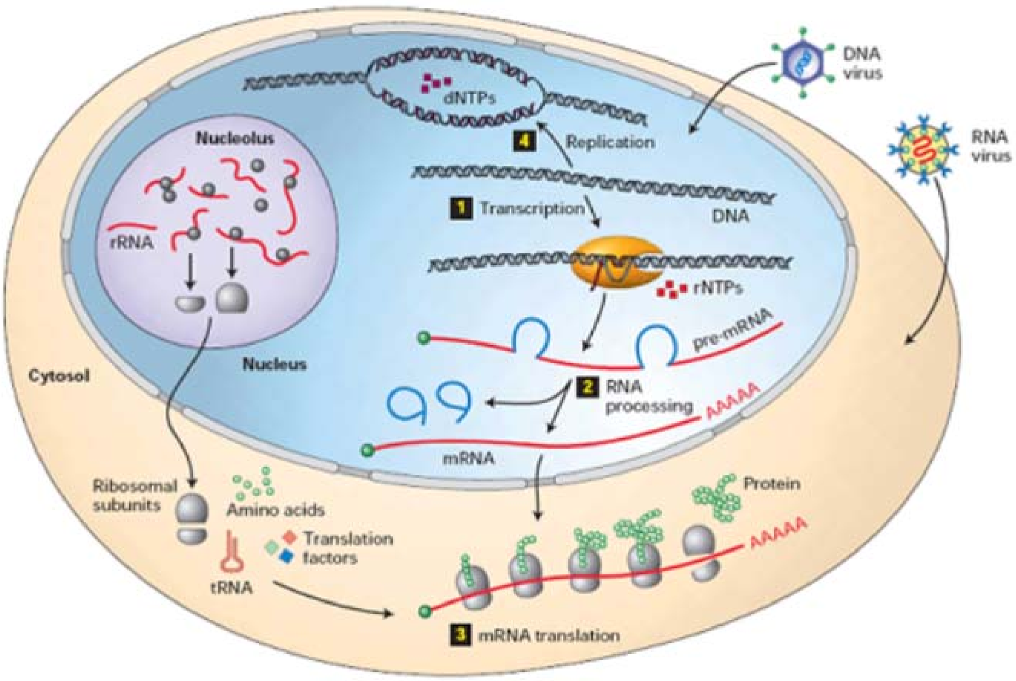
Outline

- ❖ Experimental characterization of noise in gene expression
- ❖ A two-compartment model of transport with bursty input: the chemical master equation approach
- ❖ General formulation of noise propagation in bio-networks based on time-series analysis
- ❖ Some open issues: jamming and nonlinear transport

What do cells do?

Basically, cell is a stand-alone machine that manages its own assembly, maintenance, environmental protection, and reproduction.

All done in a hurry based on noisy molecular circuits

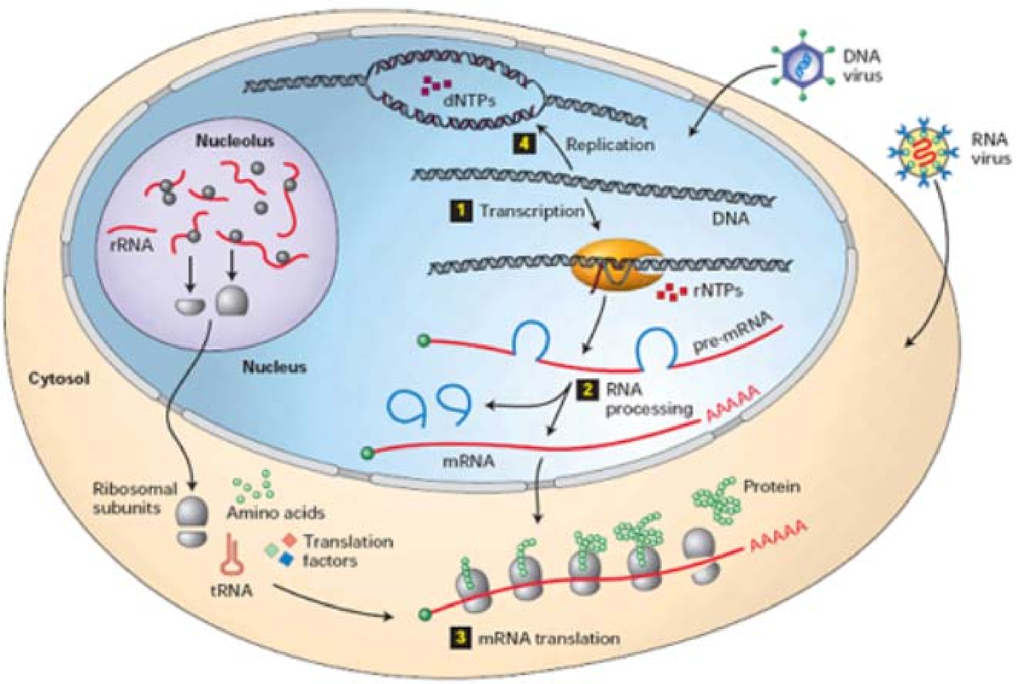


H. Lodish et al., *Molecular Cell Biology*, 5th Ed.

Noise Biology

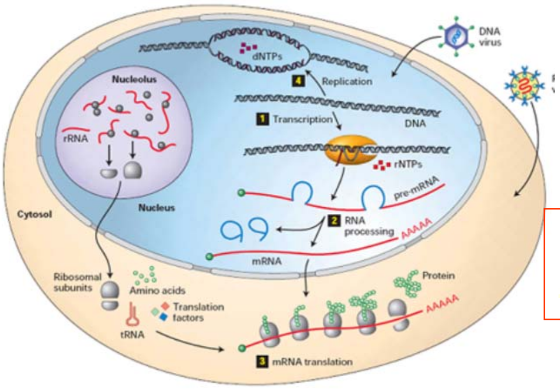
A brand new discipline that can benefit and benefit from nonequilibrium statistical physics

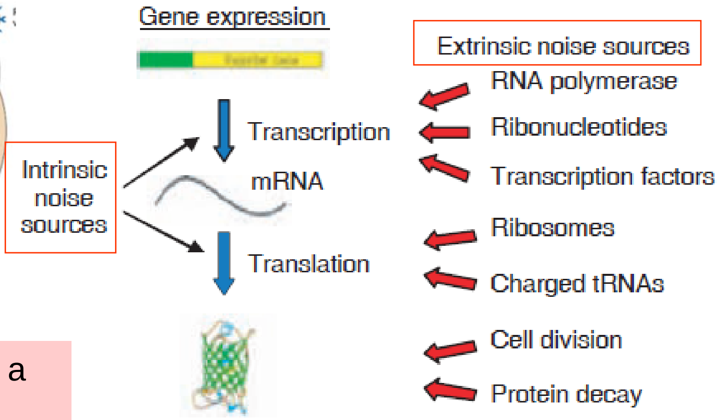
ML Simpson et al (2009) Noise in Biological Circuits, *Nanomed Nanobiotechnol* **1**: 214-225.



H. Lodish et al., *Molecular Cell Biology*, 5th Ed.

Noise Biology: it is all about numbers and timing





- Cell has a finite volume $\Rightarrow$ Space is a commodity
- Keep protein copy numbers low to:
  - Save space
  - Save the cost to synthesize proteins
- Noise/fluctuation an inconvenient truth that impacts every aspect of life

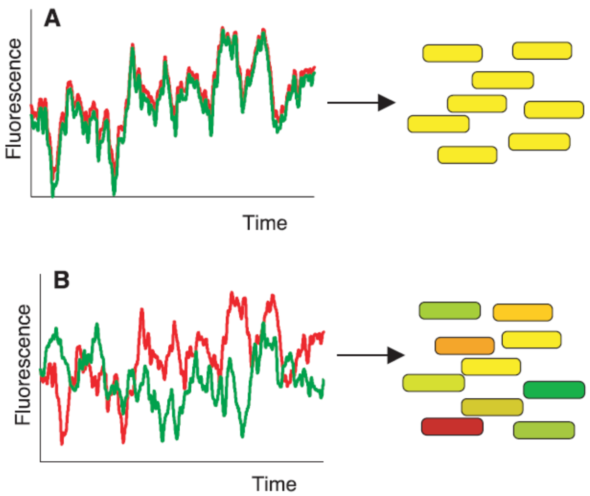
stochastic dynamics | molecular apparatus

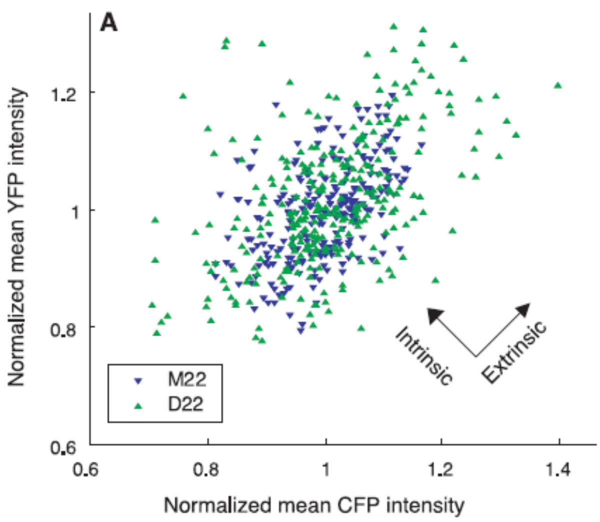
Measurement of noise in gene expression using two-color GFP

M B Elowitz, A J Levine, E D Siggia, P S Swain (2002) Stochastic gene expression in a single cell, *Science* **297**, 1183.

Two copies of the same gene under identical promoters and genomic environment in E. coli

Temporal variation of protein copy number





Bursty protein synthesis in bacteria

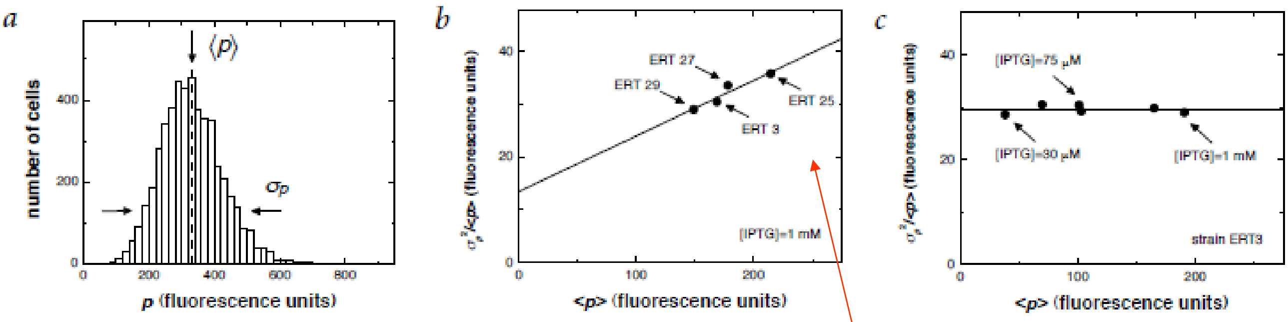
E M Ozbudak, M Thattai, I Kurtser, A D Grossman, A van Oudenaarden (2002) Regulation of noise in the expression of a single gene, *Nature Genetics* **31**, 69.

B. subtilis, chromosomal gene expression with gfp tag

Protein copy number distribution in isogenic bacterial cells

Strains with varying translational efficiency (30% difference)

One strain at different inducer concentration that controls transcriptional efficiency



Variance: $\sigma_p^2 = \langle p^2 \rangle - \langle p \rangle^2$

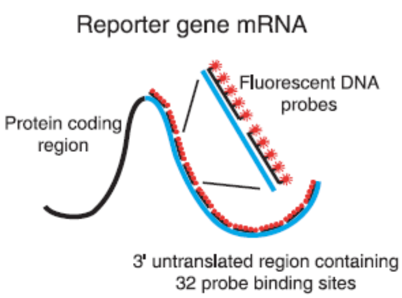
<u>Fano factor</u>: $F = \sigma_p^2 / \langle p \rangle = \begin{cases} 1, & \text{Poisson} \\ b > 1, & \text{bursty} \end{cases}$

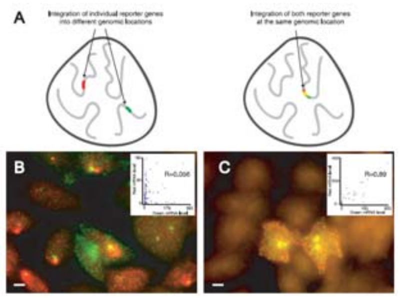
<u>Translational bursting</u>: a few tens of proteins are made from each mRNA.

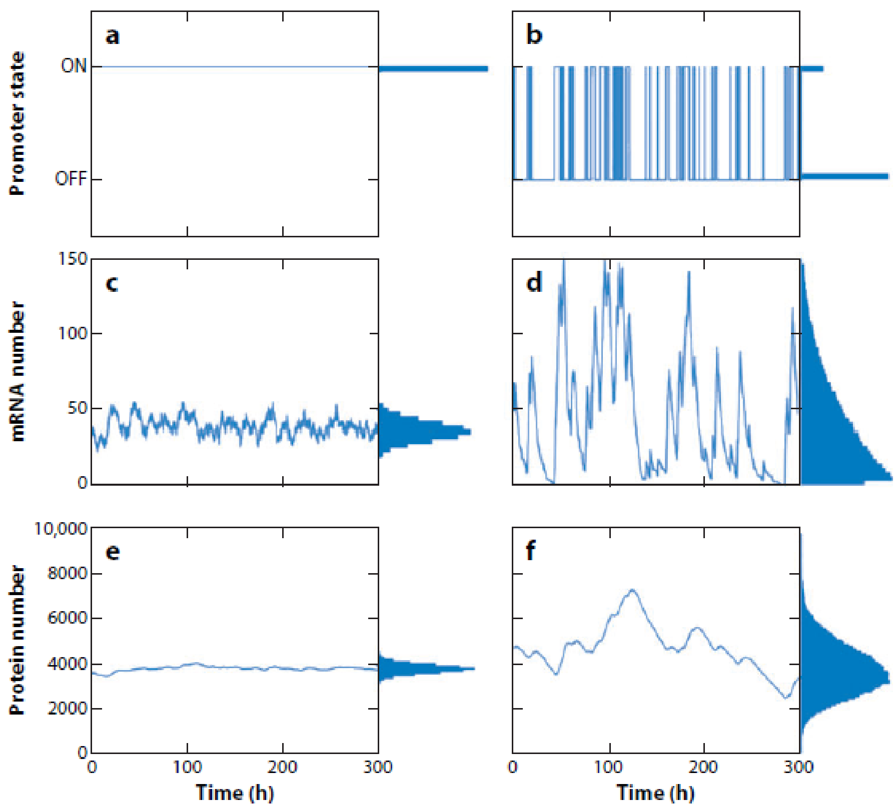
Transcriptional Bursting in Mammalian Cell Gene Expression

A Raj et al (2006)

PLoS Biology **4**: 1707







A Raj and A von Oudenaarden (2009) Single-Molecule Approaches to Stochastic Gene Expression, Annu. Rev. Biophys. **38**: 255-70.

A two-compartment model for mRNA life cycle

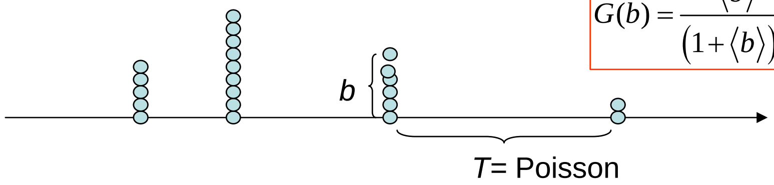
Li-ping Xiong, Yu-qiang Ma, Lei-Han Tang (2010) Attenuation of transcriptional bursting in mRNA transport, Phys Biol **7**: 016005

mRNA synthesis: $\boxed{I \xrightarrow{\lambda} A, \quad A \xrightarrow{\gamma} I, \quad A \xrightarrow{\mu} A + M_\mathrm{n}.}$

Bursty limit: $\gamma, \mu \gg \lambda, \ \langle b \rangle = \mu / \gamma = \text{constant}$

$$\boxed{G(b) = \frac{\langle b \rangle^b}{\left(1+\langle b \rangle\right)^{b+1}}}$$

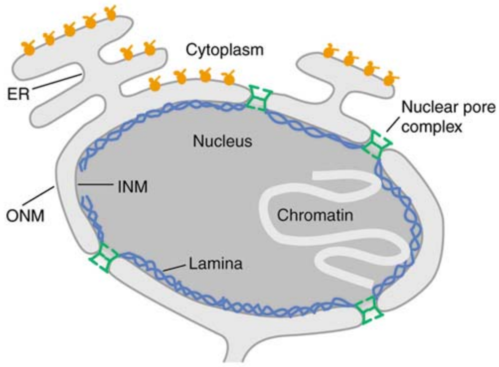
Time series of transcriptional events



mRNA transport:

Linear transport (unlimited channels) $\boxed{M_\mathrm{n} \xrightarrow{k} M_\mathrm{c}}$

Michaelis-Menten transport (queuing) $\boxed{M_\mathrm{n} + E \underset{k_2}{\overset{k_1}{\rightleftharpoons}} EM_\mathrm{n} \xrightarrow{k_3} E + M_\mathrm{c}}$



mRNA degradation: $\boxed{M_\mathrm{c} \xrightarrow{\delta} \emptyset}$

Chemical master equations for linear transport

m_n: nuclear mRNA copy number
m_c: cytoplasmic mRNA copy number

Raising/lowering operators: $\varepsilon_\alpha f(m_\alpha) = f(m_\alpha + 1)$

$\alpha = \mathrm{n,c} \qquad \varepsilon_\alpha^{-1} f(m_\alpha) = f(m_\alpha - 1)$

Prob with promoter active:

Prob with promoter inactive:

$$\begin{aligned}
\frac{dP_A(m_\mathrm{n}, m_\mathrm{c}, t)}{dt} &= \lambda P_I(m_\mathrm{n}, m_\mathrm{c}, t) - \gamma P_A(m_\mathrm{n}, m_\mathrm{c}, t) \\
&\quad + \mu(\varepsilon_\mathrm{n}^{-1} - 1) P_A(m_\mathrm{n}, m_\mathrm{c}, t) \\
&\quad + k(\varepsilon_\mathrm{n}\varepsilon_\mathrm{c}^{-1} - 1) m_\mathrm{n} P_A(m_\mathrm{n}, m_\mathrm{c}, t) \\
&\quad + \delta(\varepsilon_\mathrm{c} - 1) m_\mathrm{c} P_A(m_\mathrm{n}, m_\mathrm{c}, t), \\
\frac{dP_I(m_\mathrm{n}, m_\mathrm{c}, t)}{dt} &= \gamma P_A(m_\mathrm{n}, m_\mathrm{c}, t) - \lambda P_I(m_\mathrm{n}, m_\mathrm{c}, t) \\
&\quad + k(\varepsilon_\mathrm{n}\varepsilon_\mathrm{c}^{-1} - 1) m_\mathrm{n} P_I(m_\mathrm{n}, m_\mathrm{c}, t) \\
&\quad + \delta(\varepsilon_\mathrm{c} - 1) m_\mathrm{c} P_I(m_\mathrm{n}, m_\mathrm{c}, t),
\end{aligned}$$

Moment equations closed!

Fano factors

$$\begin{aligned}
\frac{\sigma_{m_\mathrm{n}}^2}{\langle m_\mathrm{n} \rangle} &= \langle b \rangle + 1, \\
\frac{\sigma_{m_\mathrm{c}}^2}{\langle m_\mathrm{c} \rangle} &= \langle b \rangle + 1 - \langle b \rangle \frac{\langle m_\mathrm{n} \rangle}{\langle m_\mathrm{n} \rangle + \langle m_\mathrm{c} \rangle}.
\end{aligned}$$

Noise reduced by mRNA transport

Michaelis-Menten transport

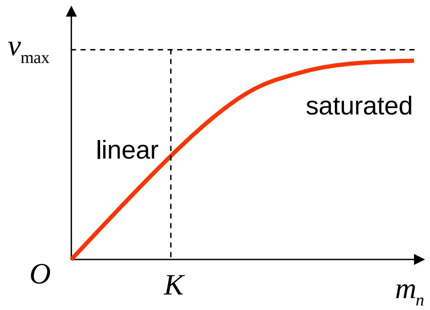
Transport reaction: $\boxed{M_\mathrm{n} + E \underset{k_2}{\overset{k_1}{\rightleftharpoons}} EM_\mathrm{n} \xrightarrow{k_3} E + M_\mathrm{c}}$

Queuing due to finite number of transport channels/enzymes

Fast equilibration: EM_n population equilibrates much faster than the time for M_n to undergo significant change

$\Rightarrow$ Effective transport rate

MM equation $\boxed{v(m_n) = k_3 \langle EM_n \rangle\big|_{m_n} \simeq v_{\max} \dfrac{m_n}{K + m_n}}$



Master equation after eliminating $\mathrm{EM_n}$:

$$
\begin{aligned}
\frac{dP(m_\mathrm{n}, m_\mathrm{c}, t)}{dt} &= \sum_{b=0}^{m_\mathrm{n}} \lambda G(b) P(m_\mathrm{n} - b, m_\mathrm{c}, t) - \sum_{b=0}^{\infty} \lambda G(b) P(m_\mathrm{n}, m_\mathrm{c}, t) \\
&\quad + (\varepsilon_\mathrm{n} \varepsilon_\mathrm{c}^{-1} - 1) \boxed{v(m_\mathrm{n})} P(m_\mathrm{n}, m_\mathrm{c}, t) \\
&\quad + \delta(\varepsilon_c - 1) m_\mathrm{c} P(m_\mathrm{n}, m_\mathrm{c}, t).
\end{aligned}
$$

Moment equations not closed due to a nonlinear $v(m_n)$

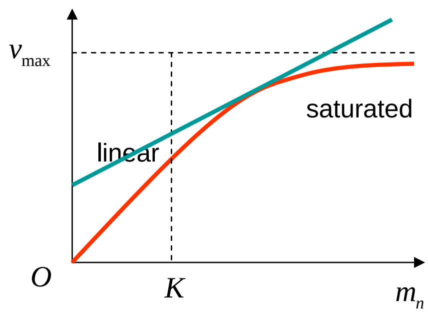
Michaelis-Menten transport (cont'd)

i) Linear approximation $\quad v \simeq k_\mathrm{eff}(m_\mathrm{n} + m_0)$

$\langle b \rangle \ll \langle m_n \rangle$ (weak noise)

Fano factors:

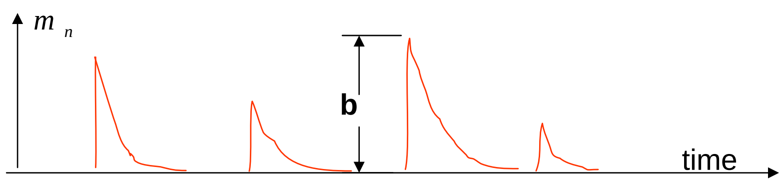
$$
\boxed{
\begin{aligned}
\frac{\sigma^2_{m_\mathrm{n}}}{\langle m_\mathrm{n} \rangle} &= \left(\frac{\langle m_\mathrm{n} \rangle}{K} + 1\right)(\langle b \rangle + 1), \\
\frac{\sigma^2_{m_\mathrm{c}}}{\langle m_\mathrm{c} \rangle} &= \langle b \rangle + 1 - \langle b \rangle \frac{\langle m_\mathrm{n} \rangle}{\frac{K \langle m_\mathrm{c} \rangle}{K + \langle m_\mathrm{n} \rangle} + \langle m_\mathrm{n} \rangle}
\end{aligned}
}
$$



ii) Independent pulse approximation

$\langle b \rangle \gg \langle m_n \rangle$ (strong noise)

$$
\begin{aligned}
m_\mathrm{n}(t) &= \sum_{t_i < t} \xi(b_i, t - t_i), \\
m_\mathrm{c}(t) &= \sum_{t_i < t} \eta(b_i, t - t_i).
\end{aligned}
$$



Fano factors:

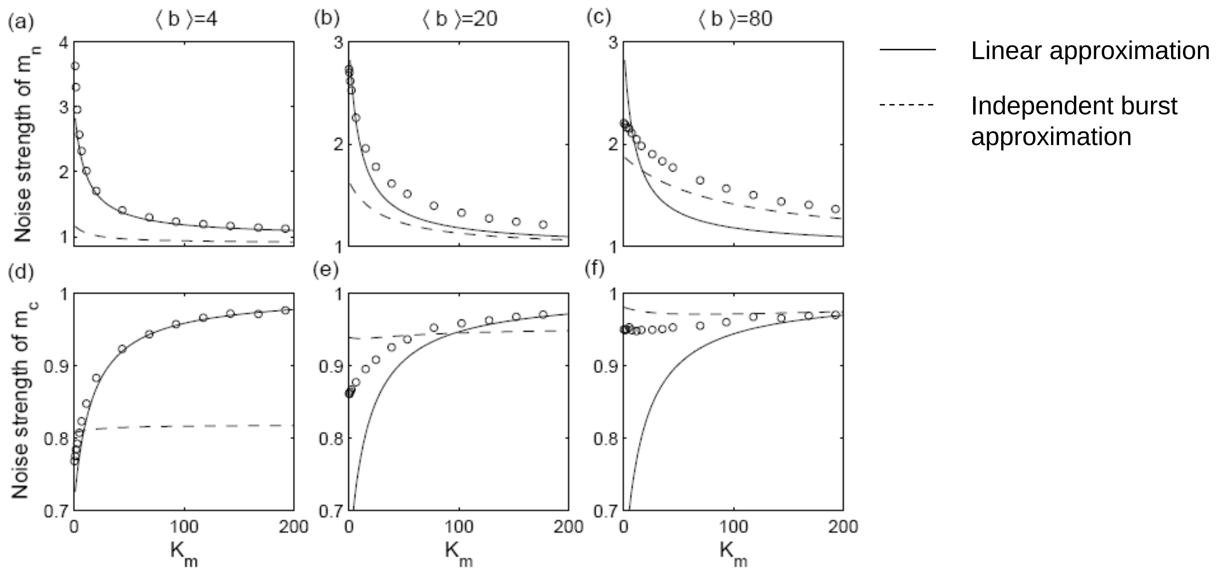
$$
\boxed{\frac{\sigma^2_{m_\mathrm{n}}}{\langle m_\mathrm{n} \rangle} = \langle b \rangle + \frac{1}{2} + \frac{\langle b \rangle^2 + \langle b \rangle + \frac{1}{12}}{K + \frac{1}{2} + \langle b \rangle}}
$$

$$
\boxed{\frac{\sigma^2_{m_\mathrm{c}}}{\langle m_\mathrm{c} \rangle} \simeq \frac{\langle b \rangle + \frac{1}{2}}{1 + \frac{\langle m_\mathrm{n} \rangle}{\langle m_\mathrm{c} \rangle} \frac{K + \langle b \rangle}{K + \langle b \rangle + \frac{1}{2}}}}
$$

Comparison with simulation results
using the Gillespie algorithm

$$e_t = 10, \langle m_n \rangle = 20, \langle m_c \rangle = 40$$

Noise strength relative to that of the linear model



weak noise

strong noise

Summary on the chemical master equation approach

- Exact representation of the Markovian dynamics
- Computation rather tedious even for a small number of molecular species and reactions
- Approximations needed when queuing is introduced. Not clear how to develop suitable expansions

Not suitable/inconvenient for

- Processes that are not markovian (e.g., those with nonexponential waiting times)
- Large networks
- Coarse-grained models

⇒ ALTERNATIVE FORMULATION by tracing the fate of individual molecules

i) Independent walkers (linear theory):

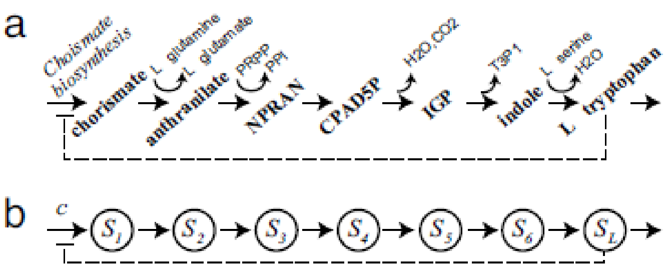
probability distribution based on waiting times

ii) Interacting walkers (nonlinear theory)

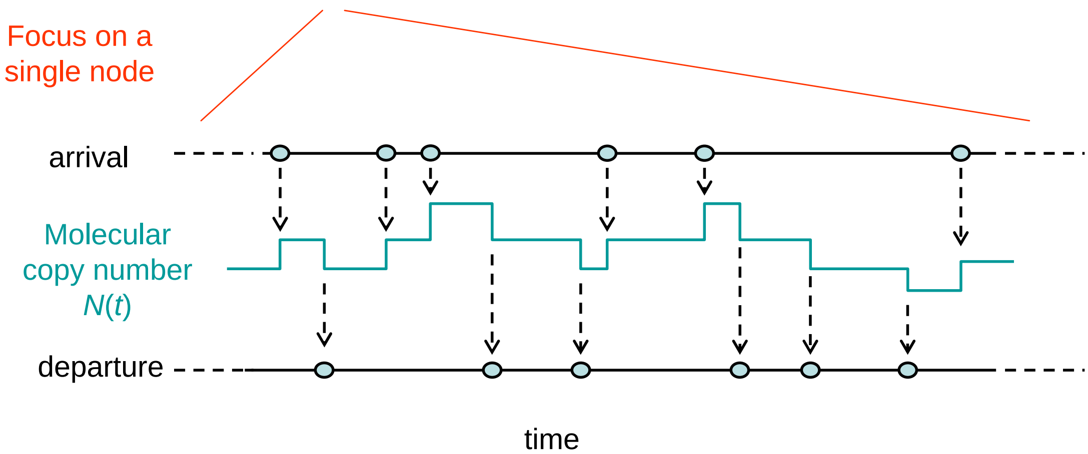
queuing, slow and fast variables, adiabatic approximation

Transport along a linear pathway

E Levine and T Hwa (2007) Stochastic fluctuations in metabolic pathways, PNAS **104**:9224-29

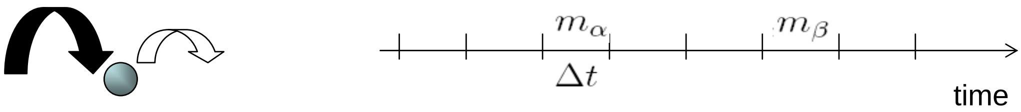


Molecular process



$$\boxed{N(t) = N_a(t) - N_d(t)}$$

Statistics of arrival time series: autocorrelation function



Discretize time axis: events specified by $m_\alpha = 0,1$

Auto-correlation function: $\boxed{\tilde{C}(t_\alpha, t_\beta) \equiv \frac{\langle m_\alpha m_\beta \rangle}{(\Delta t)^2} = \frac{\langle m_\alpha \rangle}{\Delta t}\delta(t_\alpha - t_\beta) + C(t_\alpha, t_\beta)}$

Stationary process $\boxed{C(t_\alpha, t_\beta) = \Gamma(t_\alpha - t_\beta) + J^2}$ *J*: average particle current

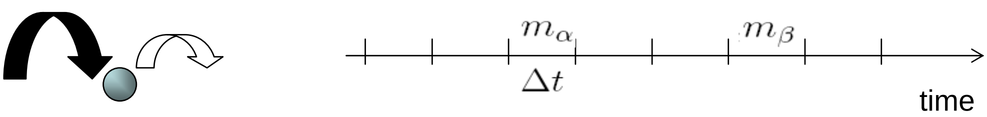
Non-Poisson component Poisson component

$\Gamma(t)$ describes clustering/anti-clustering of arrival events!

Bursty arrival: $C(t_\alpha, t_\beta) = C_B(t_\alpha - t_\beta) = \underline{J\frac{\langle b(b-1)\rangle}{\langle b \rangle}\delta(t_\alpha - t_\beta) + J^2}$

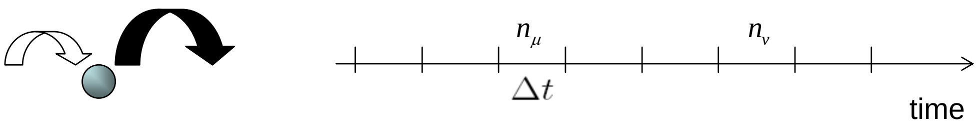
Extreme form of clustering!

Statistics of departure time series



arrival series

τ Delay/residence time with distribution $\rho(\tau)$



departure series

$$n_\mu = \sum_\alpha m_\alpha \Delta t \delta(t_\mu - t_\alpha - \tau_\alpha)$$

Non-interacting walkers: waiting times statistically independent

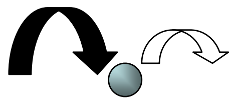
Laplace transforms of the non-Poisson component of autocorrelation functions

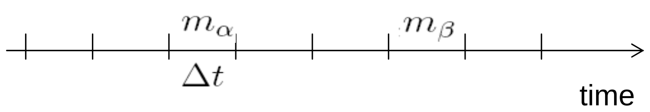
$$\boxed{\hat{\Delta}(s) = \hat{\rho}(s)\hat{\rho}(-s)\hat{\Gamma}(s)}$$

departure arrival

In particular, Poissonian input leads to Poissonian output!

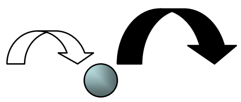
Copy number fluctuations at the node

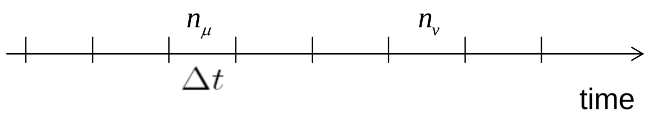




arrival series

τ Delay/residence time with distribution $\rho(\tau)$





departure series

$$\boxed{N(t) = \sum_\alpha m_\alpha [S(t-t_\alpha) - S(t-t_\alpha-\tau_\alpha)]}$$

Mean copy number: $\langle N \rangle = J\langle \tau \rangle$

arrival

dwell time

Fano factor: $$\frac{\sigma_N^2}{\langle N \rangle} = 1 + \frac{1}{J\langle\tau\rangle}\frac{1}{2\pi i}\int_{-i\infty}^{i\infty} ds\hat{\Gamma}(s)\hat{q}(s)\hat{q}(-s)$$

where $q(\tau) = \int_\tau^\infty \rho(t)dt$

Transfer matrix for linear pathways

$$\xrightarrow{\text{input}} 1 \xrightarrow{\tau_1} 2 \xrightarrow{\tau_2} 3 \cdots k \xrightarrow{\tau_k} k+1 \xrightarrow{\tau_{k+1}} \cdots$$

Iterative relation for the non-Poisson component of the autocorrelation

$$\boxed{\hat{\Gamma}_{k+1}(s) = \hat{\Delta}_k(s) = \hat{\rho}_k(s)\hat{\rho}_k(-s)\hat{\Gamma}_k(s)}$$

Noise attenuation along the pathway

Poisson input: $$\frac{\sigma_{N_k}^2}{\langle N_k \rangle} = 1$$

Bursty input: $$\frac{\sigma_{N_k}^2}{\langle N_k \rangle} \simeq 1 + \frac{\langle b(b-1)\rangle}{\langle b \rangle}\frac{\langle \tau_k \rangle}{\sqrt{4\pi}\theta_k}$$ large *k*

here $\theta_k^2 = \sum_{j=1}^{k-1}(\langle \tau_j^2 \rangle - \langle \tau_j \rangle^2)$

Interacting particles: queuing

Idea: individual reaction events proceed much faster than appreciable change in the population size of molecular species

Adiabatic approximation: rate of reaction $f(N)$

Linear expansion: $f(N) \simeq \kappa_{\text{eff}}(N+R)$

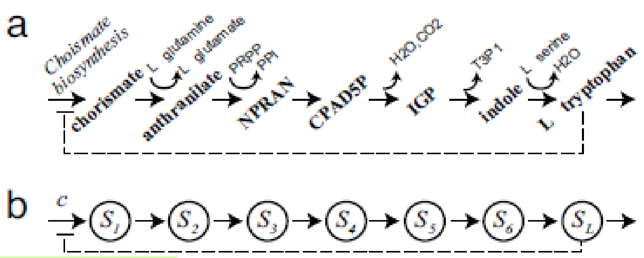
here $\kappa_{\text{eff}} = f'(\langle N \rangle)$ $R = \frac{f(\langle N \rangle)}{f'(\langle N \rangle)} - \langle N \rangle$

⇒Back to the non-interacting case

Details are being worked out

Michaelis-Menten transport along a linear pathway

E Levine and T Hwa (2007) Stochastic fluctuations in metabolic pathways, PNAS **104**:9224-29



Multiple occupancy at each node

Hopping rate

$$w_m = v_{\max} \frac{m}{m + (K + N_E - 1)}$$

Poissonian input

Steady-state distribution factorizes!

$$\boxed{\pi(m_1, m_2, \ldots, m_L) = \prod_{i=1}^{L} \pi_i(m_i)}$$

$$z = c / v_{\max} \qquad \pi(m) = \binom{m + K + (N_E - 1)}{m} (1-z)^{K+N_E} z^m$$

Mean copy number: $\langle m \rangle = (K + N_E) \frac{z}{1-z}$

Fano factor: $\frac{\sigma_m^2}{\langle m \rangle} = \frac{1}{1-z} = \frac{1}{1 - c/v_{\max}}$ Does not approach 1!

Conclusions and outlook

Statistical physics:

- Non-interacting walkers: noise propagation on a network can be studied exactly using propagation of auto-correlation function of arrival/birth time series. Strong noise associated with clustering of birth events. The formalism can be easily adapted to coarse-grained treatments (no Markovian assumption required), and provide a useful framework for data integration and extraction of kinetic parameters

- Interacting walkers: separate fast and slow variables, adiabatic approximation. Need to consolidate with exact results from zero-point processes

Biology

- Identify noise source and noise attenuation along the network
- Attenuation of noise through feedbacks etc.
- Exploiting noise to generate diversity in colonal populations
- Deeper issues w.r.t. economy vs reliable execution of biological function

Collaborate with biologists or play yourself!

Thank you for your attention!

Reference

Olaf Wolkenhauer, Systems Biology: Dynamic Pathway Modelling

http://www.sbi.uni-rostock.de/dokumente/t_sb.pdf