II Master in Biophysics
Universidad Autónoma de Madrid
Nov 8-12/2004

Noise in Gene Expression

Evolutionary Systems Biology Lab

JUAN F. POYATOS
SPANISH NATIONAL CANCER CENTRE

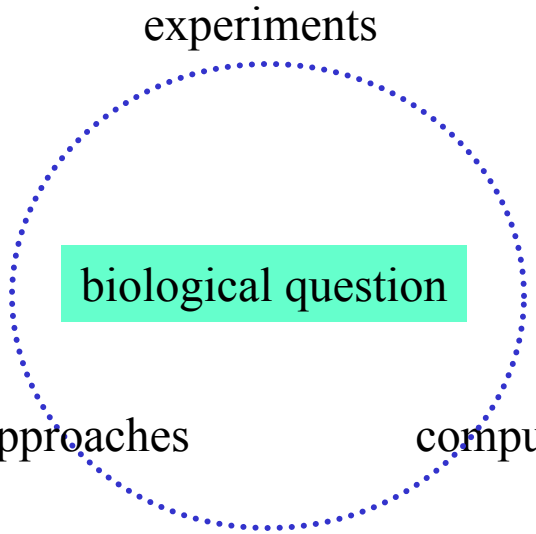
http://bioinfo.cnio.es/~jpoyatos/

Biological question

how does gene expression work in cells?

main goal of the course

a combined approach
to understand Biology

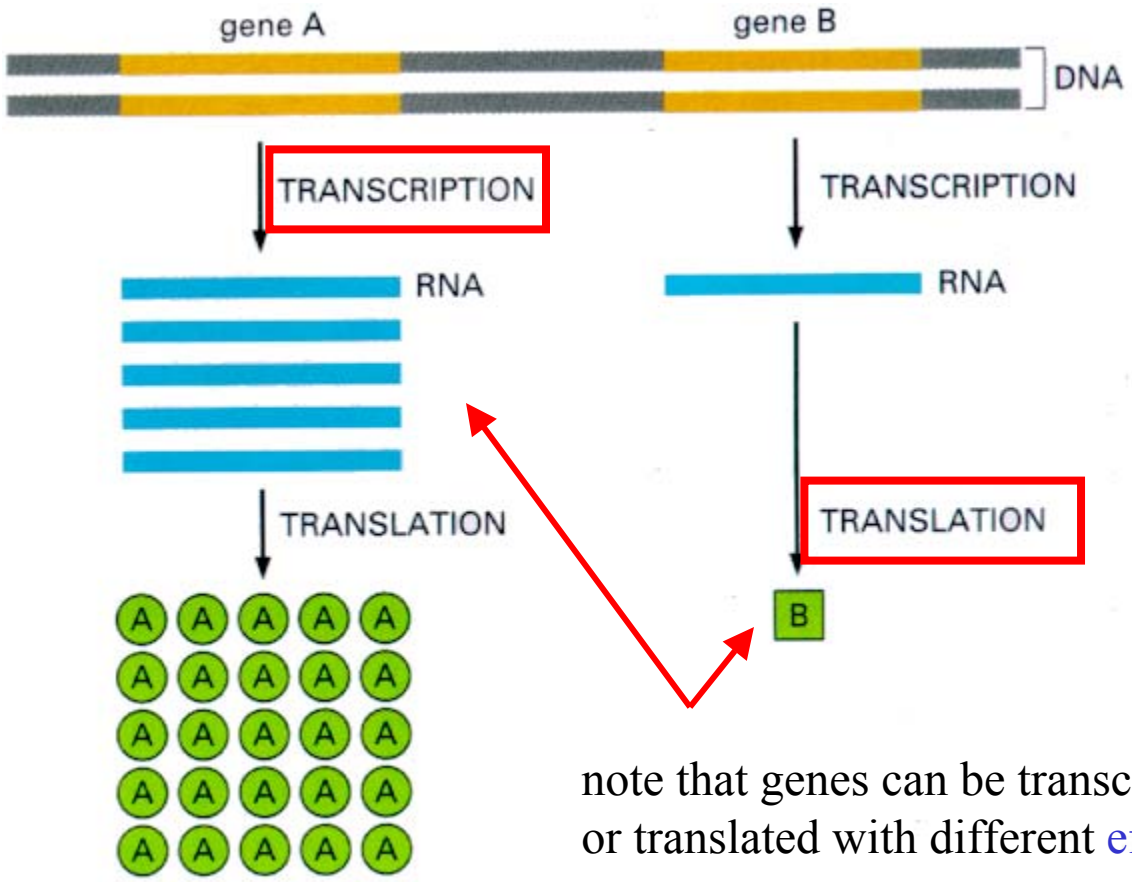


day I


0010100
001011001
001110011011
111100
01001
01011
100001
100
1
100G0AT
ATT01GCA0A
A0T01
ATTCG
TATGGATGG
TTA

What is gene "*expression*" anyway?

- A gene (a piece of DNA) expresses itself when it produces its own distinct protein product, how?, two steps are needed



note that genes can be transcripted or translated with different efficiencies/rates

how does gene expression work in cells?

such fundamental process should be well regulated,
i.e., it should be adjusted in a deterministic clockwise fashion

or ... maybe not?

is gene expression noisy?

- Many molecules that take part in gene expression (including DNA and important regulatory molecules such as the enzyme polymerase) act at extremely low intracellular concentrations (low copy numbers)

- Gene expression as a series of biochemical reactions experiences "surprising" things when one takes the discreteness of molecule number seriously

Escherichia Coli (E.coli) numbers

2μm long
1μm diameter

$V = \pi r^2 l = \pi/2\ 10^{-15}$ liters
[RNA Polymerase] ~ 100nM = 100 molecules
(1nM ~ 1 molecule)

Biochemical noise

-consider a simple gene expression system

$$\emptyset \xrightarrow{k} P, \quad \text{production}$$

$$P \xrightarrow{\delta} \emptyset, \quad \text{first order degradation}$$

a common approach is to describe these reactions by means of differential reaction-rate equations

$$\frac{d[P]}{dt} = k - \delta[P]$$

This approach assumes that the time evolution of such reaction is both continuous and deterministic

continuous? molecule number changes in discrete ways

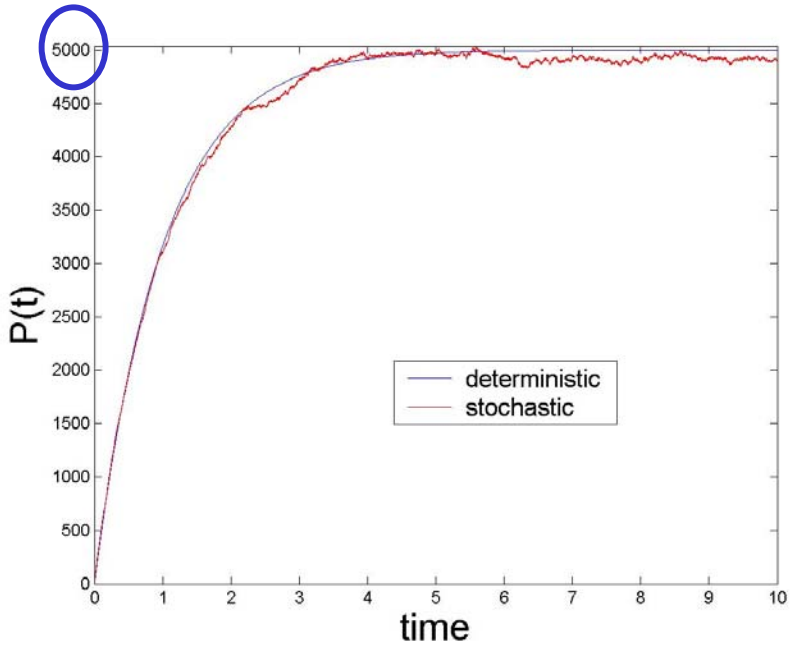
deterministic? impossible to predict the motion of (classical) molecules due to the ignorance of positions and velocities of all components of the system

however in many cases of course the time evolution of a chemically reacting system can, to a very acceptable degree of accuracy, be treated as a continuous, deterministic process

MATLAB code 1

```
% .. code1.m
% .. simple gene expression deterministic equations

clear all
k = 10;
delta = 1;

tspan = [0 10];
P0 = 0;
options = [];
[t P] = ode23(@code1equations,tspan,P0,options,k,delta);
```

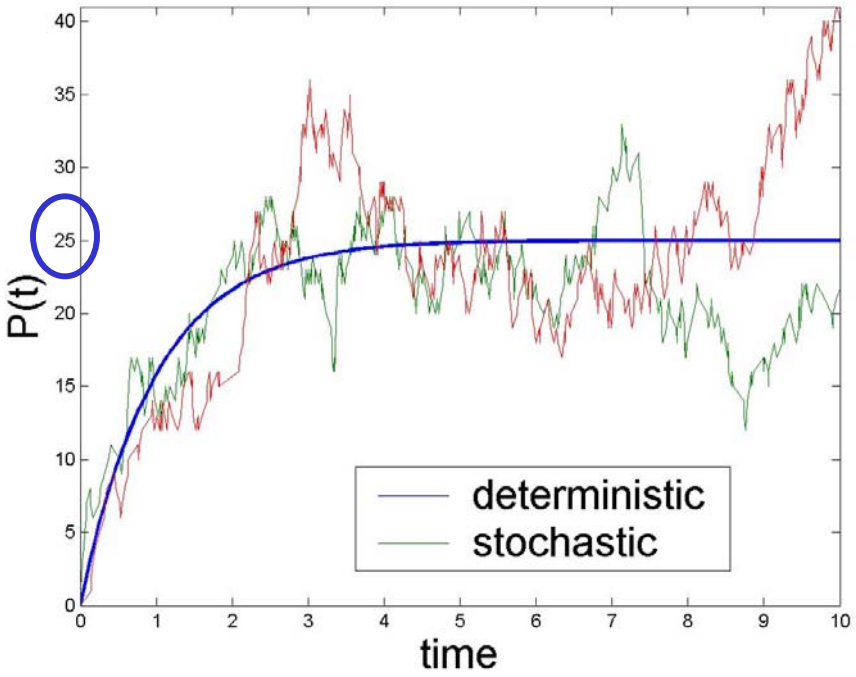
```
% .. code1equations.m
% .. rate equations for code1

function dPdt = code1equations(t,P,k,delta)

dPdt = [k - delta*P(1)];
```



large number of molecules
deterministic approximation works

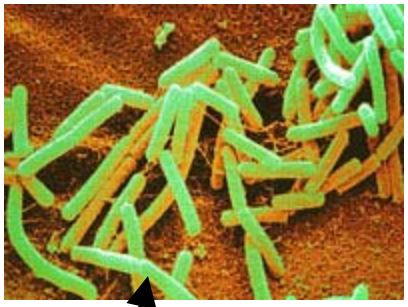
small number of molecules
deterministic approximation fails

large concentration fluctuations



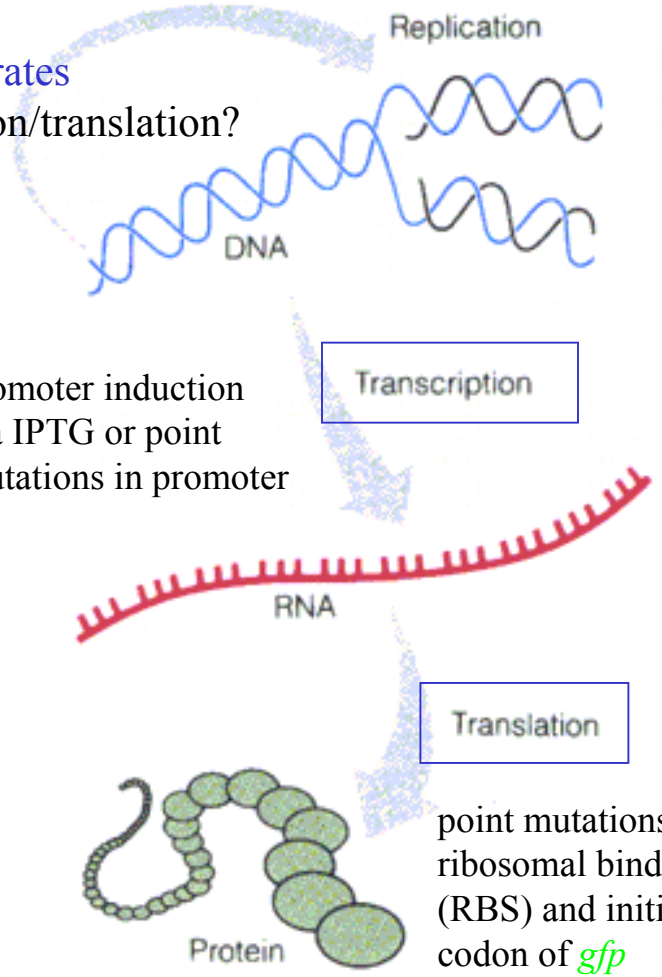
Can we measure experimentally such (intrinsic) noise?

- A single-copy chromosomal gene with an inducible promoter was introduced in Bacilus subtilis



inducible
promoter

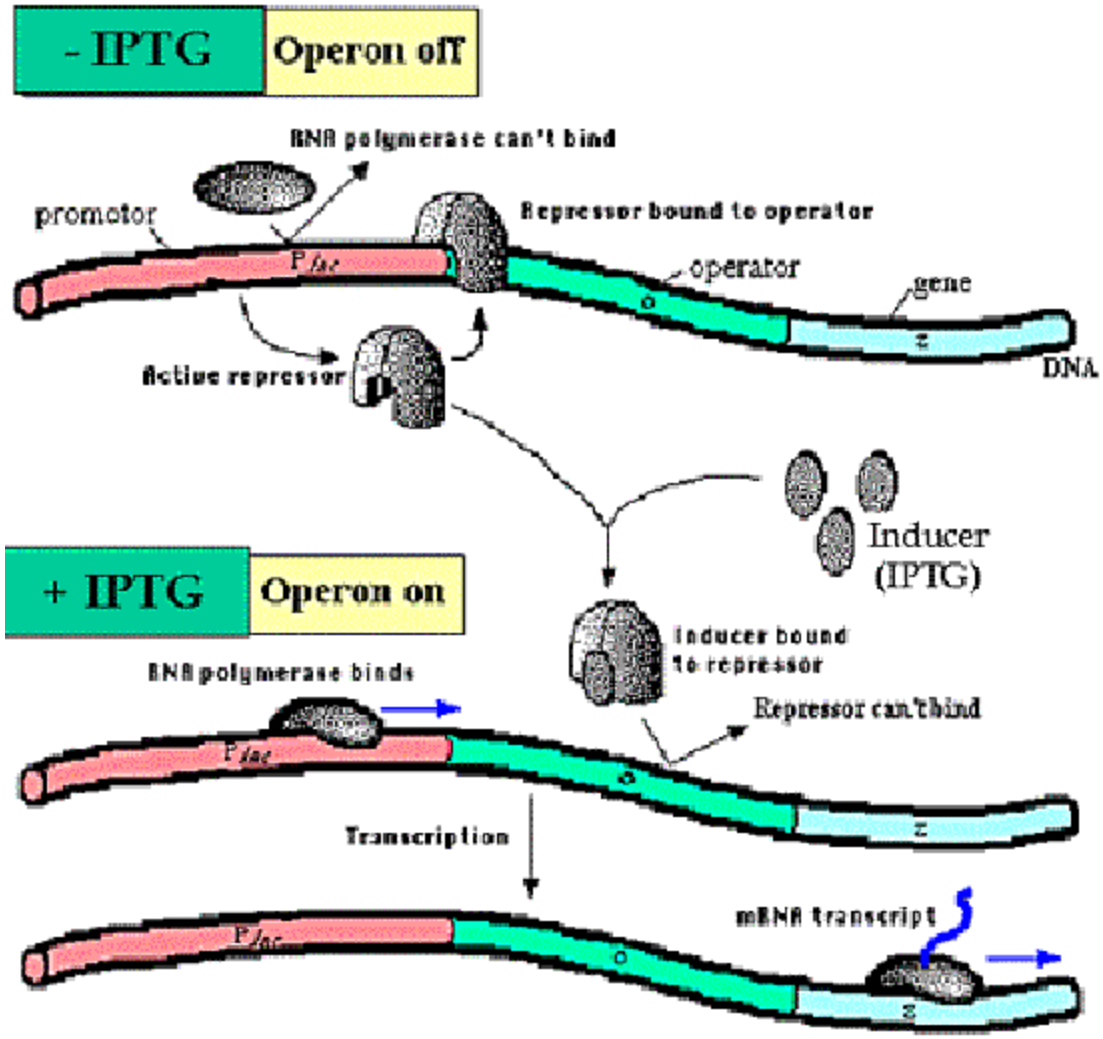
How to vary rates
of transcription/translation?



promoter induction
via IPTG or point
mutations in promoter

point mutations in
ribosomal binding site
(RBS) and initiation
codon of *gfp*

promoter induction
via IPTG or ...



Induction of the *lac* Operon

... or point mutations in promoter

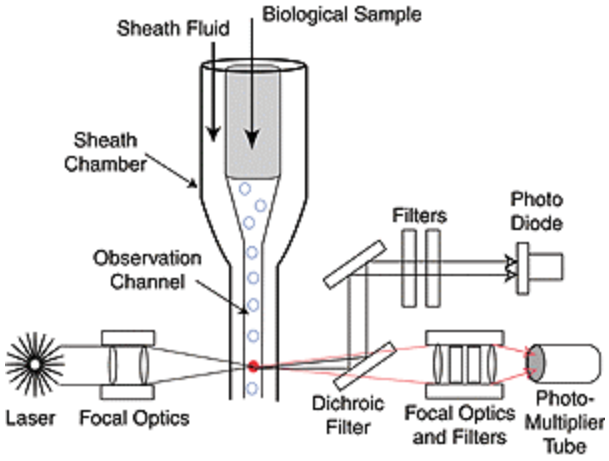
Table 1 • Translational mutants: point mutations in the RBS and initiation codon of *gfp*

Strain	Ribosome binding site	Initiation codon	Translational efficiency
ERT25	GGG AAA AGG AGG TGA ACT ACT	ATG	1.00
ERT27	GGG AAA AGG AGG TGA ACT ACT	TTG	0.87
ERT3	GGG AAA AGG TGG TGA ACT ACT	ATG	0.84
ERT29	GGG AAA AGG AGG TGA ACT ACT	GTG	0.66

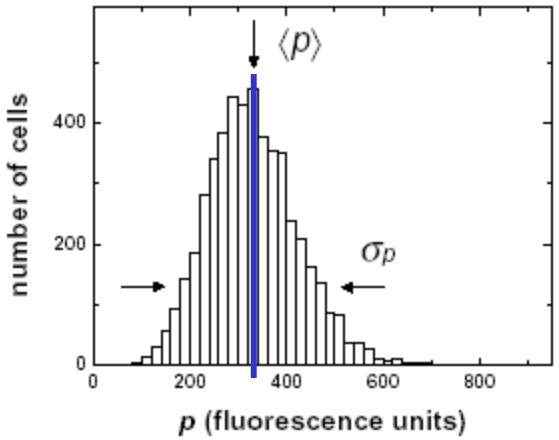
point mutations in ribosomal binding site (RBS) and initiation codon of *gfp*

Table 2 • Transcriptional mutants: point mutations in the P_{spac} promoter

Strain	−10 regulatory region −10 +1	Transcriptional efficiency
ERT57	CAT AAT GTG TGT AAT	6.63
ERT25	CAT AAT GTG TGG AAT	1.00
ERT53	CAT AAT GTG TGC AAT	0.79
ERT51	CAT AAT GTG TGA AAT	0.76
ERT55	CAT AAT GTG TAA AAT	0.76

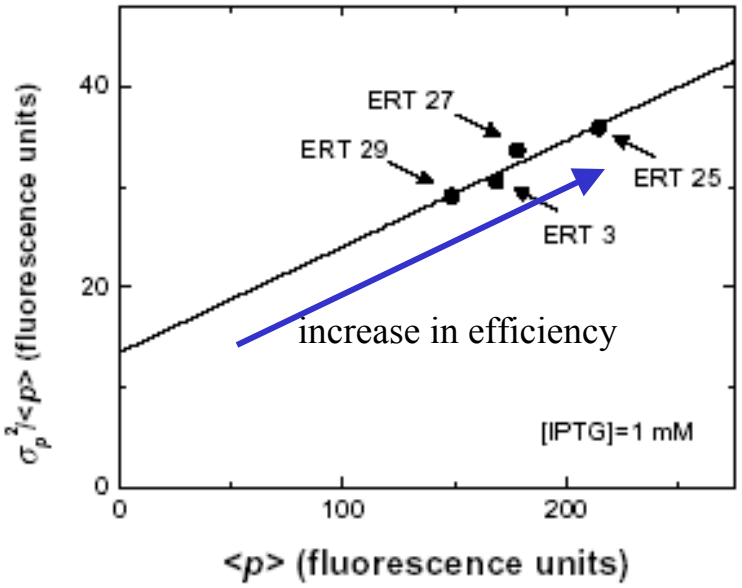


GFP expression level is measured for single cells in a bacterial population using flow cytometry



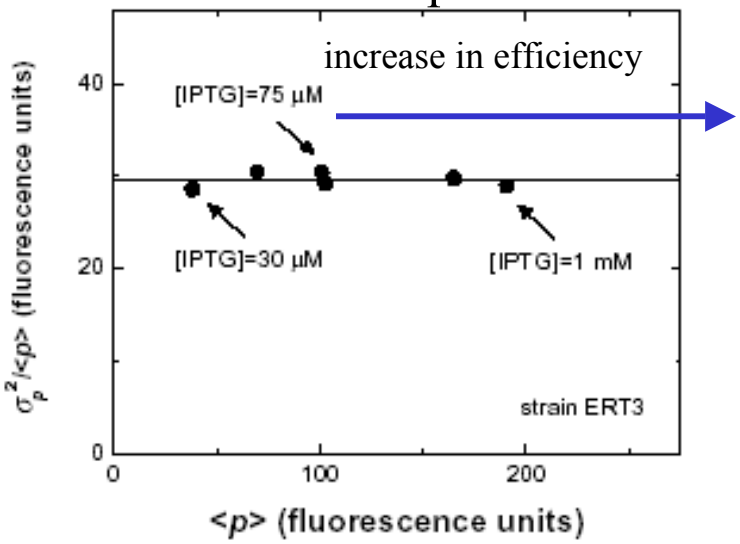
Expression level vary from cell to cell (phenotypic noise) as a consequence of molecular fluctuations within single cells (biochemical noise)

translation



translational efficiency
vs.
transcriptional efficiency

transcription



Can we understand this behaviour in theoretical terms?

day II

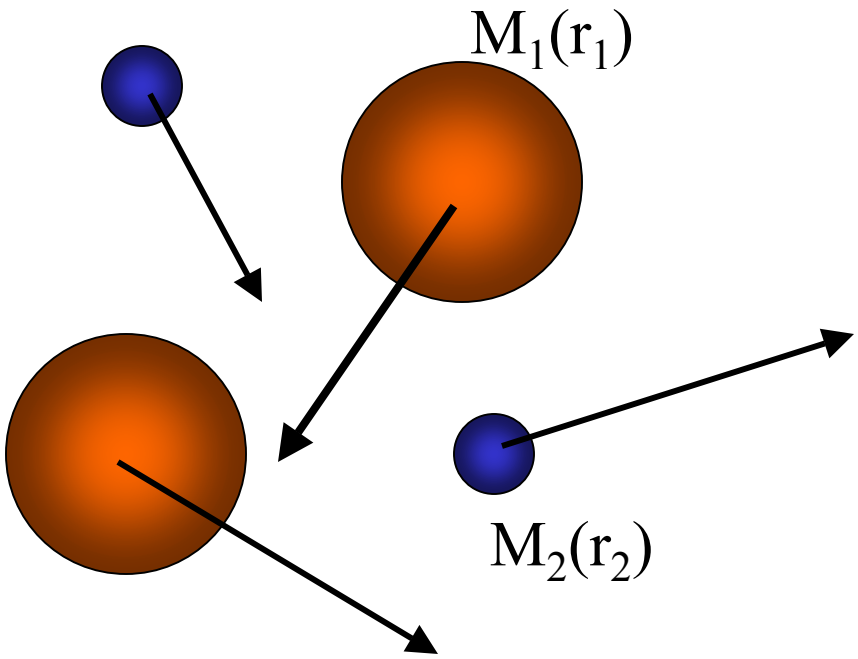

0010100
001011001
001110011011
111100
01001
01011
100001
100
1
100G0AT
ATT01GCA0A
A0T01
ATTCG
TATGGATGG
TTA

Stochastic description of chemical reactions

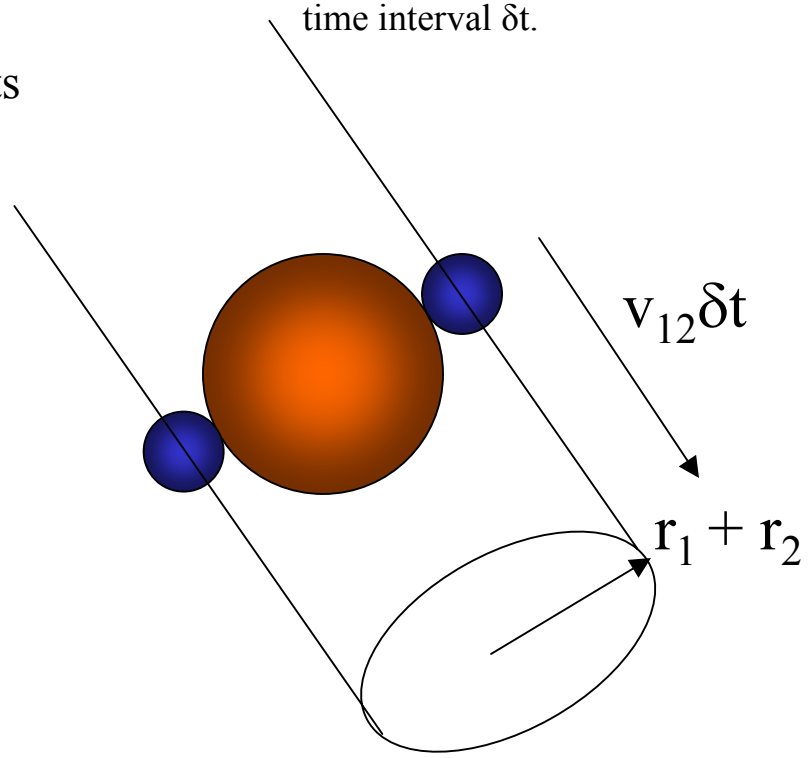
Recall: For a stochastic system it is not possible to determine exactly the state of the system at later times given its state at the current time.
We must thus deal with probabilities.

Basis of the stochastic formulation: a chemical reaction occurs when molecules collide in an appropiate way

- Molecular Collisions: random microscopic events



V_{coll} – Collision volume. The molecules M_2 which are within collision volume will be hit by a particular molecule M_1 in the next time interval δt.



$$V_{coll} = v_{12}\delta t\, \pi\, (r_1 + r_2)^2$$

$$P(\text{a given } M_1 \text{ and } M_2 \text{ collide}) = \frac{\overline{v_{12}}\,\delta t\,\pi\,(r_1 + r_2)^2}{V}$$

$$P(\text{a } M_1 \text{ and } M_2 \text{ molecule collide}) = n_1\,n_2\,\frac{\overline{v_{12}}\,\delta t\,\pi\,(r_1 + r_2)^2}{V}$$

and finally

diffusion-limited R close to one always

$$P(\text{a } M_1 \text{ and } M_2 \text{ react}) = n_1\,n_2\,\boxed{\frac{\overline{v_{12}}\,R\,\pi\,(r_1 + r_2)^2}{V}}\,\delta t = n_1\,n_2\,\mathbf{r}\,\delta t$$

probability that a given M_1 and M_2 react in unit time (**r**)

units of inverse time

this is the fundamental hypothesis from which we derive both the Master Equation and the Stochastic Simulation approaches.

The Master Equation

The stochastic framework considers the discrete number of molecules whose state changes probabilistically

Recall our previous simple gene expression model

$$\emptyset \xrightarrow{k} P,$$

P(*k* reaction) = $\mathbf{r}_k \, \delta t$

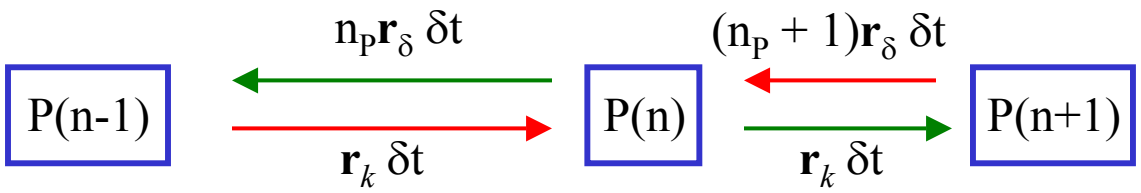
$$P \xrightarrow{\delta} \emptyset,$$

P(δ reaction) = $n_P \, \mathbf{r}_\delta \, \delta t$

$$\frac{d[P]}{dt} = k - \delta[P]$$

Thus, we go from reaction rates to reaction probabilities per unit time

How does the probability of having, say, n *P* molecules, p(n), change with time?



$$
\begin{aligned}
p(n, t+\delta t) &= p(n,t) \\
&+ p(n-1,t) r_k \delta t \\
&+ p(n+1,t)(n_P+1) r_\delta \delta t \\
&- p(n,t) n_P r_\delta \delta t \\
&- p(n,t) r_k \delta t .
\end{aligned}
$$

and thus we get in the limit $\delta t \to 0$

$$
\frac{dp(n)}{dt} = -p(n)(r_k + n_P r_\delta) + p(n-1) r_k + p(n+1)(n_P+1) r_\delta
$$

Some comments:

- All moments of the distribution p(n) can be derived from it
- It is a linear equation in p(n).
- Solving the master equation can be done for simple systems, however only normally at steady state.
- In connection with experiments, p(n) would represent the fraction of cells having n copies of some given protein

Equation of the mean; emergence of deterministic law

$$\frac{d\langle n\rangle}{dt} = \sum_n n\frac{dp_n}{dt}$$

$$= \sum_n n[-p_n(r_k + nr_\delta) + p_{n-1}r_k + p_{n+1}(n+1)r_\delta]$$

$$= -r_k\langle n\rangle - r_\delta\sum_n n^2 p_n + r_k\sum_n p_{n-1}n + r_\delta\sum_n n(n+1)p_{n+1}$$

$$= r_k - r_\delta\langle n\rangle$$

Considering that $[P] = \frac{\langle n \rangle}{V}$

We can rewrite the deterministic equation as

$$\frac{d\langle n \rangle}{dt} = Vk - \delta V[P] = Vk - \delta\langle n \rangle.$$

And thus

$$\begin{array}{rcl} r_k & = & Vk \\ r_\delta & = & \delta \end{array}$$

pseudofirst-order reaction

first order reaction

Steady State

$$\frac{dp_n}{dt} = 0 = -p_n(r_k + nr_\delta) + p_{n-1}r_k + p_{n+1}(n+1)r_\delta$$

and

$$-p_n r_k + p_{n+1} r_\delta (n+1) = -p_{n-1} r_k + p_n r_\delta n$$

then

$-p_n r_k + p_{n+1} r_\delta (n+1)$ is constant (independent of n).

further, considering that $\langle n \rangle_{ss} = \frac{r_k}{r_\delta}$ this constant is zero

thus $p_n = \frac{\langle n \rangle_{ss}}{n} p_{n-1} = \ldots = \frac{\langle n \rangle_{ss}^n}{n!} p_0.$

since $\sum_n p_n = 1$ we get $\boxed{p_n = \frac{\langle n \rangle_{ss}^n}{n!} e^{\langle n \rangle_{ss}}}$

the steady state distribution is the Poisson Distribution

Poisson distribution

mean $\langle n \rangle = \langle n \rangle_{ss}$

variance $\sigma^2 = \langle n \rangle_{ss}$

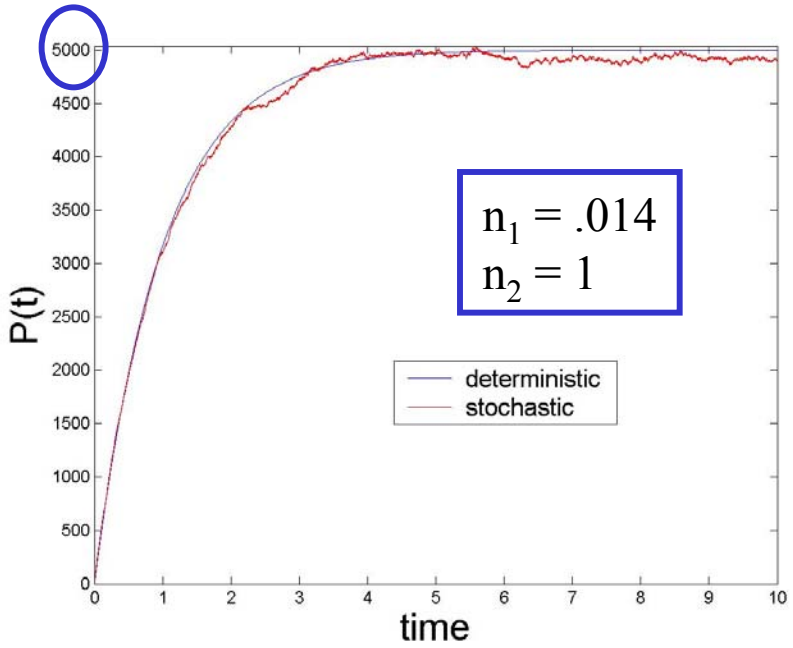
Macroscopic statistics

standard deviation

What is noise then?

definition-1 = $n_1 = \frac{\sigma}{\langle n \rangle}$ ($= 1/\sqrt{\langle n \rangle}$. Poisson distribution, noise increases as the number of molecules decreases)

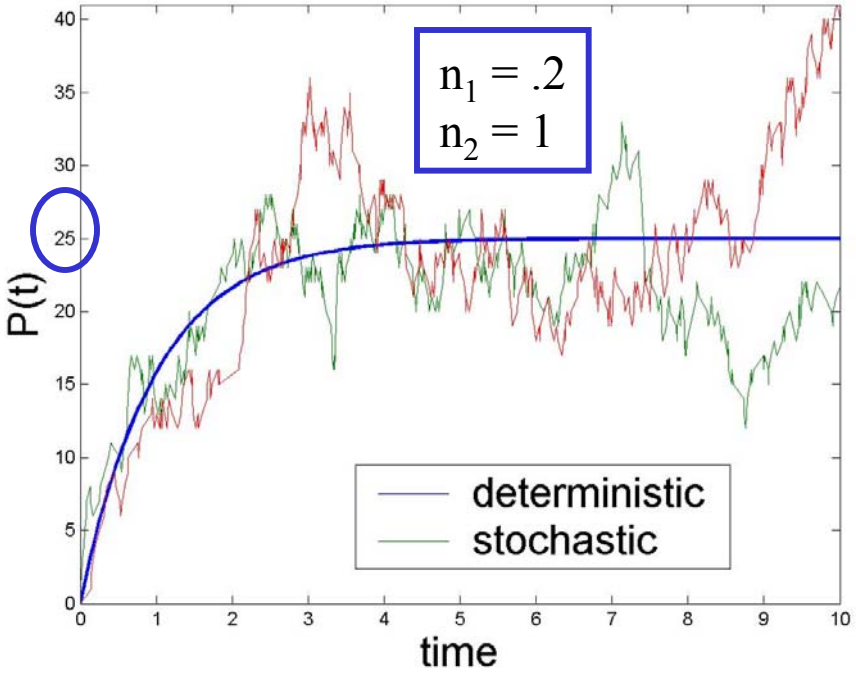
definition-2 (Fano factor) = $n_2 = \frac{\sigma^2}{\langle n \rangle}$ (= 1, Poisson distribution)



large number of molecules
deterministic approximation works

small number of molecules
deterministic approximation fails

large concentration fluctuations



Simulating Stochastic Reactions

Two key questions:

When will the next reaction occur?
What kind of reaction will it be?

$P(\tau,\mu)d\tau$ = probability that, given the state $(X_1,...X_N)$ at time t, the next reaction in V occurs in the infinitisemal time interval $(t + \tau, t + \tau + d\tau)$ and it will be an R_μ reaction.

propensity function, e.g., $n_1 n_2$ **r**

$P(\tau,\mu)d\tau = P_0(\tau)\, a_\mu d\tau$, here $P_0(\tau)$ is the probability that no reaction happens in the time interval $(t, t + \tau)$ and $a_\mu d\tau$ is the probability that reaction Rμ will happen in the time interval $(t + \tau, t + \tau + d\tau)$

The function $P_0(t)$:

$P_0(t+dt) = P_0(t)(1 - a_0\, dt) \qquad a_0 = \sum_{j=1,M} a_j$

$(P_0(t+dt) - P_0(t))/dt = -a_0\, P_0(t)$

$d\, P_0 /dt = -a_0\, P_0(t)$

$P_0(t) = \exp(-a_0 t)$

The reaction probability density function:

$P(\tau,\mu)d\tau = P_0(\tau)\, a_\mu d\tau = a_\mu \exp(-a_0\tau)\, d\tau \qquad \mu = 1,...,M \quad \tau \in (0,+\infty)$

It is possible to write $P(\tau,\mu)$ as a product of $P(\tau)$ and $P(\mu)$:

$P(\tau,\mu)d\tau = a_\mu \exp(-a_0\tau)\, d\tau = (a_\mu / a_0)\, a_0 \exp(-a_0\tau)\, d\tau$

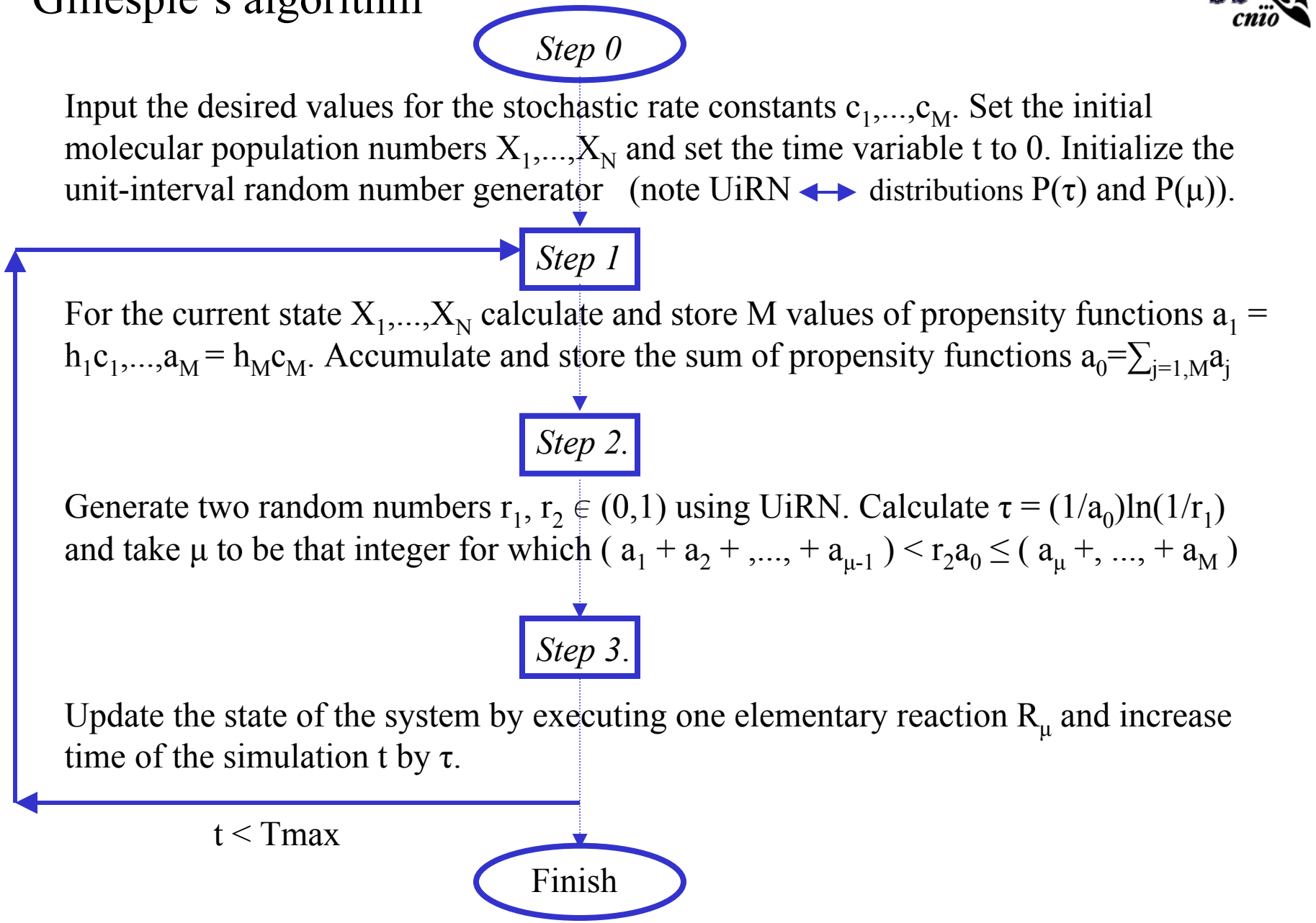
$P(\mu) = (a_\mu / a_0)$

$P(\tau) = a_0 \exp(-a_0\tau)\, d\tau$

Therefore, we may determine the waiting time for the next reaction by generating two random numbers following distributions $P(\tau)$ and $P(\mu)$.

Note that the algorithm is a rigurous consequence of the Fundamental Hypothesis

Gillespie's algorithm

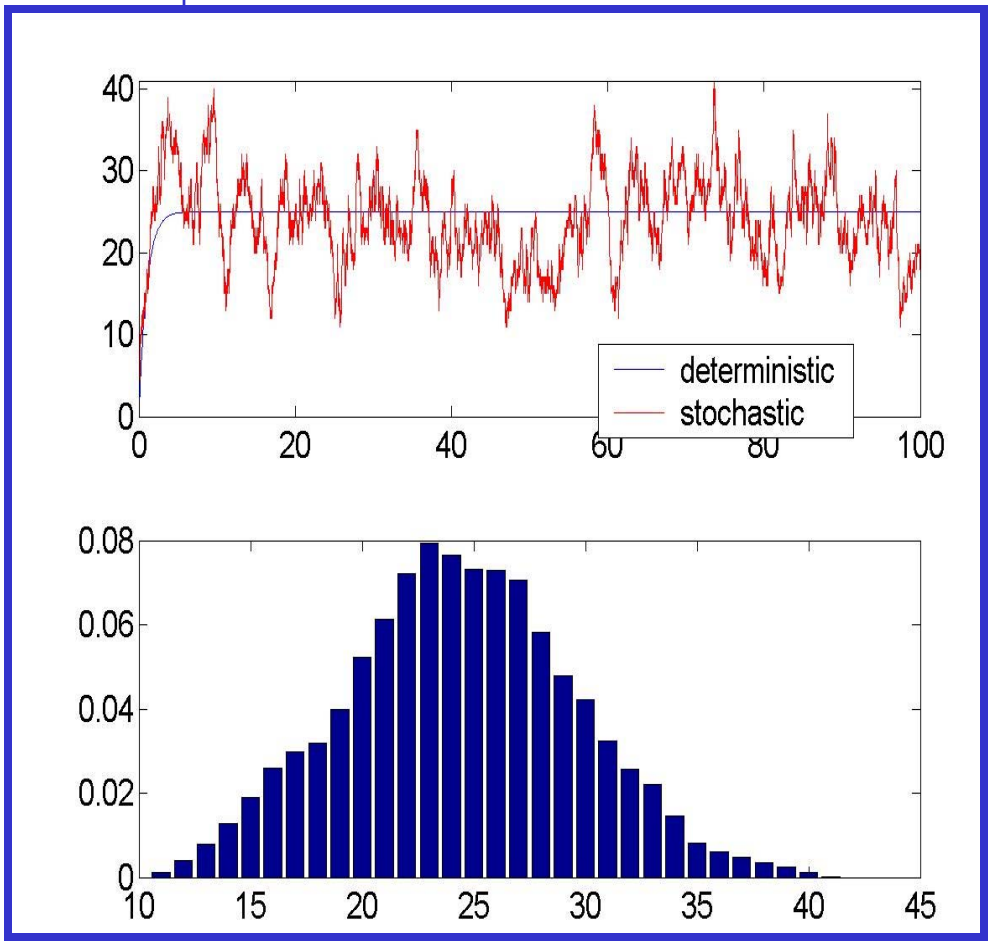


MATLAB code 2

```
% .. code1stoch.m
% .. simple gene expression stochastic and deterministic

clear all
k = 25;
delta = 1;

% .. stochastic eqs. Gillespie's algorithm
P = 0;
Pstochastic = P;
tmax = 10;
t = 0;
tspan = t;
```

```
while t < tmax

    % .. a's
    a = [k, delta*P(1)];
    a0 = sum(a);
    % .. determine time of next reaction
    r1 = rand;
    tau = -log(r1)/a0;
    t = t + tau;
    % .. determine nature of next reaction
    r2 = rand;
    acumsum = cumsum(a)/a0;
    chosen_reaction = min(find(r2 <= acumsum));

    if chosen_reaction == 1;
        P(1) = P(1) + 1;
    else
        P(1) = P(1) - 1;
    end

    tspan = [tspan,t];
    Pstochastic = [Pstochastic;P];

end
```

```
% .. deterministic eqs.
P0 = 0;
options = [];
[t P] = ode23(@code1equations,tspan,P0,options,k,delta);

% .. plot
subplot(211)
plot(t,P,t,Pstochastic,'r')
legend('deterministic','stochastic')
axis([0 tmax 0 max(Pstochastic)]);

% .. histogram, example of matlab use
subplot(212)
vv = Pstochastic(find(t>3));
his = min(vv):max(vv);
histovv = length(his);
cc = 0;
for n = his
    cc = cc + 1;
    histovv(cc) = length(find(vv == n));
end
histovv = histovv/sum(histovv);
bar(his,histovv)
meanhist = sum(his.*histovv)
varihist = sum(his.*his.*histovv) - meanhist*meanhis_
fano = varihist/meanhist
```



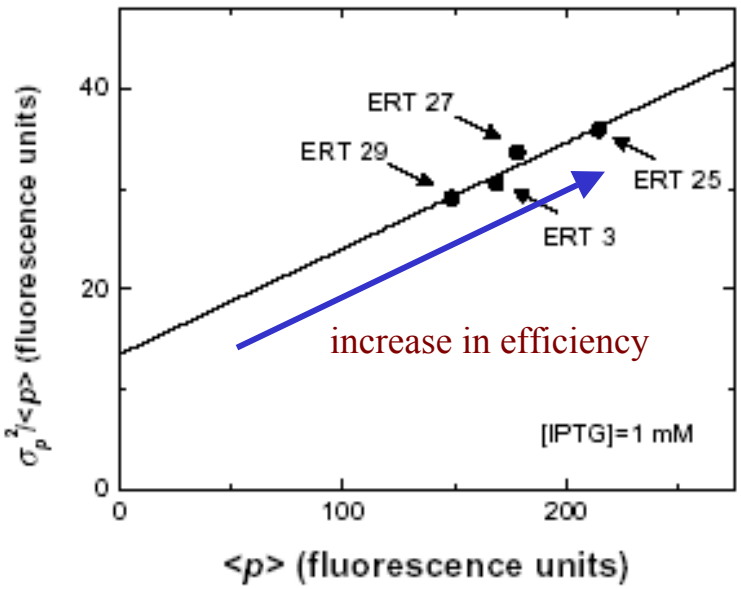

```
meanhist = 24.3909

varihist = 26.7338

fano = 1.0961
```

day III


0010100
001011001
001110011011
111100
01001
01011
100001
100
1
100G0AT
ATT01GCA0A
A0T01
ATTCG
TATGGATGG
TTA

A more detail model of gene expression

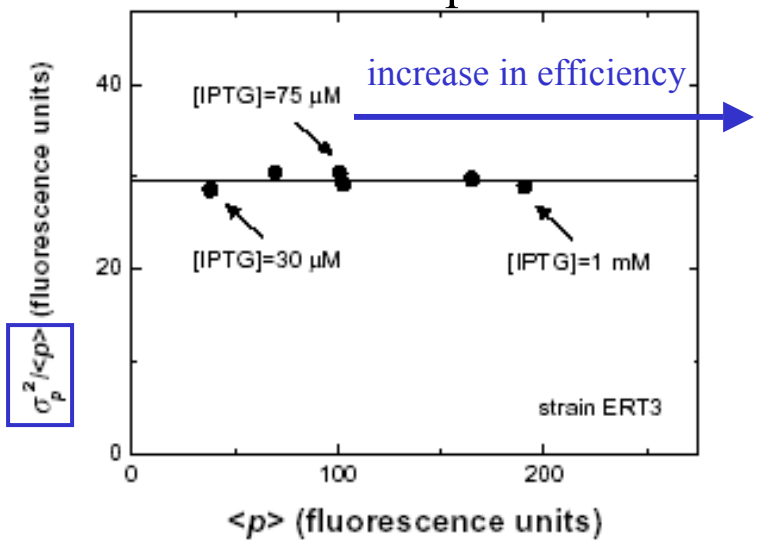
translation



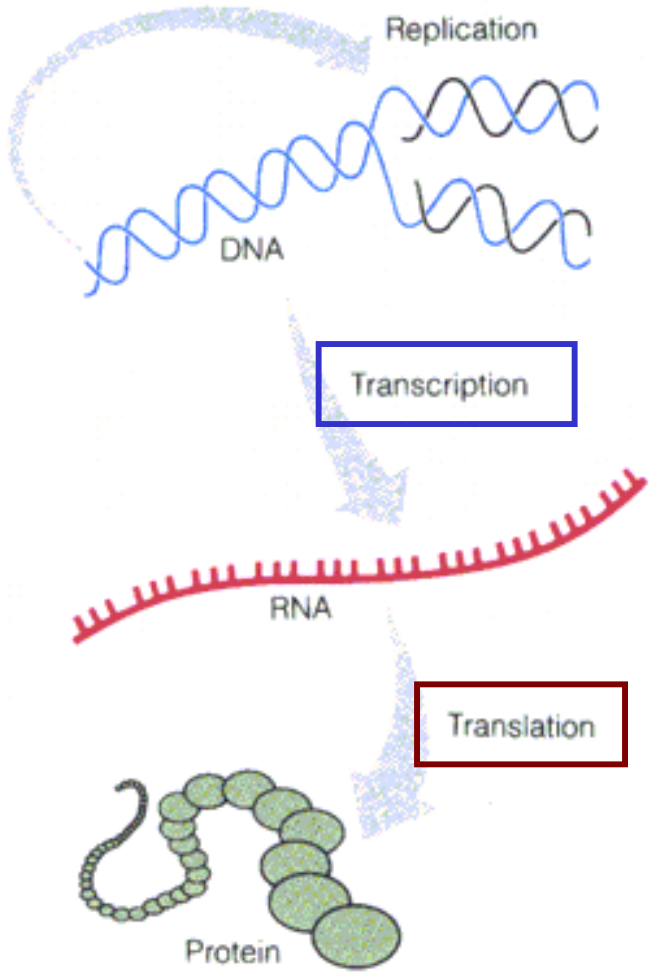
we want to understand
the separate contribution to noise of
transcription and translation

recall:

transcription



A more detail model of gene expression



Deterministic model

$$\frac{d[mRNA]}{dt} = k_R - \delta_R[mRNA]$$

$$\frac{d[P]}{dt} = k_P[mRNA] - \delta_P[P]$$

$$\emptyset \xrightarrow{k_R} mRNA, \qquad P(k_R \text{ reaction}) = \mathbf{r}_k \, \delta t$$

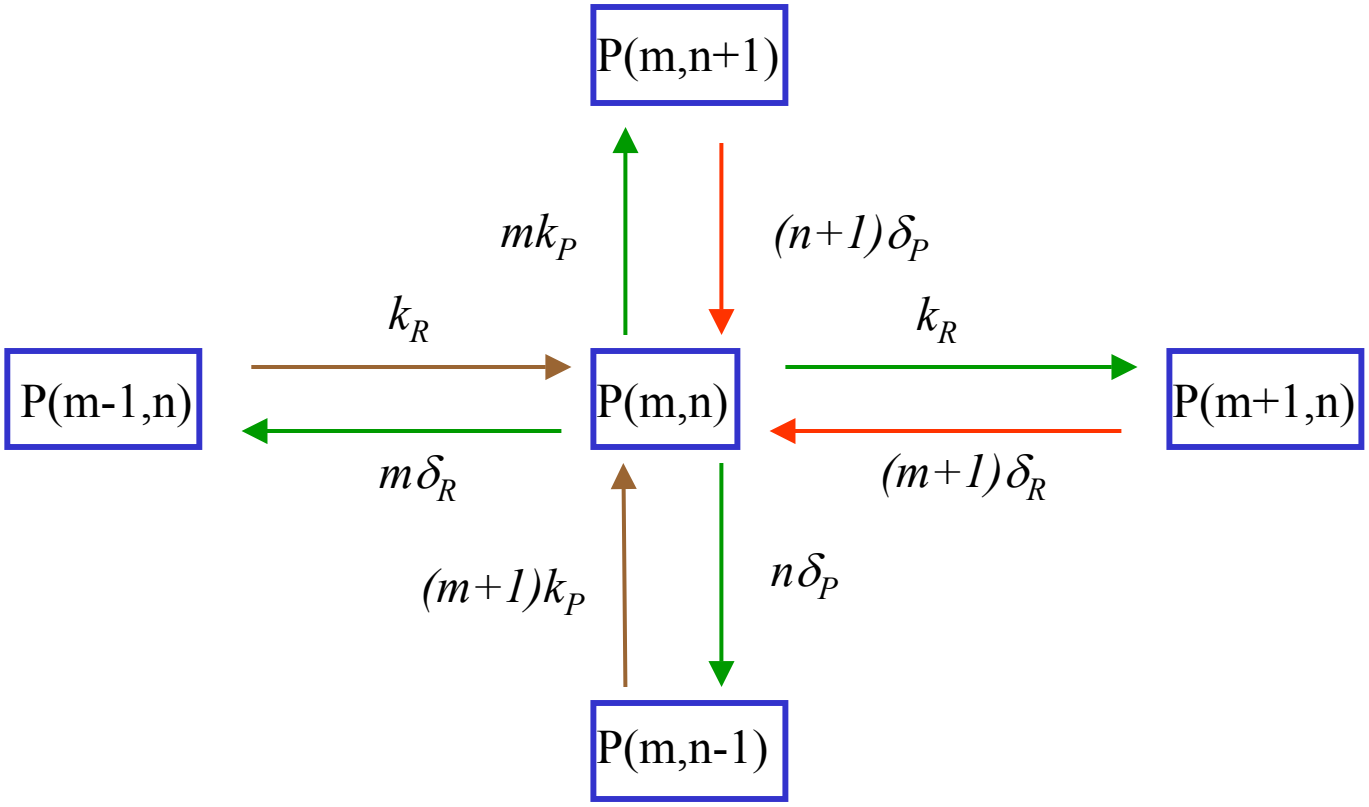
$$mRNA \xrightarrow{\delta_R} \emptyset, \qquad P(\delta_R \text{ reaction}) = m_R \, \mathbf{r}_\delta \, \delta t$$

$$mRNA \xrightarrow{k_P} P, \qquad P(k_P \text{ reaction}) = m_R \, \mathbf{r}_\delta \, \delta t$$

$$P \xrightarrow{\delta_P} \emptyset, \qquad P(\delta_P \text{ reaction}) = n_P \, \mathbf{r}_\delta \, \delta t$$

How does the probability of having, say, m mRNA molecules and n *P* molecules, p(m,n), change with time?

$\mathbf{r}_\kappa$'s as k's to simplify notation
(this could also imply that $V = 1$)



$$\frac{dp_{m,n}}{dt} = -p_{m,n}[m\delta_R + mk_P + k_R + n\delta_P]$$
$$+ \; p_{m,n+1}(n+1)\delta_P + p_{m+1,n}(m+1)\delta_R$$
$$+ \; p_{m,n-1}k_P m + p_{m-1,n}k_R$$

Equation of the mean; emergence of deterministic laws

note first, a useful equation for a given function *f(n,m)*

$$\begin{aligned}\frac{d\langle f_{n,m}\rangle}{dt} &= -\langle f_{n,m}m\rangle\delta_R - \langle f_{n,m}m\rangle k_P - \langle f_{n,m}\rangle k_R - \langle f_{n,m}n\rangle\delta_P \\ &+ \langle f_{n-1,m}n\rangle\delta_P + \langle f_{n,m-1}m\rangle\delta_R + \langle f_{n+1,m}m\rangle k_P + \langle f_{n,m+1}\rangle k_R\end{aligned}$$

thus, we get

$$\frac{d\langle m\rangle}{dt} = k_R - \delta_R\langle m\rangle$$

this is the equation the very same equation we obtained for the simple model, i.e., it implies steady state Poisson statistics for mRNA

$$\frac{d\langle n\rangle}{dt} = k_P\langle m\rangle - \delta_P\langle n\rangle$$

what kind of protein macroscopic steady state statistic characterizes protein dynamics?

we make use of the following equations ...

$$\frac{d\langle n^2\rangle}{dt} = -2\langle n^2\rangle\delta_P + \langle n\rangle\delta_P + 2\langle nm\rangle k_P + \langle m\rangle k_P$$
$$\frac{d\langle nm\rangle}{dt} = -\langle nm\rangle(\delta_P + \delta_R) + \langle m^2\rangle k_P + \langle n\rangle k_R$$

... to get the final expressions for the macroscopic statistics

Fano Protein = $\frac{\langle n^2\rangle - \langle n\rangle^2}{\langle n\rangle} = 1 + \frac{k_P/\delta_R}{1+\delta_P/\delta_R} \approx \boxed{1 + \frac{k_P}{\delta_R}}$

translation efficiency influences noise

Fano mRNA = $\boxed{1}$

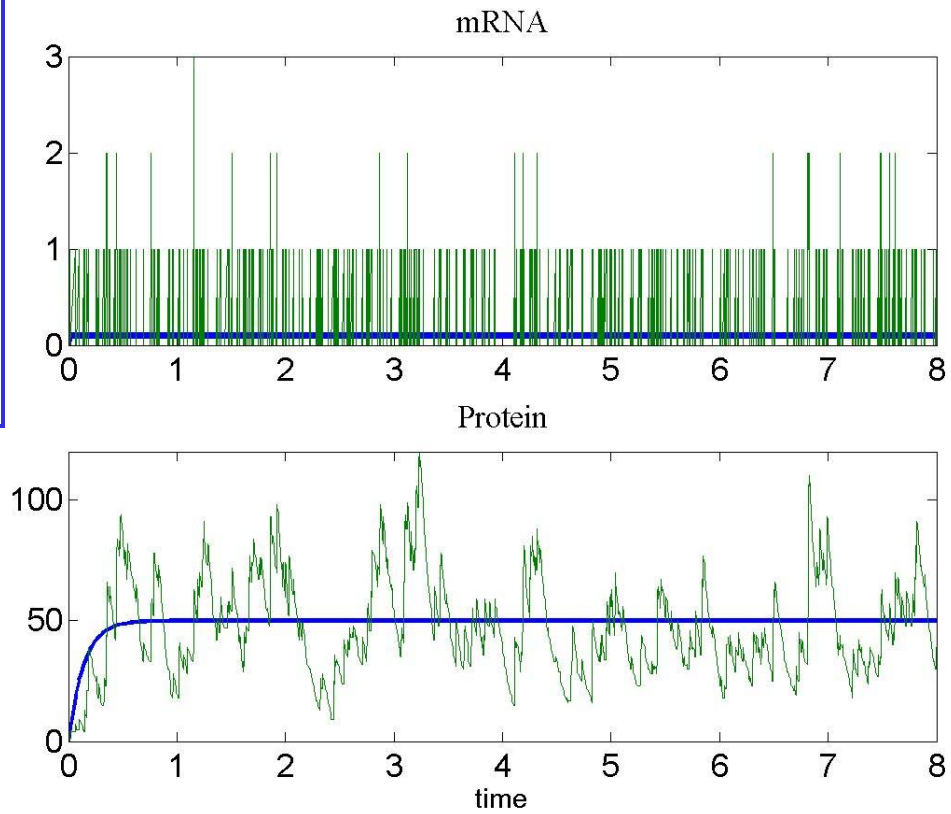
↓

transcription efficiency does not influence noise

protein half-lifetime ~ hours
mRNA half-lifetime ~ minutes
thus

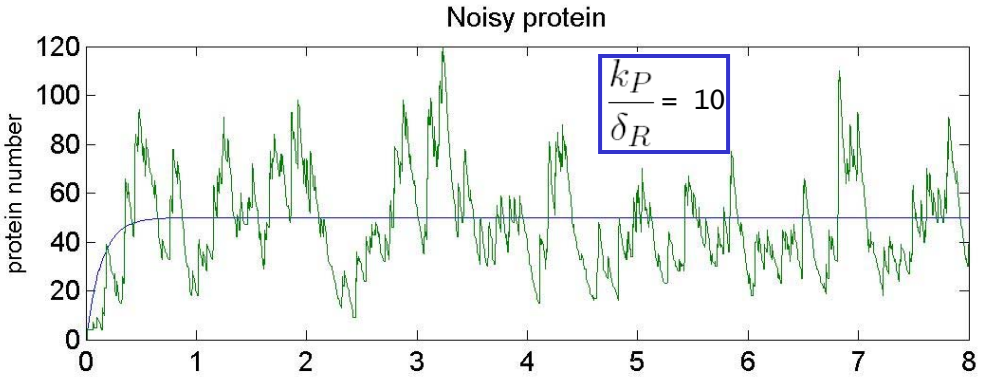
$t_{1/2} = \log 2/\delta$ and $\delta_P \ll \delta_R$

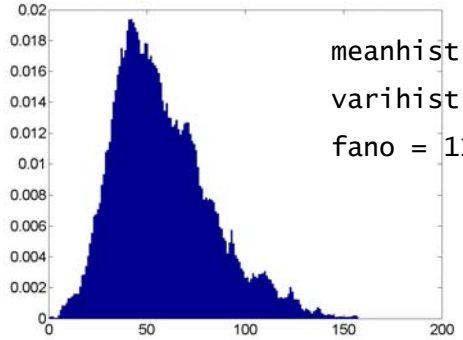
```
% .. code2stoch.m
% .. more detail gene expression stochastic and
deterministic


clear all
kR = .01;       % .. []/s
deltaR = .1;    % .. 1/s
kP = 10*deltaR; % .. 1/s
deltaP = .002   % .. 1/s


% .. stochastic eqs. Gillespie's algorithm
P = [0 0];
Pstochastic = P;
tmax = 8*60*60;      % .. hours
t = 0;
tspan = t;
```

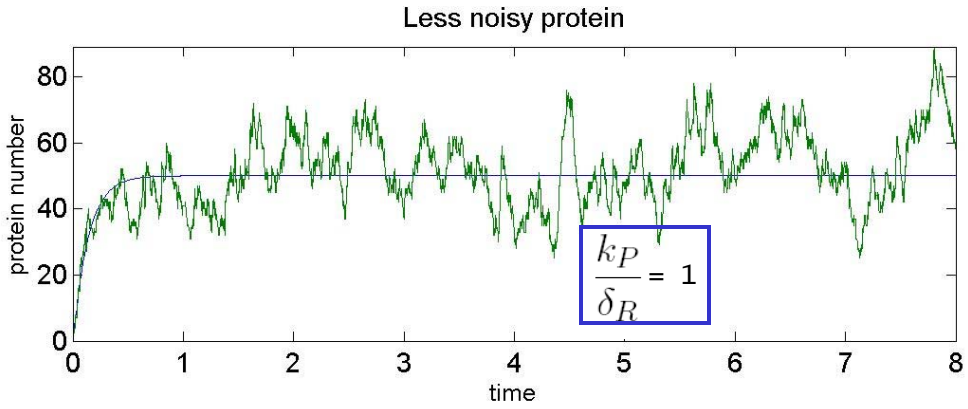


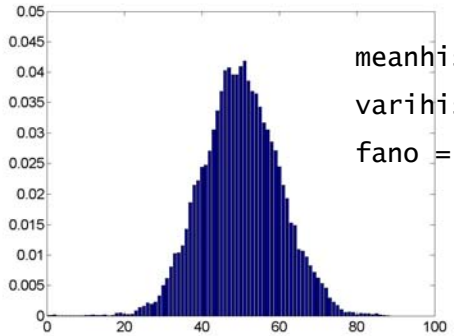
```
while t < tmax
    % .. a's
    a = [kR, deltaR*P(1),kP*P(1),deltaP*P(2)];
    a0 = sum(a);

    % .. determine time of next reaction
    r1 = rand;
    tau = -log(r1)/a0;
    t = t + tau;

    % .. determine nature of next reaction
    r2 = rand;
    acumsum = cumsum(a)/a0;
    chosen_reaction = min(find(r2 <= acumsum));

    if chosen_reaction == 1;
        P(1) = P(1) + 1;
    elseif chosen_reaction == 2;
        P(1) = P(1) - 1;
    elseif chosen_reaction == 3;
        P(2) = P(2) + 1;
    else
        P(2) = P(2) - 1;
    end

    tspan = [tspan,t];
    Pstochastic = [Pstochastic;P];

end
```

```
% .. deterministic eqs.
P0 = [0,0];
options = [];
[t P] = ode23(@code2equations,tspan,P0,options,kR,deltaR,kP,deltaP);

% .. plot
subplot(211);plot(t/60/60,P(:,1),t/60/60,Pstochastic(:,1))
axis([0 tmax/60/60 0 max(Pstochastic(:,1))]);title('mRNA');
subplot(212);plot(t/60/60,P(:,2),t/60/60,Pstochastic(:,2))
axis([0 tmax/60/60 0 max(Pstochastic(:,2))]);title('Protein')
```


Noisy protein
protein number
$\frac{k_P}{\delta_R} = 10$

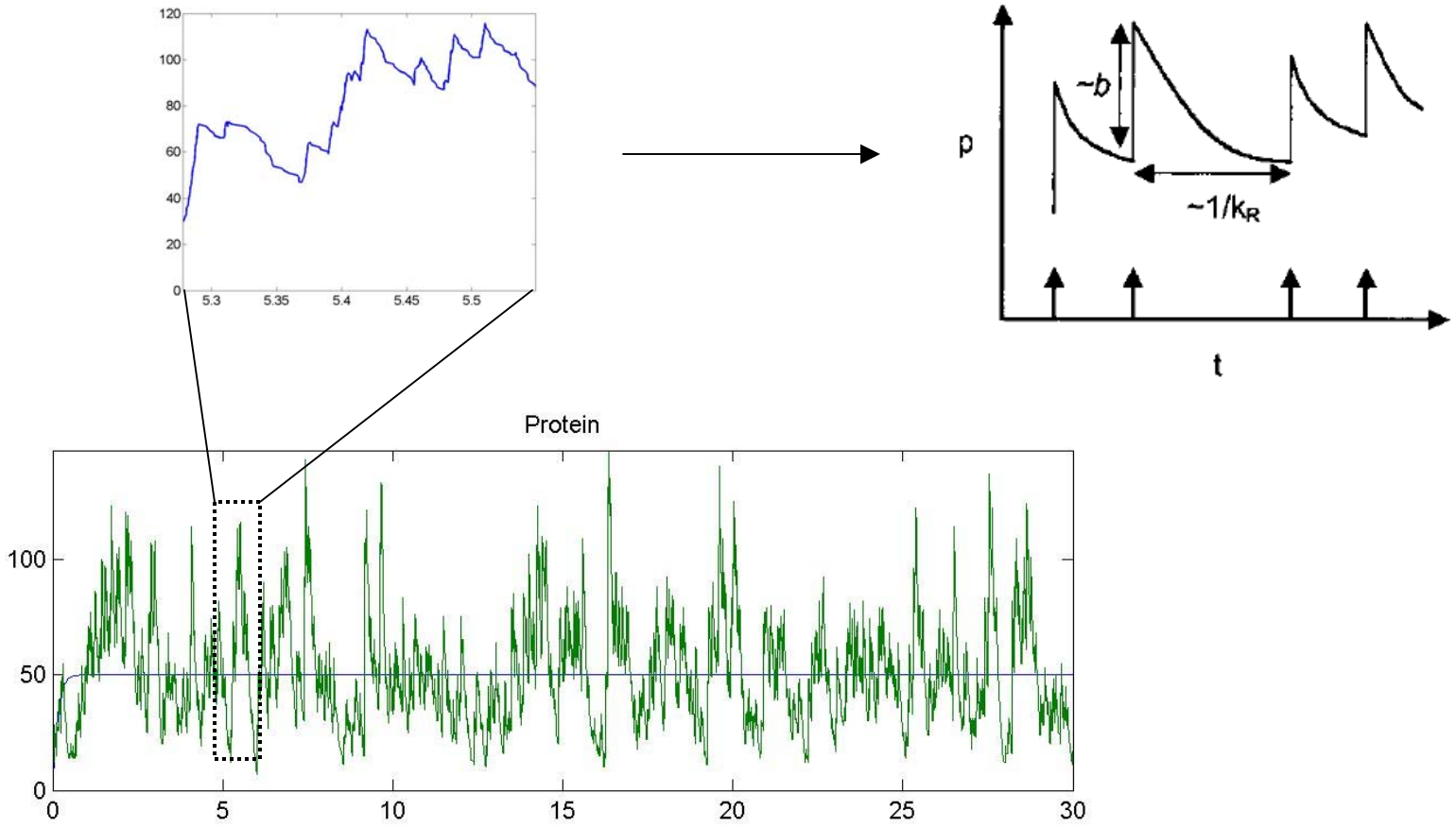

meanhist = 57.1261
varihist = 634.6932
fano = 11.1104


Less noisy protein
protein number
time
$\frac{k_P}{\delta_R} = 1$


meanhist = 50.0542
varihist = 100.2078
fano = 2.0020

“Random bursts model”

$$b = \frac{k_P}{\delta_R}$$



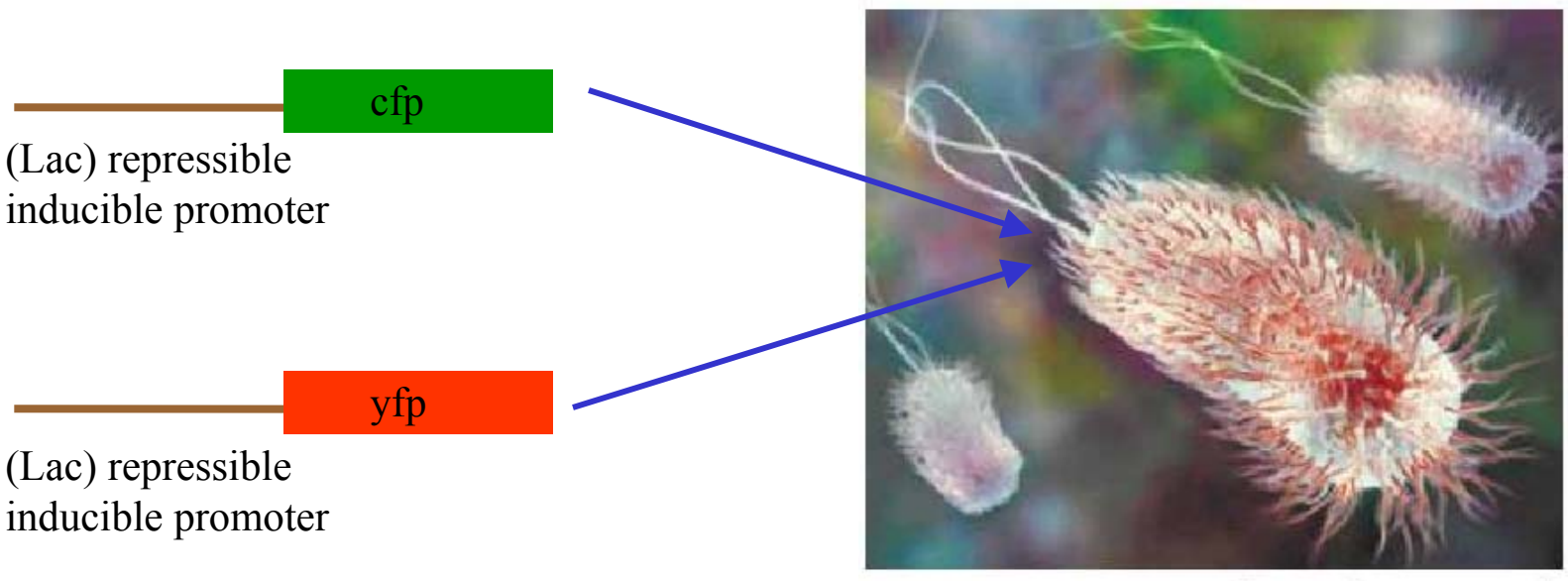
What about external sources of noise?

Intrinsic noise, even if all cellular conditions are equivalent for cells, we have seen that the reactions associated to transcription and translation originate noise

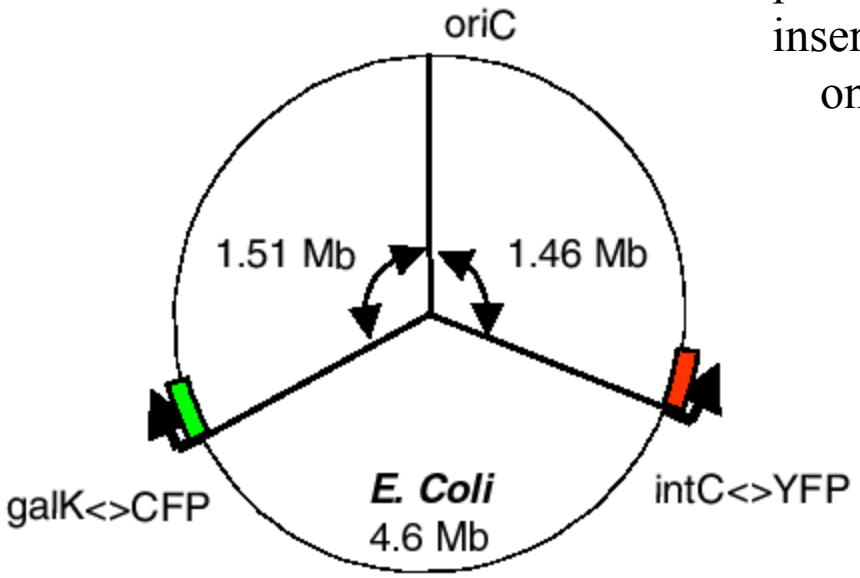
Extrinsic noise, other molecular species (genes themselves too!), e.g., RNA polymerase, originate noise too

Can we discriminate both sources of noise?

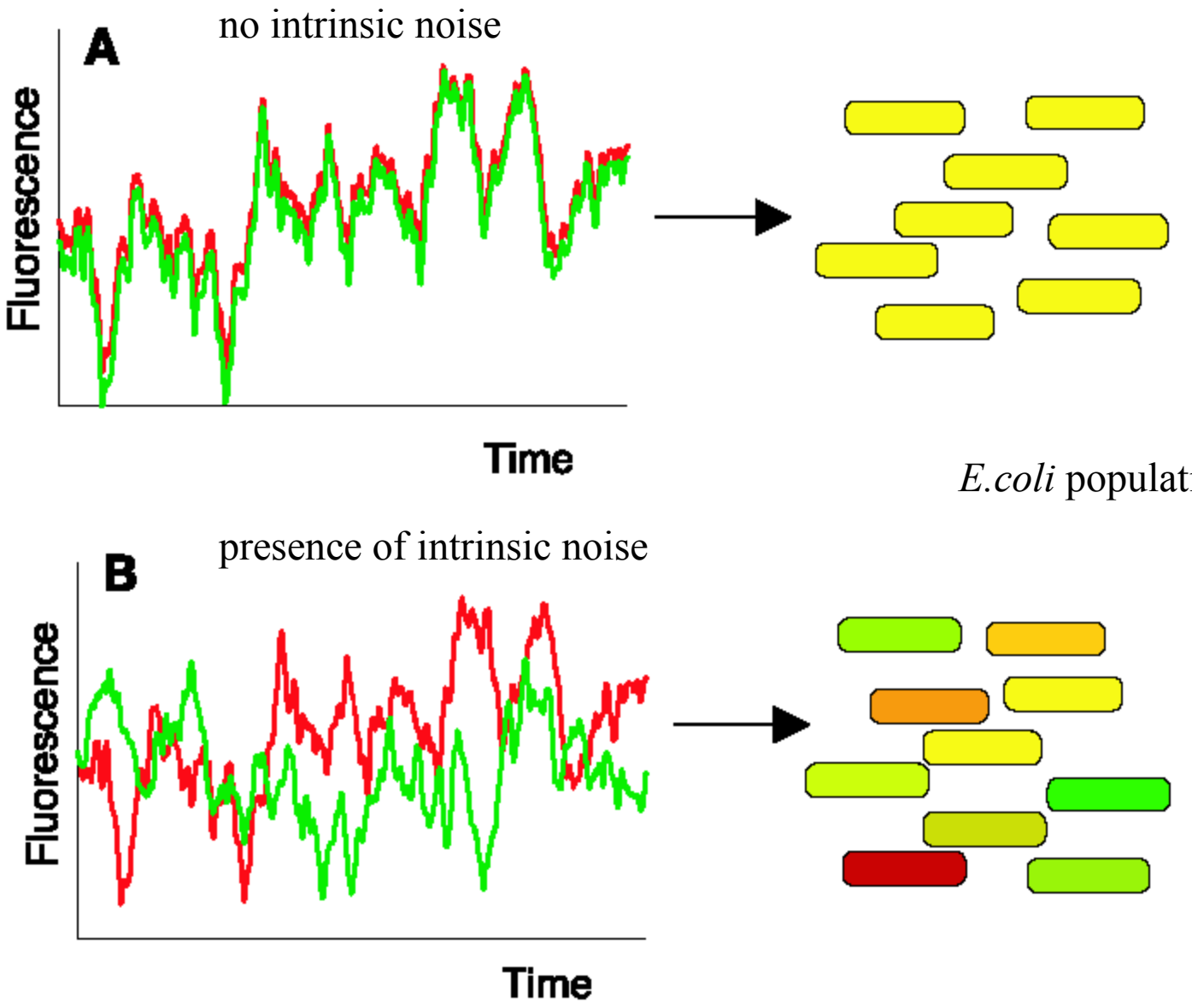
Intrinsic noise:= Difference in gene expression that arises between two identical copies of a gene expressed under precisely the same conditions



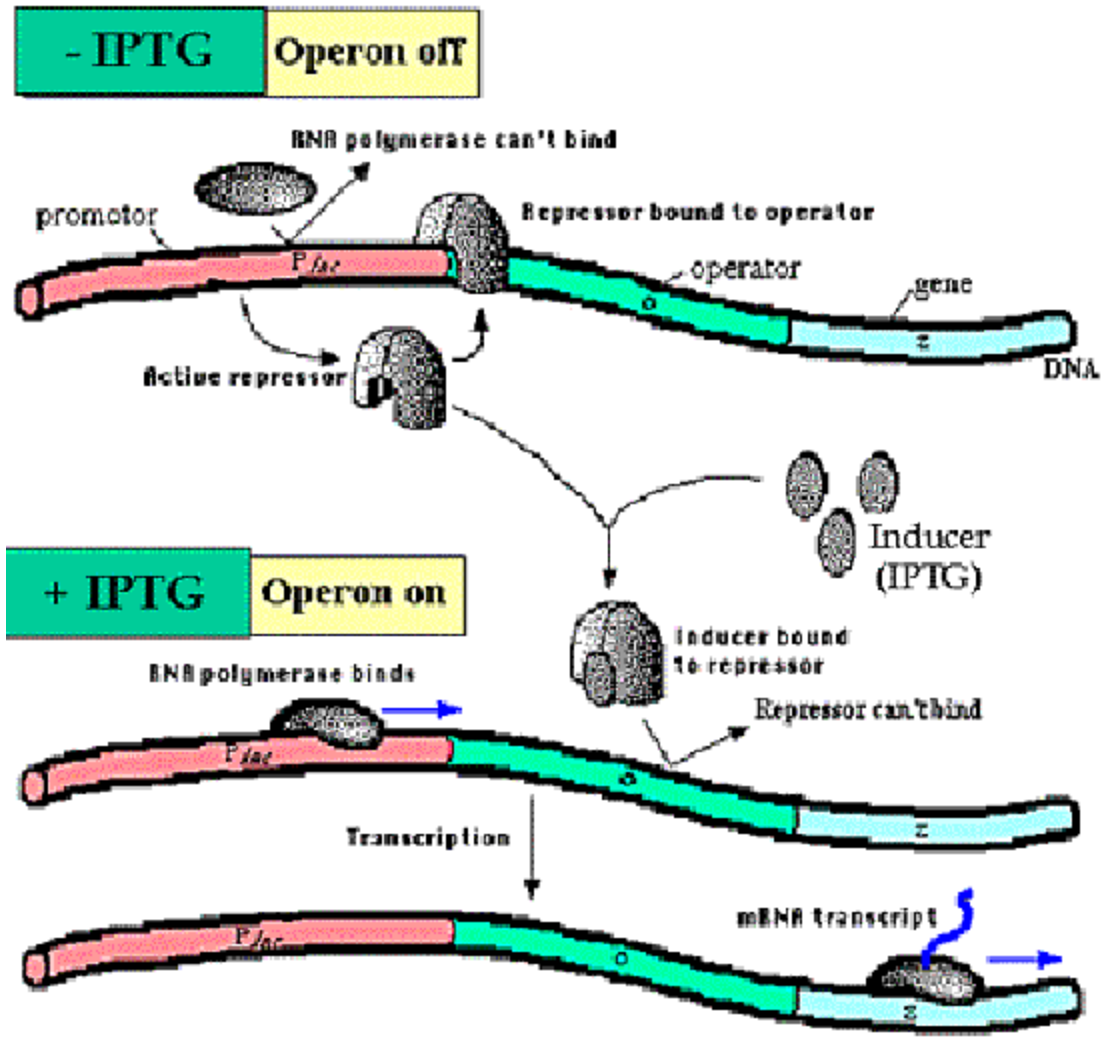
Intrinsic noise:= Difference in gene expression that arises between two identical copies of a gene expressed under precisely the same conditions



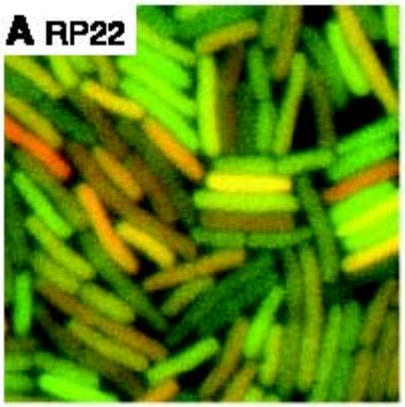
Two virtually equivalent Lac-repressible GFP reporter genes inserted in the *E.coli* chromosome on opposite sites and roughly equivalent to the origin of replication


no intrinsic noise
A
Fluorescence
Time
E.coli populations
presence of intrinsic noise
B
Fluorescence
Time

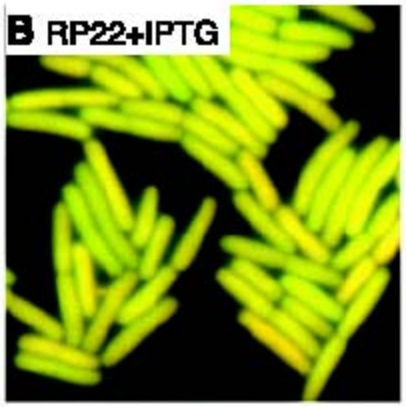
recall



Induction of the *lac* Operon

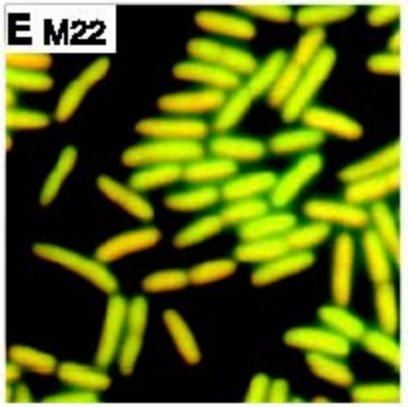


Promotors repressed by wild-type repressor (lacI) gene
(-IPTG operon OFF) low transcription, high noise

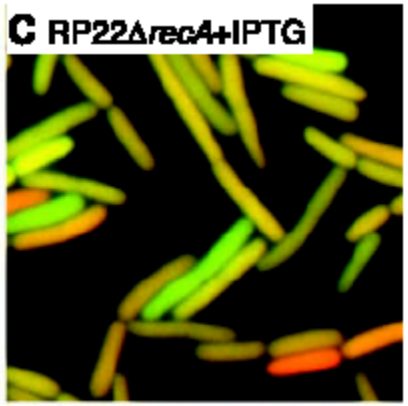


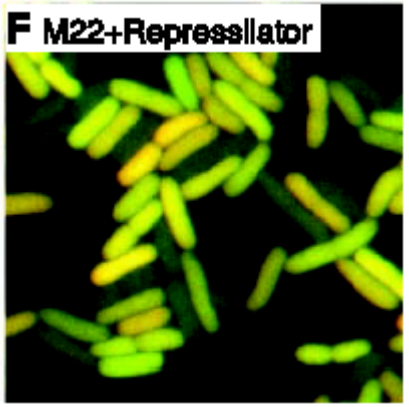
Presence of inducer (+IPTG operon ON)
high transcription, low noise

(lacI$^-$ cells)
high transcription, low noise

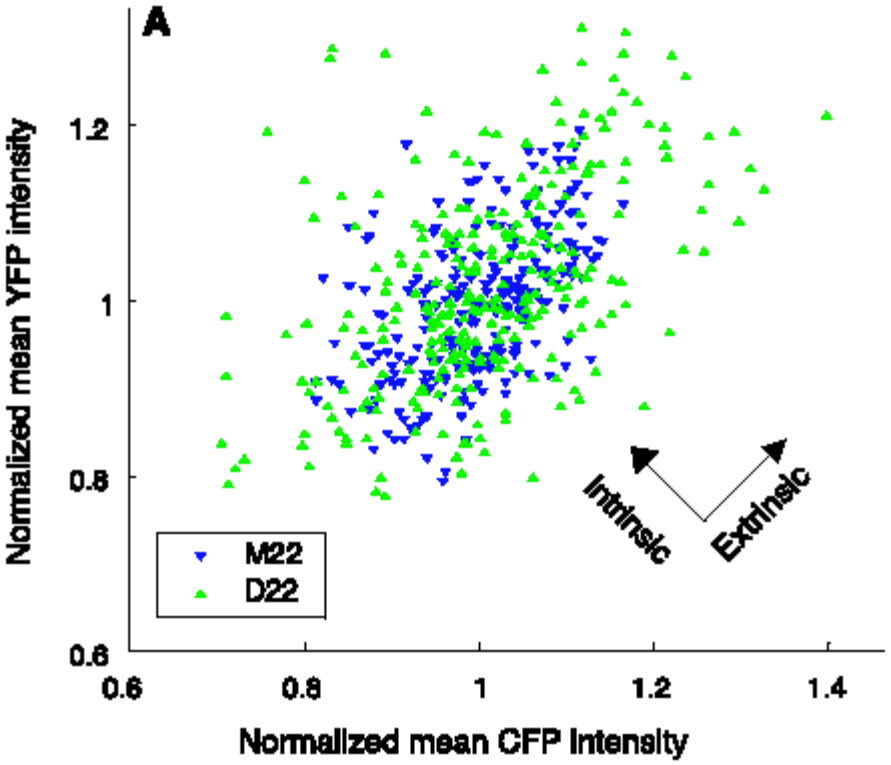


modified genetic
background
noisy





oscillating
expression
also noisy

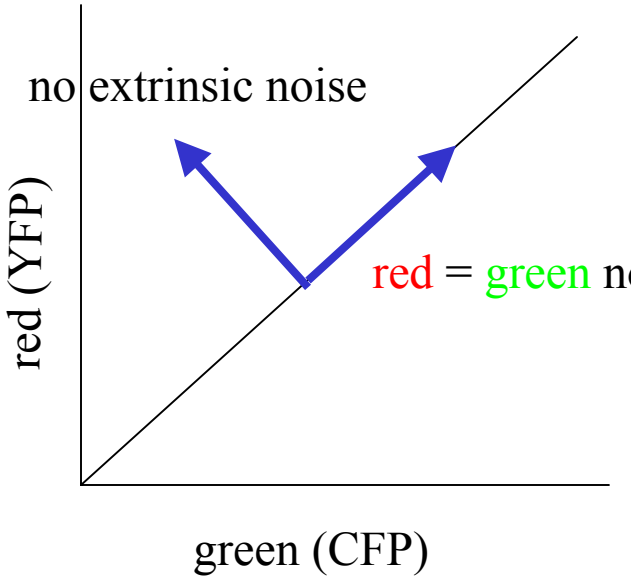


$$\text{noise} = \frac{\text{variance } (\sigma^2)}{\text{mean}^2} \; ;$$

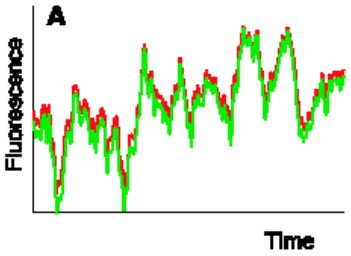
$$\text{noise}^2_{\text{total}}\,(\xi) = \text{noise}^2_{\text{intrinsic}} + \text{noise}^2_{\text{extrinsic}}$$

→ different to previous definition $n_2 = \dfrac{\sigma^2}{\langle n \rangle}$

$$\xi^2_{\text{total}} = \xi^2_{\text{intrinsic}} + \xi^2_{\text{extrinsic}} \; ?$$



red = green no intrinsic noise

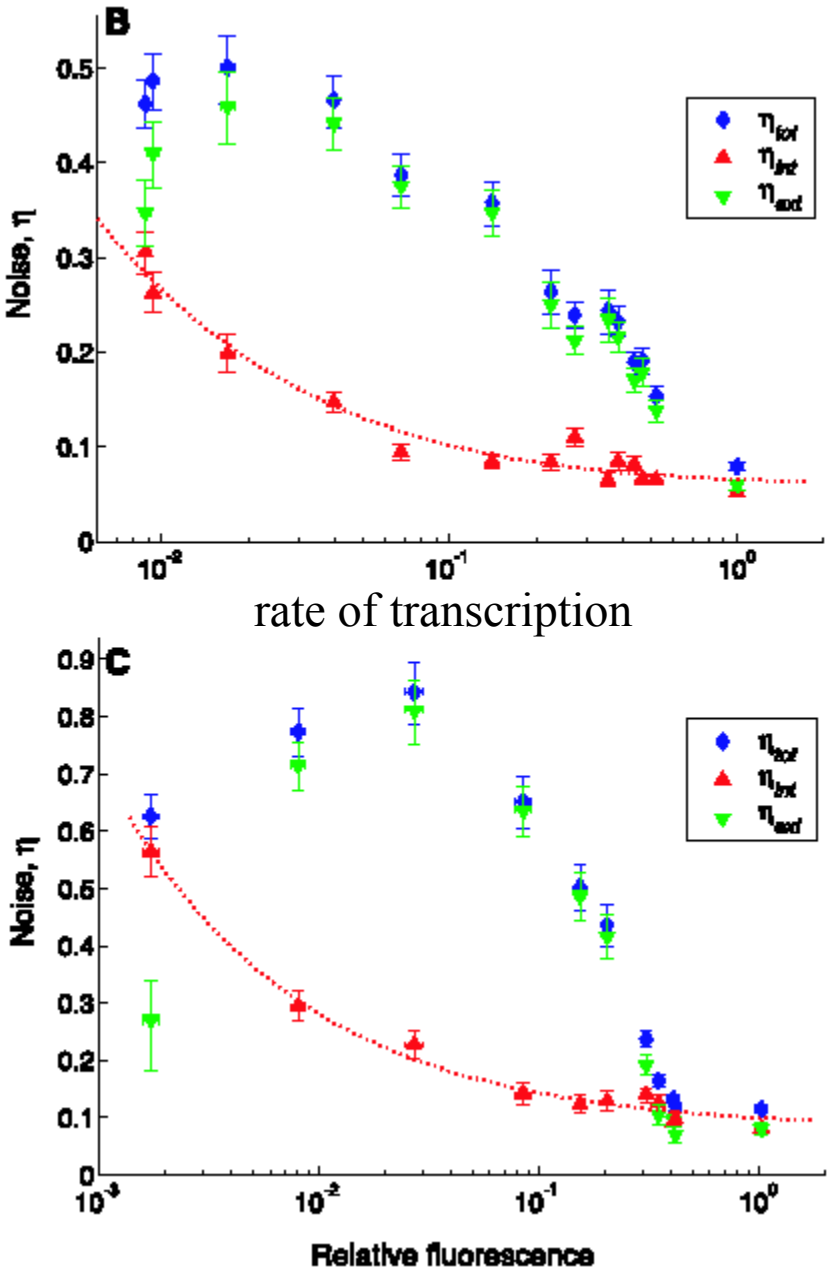


since $\frac{1}{N}\sum_{k=1}^{N} P_k^m \approx \int d\mathbf{E} d\mathbf{I} P^m(\mathbf{E},\mathbf{I}) p(\mathbf{EI}) = \int d\mathbf{E} p(\mathbf{E}) \int d\mathbf{I} P^m(\mathbf{E},\mathbf{I}) p(\mathbf{I}|\mathbf{E})$

$$= \int d\mathbf{E} p(\mathbf{E}) \langle P^m(\mathbf{E}) \rangle = \overline{\langle P^m \rangle}$$

then $\xi^2_{\text{total}} = \dfrac{\overline{\langle P^2 \rangle} - (\overline{\langle P}}{(\overline{\langle P \rangle})^2} = \dfrac{\overline{\langle P^2 \rangle - \langle P \rangle^2}}{(\overline{\langle P \rangle})^2} + \dfrac{\overline{\langle P \rangle^2} - (\overline{\langle P \rangle})^2}{(\overline{\langle P \rangle})^2}$

$$\equiv \xi^2_{\text{int}} + \xi^2_{\text{ext}}$$



intrinsic noise decreases with rate of transcription (transcription in these experiments does have an effect on noise!)

extrinsic noise peaks at intermediate levels (fluctuations in Lac repressor proteins. At high or low IPTG concentrations fluctuations are buffered by excess IPTG or excess LacI, respectively)

A glimpse on Langevin equations

$$\frac{d[mRNA]}{dt} = k_R - \delta_R[mRNA] + \xi_R$$

$$\frac{d[P]}{dt} = k_P[mRNA] - \delta_P[P] + \xi_P$$

ξ_R, ξ_P added stochastic variables.

This equations are fully specified when the probability distributions for the stochastic variables are also given.

Valid to describe an intermediate situation where fluctuations are important even though the number of particles is big enough.

Conclusions

- Phenotypic noise in a population as a consequence of protein concentration fluctuations.

-Translation and transcription leads to a control of fluctuations in protein concentration. Translation amplifies transcriptional noise.

- Some genes might have been naturally selected to have inefficient translational rates (a small rate of proteins per transcript) to avoid these fluctuations and thus avoid noise.

- In some circumstances noise can be highly desiderable as a means of creating nongenetic individuality in a population. In some other circumstances noise must be reduced (by means for instance of redundacy or negative feedback).

-Intrinsic and extrinsic sources of noise can be discriminated and measured.

-Theory + experiments + simulations a valid combined tool for biological discovery !!

References

Experiments

- Regulation of noise in the expression of a single gene
  E. M. Ozbudak et al, Nat. Gen, **31**, 69, 2002
- Stochastic Gene Expression in a single cell
  M. B. Elowitz et al, Science, **297**, 1183, 2002

Theory

- Intrinsic noise in gene regulatory networks
  M. Thattai & A. v. Oudenaarden, PNAS, **98**, 8614, 2001
- Intrinsic and extrinsic contributions to stochasticity in gene expression
  P. S. Swain et al, PNAS, **99**, 12795, 2002
- Stochastic Processes in Physics and Chemistry,
  N. G. van Kampen (North Holland, Amsterdam 1992)

Simulations

- Exact stochastic simulation of coupled chemical reactions
  D. T. Gillespie, J. Phys. Chem., **81**, 2340, 1977