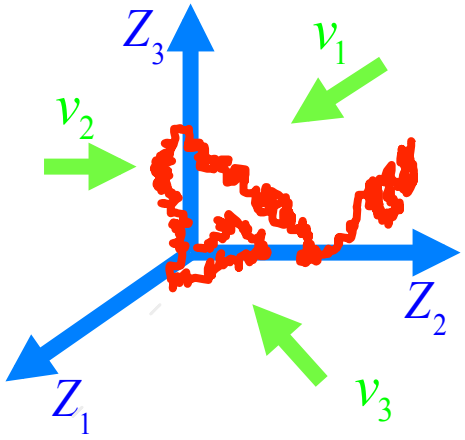
A RANDOM WALK THROUGH ANALYSIS, NETWORKS AND BIOLOGY

Lecture 3

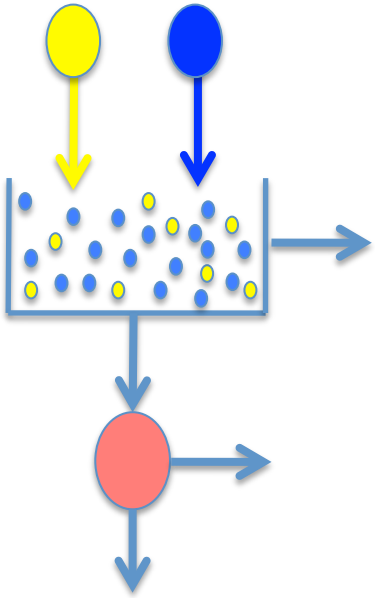


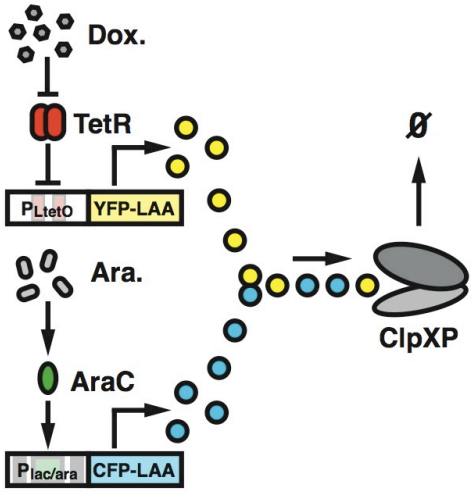
Ruth J Williams
G C Steward Visiting Fellow
Gonville and Caius College
April 2010

CONNECTIONS

- Brownian motion and analysis
- Reflecting Brownian motion and queueing networks
- Queues and biology

Queues and Biology





Ruth J. Williams
University of California, San Diego

Joint work with Natalie Cookson, Will Mather, Tal Danino, Octavio Mondragon-Palomino, Jeff Hasty, Lev Tsimring

Coupled Enzymatic Processing

Goal: To investigate the effect that processing by a common enzyme can have on correlations between numbers of proteins of different species

R A Fisher, 1890-1962

Motivation

- Measurements in gene networks (e.g., microarray data) often show correlations in levels of protein expression
- Typical explanations involve direct sources of coupling such as correlated transcription and protein-protein interactions
- Indirect coupling has not received much attention

Possible source of indirect coupling: enzymatic processing of an abundance of target molecules by a limited number of processing machines

Example

- Two uncoupled proteins X_1 and X_2

  -Processed downstream by a common enzyme *E*
- Two scenarios:

  - if *E* is abundant, X_1 and X_2 remain uncoupled

  - if *E* is limited, correlations between X_1 and X_2 appear

THEORY

Stochastic Model

Biochemical reaction network: protein species X_1, X_2

$$D_i \xrightarrow{\lambda_i} D_i + X_i \qquad \text{(production)}$$

$$X_i + E \xrightarrow{\eta} X_i E \qquad \text{(binding of enzyme)}$$

$$X_i E \xrightarrow{\mu} P_i + E \qquad \text{(degradation)}$$

$$X_i \xrightarrow{\gamma} \varnothing \quad \text{and} \quad X_i E \xrightarrow{\gamma} E \qquad \text{(dilution)}$$

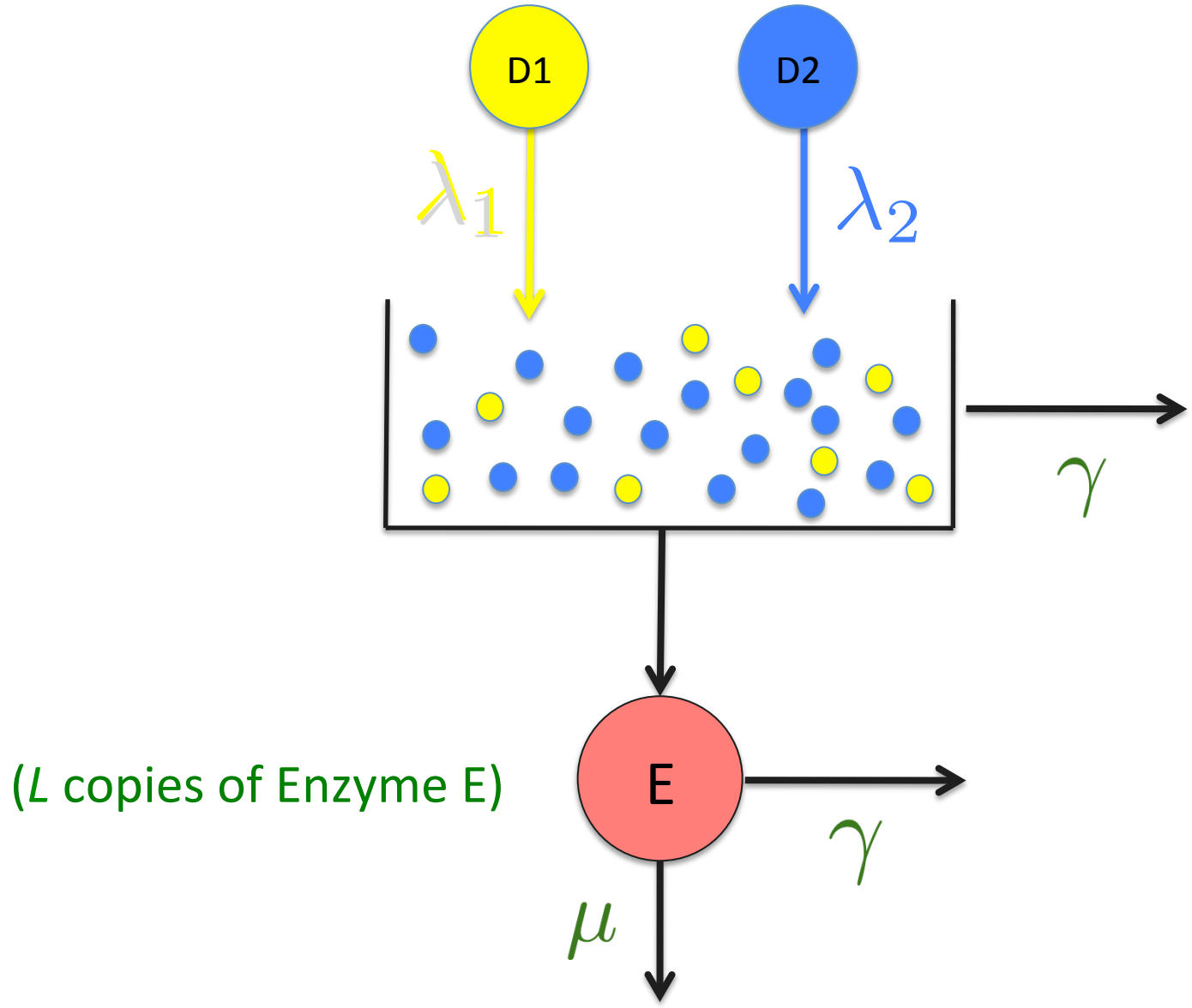
Assume: exponential reaction times and binding is instantaneous

Key stochastic processes (*i=1,2*):

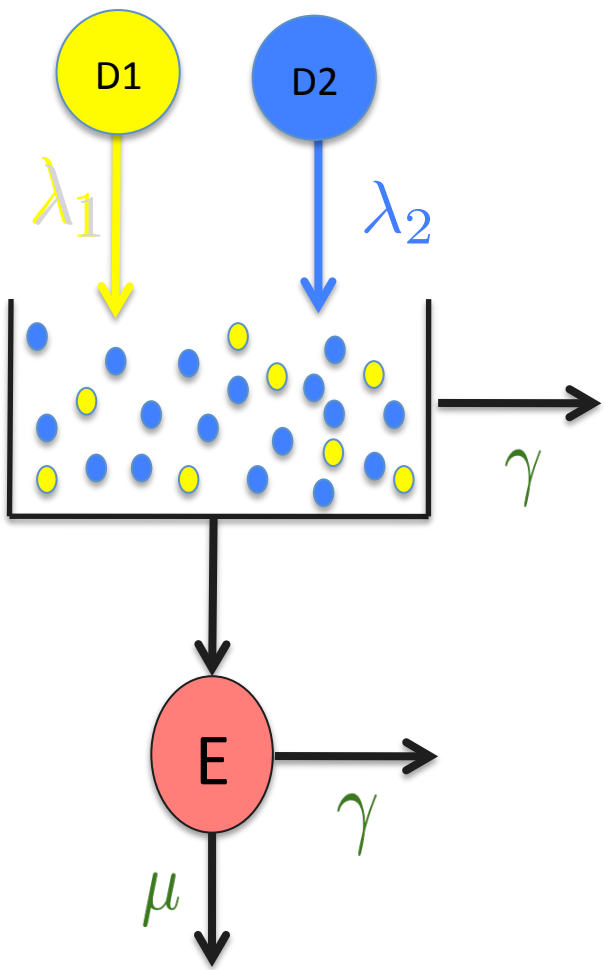
$Q_i(t)$ = total number of molecules of species i in the system at time t (includes free molecules and those being degraded)

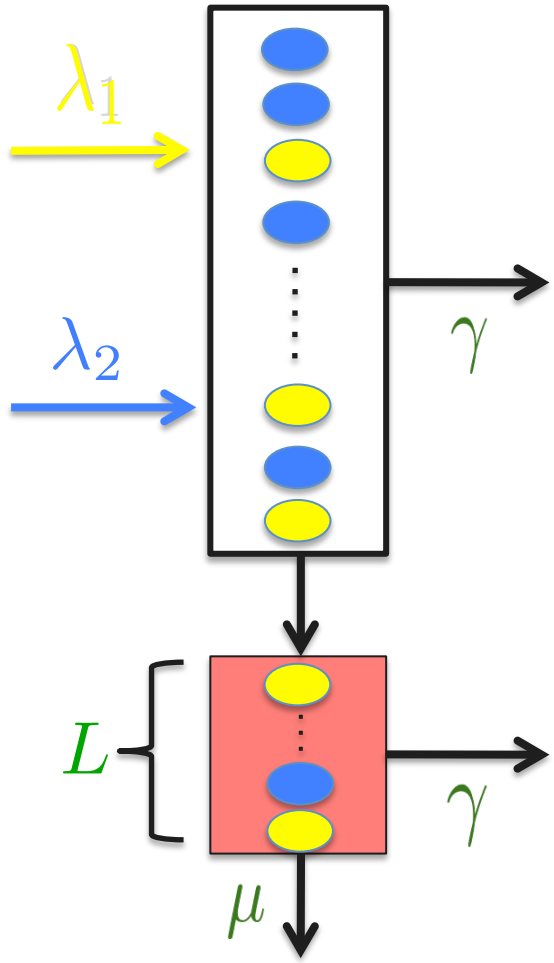
$N(t)$ = total number of protein molecules in system at time t

Stochastic Model



Multiclass Queue: Processing in Random Order + Reneging





$$\text{Total service rate} = \phi(n) = min(n, L)\mu + n\gamma$$

n = total number of protein molecules in system

Stationary Distribution (Quasireversible Queue)

Markovian state descriptor : ordered list of the types in the queue (incl. those being processed)

Theorem (Kelly '79): There is a unique stationary distribution for the "list" Markov process. The associated stationary distribution for the total number of molecules in the system, *N*, is:

$$P(N=n) = c\frac{\Lambda^n}{\prod_{\ell=1}^{n}\phi(\ell)}$$

and conditioned on *N=n*, the stationary distribution for the molecular count process *Q* is a binomial distribution with parameters $(n; p_1, p_2)$:

$$P(Q=(q_1,q_2)) = P(N=n)\frac{n!}{q_1!q_2!}p_1^{q_1}p_2^{q_2}$$

where $\Lambda = \sum_i \lambda_i, \quad p_i = \frac{\lambda_i}{\Lambda}, \quad i=1,2$

Moments and Correlations

Moments:

$$
\begin{aligned}
E[Q_i] &= p_i E[N] \\
E[Q_i^2] &= p_i(1-p_i)E[N] + p_i^2 E[N^2] \\
Var(Q_i) &= p_i^2(Var(N) - E[N]) + p_i E[N] \\
E[Q_i Q_j] &= p_i p_j (E[N^2] - E[N]) \quad \text{for } j \neq i
\end{aligned}
$$

Correlation ($j \neq i$):

$$
r_{ij} = \frac{\nu(N)}{(\nu(N) + p_i^{-1})^{\frac{1}{2}} (\nu(N) + p_j^{-1})^{\frac{1}{2}}}
$$

$$
\text{where} \quad \nu(N) = Var(N)/E[N]
$$

Moments for *N*

- Distribution: $P(N=n) = c\dfrac{\Lambda^n}{\prod_{\ell=1}^{n}\phi(\ell)}$

  where

$$\Lambda = \sum_i \lambda_i \qquad\qquad \phi(n) = \min(n, L)\mu + n\gamma$$

- Normalizing constant *c*:

$$c^{-1} = \sum_{n=0}^{L-1} \frac{\zeta^n}{n!} + \frac{\zeta^L}{L!} M(1, \beta+1, \delta)$$

$M(x,y,z) = \sum_{n=0}^{\infty} \frac{(x)_n z^n}{(y)_n n!}$

confluent hypergeometric function

$$\zeta = \frac{\Lambda}{\mu+\gamma}, \quad \beta = \frac{L\mu}{\gamma} + L, \quad \delta = \frac{\Lambda}{\gamma}$$

- Moment generating function:

$$E[e^{uN}] = c\left(\sum_{n=0}^{L-1} \frac{(e^u\zeta)^n}{n!} + \frac{(e^u\zeta)^L}{L!} M(1, \beta+1, e^u\delta)\right)$$

Moments and Correlations for Q (L=1)

$$
\begin{aligned}
E[Q_i] &= \frac{p_i \delta M(2, \beta+1, \delta)}{\beta M(1, \beta, \delta)}, \\
Var(Q_i) &= \frac{2p_i^2 \delta^2 M(3, \beta+2, \delta)}{\beta(\beta+1) M(1, \beta, \delta)} - \left(\frac{p_i \delta M(2, \beta+1, \delta)}{\beta M(1, \beta, \delta)} \right)^2 + \frac{p_i \delta M(2, \beta+1, \delta)}{\beta M(1, \beta, \delta)}, \\
r_{ij} &= \frac{h(\beta, \delta)}{(h(\beta, \delta) + p_i^{-1})^{1/2} (h(\beta, \delta) + p_j^{-1})^{1/2}},
\end{aligned}
$$

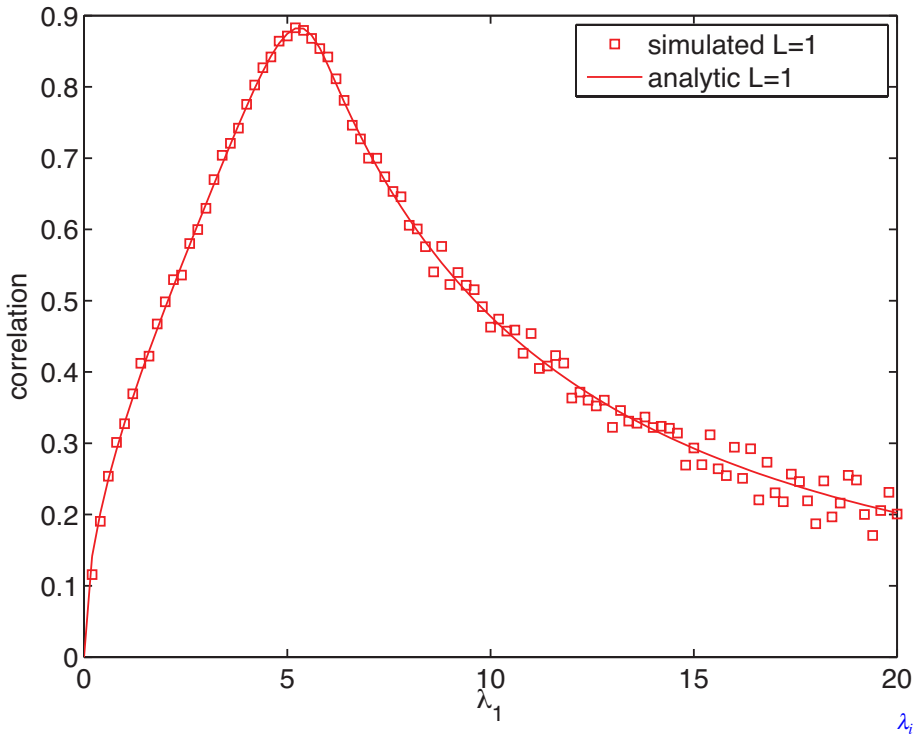
$\beta = (\mu/\gamma) + 1$, $\delta = \Lambda/\gamma$, $\Lambda = \sum_{i=1}^{m} \lambda_i$,

$$
f(\beta, \delta) = \frac{2\delta M(3, \beta+2, \delta)}{\beta+1} - \frac{\delta (M(2, \beta+1, \delta))^2}{\beta M(1, \beta, \delta)},
$$

$$
g(\beta, \delta) = M(2, \beta+1, \delta), \qquad h(\beta, \delta) = \frac{f(\beta, \delta)}{g(\beta, \delta)},
$$

Correlation Plots

- Plot of correlation as a function of λ_1



Simulation parameters:

$\lambda_2 = 5 \qquad \mu = 10$

$\eta = 10^8 \qquad \gamma = .01$

$$D_i \xrightarrow{\lambda_i} D_i + X_i \quad \text{(production)}$$

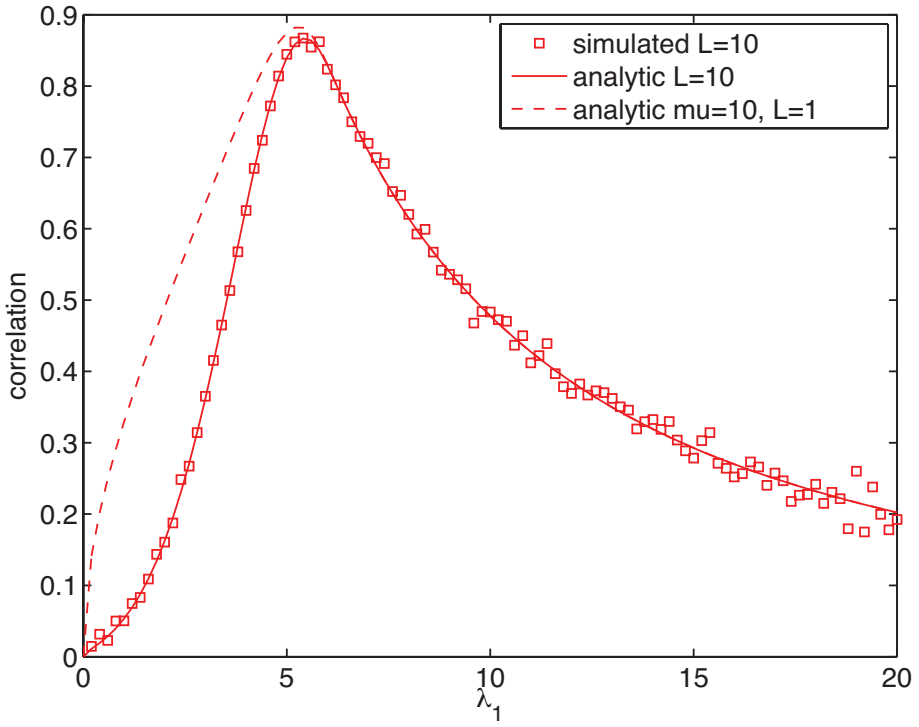
$$X_i + E \xrightarrow{\eta} X_i E \quad \text{(binding of enzyme)}$$

$$X_i E \xrightarrow{\mu} P_i + E \quad \text{(degradation)}$$

$$X_i \xrightarrow{\gamma} \varnothing \quad \text{and} \quad X_i E \xrightarrow{\gamma} E \quad \text{(dilution)}$$

Correlation Plots

- Plot of correlation as a function of λ_1



Simulation parameters:

$\lambda_2 = 5 \qquad \mu L = 10$

$\eta = 10^8 \qquad \gamma = .01$

Zero Dilution Limit (L=1)

- On letting $\gamma \to 0$ for $\rho = \Lambda / \mu < 1$ and $i \neq j$:

$$r_{ij} = \frac{1}{\left(1 + \frac{1}{p_i}\left(\frac{1}{\rho} - 1\right)\right)^{\frac{1}{2}} \left(1 + \frac{1}{p_j}\left(\frac{1}{\rho} - 1\right)\right)^{\frac{1}{2}}}$$

Here $p_i = \lambda_i / \Lambda$, $p_j = \lambda_j / \Lambda$

Uses asymptotics from Lucy Slater's book on Confluent Hypergeometric Functions

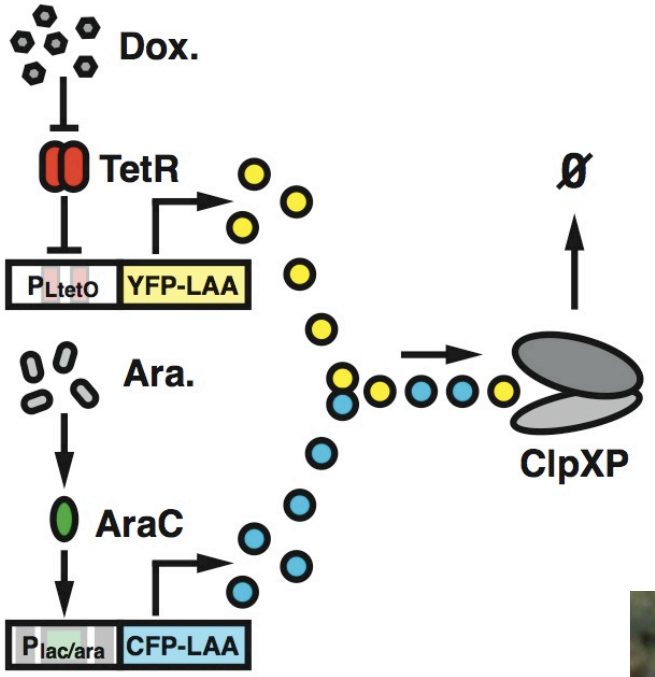
EXPERIMENT

Synthetic Genetic Network

- Enzymatic processing by *E. Coli* ClpXP machinery: targets LAA tagged proteins for degradation
- Two independently produced LAA tagged fluorescent proteins

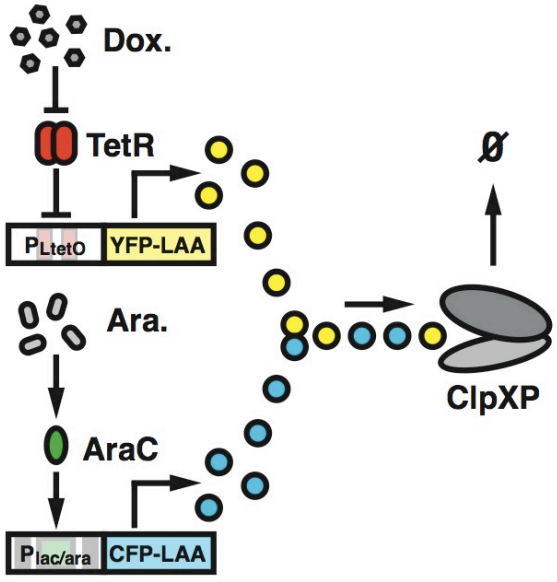
- Tet promoter driving YFP
  - Repressible by TetR
  - Tunable by Doxycycline

- Lac/Ara promoter driving CFP
  - Activated by AraC
  - Tunable by Arabinose

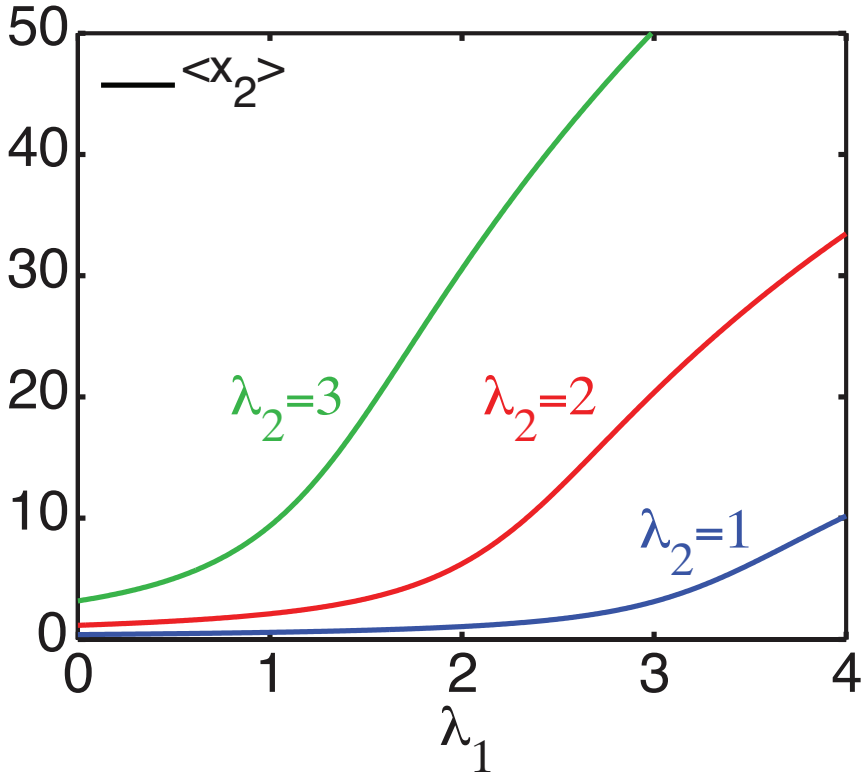


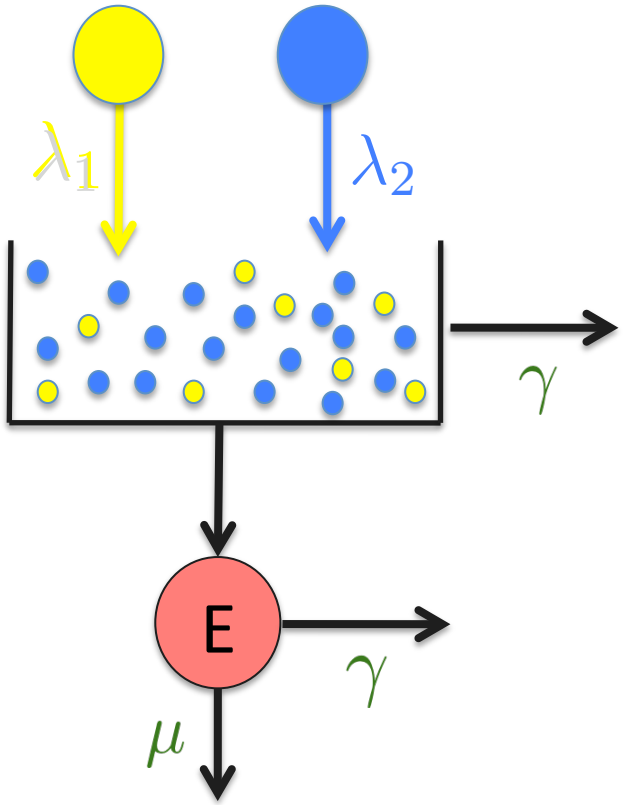
Synthetic Genetic Circuit

- Theory predicts that competition for the enzymatic processors will lead to correlations in YFP and CFP levels
- Investigated this with steady-state induction data
  - Used 2-color flow cytometry to look at YFP and CFP levels
  - Modulated inducer (doxycycline) of YFP holding inducer of CFP (arabinose) constant
  - Monitored the response of both



Effect of coupling on mean: theory





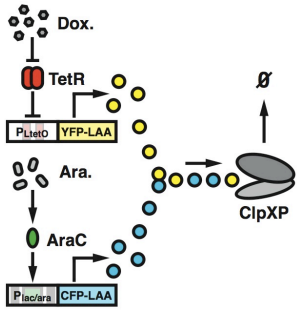
As λ_1 increases, means of X_1 and X_2 both increase and as approach "balance" point, $\lambda_1 + \lambda_2 = \mu$, have rapid increase in means.

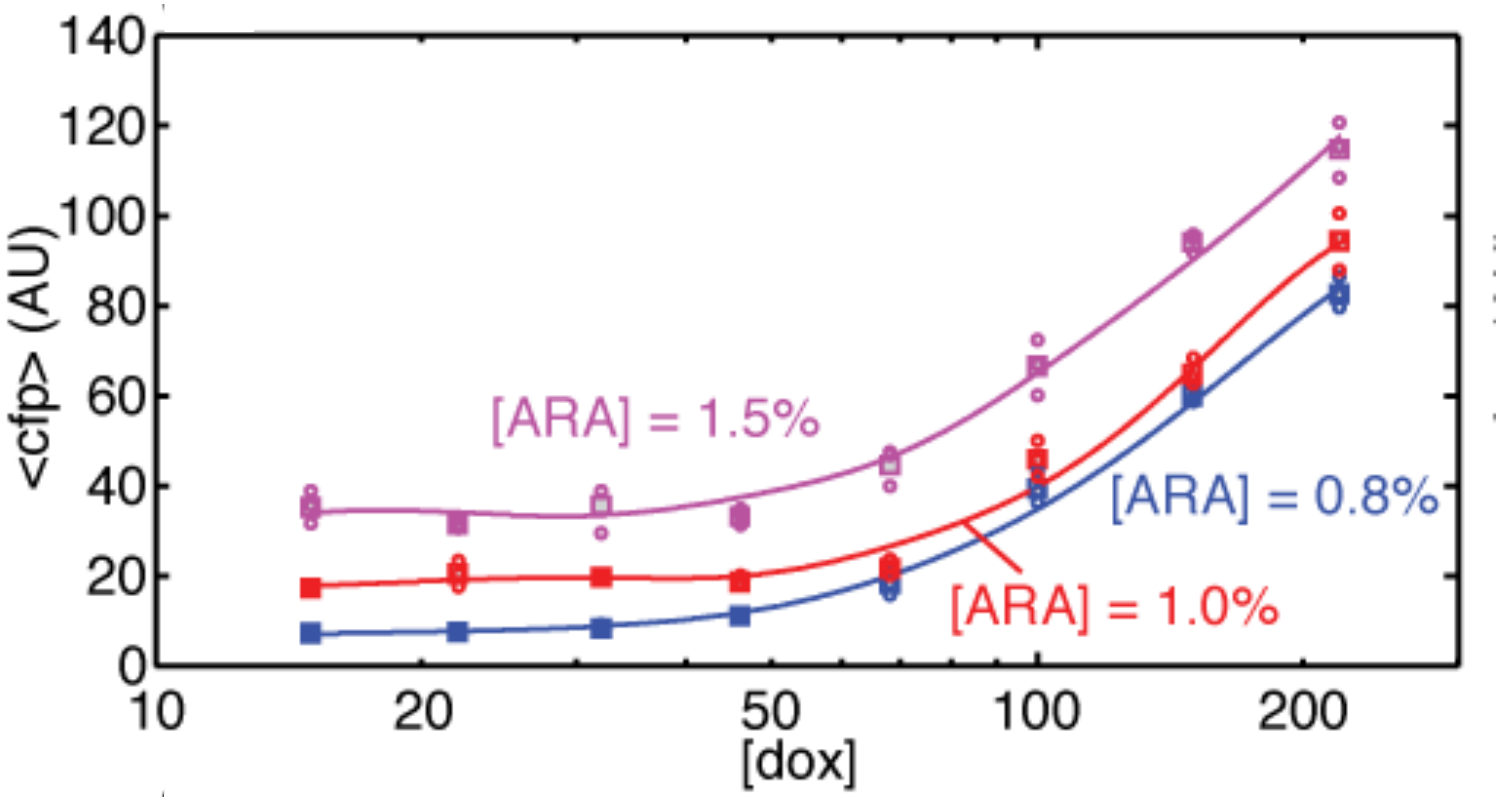
$\mu = 4 \qquad \gamma = 0.02$

$L = 1$

Effect of coupling on mean: experiment

Experiment: modulated doxycycline



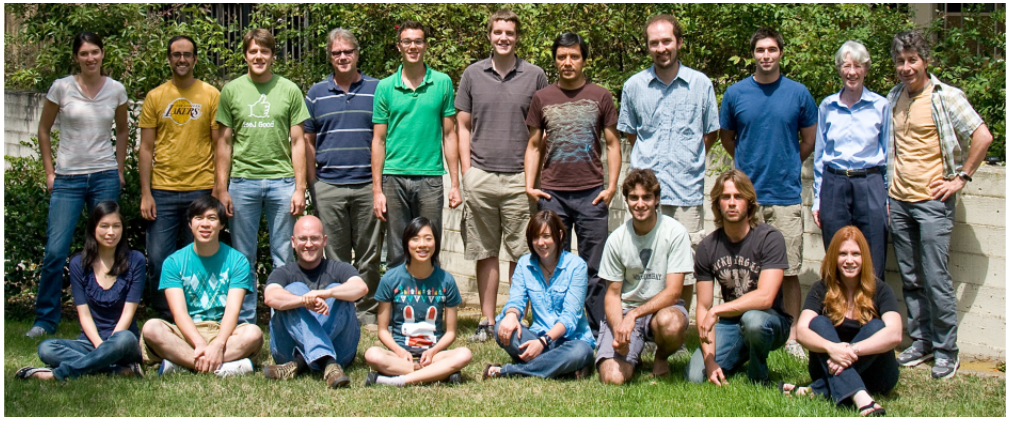


Conclusion

- Proposed a stochastic model for coupled enzymatic processing
- By mapping to a multiclass quasireversible queue, obtained stationary distribution
- Derived moments and correlations for steady-state levels of each protein
- Compared predictions with experimental results for a synthetic genetic network
- Coupled enzymatic processing produces correlated behavior that is strikingly similar to that produced by more direct sources of coupling

Acknowledgements

- Thanks to members of the UCSD Systems Biodynamics Lab.



-PIs: Jeff Hasty, Lev Tsimring

-Postdocs: Bridget Baumgartner, Natalie Cookson, Will Mather

-Graduate students: Colin Lam, Tal Danino, Mike Ferry, Martin Kolnik, Joyce Luke, Octavio Mondragon-Palomino, Sujata Nayak, Arthur Prindle, Ivan Razinkov, Jangir Selimkhanov, Phillip Samayoa, Brooks Taylor

- Thanks to NSF (IGMS), NIH, DOE, ONR for funding for this project

Getting Started in Biology

- Learn some biology and biological terminology
- Need language in which to converse with biologists
- Find a patient biologist (and perhaps a biophysicist) to talk with
- Need experiments to accompany theory to test model and predictions

THANK YOU