# Predicting Spatiotemporal Fluctuations of Gene Expression

# **Brian Munsky**

Chemical and Biological Engineering School of Biomedical Engineering Colorado State University

munsky@engr.colostate.edu

Presented at the 2015 q-bio Summer School Fort Collins (7/6); San Diego (7/22); Albuquerque (7/23)



# q-bio Summer School

## **Three Campuses:**

Albuquerque, NM (July 6-21) San Diego, CA (July 13-28) Fort Collins, CO (July 13-28)

## **Eight Focus Areas:**

Stochastic Gene Regulation Cancer Dynamics Complex Biological Dynamics Cell signaling Viral dynamics Biomolecular simulations Membrane biology Computational Synthetic Biology Experimental Synthetic Biology

## Visit us at online at: http://q-bio.org

#### The 2015 q-bio Summer School (Albuquerque, San Diego, Fort Collins)

Applications are now being accepted for the Ninth Annual q-bio Summer School. Applications will be due on **Monday, February 16 at 11:59pm (MST).** Submitted applications may be revised until that time. To apply now, please visit the application website: https://www.openconf.org/qbioss2015/openconf.php **Scholarships are available.** 

The 2015 q-bio Summer School will be held on different dates at three different campuses

- 1) July 6-21, 2015 at the Colorado State University in Fort Collins, CO 2) July 13-28, 2015 at the University of New Mexico in Albuquergue, NM
- 3) July 13-28, 2015 at the University of California in San Diego, CA.
  - July 13-26, 2013 at the oniversity of California in San Diego, (

#### School Overview:

The q-bio Summer School is an annual event intended to advance predictive modeling of cellular regulatory systems by exposing participants to a survey of work in quantitative biology and by providing in-depth instruction in selected techniques, with an emphasis on techniques useful for modeling cellular regulatory networks. Certain data analysis techniques and experimental methods will also be covered.

Lectures will be offered at **three campuses**. At the **San Diego campus**, the focus will be on synthetic biology. At the **Albuquerque and Fort Collins campuses**, the focus will be on different aspects of systems biology. Students will each work on a mentored project. Participants will attend daily core lectures, project-specific lectures, journal clubs, and computer and experimental labs. The summer school is designed for graduate students, postdocs, or anyone with a quantitative background who is new to modeling cellular regulatory systems/networks.

At the School students will attend 20-25 hours of core lectures, 20-25 hours of course-specific lectures, 10-15 hours of computational and experimental labs, and 10-15 hours of student presentations. There will also be 20-30 hours of mentored project work, which may include some simple experiments, theoretical developments and/or real data analyses.

#### The main topics of the 2015 summer school are:

Biomolecular Simulations (Albuquerque, NM), Cell Signaling (Albuquerque, NM), Membrane Biology (Albuquerque, NM), Viral Dynamics (Albuquerque, NM), Cancer Dynamics (Fort Collins, CO), Stochastic Gene Regulation (Fort Collins, CO), Complex Biological Dynamics (Fort Collins, CO), Experimental Synthetic Biology (San Diego, CA), and Computational Synthetic Biology (San Diego, CA)

#### **Organizers:**

S. Gnanakaran, *New Mexico Consortium, Los Alamos*, Jeff M. Hasty, *University of California, San Diego*, William S. Hlavacek, *New Mexico Consortium, Los Alamos*, Marek Kimmel, *Rice University, Houston*, Brian Munsky, *Colorado State University, Fort Collins*, Ashok Prasad, *Colorado State University, Fort Collins*, Douglas Shepherd, *University of Colorado, Denver*, Patrick Shipman, *Colorado State University, Fort Collins*, Mara P. Steinkamp, *University of New Mexico, Albuquerque*, Lev S. Tsimring, *University of California, San Diego* 

Cross CRASS CONTOR STUDIES CONTRACTOR

#### For inquiries about the scientific content at the summer school, please contact:

Brian Munsky (Fort Collins Campus): munsky@engr.colostate.edu Bill Hlavacek (Albuquerque Campus): wshlavacek@gmail.com Lev Tsimring (San Diego Campus): ltsimring@ucsd.edu

#### **Point of Contact:**

Shannan Yeager, q-bio Program Manager, syeager@newmexicoconsortium.org

For more information, please visit the school wiki at: http://q-bio.org/wiki/The\_Ninth\_q-bio\_Summer\_School

# Outline

## **Colorado State University**

## **1. Introduction - Information from transcript fluctuation**

- 2. Measuring and modeling single-cell and single-molecule responses
- 3. Case studies:
  - i. Kinase-activated gene transcription in budding yeast.
  - ii. Kinase-activated gene transcription in human cells.
- 4. Concluding remarks



## **Information in fluctuation**

## **Colorado State University**

Different systems (species, inputs, mechanisms, ...) may express genes at equal *average* levels.

Single-cell measurements may reveal hidden response differences.

Collective responses can exhibit distinctive "fluctuation fingerprints".



## Fluctuations may indicate gene regulation mechanisms

#### **Colorado State University**

 Consider the bursting gene expression model:



 Compute the expression mean and variability as functions of all parameters.

$$f_{on} = \frac{k_{ON}}{k_{ON} + k_{OFF}}$$
$$\mu = f_{on} \frac{k_m}{g_m}$$
$$\frac{\sigma^2}{\mu} = 1 + \frac{(1 - f_{on})k_m}{k_{ON} + k_{OFF} + g_m}$$

Expression 'Noise' versus parameters



## Fluctuations may indicate gene regulation mechanisms

#### **Colorado State University**

 Consider the bursting gene expression model:



- Compute the expression mean and variability as functions of all parameters.
- Tuning k<sub>Off</sub> or k<sub>On</sub> can increase expression, but:
- **Tuning koff increases variability.**
- Tuning kon decreases variability.





- 1. Introduction Information from transcript fluctuation
- 2. MEASURING and modeling single-cell and single-molecule responses



## **Experimental tools for single-cell analyses**

## **Colorado State University**

## **Flow Cytometry**

 Measure expression with fluorescent proteins or antibody labels for thousands of cells per second.

## **Time Lapse Fluorescence Microscopy**

 Measure spatial and temporal properties of fluorescent protein responses.







#### **Colorado State University**

- Endogenous mRNA's can be labeled with single molecule
  Fluorescence *in situ* Hybridization
  (smFISH--Femino, 1998, Raj, 2008).
- Many probes (~50) are attached to endogenous mRNA.
- High signal-to-noise ratio enables single-molecule detection.



. . .

#### **Colorado State University**

- Endogenous mRNA's can be labeled with single molecule
  Fluorescence *in situ* Hybridization
  (smFISH--Femino, 1998, Raj, 2008).
- Many probes (~50) are attached to endogenous mRNA.
- High signal-to-noise ratio enables single-molecule detection.
- Spatial localization enable inter- and intra-nuclear detection.





## **Colorado State University**

- Endogenous mRNA's can be labeled with single molecule
  Fluorescence *in situ* Hybridization
  (smFISH--Femino, 1998, Raj, 2008).
- Many probes (~50) are attached to endogenous mRNA.
- High signal-to-noise ratio enables single-molecule detection.
- Spatial localization enable inter- and intra-nuclear detection.
- Fast time resolution (1-2 min).



## **Colorado State University**

- Endogenous mRNA's can be labeled with single molecule
  Fluorescence *in situ* Hybridization
  (smFISH--Femino, 1998, Raj, 2008).
- Many probes (~50) are attached to endogenous mRNA.
- High signal-to-noise ratio enables single-molecule detection.
- Spatial localization enable inter- and intra-nuclear detection.
- Fast time resolution (1-2 min).

#### Statistics are repeatable and therefore predictable!



## **Colorado State University**



STI1 mRNA in Saccharomyces cerevisiae (budding yeast) -G. Neuert (VU)



Ysr35 sRNA in *Yersinia Pseudotuberculosis* (339nt) -D. Shepherd (CU Denver)



Traf6 mRNA in THP1 cells -D. Shepherd (CU Denver)



c-Fos mRNA (green) and p-p38 kinase (red) in U2OS cells -A. Senecal (CNRS)

smFISH has been applied to many different RNA in many different organisms



- 1. Introduction Information from transcript fluctuation
- 2. Measuring and MODELING single-cell and single-molecule responses



## A Markov description of single-cell gene regulation

- At any time, the state of the system is defined by its integer population vector:  $\mathbf{x} \in \mathbb{Z}^N$
- Reactions are transitions from one state to another:



## A Markov description of single-cell gene regulation

- At any time, the state of the system is defined by its integer population vector:  $\mathbf{x} \in \mathbb{Z}^N$
- Reactions are transitions from one state to another.
- These reactions are random, others could have occurred:



## A Markov description of single-cell gene regulation



## The finite state projection approach



## Inferring parameters from single-cell measurements.

**Colorado State University** 

Although single-cell reactions may be **Stochastic**, their statistics follow a **Deterministic** set of ODE's (*i.e., the CME*).



# **Outline**

- 1. Introduction Information from transcript fluctuation
- 2. Measuring and Modeling single-cell and single-molecule responses
- 3. Case Studies:
  - i. Predicting kinase-activated gene regulation dynamics in Saccharomyces cerevisiae (budding yeast)
  - ii. Quantitative modeling for c-Fos mRNA burst dynamics in U2OS cells.



(Osmotic shock response in yeast)

- 0.2M NaCl is added at t=0.
- Hog1 (red) activates in 1-2 min.
- ... and remains active for ~12 min.





0-20 minutes after 0.2M NaCl shock (Neuert, Munsky, et al, 2013)

#### (Osmotic shock response in yeast)

- 0.2M NaCl is added at t=0.
- Hog1 (red) activates in 1-2 min.
- ... and remains active for ~12 min.
- Stl1 mRNA appear at 4 min.
- ... and are gone by 25 min.





0-20 minutes after 0.2M NaCl shock (Neuert, Munsky, et al, 2013)

#### (Osmotic shock response in yeast)

## **Colorado State University**

- 0.2M NaCl is added at t=0.
- Hog1 (red) activates in 1-2 min.
- ... and remains active for ~12 min.
- Stl1 mRNA appear at 4 min.
- ... and are gone by 25 min.
- Stl1-GFP appear at ~30 min.







0-60 minutes after 0.2M NaCl shock (Neuert, Munsky, et al, 2013)

#### (Osmotic shock response in yeast)

#### **Colorado State University**



Our goal is to identify the mechanisms and parameters of **STL1** 



## **Possible model structures:**

## The Hog1 kinase ( ) activates STL1, but how?



## Each structure defines a hidden Markov Model

#### **Colorado State University**



**HIDDEN:** N ={2,3,...} possible gene states

#### OBSERVABLE: Integer number of mRNA

State-transition rates may vary in time, with experimental conditions, and/or with genetic mutations.

 $k_{ij} = k_{ij}(\text{Hog1}) = k_{ij}(t)$ 



# **Evaluating model structures of varying complexity**

#### Colorado State University

We fit different 2-, 3-, 4- and 5- state model structures to wild-type data at 0.4M osmotic shock.

More states (and parameters) yield better fits,...



# **Evaluating model structures of varying complexity**

- We fit different 2-, 3-, 4- and 5- state model structures to wild-type data at 0.4M osmotic shock.
- More states (and parameters) yield better fits,...
- but they also give rise to greater uncertainty.



# **Evaluating model structures of varying complexity**



## Fits and predictions for STL1 regulation



# The model can capture and predict WT mRNA dynamics for STL1



## What about other genes?

# The model can capture and predict WT mRNA dynamics for STL1, CTT1 and HSP12



# The model can capture and predict WT mRNA dynamics for STL1, CTT1 and HSP12

## It also captures STL1 mRNA dynamics in Wild Type, Hot1 over expression and Arp8 or Gcn5 deletion strains



# What about new combinations of different genes and mutant strains?

Munsky Slide: 33



:34

## Fitting and Predicting the Probability of ON Cells



Munsky Slide: 35

# **Outline**

- 1. Introduction Information from transcript fluctuation
- 2. Measuring and Modeling single-cell and single-molecule responses
- 3. Case Studies:
  - i. Predicting kinase-activated gene regulation dynamics in Saccharomyces cerevisiae (budding yeast)
  - ii. Quantitative modeling for c-Fos mRNA burst dynamics in U2OS cells.



## A glance at the activation of c-Fos



## **Dynamics of MAPK Signal Induction**

## **Colorado State University**

## Erk and p38 phosphorylation following stimulus



## Measuring c-Fos Activity at the Single Transcript Level



## **Signaling Affects Transcription Site Activation**

#### **Colorado State University**



The number of ATS's varies randomly from cell to cell. The *average* number of ATS's tracks MAPK induction dynamics.





## Signaling Affects Number of Mature mRNA.

#### **Colorado State University**

2

3

p38-P

4

5

1.0

0.8

0.6

0.4

0.2

A

Number of

Average

The *average* number of ATS's tracks MAPK induction dynamics.

The *average* number of mature mRNA tracks MAPK induction dynamics.



## Signaling does not affect Nascent mRNA numbers!

#### **Colorado State University**

The *average* number of ATS's tracks MAPK induction dynamics.

The average number of nascent mRNA per ATS is independent of MAPK.

The *average* number of mature mRNA tracks MAPK induction dynamics.



## **Models for c-Fos burst behavior**



## **Models for c-Fos burst behavior**



## **Trajectories from Alternate c-Fos Burst Models**

## **Colorado State University**



\*All models tuned to produce an average of 100 mature mRNA at equilibrium. Representative trajectories from Stochastic Simulation Algorithm (Gillespie, 1976).



## **Distributions from Alternate c-Fos Burst Models**



## Quantitative fits to signal-activated transcription dynamics

**Colorado State University** 

Burst saturation model accurately captures c-Fos dynamics.



## **Transcription Factor Modification of Burst Dynamics**

#### **Colorado State University**



The same model captures all TF activators with only one parameter change.

# **Outline**

- 1. Introduction Information from transcript fluctuation
- 2. Measuring and Modeling single-cell and single-molecule responses
- 3. Case Studies:
  - i. Predicting kinase-activated gene regulation dynamics in Saccharomyces cerevisiae (budding yeast)
  - ii. Quantitative modeling for c-Fos mRNA burst dynamics in U2OS cells.
- 4. Conclusions



**Colorado State University** 

Stochastic, temporal, and spatial fluctuations affect single-cell dynamics These complicate modeling and disrupt transcription control.

#### **But statistics of these fluctuations follow deterministic rules:**

Cells exhibit distinct, measurable `fluctuation fingerprints'. Can reveal subtle gene regulation mechanisms & parameters. Can be predicted with high accuracy and precision.

**Uncertainty Quantification** reveals when models are too simple, too complex, or just right (i.e., the Goldilocks Model).

We have identified **predictive quantitative models** of transcription regulation for many natural and synthetic genes in several organisms.

Prediction is the first step toward design, optimization and control.



## **References:**

**Colorado State University** 

...wherein dynamic stimuli and single-cell measurements reveal gene regulation mechanisms

- \*Munsky, \*Neuert, van Oudenaarden, Using Gene Expression Noise to Understand Gene Regulation, *Science*, 336:6078, 183–187, 2012.
- \*Neuert, \*Munsky, et al, Systematic Identification of Signal-Activated Stochastic Gene Regulation, Science, 339:6119, 584-587, 2013.
- Senecal, Munsky, et al, Transcription Factors Modulate c-Fos Transcriptional Bursts, Cell Reports, 8:1, 75-83, 2014.



# **Acknowledgments**

#### **Colorado State University**

#### **Hog Signaling:**

Gregor Neuert, Vanderbilt Alexander van Oudenaarden, Hubrecht Rui Zhen Tan, MIT Leonid Teytelman, MIT Mustafa Khammash, ETH

#### **Modules for Synthetic Biology:**

Chunbo Lou, Chris Voigt, MIT Brynne Stanton, MIT Ying-Ja Chen, MIT

#### **Activation of c-Fos**

Adrien Senecal, Albert Einstein Xavier Darzacq, UC Berkeley Florian Mueller, Institut Pasteur Florence Proux, ENS Nathalie Ly, ENS Floriane Braye, ENS Christophe Zimmer, Institut Pasteur

#### **sRNA** Dynamics:

Douglas Shepherd, CU Denver James Werner, LANL Elizabeth Hong-Geller, LANL Nan Li, LANL Sofiya N. Michva-Viteva, LANL

#### The PAP Switch:

Brooke Trinh, UCSB David Low, UCSB Mustafa Khammash, ETH

#### **Collaborators on Similar Projects:**

Kumkum Ganguly, LANL Babetta Marrone, LANL Taraka Dale, LANL Team at CSU: Postdocs: Vijay Gupta Graduate Students: Zachary Fox Undergrads: Lucas Suazo Michael May

Graduate Research Positions available in predictive modeling and experiment design.

#### Funding: NIH, LANL-LDRD, CSU Start-up funds

munsky@engr.colostate.edu

# q-bio Summer School

## **Three Campuses:**

Albuquerque, NM (July 6-21) San Diego, CA (July 13-28) Fort Collins, CO (July 13-28)

## **Eight Focus Areas:**

Stochastic Gene Regulation Cancer Dynamics Complex Biological Dynamics Cell signaling Viral dynamics Biomolecular simulations Membrane biology Computational Synthetic Biology Experimental Synthetic Biology

## Visit us at online at: http://q-bio.org

#### The 2015 q-bio Summer School (Albuquerque, San Diego, Fort Collins)

Applications are now being accepted for the Ninth Annual q-bio Summer School. Applications will be due on **Monday, February 16 at 11:59pm (MST).** Submitted applications may be revised until that time. To apply now, please visit the application website: https://www.openconf.org/qbioss2015/openconf.php **Scholarships are available.** 

The 2015 q-bio Summer School will be held on different dates at three different campuses

- 1) July 6-21, 2015 at the Colorado State University in Fort Collins, CO 2) July 13-28, 2015 at the University of New Mexico in Albuquergue, NM
- 3) July 13-28, 2015 at the University of California in San Diego, CA.
  - July 13-26, 2013 at the oniversity of California in San Diego, (

#### School Overview:

The q-bio Summer School is an annual event intended to advance predictive modeling of cellular regulatory systems by exposing participants to a survey of work in quantitative biology and by providing in-depth instruction in selected techniques, with an emphasis on techniques useful for modeling cellular regulatory networks. Certain data analysis techniques and experimental methods will also be covered.

Lectures will be offered at **three campuses**. At the **San Diego campus**, the focus will be on synthetic biology. At the **Albuquerque and Fort Collins campuses**, the focus will be on different aspects of systems biology. Students will each work on a mentored project. Participants will attend daily core lectures, project-specific lectures, journal clubs, and computer and experimental labs. The summer school is designed for graduate students, postdocs, or anyone with a quantitative background who is new to modeling cellular regulatory systems/networks.

At the School students will attend 20-25 hours of core lectures, 20-25 hours of course-specific lectures, 10-15 hours of computational and experimental labs, and 10-15 hours of student presentations. There will also be 20-30 hours of mentored project work, which may include some simple experiments, theoretical developments and/or real data analyses.

#### The main topics of the 2015 summer school are:

Biomolecular Simulations (Albuquerque, NM), Cell Signaling (Albuquerque, NM), Membrane Biology (Albuquerque, NM), Viral Dynamics (Albuquerque, NM), Cancer Dynamics (Fort Collins, CO), Stochastic Gene Regulation (Fort Collins, CO), Complex Biological Dynamics (Fort Collins, CO), Experimental Synthetic Biology (San Diego, CA), and Computational Synthetic Biology (San Diego, CA)

#### **Organizers:**

S. Gnanakaran, *New Mexico Consortium, Los Alamos*, Jeff M. Hasty, *University of California, San Diego*, William S. Hlavacek, *New Mexico Consortium, Los Alamos*, Marek Kimmel, *Rice University, Houston*, Brian Munsky, *Colorado State University, Fort Collins*, Ashok Prasad, *Colorado State University, Fort Collins*, Douglas Shepherd, *University of Colorado, Denver*, Patrick Shipman, *Colorado State University, Fort Collins*, Mara P. Steinkamp, *University of New Mexico, Albuquerque*, Lev S. Tsimring, *University of California, San Diego* 

Cross CRASS CONTOR STUDIES CONTRACTOR

#### For inquiries about the scientific content at the summer school, please contact:

Brian Munsky (Fort Collins Campus): munsky@engr.colostate.edu Bill Hlavacek (Albuquerque Campus): wshlavacek@gmail.com Lev Tsimring (San Diego Campus): ltsimring@ucsd.edu

#### **Point of Contact:**

Shannan Yeager, q-bio Program Manager, syeager@newmexicoconsortium.org

For more information, please visit the school wiki at: http://q-bio.org/wiki/The\_Ninth\_q-bio\_Summer\_School