Multi-modality in gene regulatory networks with slow promoter kinetics

M. Ali Al-Radhawi[1], Eduardo D. Sontag[2], and Domitilla Del Vecchio[1]

***Abstract*— We uncover the mechanism in which multiple modes emerge due to slow TF (transcription factor)-gene binding/unbinding with low molecular counts. We study a generic GRN and the associated master equation using time scale separation. We show analytically that in the limit of slow binding/unbinding the stationary distribution can be decomposed into a mixture of Poisson distributions. This offers a mathematical framework to explain phenomena such as non-genetic population heterogeneity and transcriptional bursting. The results are applied to examples including toggle switches and the pluripotency network.**

***Keywords* — Chemical Master Equation, Singular Perturbations, Gene Regulatory Networks.**

I. Motivation

Stochasticity in GRN's is a source of phenotypic variation among genetically identical (clonal) populations of cells or even organisms [1], and is considered to be one of the mechanisms facilitating cell differentiation and organism development [2]. This phenotypic variation may also confer a population advantage when facing fluctuating environments [3]. Stochasticity due to randomness in cellular components and transcriptional and translational processes have been thoroughly researched [4].

A common assumption in gene expression models is that TF-gene binding is significantly faster than the rate of protein production and decay. However, it has been proposed [4] that the emergence of new modes in stochastic systems in addition to those that arise from the deterministic model might be due to low gene copy numbers and to *slow* promoter kinetics, which means that process of binding and unbinding of transcription factors (TFs) to promoters is slow. Thus, the emergence of multi-modality may be due to the slow TF-gene binding and unbinding, especially in the case of eukaryotic cells, which involve a more sophisticated transcription machinery than prokaryotes. For example, the presence of nucleosomes makes binding sites less accessible to transcription factors and therefore TF-gene binding/unbinding is modulated by the stochastic process of chromatin opening [5], [4].

II. Results

In this work, we first introduce a formalism to model GRN's with arbitrary numbers of genes, based on continuous-time Markov chains, and we let the kinetic rate constants of the TF-gene binding and unbinding reactions approach zero and compute the stationary solution by applying the method of singular perturbations to the master equation.

Our main result is based on partitioning the state space into weakly-coupled ergodic classes which, in the limit of slow binding/unbinding, results in the reduction of the infinite-dimensional Markov chain into a finite-dimensional chain whose states correspond to ideal "gene states". In this limit, there exists non-negative weighting coefficients $\lambda_d, d = 0, .., L-1$ such that the marginal distribution $\pi(x)$ of the protein count vector $X(t)$ can be written as:

$$\pi(x) = \sum_{d=0}^{L-1} \lambda_d \pi_d(x) = \sum_{d=0}^{L-1} \lambda_d \prod_{i=1}^{N} \mathbf{P}_{a_{id}}(x_i),$$

where L is number of genetic states, $d = 0, .., L-1$ are the "gene states", a_{id} is production ratio, and $\mathbf{P}_a(x)$ is a Poission distribution with mean a.

Unlike deterministic systems, we show that the number of modes is independent of whether the TF binding is cooperative or no. However, we show that cooperative binding gives us extra degrees of freedom for tuning the weighting coefficient λ_d's in the above formula.

The proposed framework enables us to compute the number of modes, their locations, and their weights in the mixture. Hence, the proposed framework can be applied to GRNs to predict the different phenotypes that the network can exhibit with low gene copy numbers and slow promoter kinetics and can also be used for design.

III. Applications

The proposed framework is applied to a self-regulating gene, the classical toggle switch, coupled toggle switches and a pluriopotency network motif. We show that our analysis offers new insights into the behaviour of the network at slow gene binding/unbinding limit.

This work is supported by AFOSR grant number FA9550-14-1-0060.

[1]Department of Mechanical Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA. Email: {malirdwi,ddv}@mit.edu

[2]Department of Mathematics, Rutgers University, Piscataway, NJ 08854-8019, USA. Email: eduardo.sontag@rutgers.edu

References

[1] B. Munsky, G. Neuert, and A. van Oudenaarden, "Using gene expression noise to understand gene regulation," *Science*, vol. 336, no. 6078, pp. 183–187, 2012.

[2] J. X. Zhou and S. Huang, "Understanding gene circuits at cell-fate branch points for rational cell reprogramming," *Trends in Genetics*, vol. 27, no. 2, pp. 55–62, 2011.

[3] M. Stamatakis and N. V. Mantzaris, "Comparison of deterministic and stochastic models of the lac operon genetic network," *Biophysical journal*, vol. 96, no. 3, pp. 887–906, 2009.

[4] M. Kærn, T. C. Elston, W. J. Blake, and J. J. Collins, "Stochasticity in gene expression: from theories to phenotypes," *Nature Reviews Genetics*, vol. 6, no. 6, pp. 451–464, 2005.

[5] A. Paldi, "Stochastic gene expression during cell differentiation: order from disorder?" *Cellular and molecular life sciences*, vol. 60, no. 9, pp. 1775–1778, 2003.