Gene-gene cooperativity in small networks

Aleksandra M. Walczak[1] and Peter G. Wolynes[2]

Short Abstract — We infer a reduced qualitative description of interacting genes in small regulatory networks in terms of coupled binary variables. Treating both the protein and gene expression state variable stochastically and on equal footing, we propose a mapping, which connects the molecular level description of networks to the binary approach. We are able to scan phase space and determine when genes can be considered to be independent and when interactions between them cannot be neglected. We find an appropriately mapped Boolean description describes the full probabilities of gene expression states very well. We study the problem on the example of self-activating systems with multiple gene copies.

***Keywords* — stochastic gene regulation, maximum entropy models, cooperativity**

I. INTRODUCTION

Multiple copies of the same gene interacting with each other in one regulatory network can be encountered both in synthetic systems in the laboratory [1], as well as in living organisms, which multiply sections of their genomes during evolution [2]. Small numbers of molecules of a given type taking part in gene regulation are a recognized source of stochasticity in gene expression. Recently, increasing the number of copies of genes expressing proteins has been used to decrease noise in an experimental setup [1]. We discuss how the characteristics of gene-gene interactions and the resulting additional cooperativity between genes, the expression of which is mediated by one type of protein, are shaped by different sources of noise. Guided by the example of small networks with multiple gene copies, we propose a mapping between a reduced description in terms of binary variables and a full stochastic molecular description of regulatory gene-protein systems. We investigate the effects of noise arising from small protein numbers and slow gene expression state changes on the strength of gene-gene cooperativity.

II. THE MODEL

We study a group of toy systems, in which we have only one species of protein, but multiple gene copies are regulated by that type of protein. We consider examples, which have regulatory function and could be subunits of a larger network - small networks of self-activators and self-repressors producing and being regulated by one type of protein. The particular copies of the gene can be identical copies - that means we model them to have exactly the same chemical parameters; or they can be thought to be evolved versions of the gene - with different parameters. We treat both proteins and gene expression states stochastically and use a joint probability distribution, which describes the number of proteins in the system and the gene expression state to describe each gene [3].

Using a maximum entropy model we map the marginal gene expression state probability onto coupled binary variables in an effective gene field. The sign and magnitude of the effective gene field describe the given gene's preference to be in a particular expression state. The predicted expression state of the genes can be modified by protein mediated gene-gene interactions, which are quantified by the coupling constant. This approach allows us to describe the parameter regimes, in which effective gene expression units can be treated as independent and the parameter regimes, where genes cooperate to form an interacting steady state.

III. RESULTS

We find that for parameter regimes, in which all of the genes can individually maintain their own protein field and be expressed at an enhanced level, genes can be treated as independent units. However, interactions cannot be neglected when genes interact to reach a new steady state by sharing the protein field. In this case at least one of the genes alone could not sustain a protein field needed to the enhanced expression state, which is observed in the multiple gene system. In the limit of strongly non-equilibrium binding of transcription function proteins, the gene field clearly depends on the binding and unbinding rates of the gene, whereas the coupling constant is linked to the protein number.

We test the validity of the deduced Boolean approximation on a four gene system by comparing the marginal probabilities based on the maximum entropy model to probabilities from simulation studies. Two gene interactions account for a majority of interactions in the tested systems and are enough to reproduce the observed marginal probabilities.

[1]Princeton Center for Theoretical Physics, Princeton University, NJ 08540. E-mail: awalczak@princeton.edu

[2]Department of Physics, Department of Chemistry and Biochemistry, Center for Theoretical Biological Physics, UCSD, La Jolla, CA 92122 USA. E-mail: pwolynes@ucsd.edu

REFERENCES

[1] Süel G, Garcia-Ojalvo J, Liberman L, Elowitz MB, Nature (2006), **440**, 545-50

[2] Kellis M, Birren B, Lander E, Nature (2004), **428**, 617-24

[3] Sasai M and Wolynes PG, PNAS, (2003), **100(5)**, 2374 - 2379