Quantifying noise in general stochastic models of post-transcriptional regulation of gene expression

Argenis Arriojas[1], Sho Inaba[2] and Rahul Kulkarni[3]

***Short Abstract* — Gene expression is an stochastic process, and fluctuations in protein levels are often critical for generating phenotypic heterogeneity within a population of isogenic cells. There is thus considerable interest in quantifying how fluctuations (noise) in gene expression are impacted by cellular control mechanisms such as post-transcriptional regulation. In previous work, a general framework for *promoter-based regulation* has been developed [1], which leads to exact results for the moments of mRNA distributions. However, a similar framework for protein statistics in models with *post-transcriptional regulation* is currently lacking. In this work we develop an analytical framework that maps a general class of models of post-transcriptional regulation into models with promoter-based regulation, leading to exact analytical results for the moments of protein distributions. This mapping is based on the partitioning of Poisson arrivals (PPA) approach developed in recent work [2]. The proposed framework can be used to model complex schemes of post-transcriptional regulation and to evaluate its effects on variability in protein distributions.**

***Keywords* — Gene expression, post-transcriptional regulation, Markovian Arrival Process, stochastic modeling, promoter-based regulation, protein variability**

I. INTRODUCTION

THE stochastic nature of gene expression leads to fluctuations of mRNA and protein levels among isogenic cell populations. Modeling and quantifying this variability is important, since it has a significant impact on the resulting phenotypic variability of cell populations (e.g. emergence of bacterial persisters, drug-tolerant cancer cells). To address this, an analytical framework for mRNA distributions have been previously developed, in particular for general models of promoter-based transcriptional regulation [1]. In this framework, mRNA creation is represented by a Markovian Arrival Process (MAP) and obtaining exact results for moments only requires finding solutions for linear algebraic equations. However, a similar framework is currently lacking for determining protein statistics in general models of post-transcriptional regulation. Meanwhile, recent work has developed a novel approach that enables to obtain exact result for protein distributions in coarse-grained stochastic models of gene expression. This is achieved through the so called Partitioning of Poisson Arrivals mapping (PPA mapping) [2]. By using this approach, we show that a general class of models of post-transcriptional regulation can be mapped onto reduced models with promoter-based regulation. By combining the two previous approaches, we have developed a framework that can be used to derive exact analytical expressions for the moments of protein distributions for general models of post-transcriptional regulation.

[1,2]University of Massachusetts, Boston, Massachusetts USA.

[1]Email: ArriojasMaldonado001@umb.edu

[2]Email: Sho.Inaba001@umb.edu

[3]Department of Physics, University of Massachusetts, Boston, Massachusetts USA. Email: Rahul.Kulkarni@umb.edu

II. MAPPING TO REDUCED MODELS

A general class of models with post-transcriptional regulation can be mapped to a reduced model by means of a PPA mapping [2, 3]. The generating function for proteins can be written as in equation (1),

$$G(z,t) = \lim_{N\to\infty} \exp\{N[g(z,t)-1]\} \tag{1}$$

where $g(z,t)$ is the generating function for proteins in the corresponding reduced model. In the reduced model, a single mRNA can transition through a finite set of states, with each state having its own protein production rate. This is now equivalent to a promoter-based model with the mRNA transitions analogous to promoter-state transitions. Correspondingly, its MAP representation can be used to determine the moments for the protein distribution using linear algebra.

III. CONCLUSIONS

The PPA mapping is a powerful tool that can be used to derive and compute moments of protein distributions, given a model of post-transcriptional regulation. This general class of models can be used for analyzing combinations of different processes, i.e. mRNA senescence and interactions with different RNA-binding proteins. The developed framework is general, and can be used to design strategies for controlling variability of protein levels. This work is currently being extended to consider more general cases wherein mRNA arrivals are not a simple Poisson process (e.g. to include production of mRNAs in bursts) and to analyze rare events and large deviations in the rate of protein production [4].

REFERENCES

[1] Álvaro Sánchez and Jané Kondev. Transcriptional control of noise in gene expression. *Proceedings of the National Academy of Sciences*, 105(13):5081–5086, 2008.

[2] Hodjat Pendar, Thierry Platini, and Rahul V. Kulkarni. Exact protein distributions for stochastic models of gene expression using partitioning of Poisson processes. *Phys. Rev. E*, 87:042720, Apr 2013.

[3] Sheldon M. Ross. *Introduction to Probability Theory*. Academic Press, Boston, eleventh edition, 2014.

[4] Jordan M Horowitz and Rahul V Kulkarni. Stochastic gene expression conditioned on large deviations. *Physical Biology*, 14(3):03LT01, 2017.