Analytical solutions of separable Master Equations and applications to gene regulatory networks

Martin Hemberg[1], Justine Dattani[2*], and Mauricio Barahona[2**]

***Short Abstract* —We present a method for obtaining analytical solutions for a large class of MEs that are relevant to gene regulatory networks. Our method provides the time-dependent solution as well as the stationary distribution, and it can be applied to networks of arbitrary size that fulfill a certain set of broad conditions. We also demonstrate an equivalence between stochastic reaction rates, non-exponential waiting times and memory kernels, thereby extending a classic result by Montroll and Weiss [1]. This result is helpful when analyzing different types of experimental data and for the biological interpretation of non-linear reaction rates and waiting times.**

***Keywords* — Solution to nonlinear Master Equations, gene regulatory networks, waiting times, memory kernels.**

I. INTRODUCTION

MOST cellular processes are orchestrated by complex gene networks, which are subject to significant expression noise. Recent experiments in eukaryotes [2-4] and prokaryotes [5-7] have shown that mRNA and protein levels can vary by several orders of magnitude, and fluctuations can have an auto-correlation time that is of at least the same order of magnitude as the cell-cycle, even between genetically identical cells in a homogeneous environment [8,9].

Despite an ever-increasing amount of evidence suggesting that cells possess strategies to regulate the noise levels in different gene networks, the precise sources of variability, how the noise is harnessed and the possible functional roles for noise remain poorly understood.

II. RESULTS

The most rigorous stochastic models are based on the Master Equation (ME), but solving the ME is very challenging and exact solutions exist for only a handful of systems [10-13]. We present full analytical solutions for *separable MEs,* the class of MEs where the creation and degradation rates can be arbitrary functions of time but do not depend directly on the state of the ME.

An important example of non-linear separable MEs that can be solved using our framework are mixture models, where the reaction rates are modeled as stochastic processes. We demonstrate how these types of non-linearities can be represented as non-exponential waiting times and/or memory kernels.

Importantly, the availability of the full solution to the ME allows us to directly compare different models of gene regulation in an unbiased manner. We present several different models, and we find that the best fits for single-cell RT-qPCR data and genome-wide single-cell RNA-Seq data [13] are obtained for a Poisson-Jacobi (PJ) model, a model that generalizes previous results [14]. Furthermore, we derive a generic scaling law, demonstrating that the magnitude of the noise is proportional to the square root of the time-scale of the fluctuations for both mRNAs and proteins. We also find that this scaling is supported by fluorescent protein time-series data [9].

Acknowledgements: We warmly thank Anders Ståhlberg for his helpful discussions.

[1]Department of Ophthalmology, Children's Hospital Boston, Center for Brain Science and Swartz Center for Theoretical Neuroscience, Harvard University, 300 Longwood Avenue, Boston, Massachusetts 02115, USA. E-mail: martin.hemberg@childrens.harvard.edu

[2]Department of Mathematics, Imperial College London, South Kensington Campus, London SW7 2AZ, United Kingdom. E-mail: *j.dattani11@imperial.ac.uk; **m.barahona@imperial.ac.uk

REFERENCES

[1] Montroll EW, Kenkre VM, Schlesinger MF (1973) Generalized master equations for continuous-time random walks. *Journal of Statistical Physics* **9**(1), 45-50.

[2] Raj A, et al. (2006) Stochastic mRNA synthesis in mammalian cells. *PLOS Biology,* **4**(10), 1707-1719.

[3] Bengtsson M, et al. (2005) Gene expression profiling in single cells from the pancreatic islets of Langerhans reveals lognormal distribution of mRNA levels. *Genome Research* **15**, 1388-1392.

[4] Raj A, et al. (2010) Variability in gene expression underlies incomplete penetrance. *Nature* **463**, 913-918.

[5] Elowitz MB, et al. (2002) Stochastic gene expression in a single cell. *Science* **297**, 1183-1186.

[6] Elf J, Li G-W, Xie XS (2007) Probing transcription factor dynamics at the single-molecule level in a cell. *Science* **316**, 1191-1194.

[7] Golding I, et al. (2005) Real-time kinetics of gene activity of individual bacteria. *Cell* **123**, 1025-1036.

[8] Cai L, Friedman N, Xie XS (2006) Stochastic protein expression in individual cells at the single molecule level. *Nature* **440**, 358-362.

[9] Sigal A, et al. (2006) Variability and memory of protein levels in human cells. *Nature* **444**, 643-646.

[10] Gadgil C, Lee C-H, Othmer HG (2005) A stochastic analysis of first-order reaction networks. *Bulletin of Mathematical Biology* **67**(5), 901-946.

[11] Jahnke T, Huisinga W (2007) Solving the chemical master equation for monomolecular reaction systems analytically. *Journal of Mathematical Biology* **54**(1), 1-26.

[12] Van Kampen NG (1992) *Stochastic processes in physics and chemistry*. Elsevier, 2nd edition.

[13] Islam S, et al (2011) Characterization of the single-cell transcriptional landscape by highly multiplex RNA-seq. *Genome Research* **21**, 1160-1167.

[14] Iyer-Biswas S, Hayot F, Jayaprakash C (2009) Stochasticity of gene products from transcriptional pulsing. *Physical Review E* **79**(3), 03191.