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

Martin Hemberg1, Justine Dattani2 and Mauricio Barahona2

1 Department of Ophthalmology, Children’s Hospital Boston, Harvard Medical School, Boston, MA; 2 Department of Mathematics, Imperial College London, UK

Motivation

- Most cellular processes are subject to significant expression noise, even between genetically identical cells in a homogeneous environment.
- The precise sources of variability, how the noise is harnessed, and the possible functional roles for noise remain poorly understood.
- Noise is fundamentally important in gene regulation.
- One source of stochasticity in gene transcription is the small number of molecules involved. In this sense, it is intrinsic.
- Additional variability is caused by the intracellular environment and the uncertainty of the biochemical processes involved, leading to cell-to-cell variability. In this sense, it is extrinsic.

Analytical solution of separable Master Equations

The most fundamental way to represent gene transcription is with the Master Equation (ME), but it is difficult to solve. Most researchers simulate the ME or consider only linear cases. Here we consider systems of the form:

\[
\frac{dP_n}{dt} = -\kappa(t)P_n + \kappa(t)\sum_{n_1+n_2=n} P_{n_1}P_{n_2} - n\kappa(t)P_n.
\]

Definition 1 (Separable ME): A ME of the form above is separable if the reaction rates \(\kappa(t)\) and \(\lambda(t)\) do not depend on \(n\).

Theorem 2 (Solution of the separable ME): The solution of a separable ME with initial conditions \(P_n(0) = 1\) is given by:

\[
P_n(t) = \sum_{\pi=0}^{\infty} \binom{n}{\pi} \exp(-\int_{0}^{t} \kappa(s) \, ds - \int_{0}^{t} \lambda(s) \, ds) \frac{\kappa^\pi(t)}{\pi!}.
\]

This theorem generalizes previous results for linear ME’s, and for the case \(n_0 = 0\).

Theorem 3 (Networks of separable reactions): Consider a network of anto-molecular reactions, as above. If \(N\) does not contain directed cycles, and if the reaction rates for each molecular species \(s_i\) are independent of \(s_j\), then we can solve the ME for the network sequentially using Theorem 2.

Acknowledgements

We thank Anders Stählinger for helpful discussions. JD is funded through a Doctoral Grant of the Engineering and Physical Sciences Council (EPSRC) of the UK.

References


Coupled stochastic dynamics and mixture models

- We need to include the extra variability surrounding the process of transcription. Specifically, we model the stochastic reactions at the promoter level.
- Treating the reaction rates as Markovian is a gross oversimplification.
- In particular, Markovian models provide poor fits to data.
- Doubly stochastic processes mixture models take this complexity into account.

Definition 4 (Mixture models): A ME with the rate \(\kappa(t)\) is governed by a stochastic process with density \(\pi(k, t)\). If \(\kappa(t)\) is governed by a stochastic process with density \(\lambda(k, t)\), it is reffered to as a degradation mixture model.

The (Markovian) Berg model

The seminal gene transcription model by Berg [1] can be represented as a Markovian mixture model:

- The promoter toggles via a telegraph process between a bound (B) and an unbound (U) state with constant rates \(\kappa_B\) and \(\kappa_U\).
- The mRNA creation rate is \(\kappa(k) = \kappa(B)\), where \(\kappa\) is a random indicator variable that takes the value 1 when the promoter is in state \(B\), and 0 otherwise.

One can instead describe the reaction rates as stochastic variables governed by a wide class of solvable stochastic processes known as Poisson diffusions [2].

\[
dK(t) = -\kappa(K(t)) \, dt + \sqrt{2\alpha(K(t))^2 + \beta(K(t))} \, dW(t),
\]

We take this approach for the two models below.

The Poisson-Jacobi (PJ) model

We model the binding and unbinding rates as CLR processes (a particular case of Pearson diffusion):

\[
dK_B(t) = -\alpha(K_B(t)) \, dt + \sqrt{\alpha(K_B(t))} \, dW_B(t),
\]

\[
dK_U(t) = -\alpha(K_U(t)) \, dt + \sqrt{\alpha(K_U(t))} \, dW_U(t).
\]

The stochastic process governing the promoter occupancy, \(\pi(t) = K_B(t)/K_B(t) + K_U(t)\), is a Jacobi process [3] and the creation rate in the ME is \(\kappa(k) = \kappa(B)\).

The stationary distribution of the mRNAs is a Poisson-Beta (PB) mixture distribution:

\[
P_B(n, k, \alpha, \beta) = \binom{n}{\alpha} \frac{\alpha^\alpha \beta^\beta}{(\alpha+\beta)^{\alpha+\beta+n}} K_{\alpha}(\alpha, \beta, k),
\]

where \(\frac{\alpha}{\beta}\) is the Poisson symbol and \(K_{\alpha}(\alpha, \beta, k)\) is the confluent hypergeometric function.

The Poisson-Inverse-Gamma (PIG) model

- Model the degradation rate \(\lambda(t)\) as governed by a CLR process.
- The stationary distribution of the mRNAs is a Poisson-Inverse-Gamma (PIG) mixture distribution:

\[
P_{PIG}(n, a, b, \beta) = \frac{1}{\beta n!} \binom{n}{a} \frac{\beta^a}{(\beta+1)^{a+n}} K_{\beta}(a, n, b),
\]

where \(K_{\beta}(a, n, b)\) is the Bessel function of the third kind.

Stochastic Reaction Rates, Non-Exponential Waiting Times and Memory Kernels

Stochastic reaction rates can be interpreted as other types of non-linearities: non-exponential waiting times and memory-kernels. Kenkre et al [4] showed that the description of the ME using non-exponential waiting times or memory-kernels are equivalent. Using a memory kernel, the ME can be written as:

\[
P_n(t) = \int_0^t \int_0^{n-1} \int_0^{n-2} \cdots \int_0^{n-k} \kappa_j(t) \, dP_{n-k}(t) - n\kappa(t)P_n(t) \, dt.
\]

Assuming stationarity of the mixing process, \(\pi(k, t)\), we can extend the equivalence.

Below, \(C\) is the Carson-Laplace transform.

Data fitting and temporal characteristics

(Left) Simulation of the PI model with \(n = 10, h = 10, \alpha = 7, \beta = 2, \gamma = 1, \lambda = 1\). The magenta markers at the bottom show when the promoter was active. (Right) Fits of the Berg, PJ and PIG data from the PI/S1 or OCI15 gene from single-cell human embryonic stem cells [3]. The PI model is clearly superior in fitting the data, as shown by the Akaike Information Criteria.

Summary

- We obtain full analytical solutions for a large class of MEs, including networks of arbitrary size that fulfill certain broad conditions.
- Applying the method to mixture models provides a general class of solutions that:
  - fit experimental data exceptionally well
  - are biologically interpretable
  - are analytically tractable.
- We demonstrate: stochastic reaction rates \(+\) non-exponential waiting times \(+\) memory kernels.