Challenging the concept of a steady-state in stochastic models of gene expression

Ramon Grima
School of Biological Sciences,
University of Edinburgh, UK

30th April 2019
Contents

- Central dogma of molecular biology & deterministic modeling of gene expression.
- The inherent stochasticity of gene expression.
- Constructing a stochastic description: chemical master equation.
- A simple stochastic model of gene expression.
- What is a biologically meaningful steady-state of this stochastic model?
A very quick introduction to gene expression
Central dogma of molecular biology

A simple deterministic model of gene expression

\[
\frac{dn_M}{dt} = v_0 - d_0 n_M
\]
\[
\frac{dn_P}{dt} = v_1 n_M - d_1 n_P
\]

**Note.** mRNA (messenger RNA) is NOT destroyed when it is translated into protein and hence there is not a term proportional to \(v_1\) in the equation for the mRNA.

Shahrezaei & Swain, PNAS (2008)
But gene expression is stochastic!!!

Gene expression is stochastic because of two main factors:

(1) *We cannot precisely tell how long we have to wait before the next transcription, translation or decay event occurs.*

(2) *We cannot tell which reaction event will occur next – will it be transcription, translation, mRNA decay or protein decay?*

These imply that given some initial conditions, the protein and mRNA numbers cannot be exactly predicted at a future time.

In other words the dynamics are **stochastic** and not deterministic.

The fluctuations in molecule numbers arising from the above processes are collectively called **INTERNAL (intrinsic) noise**.
Plenty of experimental evidence that gene expression is stochastic.

Each trajectory corresponds to a different cell. Fluorescence intensity is proportional to number of proteins.

Austin et al., Nature 2006
Modeling stochastic dynamics in single cells

The Chemical Master Equation
Discrete and stochastic description: Chemical Master Equations (CMEs)

Major Assumption:

The system is well-mixed – guarantees that the molecules are randomly distributed throughout the volume and that hence we can describe them just by the molecule numbers, i.e. ignore molecular position & velocity.

The description of the system:

Since we have noise, we here describe the state of the system at any time by the probability that the system has so many molecules of this species, that species, etc ...

\[ P(n_A, n_B, \ldots, t) = \text{probability that the system has } n_A \text{ molecules of } A, \]
\[ n_B \text{ of } B, \ldots, \text{ at time } t. \]

The chemical master equation is a differential equation telling us how this probability changes with time.
A master equation description of mRNA dynamics inside a cell (Part 1)

Consider the transcription (production) of mRNA with probability per unit time $k_0$. Furthermore say that each mRNA molecule decays independently from all others with with probability per unit time $k_1$.

\[ P(n_1,t) = \text{probability of having } n_1 \text{ molecules at time } t \]

\[ P(n_1,t + dt) = k_0 dt \ P(n_1 - 1,t) + k_1 (n_1 + 1)dt \ P(n_1 + 1,t) \]

\[ + (1 - k_0 dt - k_1 n_1 dt) \ P(n_1,t) \]

\[
\frac{dP(n_1,t)}{dt} = k_0 (P(n_1 - 1,t) - P(n_1,t)) \\
+ k_1 ((n_1 + 1)P(n_1 + 1,t) - n_1 P(n_1,t))
\]
A master equation description of mRNA dynamics inside a cell (Part 2)

\[
\frac{dP(n_1, t)}{dt} = k_0(P(n_1 - 1, t) - P(n_1, t))
\]
\[
+k_1((n_1 + 1)P(n_1 + 1, t) - n_1P(n_1, t))
\]

The mean and variance of mRNA numbers can be found from the CME by using their definitions:

\[
Mean = \mu(t) = \sum_{i=0}^{\infty} n_1 P(n_1, t)
\]
\[
Variance = V(t) = \sum_{i=0}^{\infty} n_1^2 P(n_1, t) - \mu(t)^2
\]
A master equation description of mRNA dynamics inside a cell (Part 3)

\[
\frac{dP(n_1,t)}{dt} = k_0(P(n_1 - 1,t) - P(n_1,t)) \\
+ k_1((n_1 + 1)P(n_1 + 1,t) - n_1P(n_1,t))
\]

Using the CME and the mean and variance definitions, the differential equations for the mean and variance of mRNA numbers can be shown to be given by:

\[
\frac{d\mu(t)}{dt} = k_0 - k_1\mu(t)
\]

\[
\frac{dV(t)}{dt} = k_0 + k_1\mu(t) - 2k_1V(t)
\]
A master equation description of mRNA dynamics inside a cell (Part 4)

\[
\frac{dP(n_1,t)}{dt} = k_0(P(n_1-1,t) - P(n_1,t)) \\
+ k_1((n_1+1)P(n_1+1,t) - n_1P(n_1,t))
\]

CME

In (conventional) steady-state, the equations can be easily solved:

\[
\frac{d\mu(t)}{dt} = k_0 - k_1\mu(t) = 0 \\
\frac{dV(t)}{dt} = k_0 + k_1\mu(t) - 2k_1V(t) = 0
\]

\[\mu = V = \frac{k_0}{k_1}\]
The Big Question/s:

In vivo, does mRNA ever reach a steady-state?

Does it make sense to use the steady-state mRNA equation to understand experimental data?
Taking a deeper look at how experimental measurements of mean and variance of mRNA numbers are made
Measurements of mean and variance are typically made over a population of cells

◆ The number of molecules in each cell is counted at a certain observation time. The mean and variance are then computed from the individual cell measurements.

◆ Cells are not synchronized with each other, i.e. cells of all ages are distributed according to a probability distribution. The cell age can vary between 0 and $T$ where $T$ is the cell-division time. The cell age distribution is approximately given by:

$$f(t) = \frac{\ln(2)}{T} 2^{1-t/T}, \quad \text{where} \ t \in [0, T]$$
Cell growth, cell division and the cyclo-stationary limit (Part 1)

1. A cell is born
2. The cell grows
3. The cell divides into two new cells after a time $T$
4. Molecules inside a the original cell are partitioned into either of the 2 new cells with probability $= \frac{1}{2}$
Cell growth, cell division and the cyclo-stationary limit (Part 2)

The mean & var of molecule numbers in a cell of age $t$ (where $t$ varies between 0 and $T$) is independent of which generation the cell belongs to.

Biological definition of Steady-State

Problem 1: Is conventional steady-state a good approximation to cyclo-stationary limit for mRNA in a population of cells?

(i) Modify the mRNA mean & var eqs on Slide 12 to include binomial partitioning at cell division time T.

(ii) Solve the modified eqs to obtain formulae for the mean & var in the cyclo-stationary limit.

(iii) Use the solution of (ii) to obtain the population average of the mean and var in the cyclo-stationary limit (using the cell-age distribution).

(iv) Under what conditions, the solution of (iii) agrees with the conventional steady-state solution on Slide 13?
Problem 2: Is conventional steady-state a good approximation to cyclo-stationary limit for protein in a population of cells?

(i) You will need to first formulate the chemical master equation for both mRNA and protein taking into account transcription, translation and protein/mRNA decay, i.e., modify the derivation on Slide 10 to model the reaction system on Slide 5.

(ii) Find the conventional steady-state solution for the mean and variance of protein numbers.

(iii) Modify the equations as in Problem 1 to account for binomial partitioning of protein at cell division and asynchronous cells.

(iv) Under what conditions the conventional steady-state solution for the mean and variance agree with those found in (iii)?
Problem 3: Use physiological parameter values to calculate the error between the conventional steady-state & the cyclo-stationary limit of the population models

(i) The four parameters and the cell-cycle time have been estimated for thousands of mouse genes in the paper (see its Supplementary Methods/Tables):


(ii) Use the parameter data to calculate the percentage error between the steady-state mean & var and the mean & var calculated in Problems 1, 2 for each gene.