

Simulation for the distributions of stochastic gene expression: Runge-Kutta Method

Adarsh Vishwakarma

The Institute of Mathematical Sciences

2023

Goal

- 1 Simulate steady-state protein distribution.
- 2 Verify with the analytically obtained distributions.

Two-Stage Model of Gene Expression

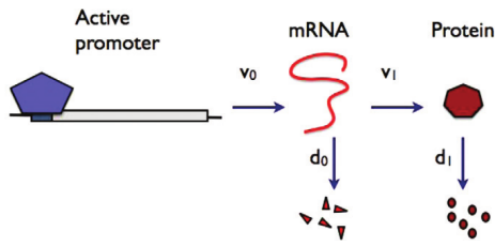


Figure: Two stage model of transcription and translation.

Two-Stage Model of Gene Expression

- The promoter is always active.
- P_{mn} : probability of m mRNAs and n proteins.

$P_{mn}(t)$ satisfies the master equation,

$$\begin{aligned}\frac{\partial P_{m,n}}{\partial t} = & \nu_0(P_{m-1,n} - P_{m,n}) + \nu_1 m(P_{m,n-1} - P_{m,n}) \\ & + d_0[(m+1)P_{m+1,n} - mP_{m,n}] \\ & + d_1[(n+1)P_{m,n+1} - nP_{m,n}],\end{aligned}\tag{1}$$

ν_0 - probability per unit time of transcription,

ν_1 - probability per unit time of translation,

d_0 - probability per unit time of degradation of mRNA,

d_1 - probability per unit time of degradation of protein.

Analytic Steady-State Solution

$$\begin{aligned} P_n(\tau) &= \frac{\Gamma(a+n)}{\Gamma(n+1)\Gamma(a)} \left(\frac{b}{1+b}\right)^n \left(\frac{1+be^{-\tau}}{1+b}\right)^a \\ &= \times {}_2F_1\left(-n, -a, 1-a-n; \frac{1+b}{e^\tau+b}\right), \end{aligned} \quad (2)$$

n - number of proteins,

$\gamma = d_0/d_1 \gg 1$,

$\tau = d_1 t \gg 1/\gamma$ ($t \gg 1/d_0$).

Numerical vs Analytical

Two steady-state solutions are plotted for different values of the parameters a and b and are compared with analytical distributions.

Numerical vs Analytical

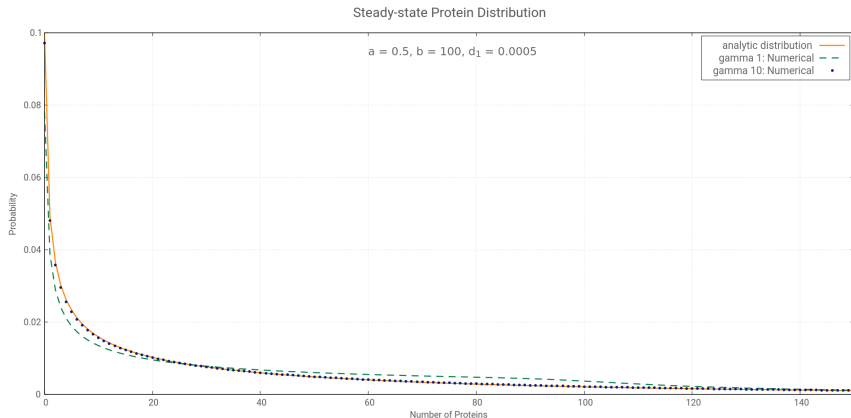


Figure: Comparison of Numerical simulation of eq(1) with the analytical result obtained using eq(2)

Numerical vs Analytical

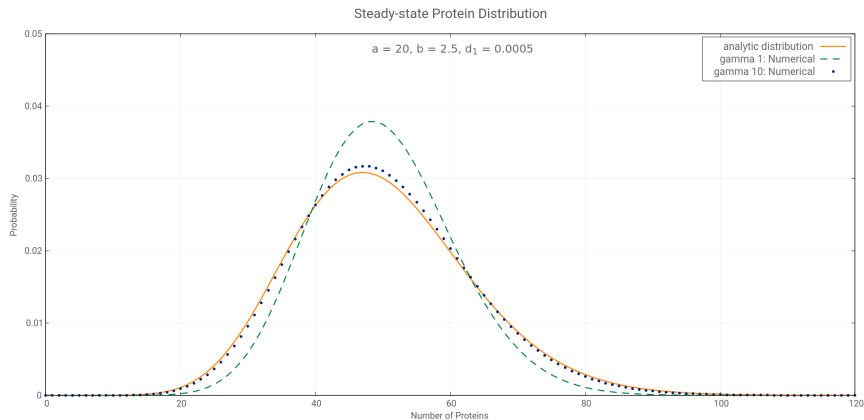


Figure: Comparison of Numerical simulation of eq(1) with the analytical result obtained using eq(2)

Time-dependent solution

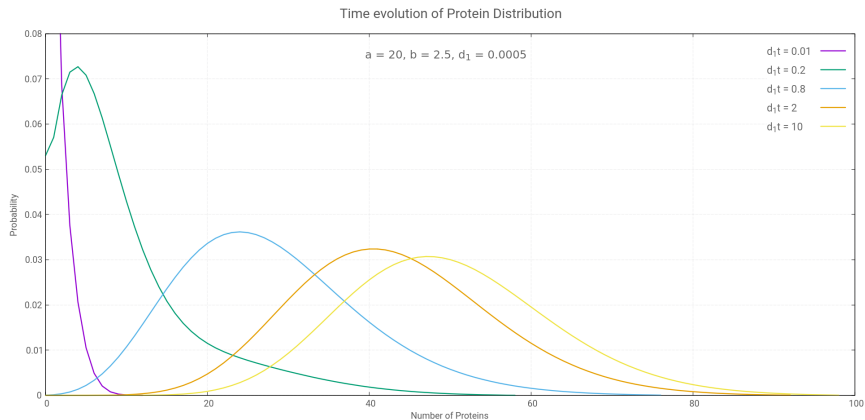


Figure: Time evolution of the gaussian distribution is shown. Initial protein is zero.

Time-dependent solution

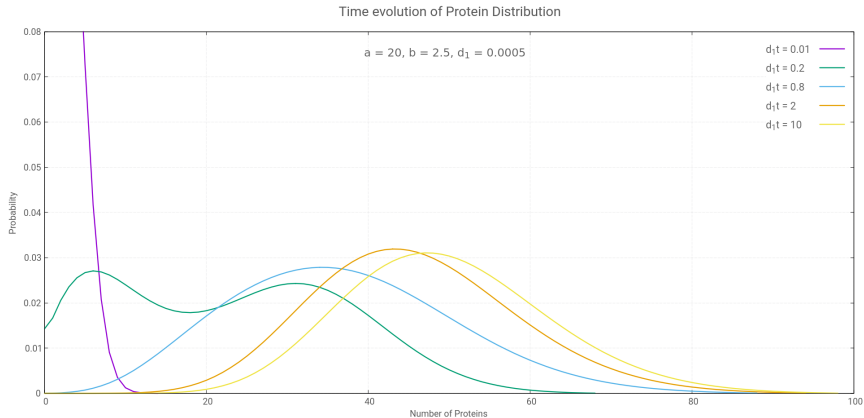


Figure: Time evolution of the gaussian distribution with different initial mRNA distribution. Initial protein is zero.

Time-dependent solution

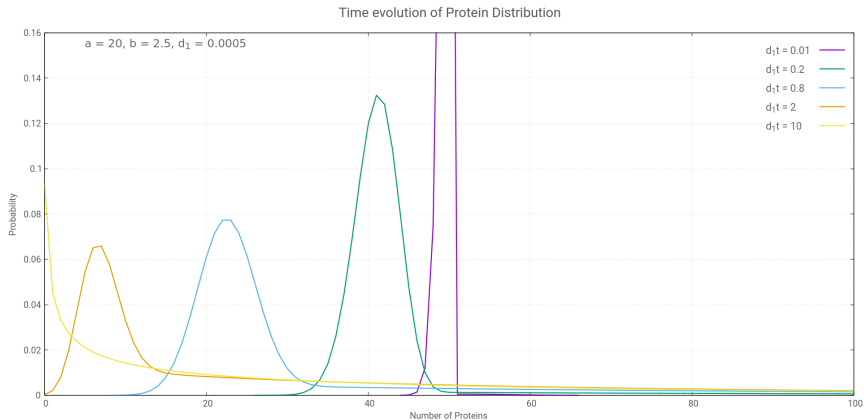


Figure: The time evolution of the poisson distribution is shown.

Figures (2) and (3) show that $\gamma = 10$ results match better with the analytical distribution (2), which is valid for large γ .

Three-Stage Model of Gene Expression

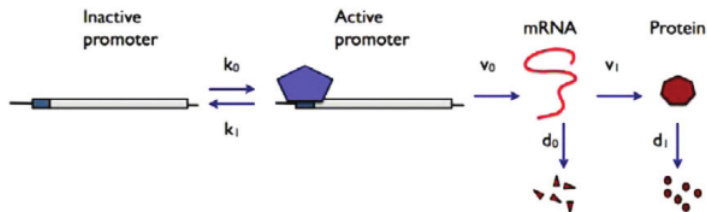


Figure: Three stage model of transcription and translation. The promoter switches between active and inactive states.

Three-Stage Model of Gene Expression

- The promoter switches between active and inactive states.
- P_{mn}^0 : probability of m mRNAs and n proteins when the DNA is inactive.
- P_{mn}^1 : probability of m mRNAs and n proteins when the DNA is active.

Master equations

$$\begin{aligned}\frac{\partial P_{m,n}^0}{\partial \tau} = & \kappa_1 P_{m,n}^1 - \kappa_0 P_{m,n}^0 + (n+1)P_{m,n+1}^0 - nP_{m,n}^0 \\ & + \gamma[(m+1)P_{m+1,n}^0 - mP_{m,n}^0 + bm(P_{m,n-1}^0 - P_{m,n}^0)], \quad (3)\end{aligned}$$

$$\begin{aligned}\frac{\partial P_{m,n}^1}{\partial \tau} = & -\kappa_1 P_{m,n}^1 + \kappa_0 P_{m,n}^0 + (n+1)P_{m,n+1}^1 - nP_{m,n}^1 \\ & + a(P_{m-1,n}^1 - P_{m,n}^1) \\ & + \gamma[(m+1)P_{m+1,n}^1 - mP_{m,n}^1 + bm(P_{m,n-1}^1 - P_{m,n}^1)], \quad (4)\end{aligned}$$

$$\kappa_0 = k_0/d_1, \quad \kappa_1 = k_1/d_1,$$

$$\tau = d_1 t.$$

Analytic Steady-State Solution

$$\begin{aligned}P_n &= \frac{\Gamma(\alpha + n)\Gamma(\beta + n)\Gamma(\kappa_0 + \kappa_1)}{\Gamma(n + 1)\Gamma(\alpha)\Gamma(\beta)\Gamma(\kappa_0 + \kappa_1 + n)} \\&= \times \left(\frac{b}{1 + b}\right)^n \left(1 - \frac{b}{1 + b}\right)^\alpha \\&= \times {}_2F_1\left(\alpha + n, \kappa_0 + \kappa_1 - \beta, \kappa_0 + \kappa_1 + n; \frac{b}{1 + b}\right),\end{aligned}\quad (5)$$

$$\alpha = \frac{1}{2}(a + \kappa_0 + \kappa_1 + \phi),$$

$$\beta = \frac{1}{2}(a + \kappa_0 + \kappa_1 - \phi),$$

$$\phi^2 = (a + \kappa_0 + \kappa_1)^2 - 4a\kappa_0,$$

$$\gamma \gg 1.$$

Numerical vs Analytical

Three steady-state numerical solutions are plotted for different values of the parameters a and b and are compared with analytical distributions.

Numerical vs Analytical

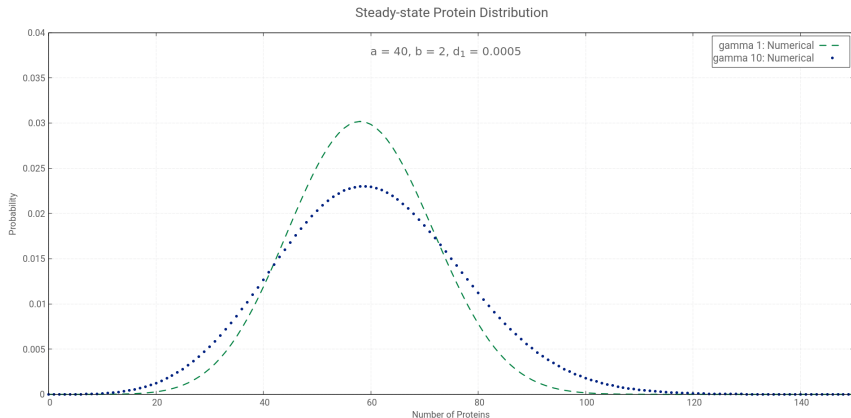


Figure: Comparison of Numerical simulation of eqs(3) and (4) with the analytical result obtained using eq(5)

Numerical vs Analytical

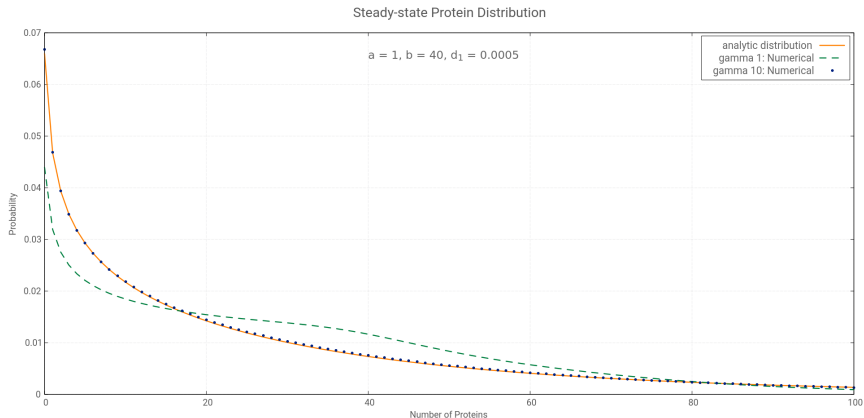


Figure: Comparison of Numerical simulation of eq(3) and (4) with the analytical result obtained using eq(5)

Numerical vs Analytical

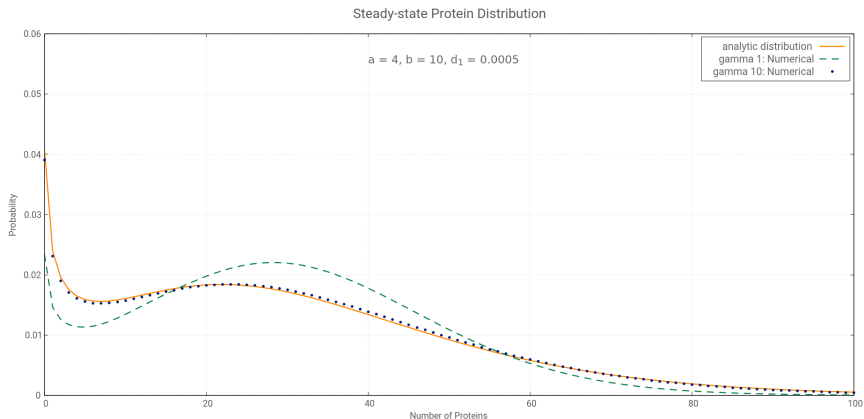


Figure: Comparison of Numerical simulation of eq(3) and (4) with the analytical result obtained using eq(5)

The figures (8), (9), and (10) show that $\gamma = 10$ match better with the analytical result for large γ .

Conclusions

- 1 $\gamma = 10$ is sufficiently large for the analytical distribution to be valid.
- 2 The two-stage protein distribution either peaks at zero or at some non-zero value.
- 3 The three-stage protein distribution can have bimodal peaks.

Conclusions

- 1 $\gamma = 10$ is sufficiently large for the analytical distribution to be valid.
- 2 The two-stage protein distribution either peaks at zero or at some non-zero value.
- 3 The three-stage protein distribution can have bimodal peaks.

Conclusions

- 1 $\gamma = 10$ is sufficiently large for the analytical distribution to be valid.
- 2 The two-stage protein distribution either peaks at zero or at some non-zero value.
- 3 The three-stage protein distribution can have bimodal peaks.

- 1 Analytical distributions for stochastic gene expression. Vahid Shahrezaei and Peter S. Swain