Genetic Switches between two population with regards to mRNA and proteins

Qin He[1,2,3] Jukka Peltola[3,4] Susanna Narkilahti[4] Sampsa Pursiainen[2] RubinWang[5]
XiaochuanPan[5]
1.Signal Processing Department,Tampere Universities, 2.Math Department,Tampere Universities
4.Medicine and Health Technology department,Tampere Universities, 3.TAYs
Tampere,Finland
5.Neural-Dyamics Lab, East China University of Science and Technology
Shanghai, China

Abstract

As Arc, one virus-like gene, crucial for learning and memory, was discovered by researchers in neurological disorders fields, Arc mRNA's single directed path and allowing protein binding regional restrictively is a potential investigation on helping shuttle toxic proteins responsible for some diseases related to memory deficiency. To study especially the transform between mRNA and proteins, the switching function of the phenotypes, 'normals' multiplying populations and 'persisters', resilient to stress instead of multiplying is of our interest. Mean time to switching (MTS) is calculated explicitly quantifying the switching process in statistical methods combining Hamiltonian Markov Chain. The model derived from predator and prey with type II functional response studies the mechanism of normals with intrinsic rate of increase and the persisters with the instantaneous discovery rate and converting coefficients. During solving the results, since the numeric method is applied for the 2D approximation of Hamiltonion with intrinsic noise induced switching combining geometric minimum action method. The MTS, trajectories and Hamiltonian dynamics demonstrate the practical and robust advantages of our model on interpreting the switching process of genes (IGFs, Hax Arcs and etc.) with respects to memory deficiency in aging process which can be useful in further drug efficiency test and disease curing.

1 Introduction

In cell biology, non-equilibrium stochastic process is of interest since the observation of experimental results are becoming of higher resolution, studying the molecules both with imaging and expression data are often conducted in both single and population (thousand) order, which basically described in stochastic process whether on a discrete or continuous scale with status changes either genotypically or phenotypically. Many problems are thus studied related to status switching, including cell regulatory networks[1], signal response on excitability and inhibition[2], (convinced by translational and transcriptional burst of expression for instances.), metastability among populations, (binding of ligands and proteins, forming of polymerases and etc.).In this paper, we focus on the interaction among genes, mRNA, proteins and etc. To be more specific, while the switching problem among molecules can be studied on genotype, including sequencing for single RNA, alignments and binding considering condons and etc, we stay on the switching with expression (concentration) only, which is simplified as modified population problem using Lotka-Volterra equations[3] of two populations only. Thus, rather than the competitor model(for instances, cell bifurcations.), we applied simulation of switching on predator model. The model is based on the following basic assumptions: Prey population(promoters) is fed with enough food all the time while the predator population of the predator(the persisters) depends on the size of prey(promoters).In our paper, we mainly study the interaction of DNA and its interaction with the associated proteins.(Clinical data of Hax1 and HS_1 is downloaded from Ensmbl gene database[4]). On one hand, the switching model is calculated under the large deviation theory(LDT)[5] combining the least actions. The Markov chain[6] consider the states of the 2D coordinates (x,y) of mRNA numbers and protein numbers referencing the distribution of x, which follows the order O(1) while P_X follows the time scale on O(1/ε) and guaranteeing the variant of LDT hold with the transform of the expressions in single population. Only considering the process of diffusion case, we study the binding of hax1 with simple switching between on and off status under its interaction with HS_1 seen as in the constant environment, i.e. the closed system at mean field. The dimer which can be cancelled out connect the binding between two single population. On the other hand, one numeric method is applied to solve the problem, making compare with the stochastic process[7] on approximation equation of the mean switching time(MST) with the transform between two status (we studied the switching time with four situations, both multiplicative and asymptotic of single population and the binding and degradation between two population.) Again, this method is also calculated based on the Hamiltonians. We give out the MST with respect to N/Nc denoting N as the population number of interest and Nc as the threshold of certain status(either of that population or the other population). Since our study only based on data in the process of transforming in the constant environment, extinction is not considered in this paper. To study both intrinsic and extrinsic noise with the exciting and inhibiting bursts is the potential topic in the future. In the following contents, the first chapter is the proposition of the model, based on least action with LDT and MTS approximation with one stochastic differential equation (SDE) [8]separately; And the second chapter gives numeric experiments based on Hamilton Markov Chain computation of the expression data of hax1 and $HS-1$; In the last chapter, the results is analysed with both hamiltonian, realization size, convergence, the rewards computation taking the continuous markov chain as Poisson process[9] and etc. In the appendix, there also includes the complete proof of model with action S based on hamilton not only based on the explicit equation in this paper. Some descriptive Statistics and pre-computation based on the data can be accessed through link in availability. As the process related to motor coordination and function, the Hax's function in regulation, B cell's signal transduction can be further studied with more data considering its excitability and metastability functions with stimulation of drugs for instance in the future as well.

2 Proposed Model

Molecular interactions are studied on phenotypic data of the mRNA

and its associated protein in this paper, especially the trajectory of the production of hax1 and HS_1 with interaction with each other through least action method combining diffusion process[10]. Furthermore, in solving the equation, one stochastic differentiation equation approximates the analytic solution and calculation of MST[11] based on converging with Hamiltonian quantities, finding three convergence points through eigenvalue of position quantities as well as satisfying H = 0 and $H_\theta = 0$ where $\theta(P_X, P_Y)$ are momentum quantities.

2.1 switching model with least action

First of all, we consider the dynamics of population of the interaction involved systems as diffusion[12], and thus the Hamiltonian $H(x,\theta)$ is computed with the minimization of action (quasi-potential)[13] instead of some other methods, for instance WKB[14]. With the Lagrangian denoted with respect to Hamiltonian according to LDT: $L(x,y) = sup_{\theta\in R^n}(< y,\theta > -H(x,\theta)) =< y,\theta(x,y) > -H(x,\theta(x,y))$ Due to the maximizer $\theta(x,y)$ being implicitly defined by $H_\theta(x,\theta(x,y)) = y$, we calculate the action from quasi-potential: $V(x1,x2) = inf_{T>0} inf_{\psi\in C^{x2}_{x1}(0,T)S_T(\psi)} = inf_{T>0} inf_{\phi\in C^{x2}_{x1}(0,1)} inf_{\psi\in C_\phi(0,T)} S_T(\psi) = inf_{\phi\in C^{x2}_{x1}(0,1)S(\phi)}$
So that for any $\phi \in C(0,1)$ the action $S(\phi)$ is given by the equivalent four formula:
$S(\phi) = inf_{T>0} inf\phi \in C_\phi(0,T) * S_T(\psi)$
$S(\phi) = sup_{\theta:[0,1]->R^n_{H(\phi,\theta)=0}} \int_0^1 <\phi',\theta> d\alpha$
$S(\phi) = \int_0^1 <\phi', \hat{\theta}(\phi,\phi')> d\alpha$
$S(\phi) = \int_0^1 \frac{L(\phi,\lambda*\phi')}{\lambda} d\alpha, \lambda = \lambda(\phi,\phi')$
Note that L(x, y) is the Lagrangian associated with the Hamiltonian $H(x,\theta)$ with function $\theta(x,y)$and $\lambda(x,y)$ are implicitly defined for all $x \in D$ and $y \in R^n/0$ as the unique solution (solution$(\theta,\lambda) \in R^n X[0,\text{inf})$ of the system possessing zero value when $\phi' = 0$ or $\lambda(\phi,\phi') = 0$ setting the integrands to zero with: $H(x,\theta) = 0, H_\theta(x,\theta) = \lambda y, \leq lambda$
where the lower bounds for $S(\phi)$ is directly achieved :
$S(\phi) = inf_{T>0} inf\phi \in C_\phi(0,T) S_T(\psi)$
$\geq \int_0^1 sup^{\theta\in R^n}_{H(\phi,\theta)=0} <\phi',\theta> d\alpha$
$\geq \int_0^1 <\phi', \theta(\phi,\phi')> d\alpha,$
utilizing the first equation of the four. Furthermore, $S(\phi)'s$ upper bound can also be obtained through defining a minimizing sequences $(T_k, \psi_k)_{k\in N}$ with the following rescaling process: For every $k \in N$ let: $\lambda_k(\alpha) = max(\lambda(\phi(\alpha),\phi'(\alpha)), \frac{1}{k}), \alpha \in [0,1], B_k(\alpha) = \int_0^{/alpha} \frac{1}{\lambda_k} da, \alpha \in [0,1], T_k(\alpha) = B_k(1), \psi_k(t) = \phi(B_k^{-1}(t)), t \in [0, T_k]$ Specifically, the inverse of B_k is approximated with the Brownian standard σ_x satisfying the $\alpha'(t) = \lambda_k(\alpha(t))$ and thus $\frac{1}{k} \leq \alpha'(t) \leq |\lambda_k|_{\text{inf}} \leq \text{inf holds for all } t \in [0, T_k]$ with the absolute continuity of $\alpha(t)$. And thus, the ψ_k is continuous in the whole time sequence (0, T_k), enabling the inverse process: $t = t(\alpha) = G_k(\alpha)$ with $dt = d\alpha/\lambda_k$ and $\phi'(alpha) = \psi'_k(t)G'_k(\alpha) = \frac{\psi_k(t)'}{\lambda_k(\alpha)}$, Thus,
$S_{T_k}(\phi_k) = \int_0^{T_k} \frac{L(\phi,\phi'\lambda_k)}{\lambda_k} d\alpha$
leading to the upper bound switching the integrate and limitation with $k- > \text{inf}$, and with the proof in appendix B(in another work with landscape model) fufilling the first order and second order conditions: $\phi' = \frac{H_\theta(\phi,\theta)}{\lambda}$ is negative definite during the θ maximizing process: $\frac{L(\phi,\lambda*\phi')}{\lambda} = sup_{\theta\in R^n}(< phi', \theta > -\frac{H(\phi,\theta)}{\lambda})$ and guaranteeing them both fufiled by $\theta = \theta(\phi,\phi')$ with the second equation, so that upperbound here is the same as the integrands of the lower bound as well as holds the $\theta = 0$ when the $\lambda = 0$ is satisfied, and therefore:
$\frac{L(\phi,\lambda*\phi')}{\lambda} =< \phi' theta > -\frac{H(\phi,\theta)}{\lambda} =< \phi', \theta >, \theta = \theta(\phi,\phi')$
The calculation can be found completely in Appendix B.

2.2 diffusion Approximation with numerical methods on the convert ratio referencing bacteria sensoring and MTS on difference mapping

As we want to study the switching model interpreting the process explicitly, we thus combine the deterministic[15] background of the switching between on and off and give out one stochastic model based on the explicit (ordinary differential equation) ODE of the numbers of mRNA and proteins. Although the final model(referencing the quorum sensing model of bacteria in changing environment[16]) removes the dimers but it is used in the first place while cancelled out the in the quasi steady state according to its far more faster production and degradation rate comparing to transcription and translation.(Simplified mechanism sees Appendix A). Start from the bistability of the metastability[17] of the two state model, with the absorbing boundary conditions, $\rho_0(x*,t) = 0$ and the identification of mean transition rate with principal eigen value λ^0_ε, the quasi-stationary approximation of $\rho_n(x,t) = C_0 * \exp(-\lambda^0_\varepsilon t) * \phi^0_\varepsilon(x,n)$. Furthermore, with the quasi-potential satisfying
$\Sigma_{n\in 0,1} S_n(x)(A_{n,m}(x) + \phi'_0(x)\delta_{n,m})F_m(x) = 0,\ H = 0.5 * (g_x^2 p_x^2 + g_y^2 p_y^2) + p_1 * \phi_1 + p_2 * \phi_2$
where p_i is the momentum conjugate to the generalized coordinate x_i, where $g_i = \sqrt{(S_i^2 f_i + X_i/\tau i)}$ (For more specific study of the ϕ_1 and ϕ_2 as the interacted diffusive speed, most studies applies WKB equations.) Since we focus on the transform between two status of the two populations, mRNA (of HS-1) X_n and proteins Hax1 Y_n as the system.(with dimer Z of production rate k_{XY} and degradation rate k_P)[18] and the degradation rate of HS_1 and hax1, as K_X and K_Y, separately. From the original ODES:
$\frac{dZ}{dt} = k_{XY}XY - k_Z Z$
$\frac{dX}{dt} = -k_{XY}XY + k_Z Z - k_X X + V_X * \frac{Z}{(K_X+Z)+X_0}$
$\frac{dY}{dt} = -k_{XY}XY + k_Z Z + V_Y \frac{Z}{(K_X+Z)} + V_Y * \frac{Z}{(K_Y+Z)} + Y_0 - k_X X$
where X_0 and Y_0 are the initial volumes or baseline volumes of these two populations and with instant volume as V_X and V_Y and due to the zero value of $\frac{dP}{dt}$, the term of P can be replaced through:
$P = \frac{K_{XY}}{K_P} XY$
$\frac{dX}{dt} = -k_X X + V_X * \frac{k_{XY}}{k_P} XY + X_0 = -k_X X + V_X * \frac{1}{(1+\frac{K_X k_P}{k_{XY}})} + X_0$
$\frac{dY}{dt} = -k_Y Y + V_Y * \frac{K_{XY}}{k_P} XY + Y_0 = -k_Y Y + V_Y * \frac{1}{(1+\frac{K_Y k_P}{k_{XY}})} + Y_0$
Considering the transform of X(Upstream only), in the first step as degradation as the first term of right of the upper formula, the degradation part of X with k_X which can be interpreted as the Poisson process and rewrite into $-\frac{\mu 1}{exp^{P_X}}$, and in the second term, the coefficient of degradation part of X , $C1$ is denoted as $\frac{V_X * k_X}{\mu 1}$. Mean while with the assumption of continuous Markov chain, where the convert ratio of Y is n, the $\frac{k_{XY}}{k_P} * X * Y$ is equivalent to $(\frac{Y}{X+Y})^n$ so that the whole degradation part becomes $\frac{C1*\mu 1}{1+(\frac{Y}{X+Y})^n} exp^{P_X}$, $\frac{C2*\mu 2}{1+(\frac{X}{X+Y})^n} exp^{P_Y}$ (*1) and the final transform rate of mRNA number X and proteins Y are:

$\frac{C1}{1+(\frac{Y}{X+Y})^n}(exp^{P_X} - 1) - \mu 1 * X(exp^{-P_X} - 1)$and $\frac{C2}{1+(\frac{X}{X+Y})^n}(exp^{P_Y} - 1) - \mu 2 * Y * (exp^{-P_Y} - 1)$ where, the coefficient of degradation part of Y $C2$ denoted as $\frac{V_X * k_Y}{\mu 2}$. as the reciprocal of the other population ratio. And as Y stands for the number of the proteins, X for the number of the mRNA separately with m and n as their translation and transcription rate. With the total sum of the system mocule numbers assumed as X+Y, we have the Hamiltonian:
$\frac{C1}{1+(\frac{Y}{X+Y})^n}(exp P_X - 1) - \frac{\mu 1 * X}{(exp - P_X - 1)} + \frac{C2}{1+(\frac{X}{X+Y})^n}(exp P_Y - 1) - \frac{\mu 2 * Y}{(exp - P_Y - 1)}$
,where p_X, p_Y are calculated setting H = 0 and $H_\theta = 0$, and convertion rate which can be calculated as $\frac{dy}{dx}$ specifically here, letting the first term equals the second and third equals the fourth term. (Complete see Appendix B) Note that: each single DNA population($hax1$ and HS_1) has its own degradation rate when considering about its mRNA computation, and the other population's protein is taken as the intake, promoting its population when as normals binding onto the according site of persisters, activating it. Vice versa, thus, the two populations have similar structured formula describing each degradation and population under the dual interacted population. The mean switching time is calculated based on the solution of the SDE: $z' = z + \sqrt{(\frac{N_c}{N})} * \sqrt{1+2} * \varepsilon\text{-}z^2 * \eta$,

where $N_c = \frac{1}{}$ and η $N(0,\delta)$ is the white noise with correlator $<\eta(\tau)\eta(\tau')>=\delta(\tau-\tau')$. Note that it is the span of the master equation in powers of the inverse population size N^{-1} re-scaling with $\tau = 2**t/N$, and $z = x_1 - x_2$ ranges over the interval [-1, 1][20], leading to the solution $\tau_0* = \frac{2*\lambda_*}{1-2^*} * cot(\pi)$. Thus, we have the algorithm:

Input: maximum time scale size T, mRNA numbers y, proteins x, maximum steps Steps, tolerance Tol, parameters of the sensing model(coefficients of convertion c1, c2, transcription and translation rate m, n, degradation rate k1, k2, formation coefficients mu1, mu2, diffusion rate b1, b2), dt as each time increasement

Initialize:`maximum time scale`, T, maximum step number steps, tolerance Tol, numbers of mRNA after the first diffusion process xhat if necessary, initialized as one random the number in the first status and thus we start with the largest interval to cover higher possibilities, i.e. [x(0),x(0)+1,...,x[1]-1],

for do

xhat(length(xhat))< x(i+1): **repeat**

record the size T, time t, steps Steps - steps +1

set the sequence according to size T (the interval for mRNA numbers) x(1), x(2),...x(T) and generate the population number of proteins according data distribution, y(1), y(2),...y(T). T initialized as the X(i+1)-X(i).

consider Hamilton Markov (Hierarchical)[19]

if *xhat* exists(iterated from previous status) **then**

: segment the interval into several sub-sequences (X0 as the new current status, X1 as the previous status.)

end if

Note that: As we only consider up streaming, down regulation into those before the previous status is not included.

function dynamics Inputs: x and fitted y (or $x0, y0$ or $x1, y1$)

calculate degradation term $w1$ and $w2$ according to the (*1)

calculate px, py, dx, dy, conversion rate, H_θ and H_x, s, according to Appendix 2

predict multiplied mRNA and protein numbers xhat, yhat, and other Hamiltonians.

calculate updated gamma, delta

calculate tolerance for further stopping criteria as the residule of gamma and cell numbers with: $tol = abs(Gamma - gamma)/gamma; toll = mean(abs(xhat - x1)./x1 + abs(yhat - y1)./y1)$

Output: H, Hhat, HthetaX, HthetaY, HxX, HxY, HamilX, HamilY, HamilXhat, HamilYhat, xhat, yhat, sX, sY, px, py, pxhat, pyhat, actionratio, delta, gamma, Delta, Gamma, cr1, c1, c2, crhat, c1hat, c2hat, tol, toll, Txc, Tyc, Txchat, Tychat

concatenate results:

if X0, X1 exist **then**

xhat = [X0hat, X1hat]

end if;

if only X0 exist **then**

xhat = X0hat.

end if

do similar prediction regenerate the mRNA numbers X according to Y with **function dynamics** again for comparison. Results are with postfix 'L'

store the quantities of 'successful' moves with smaller tolerance and action for either from mRNA or protein numbers.

if satisifies the configuration condition **then**

$Tol = min(tol * toll, tolL * tollL)$

else

Fail++

end if

---

3 results and discussion

0.1 To have a clearer understanding of the switching process combining the binding with increasing and decreasing speed both of *hax*1 and HS_1, the two population are regarded as promoters and resistors both when activating and deactivating each other's production.(The coarse process can be briefly described as in 1, and it is briefly introduced in the previous chapter.)

3.1 Finding three critical points and explore their stability

As we have data (see Appendix C) of 15 status in all both for hax1 and HS_1 with their different cell numbers taken as X and Y in our model. For reward computation for their Markov chain, we pre-compute the their Hamiltonians, Action Potentials, mean switching time and related dynamics in the form (see availability), and the 6 upstreaming status, which is the focus of the experiment application of our model. Using the pre-computation results, we are able to discuss about some practical problems about the current model. There are three groups of quantities studied combining the action potential as well as Hamiltonian inspired by bacterial quorom sensoring, 'momentum and cell numbers', 'MTS with the SDE', and 'corresponding Hamiltonians', of each transform status in Appendix A. First, we use the taylor expansion to simplify the four ODE achieved in Appendix B: to

$dx = \frac{C1}{1+(\frac{y}{x+y})^m} P_x - \mu1 * x * P_X)$

$dy = \frac{C2}{1+(\frac{x}{x+y})^n} P_y - \mu2 * y * P_Y)$

$dP_X = \frac{C2*m*(\frac{x}{x+y})^{m-1}}{(1+(\frac{x}{x+y})^m)^2}(P_Y - 1) - \mu1 * P_X - \mu1$

$dP_Y = \frac{C1*n*(\frac{y}{x+y})^{n-1}}{(1+(\frac{y}{x+y})^n)^2}(P_X - 1) - \mu2 * P_Y - \mu2$ Note that our model here simplify the origin model where $c_i = \frac{a_i}{b_i}$, with b_i = 1 as the burst size of protein i, $\frac{x}{(x+y)} = \frac{x}{K_2*(x+y)}$ as $k_2 = 1$ is the dissociation constants standing for gene x binding on y's protein binding site. Regarding x and y as leading order variable, we apply phase analysis to consider the solution's stability around the three zero-energy points, which achieved through setting dx, dy, dP_X and dP_Y all to zero and combine the Hamiltonian's special case when H= 0 (and $H_P = 0$): P1(x, y, $\frac{\mu2*X}{C1}$, $\frac{\mu1*Y}{C2}$, where x and y are the solution of $x = \frac{C1}{\mu1*(1+(\frac{y}{x+y})^n)}$ and $y = \frac{C2}{\mu2*(1+(\frac{x}{x+y})^m)}$, P2(x, y, 0, 0), where x and y are the solution of $x = \frac{C1}{\mu1(1+(\frac{y}{x+y})^n} = -y = -\frac{C2}{\mu1(1+(\frac{x}{x+y})^m}$, and P3(0,0,0,0) As P_X and P_Y are either zero or formula can be replaced by x and y around those three convergence points. We here, consider the analysis on x and y as following: denote $dx = f(x,y)$ and $dy = g(x,y)$, and the we try to find x* and y* satisfy the $f(x,y) = 0$ and $g(x,y) = 0$ as well as holding the zero-energy points for their momentum. Thus with approximation: $dx = f_X(x*,y*)(x-x*) + f_Y(x*,y*)(y-y*)$, and $dy = g_X(x*,y*)(x-x*) + g_Y(x*,y*)(y-y*)$. We have $A = \begin{bmatrix} f_X & f_Y \\ g_X & g_Y \end{bmatrix}$

$= \begin{bmatrix} -\mu1 * P_X & \frac{C2*n*(\frac{y}{x+y})^{n-1}}{(1+(\frac{y}{x+y})^n)^2} * P_X \\ -\mu2 * P_Y & \frac{C1*m*(\frac{x}{x+y})^{m-1}}{(1+(\frac{x}{x+y})^m)^2} * P_X \end{bmatrix}$

where there exists the a>0, b>0 for the eigenvalue λ:

$$\lambda^2 + a*\lambda + b = 0 \quad (1)$$

$$a = -(f_X + g_Y)|_{(x*,y*)} \quad (2)$$

$$b = |A| \quad (3)$$

so that point(x*, y*) is the convergence points. Thus, we discuss about the stability of the three points as following: we denote $X = (1 + (\frac{x}{x+y})^m)$ and $Y = (1 + (\frac{y}{x+y})^n)$, compute the a and b as: $a = -(f_X + g_Y)|_{(x*,y*)} = \mu1 * P_X + \frac{C2*m(\frac{x}{x+y})^{m-1}}{(1+(\frac{x}{x+y})^m)^2} * P_Y, b = |A||_{(x*,y*)} = \mu2 * P_X * P_Y * \frac{C1*n(\frac{y}{x+y})^{n-1}}{(1+(\frac{y}{x+y})^n)^2} - \mu1 * P_X * P_Y * \frac{C2*m(\frac{x}{x+y})^{m-1}}{(1+(\frac{x}{x+y})^n)^2}$

[1] for $P1(x, y, \frac{\mu2*X}{C1}, \frac{\mu1*Y}{C2})$, where x and y are the solution of $x = \frac{C1}{\mu1*(1+(\frac{y}{x+y})^n)}$ and $y = \frac{C2}{\mu2*(1+(\frac{x}{x+y})^m)}$, $a = \mu1 * \frac{\mu2*x}{C1} + \frac{C2*m*(\frac{x}{x+y})^{m-1}}{(1+(\frac{x}{x+y})^m)^2} * \frac{\mu1*y}{C2} = \frac{\mu2*B^2}{B} + \frac{\mu1*C2*m*(\frac{x}{x+y})^{m-1}}{\mu2*A}$

$= \frac{\mu2^2*B^2 + \mu1*C2*A*m*(\frac{x}{x+y})^{m-1}}{\mu2*A*B}$ As in out model, $\mu1 = \mu2 = 1$ and $m > 0, C2 > 0$, we have $fracxx+y > 0$ and thus $m*(\frac{x}{x+y})^{m-1} > 0, A > 0, nominator = B^2 + C2*A*m*(\frac{x}{x+y})^{m-1} > 0$ while denominator $= A * B > 0$, which gives out $a > 0$, $b = \mu2 * P_X * P_Y * \frac{C1*n*(\frac{y}{x+y})^{n-1}}{B^2} - \mu1 * P_X * P_Y * \frac{C2*m*(\frac{x}{x+y})^{m-1}}{A^2}$

$= \frac{\mu1*C1*n*\mu2*C2*m*P_X*P_Y}{(A^2*B^2)} * (A^2(\frac{y}{x+y})^{n-1}\mu1*C2*m - (B^2(\frac{x}{x+y})^{m-1}\mu2*C1*n), = \frac{\mu2*C1*n*A^2*(\frac{y}{x+y})^{n-1} - \mu1*C2*m*B^2*(\frac{x}{x+y})^{(m-1)}}{\mu1*C2*m*\mu2*C1*n}$

nominator = $\frac{dx}{dy} - \frac{dy}{dx}$ while denominator > 0

Thus, P1 is stable if and only if $\frac{dx}{dy} > \frac{dy}{dx}$, here means the production of hax1 is faster than HS_1.

[2] for P2(x, y, 0, 0), where x and y are the solution of $x = \frac{C1}{\mu1(1+(\frac{y}{x+y})^n} = -y = -\frac{C2}{\mu1(1+(\frac{x}{x+y})^m}$,and $P_X = P_Y = 0$

Thus, for P2, a = b = 0. It's unstable.
[3] for P3(0, 0, 0, 0), same as P2, a = b = 0 and it's unstable.

3.2 HMC dynamics

In the second part here, with regard to the detailed behavior of mRNA and protein dynamics, we look into their momentum and numbers with 6 status(only the first2(a)-2(c) and the last 2(d)-2(f) transition examples of the origin 3groups*5transition statuses figures) in all are studied detailedly while the whole data based on 15 status. As we only investigated the positive direction, the red ones(top left) the application on clinical data while green one(top right) in the larger scaled simulation with more transition status(blue dashed line is the predicted dynamics).Note that the persisters and normals are the roles they take in the whole process(considering from bifurcation to catastrophe and extinction) where here they can be all considered as promoters as their numbers both grows in this process until the last status as their interaction in constant environment is of our main interest as we mentioned before. Generally, with small change studied in one status, the trend is more significant than the larger scale transition. For instance, the green simulation are always more sensitive to the momentum change and shows them more significantly on the cell trajectory comparing to the red clinical transition(we manually break one clinical status into sub-status in simulations.)
Specifically, in the$1->2$ transition, the production of the HS_1 is slightly faster than $hax1$ with the accelerate from faster to slower as well as the $hax1's$ momentum decreases from fast to slow while $HS_1's$ momentum increases from fast to slow similarly. The larger scaled simulation show the trend similarly but with larger momentum difference and thus gives out the curve trajectory instead of straight line in the top left figure; On contrary, in $2->3$ transition,both the clinical application and larger simulation give totally the same behavior according to the dynamics , where proteins products faster than mRNA but with similar acceleration. Other transition can be similarly analysis. Note that from the $4->5$ of the larger scale simulation, there starts to show the switching where the protein changes into persisters with degradation instead of production which can be both detected from cell numbers figure in the top left and momentum figures in the right bottom although the fewer status contained clinical data does not show this behavior yet. In the last transition status, the switching of proteins becoming into persister is detected in both clinical process and simulation, where in the clinical data, the momentum change of proteins and mRNA are both linear process while in simulation, the momentum of the mRNA grows slightly from faster to slower and proteins degradate slightly from faster to slower as well and in the last short time, proteins go back to normals again which according to the rising number change in the top left and increase in the momentum both relatively to mRNA(left bottom) and absolutely (right bottom.))
In the second series of figures2(c),2(f), we compute the mean time to switch approximation with the solution based on mapping to their difference space where we choose the object as 1) sigle population of mRNA to the end of the transition(top left);2) sigle population of proteins to the end of the transition (top right); 3) mRNA population to the end status of protein(left bottom) and 4) proteins population to the end status of mRNA.(right bottom.)There gives some different patterns, as in the $1->2$, both the mRNA and proteins has the mean time to switch increase linearly with their number change while there exists one significatly longer time at 0.4 for the proteins compare to the final status of mRNA and one totally unstable transition recorded; In the $2->3$, all the MTS increase linearly with the cell number growth; In the $4->5$, as there exhists the decrease of proteins thus there exhists one negative MTS stands for the status; And in $5->6$, the last status for the proteins again, compared to the final status where the number back to increase, the previous degradation status also leads to the minus MTS but positive to the mRNA as they both grow in the end.
In the last part, **Further application using the transition matrix of the model**, we compute some basic markov chain quantities based on the stochastic process as following with the pre-computation result(in availability):

3.3 Conclusion

In general,Hamiltonian markov chain advantage over the markov chain random walk with its faster convergence. As in 2(g) and 2(h), the convergence(variation to mean) of the markov chain hamilton is in blue line and the red line for clinical data and simulation on more possible transition status, giving different convergence but similar phase interval(according to 2(i)), intrestingly. The last status transition converge the worst followed by the fist transition. And the result simulated with more markov chain status converges better than the clinical results. And according to the convert rate, the mRNA to Protein transfer ratio should be the highest when starting, and goes especially lower in the last two status which is in assistance to the protein binding as we cut off the process around the convergence point where the two population has reached metastability $F_M[5]$. According to the simulation result, the protein has gone through the switching process changing from normals to persisters and back to normals(bursts in optimal time in 2(h) might also due to the switch.). Mean while, as the second population providing food(protein) to the other's binding site and either activiate or deactivate it, it works as the extrinsic noice induced the excitability or exhibition of the other gene. Here, as we choose hax1 and HS_1, they work as promoters for each others. One noticable computation is the reward computaiton based on stochastic model selection which is useful in predict the possible status of the cell numbers easily with precomputation. And we can consider correct the transition matrix with simualted clinical tested results to improve the prediction as well. On the other hand, the most important calculation action potential is easier to be achieved through Hamilton as we proved with geometric minimum action and stochastic approximation. Ohter methods can cover

---

The first step of the further computation of rewards of continuous Markov Chain is to prepare the matrix as follows: use conversion rate computed as transition rate in matrix R: R =

$$\begin{bmatrix} 0 & \frac{ConvertRate1}{A} & 0 & 0 & \frac{ConvertRateL5}{B} & 0 \\ \frac{ConvertRateL1}{B} & 0 & \frac{ConvertRate2}{A} & 0 & 0 & 0 \\ 0 & \frac{ConvertRateL2}{B} & 0 & \frac{ConvertRate3}{A} & 0 & 0 \\ 0 & 0 & \frac{ConvertRateL3}{B} & 0 & \frac{ConvertRate4}{A} & 0 \\ 0 & 0 & 0 & \frac{ConvertRateL4}{B} & 0 & \frac{ConvertRate5}{A} \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

where A = ConvertRate1:5; B = ConvertRateL1:5; we have: R =

$$\begin{bmatrix} 0 & 0.0095 & 0 & 0 & 0 & 0 \\ 0.9634 & 0 & 0.0015 & 0 & 0 & 0 \\ 0 & 0.0001 & 0 & 0.0005 & 0 & 0 \\ 0 & 0 & 0.0002 & 0 & 0.8678 & 0 \\ 0 & 0 & 0 & 0.0362 & 0 & 0.1207 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Then, we get the approximated marginal distribution by summation of R: E = sum(R, 2), and the embedded probability: $Pemb = -\frac{R}{repmat(E',6,1)'} + diag([E'])$ =

$$\begin{bmatrix} 0.0095 & -0.9999 & 0 & 0 & -0.0001 & 0 \\ -0.9985 & 0.9649 & -0.0015 & 0 & 0 & 0 \\ 0 & -0.1658 & 0.0006 & -0.8680 & 0 & 0 \\ 0 & 0 & -0.0002 & 0.8680 & -0.9998 & 0 \\ 0 & 0 & 0 & -0.2309 & 0.1570 & -0.7691 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$Q = R + diag([E']) = \begin{bmatrix} 0.0095 & 0.0095 & 0 & 0 & 0 & 0 \\ 0.9634 & 0.9649 & 0.0015 & 0 & 0 & 0 \\ 0 & 0.0001 & 0.0006 & 0.0005 & 0 & 0 \\ 0 & 0 & 0.0002 & 0.8680 & 0.8678 & 0 \\ 0 & 0 & 0 & 0.0362 & 0.1570 & 0.1207 \\ 0 & 0 & 0 & 0 & 0 & 2 \end{bmatrix}$$

$$P_{unif} = eye(size(Q)) + \frac{Q}{max(E)} = \begin{bmatrix} 1.0095 & 0.0095 & 0 & 0 & 0 & 0 \\ 0.9634 & 1.9649 & 0.0015 & 0 & 0 & 0 \\ 0 & 0.0001 & 1.0006 & 0.0005 & 0 & 0 \\ 0 & 0 & 0.0002 & 1.8680 & 0.8678 & 0 \\ 0 & 0 & 0 & 0.0362 & 1.1570 & 0.1207 \\ 0 & 0 & 0 & 0 & 0 & 3 \end{bmatrix}$$

1)$takingI0 = [0,2], froms0tos4:$
$Prob = P_{emb}(1,5) * (exp^{(-E(1)*0)} - exp^{(-E(1)*4)}) =$
$[0.0095, 0.9649, 0.0006, 0.8680, 0.1570, 1];$
2)$Prob(trueU[0,2], s4) : stat = max(E) * 2$
$n = 4;$
$Probc = 0$
$FOR : i = 1 : n$
$Probc = Probc + exp^{\frac{(-stat)*stat(i)}{factorial(i)}}$
END
$Probc = Probc * [0,0,0,0,1,0]' = [0,0,0,0,0.8120,0];$
3) The number of cells expected after 6 time units have inactivated can be given by
$Exp^C(s, X_c <= t) = \frac{1}{q} * \Sigma_{i=0}^{\inf} \exp^{(-qt)} * \frac{(qt)^i}{factorial(i)} * (P_{unif})^i * (q * P_{unif} * [1,1,1,1,1,1,1]')$

$$Probc2 = 1.0e+03 * \begin{bmatrix} 0.0487 & 0.0024 & 0 & 0 & 0 & 0 \\ 0.2475 & 0.2941 & 0.0004 & 0 & 0 & 0 \\ 0 & 0.0000 & 0.0464 & 0.000 & 0 & 0.0001 \\ 0 & 0.0001 & 0.0000 & 0.2555 & 0.2258 & 0.0393 \\ 0 & 0 & 0 & 0.0094 & 0.0704 & 0.0827 \\ 0 & 0 & 0 & 0 & 0 & 1.3212 \end{bmatrix}$$

Probc2 =1.0e+03 *[0.0526, 0.6184, 0.0471, 0.6051, 0.1983, **1.8000];**
Thus, if we want to know:the status when after 6 unit times products mRNA cells over 1000 the only satisfied status is the last one which might product 1800 mRNAs.

Hamilton Jacobian matrix, WKB and etc.[21]. As we also improve the algorithm with adding hierarchical markov in calculating number of cells in different status and only record 'sucessful move' according to the tolerance based on action potential and residule of predition numbers both, the convergence of the algorithm is guaranteed. And further research can be conducted on the whole process from bifurcation to catastrophe and extinction as well. Problem with multi population is also possible. As hax1 is observed to have function in signaling and regulating of genes especially in learning systems and motor related brain function, this switching model study related to its binding might help to predict the cell numbers and production or degradation rate especially later with further study into both with promoters and persisters as to test different drug and their efficiency on the aging process related disease.

Appendix

A figures

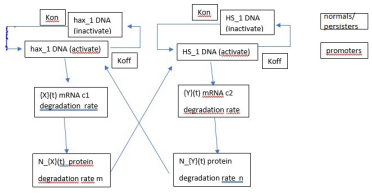


Figure 1: Caption

B Proof

As our assumption of diffusion process is based on the proposition of the geometric quasi-potential:

$$V(x1,x2) = 2 * inf \int_{\gamma} |b|_a sin^2(\frac{1}{2})\eta ds \tag{4}$$

which later can be used in computing quasi-potential under the case without SDE: $V(x1,x2) = inf_{\phi \in C^{x2}_{x1}(0,1)} S(\phi) with S(\phi) = sup^{\theta:[0,1]->R^n}_{H(\phi,\theta)=0} \int_0^1 < \phi', \theta > d\alpha$ As we have proved the most important invariance of the key four representation of the action S through the same value of upper bound and lower bound of S. The only supplementary required is after the rewrite of V(x1, x2) with S, the inf is computed with the lower-semicontinuity of S and the compactness of certain functions: Consider the functions $\phi_k \in C_{\phi_k}(0,1)$ satisfies the T_k and ψ_k such that with some normalized unit speed($|\phi'_k| = L_k(0,1)$),they cross through the curves and the L_k is the length of the curve ψ_k: On $[0, T_k]$->[0,1], α_k is defined as $(\frac{1}{L_k}) int'_0 |d\phi_k(t)| d\tau$ here $L_k = int^{T_k}_0 |d\phi_k(\tau)| d\tau$ and with inverse as $inf t \in [0,1] | \alpha_k(t) >= \alpha$, and the set $d\phi_k(\alpha_k(t)) = \psi_k(t)$ for all $t \in [0, T_k]$, and both $\alpha'_k = \frac{|d\phi_k|}{L_k}$ and $L_k = |\phi'| = \frac{|d\phi|}{\alpha'_k} o d\alpha^{-1}_k$. so that $\gamma(\phi_k) = \gamma(\psi_k)$, i.e., $\psi_k \in C_{\phi_k}(0, T_k)$. With the first equation of the four we proved, we have the estimate of S as: $int_{\phi \in C^{x2}_{x1}(0,1)} S(\phi) <= S(\phi_k) = inf_{T>0} \psi \in C_k(0,T) S_T(\Psi) <= S_{T_k}(\psi_k)$ where when k -> inf, the right-hand side converges to the left-hand side, and it follows that

$$lim_{k->\inf} S(\phi_k) = inf_{\phi \in C^{x2}_{x1}(0,1)} S_T(\phi) \tag{5}$$

With $M = sup_k |\phi'_k| = sup_k L_k <$ inf, and thus ϕ_{k_l} that converges uniformaly to one limiting function $\phi * \in C_{x1}, M$, by we have:

$$S_(\phi*) = inf_{\phi \in C^{x2}_{x1}(0,1)} S(\phi); \tag{6}$$

and thus $\phi/*$ is a minimizer of S. Since the functions ϕ_{k_l} are time-rescaled ϕ_{k_l} and converge uniformly to $\phi_k *$, leading to

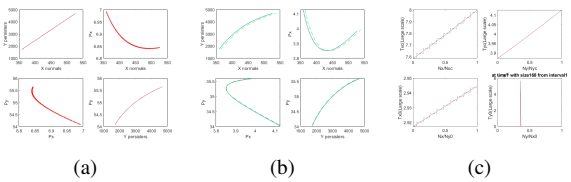
$$\rho(\phi_{k_l}, \phi*) = \rho(\phi_{k_l}, \phi*) <= |\phi_{k_l} - \phi * |_{[0,1]} - > 0 asl - > \inf \tag{7}$$

In addition, with the minimizer of S unique to reparameterization and let $\phi*$ be the limit of some converging subsequence of $(/phi_k)$from the assumption that $\rho(\psi_k, \phi*) - > 0 ask - > \inf$ and with ϕ_{kl} s.t. $inf_{l \in N} \rho(\phi_{kl}, \phi*) > 0$ enables its subsequence $\phi_{k_{lm}}$ that converges in Frecht metric to some limit ϕ that is a minimizer of S. And thus, $\gamma(\phi) = \gamma(\phi_{k*})$ where $/rho(\psi_k lm, \phi*) = \rho(\psi_{K lm}, \phi) - > 0$ as $m - > /inf$, which contradicts. Some basic quantities of Hamiltonians: From our model after the Hamiltonian as follows: $\frac{C1}{1+(\frac{y}{x+y})^n}(expP_X - 1) - \mu 1 * x(exp - P_x - 1) + \frac{C2}{1+(\frac{x}{x+y})^n}(expP_y - 1) - \mu 2 * y(exp - P_y - 1)$ we calculate

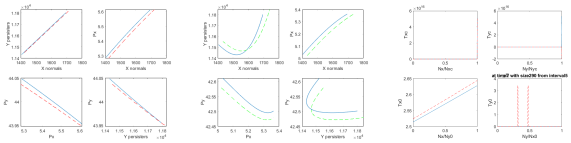
H_X, H_Y, H_{P_X}, H_{P_Y}, P_X, P_Y, HX, HY $H_X = \mu 1 * (exp - P_X - 1) - \frac{C2 * m * (\frac{x}{x+y})^{m-1}}{(1+(\frac{x}{x+y})^m)^2}(expP_Y - 1)$

$H_Y = \mu 2 * (exp - P_Y - 1) - \frac{C1 * n * (\frac{y}{x+y})^{m-1}}{(1+(\frac{y}{x+y})^n)^2}(expP_X - 1)$ $H_{P_X} = \frac{C1}{1+(\frac{y}{x+y})^n} expP_x - \mu 1x * (exp - P_X$

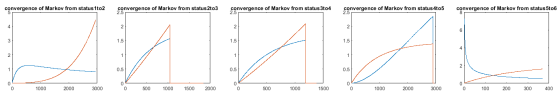
$H_{P_Y} = \frac{C2}{1+(\frac{x}{x+y})^m} expP_y - \mu 2y * (exp - P_Y$ $HX = \frac{C1}{1+(\frac{x}{x+y})^m}(expP_x - 1) + \mu 1 * x(exp - P_x - 1)$



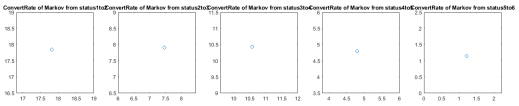
(a) (b) (c)



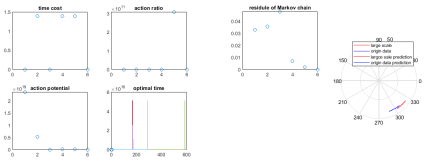
(d) (e) (f)



(g)



(h)



(i) (j)

$HY = \frac{C2}{1+(\frac{y}{x+y})^n}(expP_y - 1) + \mu 2 * y(exp - P_y - 1)$ $dx = \frac{C1}{1+(\frac{y}{x+y})^m} expP_x - \mu 1 * x(exp - P_X)$

$dy = \frac{C2}{1+(\frac{x}{x+y})^n} expP_y - \mu 2 * y(exp - P_Y)$

$dP_X = \frac{C2 * m * (\frac{x}{x+y})^{m-1}}{(1+(\frac{x}{x+y})^m)^2}(\exp P_Y - 1) + \mu 1 * (exp - P_X) - mu1$

$dP_Y = \frac{C1 * n * (\frac{y}{x+y})^{n-1}}{(1+(\frac{y}{x+y})^n)^2}(\exp P_X - 1) + \mu 2 * (exp - P_Y) - mu2$ Finally, because of the approximation with diffuse process[15], the Hamiltonian quantities are: $H_{theta}(a,b) = a(x)^{-1}(\frac{|b(x)|_a}{|y|_a} * y - b(x))$ $\lambda(x,y) = \frac{|b(x)|_a}{|y|_a}$

As a result, $S(\phi)$ can be calculated explicitly with the Hamiltonians in a diffuse process as following: $S(\phi) = \int_0^1 (|\phi'|_a |b(\phi)|_a - < \phi', b(\phi) > a) d\alpha$

Acknowledgments

Although I have graduated with excellence graduates awards,I am still working in the university currently cooperated with university hospital. Thanks to all the supervisors and colleagues I have worked with.

Availability

Please find supplementary data, code and experimental results in this link: *https : //github.com/LilyHeAsamiko/Biophysics/tree/master/usenix*

Discussion

(a) "what kind of feedback you are looking to receive" is the topic I want to add here. As my major is machine learning and although I have my research from bachelor to master both related to neuroscience and working with clinical data for a while mainly with Bayes Inference Model, genetic field is still quite new to me and I do know it is a heated field either with genotypical or phenotypical study. I would really like to know how often do predictive model, especially those based on causality or reinforcement machine learning are utilized in biology related problem. Meanwhile, how practical do you think my model can be conducted and utilized? What kind of question do you hold after listening to my presentation. I would appreciate a lot if there are professionals who would like to give some similar experience or specific project example on how they deal with the similar problems. For instance, just about the protein binding, what model and how they applied those on their clinical data. Thank you!

References

Elijah Roberts, Shay Be er, Chris Bohrer, Rati Sharma, Michael Assaf, Dynamics of simple gene-network motifs subject to extrinsic fluctuations, 2015

Freedman,H.I.,Dynamics of simple gene-network motifs subject to extrinsic fluctuations,1980

Noufe H. Aljahdaly,Analytical Solutions of a Modified Predator-Prey Model through a New Ecological Interaction,2019

ensembl,https://www.ensembl.org/index.html

Ingo Lohmar, Baruch Meerson, Switching between phenotypes and population extinction, Racah Institute of Physics,2018

Marta Kwiatkowdka, Gethin Norman, David Parker, Stochastic Model Checking, Oxford, 2019

David Martínez-Rubio, Varun Kanade, Patrick Rebeschini, Decentralized Cooperative Stochastic Bandits, Oxford, 2005

Paul C Bressloff, Stochastic swiching in biology: from genotype to phenotype, Department of Mathematics, University of Utah, 2017

Baier, Christel, Principles of Model checking, Massachusetts Instititute of Technology,2008

Cheng Lv, Xiaoguang Li, Fangting Li, Tiejun Li, Constructing the Energy Landscape for Genetic Switching System Driven by Intrinsic Noise, Peking University, 2014

C.W. Gardiner, Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences, 4th ed, Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences, 4th ed.,Springer, 2009

Yasar Demirel, in Nonequalibrium Thermodynamics(Third Edition), 2014

Matthias Heymann, Eric Vanden-Eijnden, Geometric Minimum Action Metod: A Least Action Principle on the Space of Curves, Courant Institute, 2007

authorRobert Balson Dingle, Their Derivation and Interpretation, Academic Press, 1973

J.C., Gittins, Keble College, Oxford, 2010

Ali, Isra, Alfarouk, Khalid O., Reshkin, Stephan J., Ibrahium, Muntaser E., Doxycycline as Potential Anti-cancer Agent, Anti Cancer Agents, 2018

G.S. Skone, Irina Voiculescu, Stratagems for effective function evaluation in computational chemistry, Oxford, 2010

Zahid Ur Rehman, Quorum-Quenching Bacteria Isolated From Red Sea Sediments Reduce Biofilm Formation, KAUST,2016

Peter Minary, Charlotte M Deane, Explorong peptide/MHC detachment process using hiearchical natural move Monte Carlo, Oxford, 2015

Biancalani, E. Giamperi, A. Bazzani, G. Catellani, and A. Maritan, Phys.Soc, Jpn., Physics, 2015

D.M. Roma, R.A. P'Flanagan, A.E. Ruckenstein, A.M. Sengupta, R., Optimal path to epigenetic switching, Rarah Institute of Physics,2015