Branching and oscillations in the epigenetic landscape of cell-fate determination

Authors: Jomar Fajardo Rabajante[a,*], Ariel Lagdameo Babierra[a]

[a]Institute of Mathematical Sciences and Physics, University of the Philippines Los Baños, College, Laguna 4031 Philippines

[*]Corresponding author. E-mail address: jfrabajante@up.edu.ph. Present address: Shizuoka University, Hamamatsu, Japan.

Summary. Waddington's epigenetic landscape is a famous sketch of cell-fate determination. This landscape is not static but varies because of the dynamic kinetics of gene regulation during development. Mathematical models of gene regulation usually contain fixed parameters and limited number of state variables, hence cannot fully characterize the landscape's temporal transformation. Here we simulate a high-dimensional model of gene regulation with time-varying repression links among gene regulatory factors. We are able to show sequential multi-lineage differentiation at different timescales that portrays the branching canals in Waddington's landscape. We also show that sustained oscillations by a repressilator-type network can activate suppressed genes while cells trail a flattened landscape. Gradient-based dynamics dampen the oscillations resulting in dedifferentiation to multipotency or pluripotency. Our theoretical results provide insights into the possible mechanisms of cell differentiation and associated diseases, such as cancer, as well as contribute in devising strategies for cellular reprogramming.

Keywords. gene regulation, stem cells, induced pluripotency, synthetic biology, evolution, cancer

Main text:

Waddington's epigenetic landscape illustrates the canalization in the cell differentiation and fate determination process[1-4]. The topography of Waddington's illustration represents the developmental pathways of tissues formed from totipotent and pluripotent cells to terminally-differentiated specialized cells (Fig. 1a). Various theoretical studies have quantified Waddington's epigenetic landscape and are able to predict bistability in gene regulatory networks (GRNs)[5-9]. However, many of the mathematical models consider only two regulatory factors and focus on static epigenetic landscape represented by fixed parameter values. In reality, the topography of Waddington's illustration is dynamic and high-dimensional, and the parameters that represent gene regulation are changing during the development of an organism[10-16]. Mathematical models with two regulatory factors and fixed parameter values only describe a particular temporal scenario in cell differentiation.

The mechanisms that regulate gene expression, such as kinetics of gene regulatory factors (GRFs) and the structure of GRF-GRF interaction, influence the outcome of cell-fate determination[7,10,11,14,15,17]. Waddington observed that changes (e.g., mutations) in these mechanisms can modify the epigenetic landscape leading to cell-lineage switching[1]. Mathematically, these modifications can be represented by variations in the parameter values of the quantitative models. Bifurcation analyses of existing models have been done[5,7,11,15,16], but most of them do not provide elaborate illustrations of cells trailing the high-dimensional dynamic pathways. Here we present numerical illustrations of cells trailing different epigenetic routes such that the pathways transform due to changes in the strength of repressive interaction among multiple GRFs (Box 1).

Our main assumption is that gradient-based optimization[18,19] governs the transformations of the pathways leading to the essential cell type (cell fate that is essential for proper normal development, or desired cell type during cellular engineering). This

assumption assures that the cells trail the epigenetic landscape via the steepest pathways shaped by the time-varying antagonistic interaction among GRFs (Fig. 1b). The gradient-based method can be considered as a cell-fate induction strategy such that the cells move from a pluripotent state, which has higher entropy, towards differentiated state with lesser entropy[20]. We show that the dynamic GRF-GRF interaction can illustrate the cascade of branching canals in Waddington's illustration. It can also describe cell plasticity by allowing cell-lineage switching, such as by transdifferentiation and dedifferentiation.

Box 1: Mathematical model

We consider a gene regulatory network of the form shown in Figure 1b, where the number of nodes (n) is arbitrary[10,24,29,36]. One of the simple high-dimensional models that can represent the dynamics of this GRN is the following system of differential equations[10,12]:

$$\frac{dX_i}{dt} = \frac{\beta X_i^{\,2}}{1 + X_i^{\,2} + \sum_{j \neq i} \gamma_{ij} X_j^{\,2}} - \rho X_i + g, i = 1,2,...,n. \qquad \text{(Box Eq. 1)}$$

The state variable X_i represents the strength or concentration of the GRF involved in expressing the gene towards the i-th cell type. The parameters β, ρ, g and γ_{ij} are the efficiency of GRF in expressing the corresponding gene, degradation rate, basal constitutive growth rate, and time-varying coefficient associated with the inhibition of X_i by X_j, respectively. The kinetics of GRF auto-activation is a sigmoidal increasing function which is negatively influenced by the strength of the other GRFs.

Furthermore, we assume that the goal of gene regulation is to maximize the strength of all GRFs so that the outcome is moving towards the steepest canal and possibly towards the deepest valley in the landscape. Here we apply a gradient-based optimization method to the time-varying interaction coefficient γ_{ij} (see Methods).

The equilibrium points and sustained oscillations of this model lie on the hyperspace

$$\left[\frac{g_1}{\rho_1},\frac{g_1+\beta_1}{\rho_1}\right)\times\left[\frac{g_2}{\rho_2},\frac{g_2+\beta_2}{\rho_2}\right)\times\cdots\times\left[\frac{g_n}{\rho_n},\frac{g_n+\beta_n}{\rho_n}\right), \rho_i>0 \text{ for all } i. \qquad \text{(Box Eq. 2)}$$

The model has at most 3^n equilibrium points[22].

The qualitative dynamics arising from the model are possible to arise in a more complex gene regulatory system. This model is straightforward so that any peculiar dynamics can be clearly interpreted. However, this model is an abstraction of the cell differentiation process. Hence, adjustment should be done before applying the model to a particular scenario.

(End of Box 1)

Results. There are two significant dynamics that we observe in our simulations. The first one is multi-lineage differentiation via sequentially branching developmental pathways. This sequential branching portrays the canalization in Waddington's landscape at different timescales. The pathways trailed by the differentiating cells depend not only on the structure of the GRN and parameter values but also on the initial condition (see Supplementary Fig. 1). The second one is flattening of the epigenetic landscape which eliminates the deep valleys, resulting in sustained oscillations. The GRN that generates this oscillatory behavior is an attracting oscillator where it can lift the strength of a suppressed GRF.

Mathematically, the sequential branching in the epigenetic landscape towards different cell types represents convergence to one of the equilibrium points (Supplementary Fig. 1). The variations in the interaction coefficients drive changes in the topography of the landscape (Supplementary Fig. 2) and eventually stabilize at different timescales, hence sequential branching become possible (refer to the timescale factor in Methods). The pathway bifurcates from one branch (a primed state) to multiple branches, which may further have sub-branches that culminate in branch endpoints (Figs. 2a, 2c-2e). The final structure of the GRN, which is determined by the parameters stabilized at different timescales, dictates the

number and location of the branch endpoints (Figs. 2a-2e). Generally, in order for sequential branching to arise, the initial structure of the GRN and parameter values should allow bistability or multistability which is a property of multipotent and pluripotent cells (Supplementary Fig. 1). The entropy and the quasi-potential of the landscape are lower in differentiated cells compared to the undifferentiated state as expected (e.g., Fig. 2f and see the quasi-potential axis in Figs. 2a and 2b). In addition, the pathways with two endpoints that are distant from each other are usually more robust against stochastic noise compared to the pathways with many endpoints (Supplementary Figs. 3, 6, 8 and 10). This implies that cells trailing the pathways with more endpoints can be candidates for stochastic direct reprogramming.

There are cases where a stable equilibrium point vanishes and a stable limit cycle emerges, especially when the repressive interaction among GRFs is asymmetric. This stable limit cycle is generated by an oscillator that attracts suppressed genes in partially or terminally-differentiated cells, resulting in activation with fluctuating kinetics (Figs. 3b and 3c). One example of an oscillator is a repressilator-type network (Fig. 3a), which is similar to the repressilator proposed by Elowitz and Leibler[21]. In this repressilator-type network, the strength of repression in one loop is stronger than the strength of repression in the reverse loop. The sustained oscillations generated by this repressilator-type network can arise in a GRN with three or more nodes (odd or even number of GRFs; e.g., Figs. 3b and 3c). The dynamics of this oscillator can be illustrated by an epigenetic landscape with flattened topography, that is, there are no deep valleys in the route of the differentiating cells, and the cells are continually sliding in zigzag canals without endpoint.

The oscillations drive the strengths of the GRFs to have alternating positive and negative rates, whereas the gradient-based optimization method forces the dynamics towards positive rates only. Thus, oscillatory behavior is not optimal in the sense of differentiation

towards cell types located at deep valleys in the landscape. We then expect that gradient-based dynamics which are persistent for some period result in damped oscillations (Figs. 3d and 3e). There are cases where the damped oscillations illustrate multipotency (or pluripotency depending on the GRN), which is represented by the equal probabilities of differentiating towards all the considered cell types (Fig. 3d). In some cases, the interaction coefficients vary with different timescales, resulting in partial differentiation and sometimes in the reversal of the status of the initially dominant GRF (Fig. 3e). Note that if the dampening of the oscillations is fast, the initial oscillations can be unnoticeable yet can still activate suppressed genes (Supplementary Fig. 19).

Discussion. Various studies have attempted to model the cell differentiation process, but there are still more to uncover in epigenetics. Further theoretical prediction and experimental validation are needed to fully explain cell-fate determination and reprogramming. Varying the efficiency of GRF in expressing a gene (β), the degradation rate (ρ), or the constitutive growth rate (g) is a straightforward technique in stimulating the activation or deactivation of a GRF and its corresponding gene[6,7,10,17, 22]. However, regulating the repression strength of GRFs (γ_{ij}) has not been explored, and we have shown numerical illustrations where variations in these GRF-GRF repression affect the qualitative behavior of the cell differentiation system. We are able to replicate Waddington's model using a single set of equations with many GRFs involved. A GRN with only two nodes generally cannot describe the sequential bifurcation of canals and the oscillations in cell-fate determination. The different timescales involved in gene interaction influence the outcome of cellular regulation[5,23-26].

One of the aims of this study is to spur more discussions on non-equilibrium dynamics and oscillations arising from high-dimensional asymmetric systems, which can broaden our understanding about the mechanisms of gene regulation. Reversal of the route

from differentiated state to pluripotency is previously thought to be impossible but now many dedifferentiation techniques have been proposed and tested by experiments[9,27,28]. We propose another alternative technique for cellular reprogramming, which is by rewiring the GRN to have a repressilator-type network, possibly with the aid of external stimulus and stochastic noise[17]. External stimulus can be introduced to weaken the repression in one loop (Fig. 3a). Our numerical predictions can help design cellular engineering strategies for generating induced multipotent stem cells (or pluripotent stem cells depending on the GRN[17,29,30]) using the oscillations that can activate silenced genes. Indeed, oscillating GRF expression is an attribute of progenitor cells[31-33]. However, note that in reprogramming back to pluripotency, we also need to assure activation of defined factors, such as *Oct4*, *Sox2* and *Nanog*[17,27].

The oscillator motifs (e.g., repressilators) which are part of a larger GRN contribute to the fluctuations observed in gene regulation dynamics. In fact, there are many types of oscillators[34-36]. The oscillations generated by each oscillator motif when combined are often interpreted as stochastic noise or as chaos. However, note that the combined large and small oscillations are not entirely stochastic fluctuations, especially when the detected noise is part of the gene regulation system and not just coming from random sources[37,38]. In addition, the oscillator motif of a larger GRN can be used for artificial transdifferentiation by generating oscillations that can prompt cell-lineage switching, similar to what stochastic fluctuations can do[38,39]. Transdifferentiation between related cell types branching from one lineage can be more straightforward compared to dedifferentiation to pluripotency[40].

From our simulations, we formulate some conjectures: (i) The dedifferentiation caused by abnormal oscillators (e.g., aberrant repressilator-type network) play a role in the existence of cancer stem cells and mutator phenotype[41-45]. Abnormal changes in the structure of GRF-GRF interaction, such as abnormal timescale factor and abnormal weakening of repression links, can lead to disease. Indeed, partially reprogrammed cells and excessive

plasticity can cause cancer[40,44,46,47]. Moreover, it is also possible that these oscillators play a part in epigenomic reprogramming and influence transgenerational epigenetic inheritance[48]. Abnormalities in the GRF-GRF interaction could be passed-on to offspring. (ii) We can reprogram cells back to pluripotency by regulating the wiring of the GRN. This implies that there are no unique reprogramming factors, and we can reprogram cells using any regulatory factor as long as it can lower the "gravity" of the epigenetic landscape[40]. One approach to demonstrate our numerical predictions is by employing synthetic biology techniques[21,49].

In reality, the temporal transformation in the epigenetic landscape are due to multiple intrinsic and extrinsic factors. Here we only consider changes in the GRF-GRF interaction coefficient γ_{ij} but we should not disregard that gene regulation consists of the interplay among many factors and processes. For example, GRF-GRF interaction can be regulated not only through γ_{ij} but also through the modifications in the maximal growth rate (β+g) or through the degradation rate (ρ)[6,7,10,17,22]. Increasing the maximal growth rate or decreasing the degradation rate of a certain GRF enhances the steady state strength of the GRF, which in turn intensifies the repression of the other GRFs. Furthermore, the dynamics observed from empirical data combine the various effect of many parameters. For example, a decline in the strength of a GRF suggests various possible reasons, such as due to an increased degradation rate or due to an increased repression by an antagonist GRF. Hence, we need to interpret data by considering all possible factors.

In summary, our simulations predict the following outcomes: First is the sequential branching of lineages in cell-fate determination that portrays differentiation from pluripotent state to transient states (lineage progenitors) towards specialized cell types. Second is the dedifferentiation driven by oscillations generated by a repressilator-type network, which is a possible mechanism for reprogramming cells back to pluripotency. A two-variable switch-like model usually cannot illustrate the branching phenomena and sustained oscillations, but a

high-dimensional model with asymmetric reciprocal interaction between GRFs can. Oscillatory behavior cannot be taken for granted because this could explain peculiar dynamics related to the epigenetic machinery of organisms, such as dedifferentiation as part of regenerative process. Oscillations are also involved in pattern formation, circadian rhythms, and the progression of diseases[32,35,39,50-54]. Network motifs, such as the repressilators, that are part of a larger GRN induce functional fluctuations necessary for tissue development and cellular engineering. Investigating the dynamics of these network motifs can be helpful in drug discovery[55].

Methods. In our simulations, we use the following differential equation model (see Box 1):

$$dX_i = \left(\frac{\beta X_i^{\,2}}{1 + X_i^{\,2} + \sum_{j \neq i} \gamma_{ij} X_j^{\,2}} - \rho X_i + g \right) dt + \sigma_A dW, i = 1,2,...,n. \quad (1)$$

Note that we restrict our simulations to specific parameter values such as $\beta=1$, $\rho=0.05$ and $\gamma_{ij} = {}^{a_{ij}}\!/_{1+u_i^{\,2}}$. We suppose all GRFs have the same value of β and ρ to highlight the effect of time-varying γ_{ij}. The term $\sigma_A dW$ represents Gaussian white noise with amplitude σ_A. Let $\sigma_A=0$ and $\sigma_A=0.5$ for deterministic and stochastic simulations, respectively. The noise term approximates multiple heterogeneous sources of additive random fluctuations.

Time evolution of the interaction coefficient. The value of $\gamma_{ij} = {}^{a_{ij}}\!/_{1+u_i^{\,2}}$ is updated using the following gradient function:

$$du_i = exp(-\varepsilon_i t)\left\{\frac{\partial}{\partial u_i}\left(\frac{\beta X_i}{1+X_i^{\,2}+\sum_{j\neq i}\frac{a_{ij}}{1+u_i^{\,2}}X_j^{\,2}}\right)\right\}dt+\sigma_u dW \qquad (2)$$

$$\approx \frac{exp(-\varepsilon_i t)}{2\Delta}\left(\frac{\beta X_i}{1+X_i^{\,2}+\sum_{j\neq i}\frac{a_{ij}}{1+(u_i+\Delta)^2}X_j^{\,2}}-\frac{\beta X_i}{1+X_i^{\,2}+\sum_{j\neq i}\frac{a_{ij}}{1+(u_i-\Delta)^2}X_j^{\,2}}\right)dt+\sigma_u dW.$$

Equation (2) is used for finding relative optimum and is similar to the trait dynamics frequently used in evolutionary biology[19]. Increasing the value of u_i decreases the value of repression coefficient γ_{ij} ($j\neq i$). Hence, the variable u_i can be defined as the time-varying attribute for maximizing the strength of the GRF X_i. However, note that increasing the value of u_i does not always result in an increased equilibrium value of X_i, especially when the initial value and other parameter values do not allow significant changes in the epigenetic landscape in favor of X_i.

The timescale factor[5,11] is described by $exp(-\varepsilon_i t)$ with decline rate ε_i. As time progresses (e.g., as cell matures), the time scale factor declines and the value of u_i leads to equilibrium. In addition, the dynamic parameter u_i is initialized with value $u_i(0)=0.001$ for all i. Let $\sigma_u=0$ and $\sigma_u=0.01$ for deterministic and stochastic simulations, respectively. For simplicity, we approximate the partial derivative in Equation (2) using central difference formula with $\Delta=0.001$.

Quasi-potential, cell type probability and entropy. The quasi-potential (Φ) of the landscape is computed as

$$\frac{d\Phi}{dt} = -\sum_i\left(\frac{dX_i}{dt}\right)^2 \qquad (3)$$

where the dX_i/dt is deterministic[6,16]. The probability of differentiating to cell type i or the expected proportion of cells committed towards cell type i is $P_i = X_i/\sum_{k=1}^{n} X_k$. To visualize the

canalization in Waddington's illustration, we use three coordinate axes: time, quasi-potential, and cell type probability. We also compute for the entropy[56] defined by

$$E = -\sum_i P_i \log(P_i), P_i \neq 0. \quad (4)$$

Numerical method. The ordinary and stochastic differential equations are solved using Runge-Kutta 4 and Euler-Maruyama with 0.01 as step size, respectively. For supplementary mathematical discussions about the differential equation model (Box Eq. 1), refer to related literatures[10,17,22].

References

1. Waddington, C. H., *The Strategy of the Genes*. (George Allen & Unwin., London, 1957).

2. Graf, T. & Enver, T., Forcing cells to change lineages. *Nature* **462**, 587-594 (2009).

3. Zhou, J. X. & Huang, S., Understanding gene circuits at cell-fate branch points for rational cell reprogramming. *Trends Genet*. **27**(2), 55-62 (2011).

4. Bogdan, P., Deasy, B. M., Gharaibeh, B., Roehrs, T. & Marculescu, R., Heterogeneous structure of stem cells dynamics: statistical models and quantitative predictions. *Sci. Rep.* **4**, 4826 (2014).

5. Wang, J., Xu, L., Wang, E. & Huang, S., The potential landscape of genetic circuits imposes the arrow of time in stem cell differentiation. *Biophys. J*. **99**(1), 29-39 (2010).

6. Bhattacharya, S., Zhang, Q. & Andersen, M. E., A deterministic map of Waddington's epigenetic landscape for cell fate specification. *BMC Syst. Biol*. **5**, 85 (2011).

7. Wang, J., Zhang, K., Xu, L. & Wang, E., Quantifying the Waddington landscape and biological paths for development and differentiation. *PNAS* **108**(20), 8257-8262 (2011).

8. Ferrell Jr., J. E., Bistability, bifurcations, and Waddington's epigenetic landscape. *Curr. Biol.* **22**(11), R458-R466 (2012).

9. Shu, J. et al., Induction of pluripotency in mouse somatic cells with lineage specifiers. *Cell* **153**(5), 963-975 (2013).

10. Cinquin, O. & Demongeot, J., High-dimensional switches and the modelling of cellular differentiation. *J. Theor. Biol.* **233**(3), 391-411 (2005).

11. Huang, S., Guo, Y., May, G. & Enver, T., Bifurcation dynamics in lineage-commitment in bipotent progenitor cells. *Dev. Biol.* **305**(2), 695-713 (2007).

12. Aguda, B. D. & Friedman, A., *Models of Cellular Regulation.* (Oxford Univ. Press, NY, 2008).

13. Foster, D. V., Foster, J. G., Huang, S. & Kauffman, S. A., A model of sequential branching in hierarchical cell fate determination. *J. Theor. Biol.* **260**(4), 589-597 (2009).

14. Kim, M., Kim, J. & Cho, K., Dynamic network rewiring determines temporal regulatory functions in *Drosophila melanogaster* development processes. *Bioessays* **32**(6), 505-513 (2010).

15. Li, C. & Wang, J., Quantifying Waddington landscapes and paths of non-adiabatic cell fate decisions for differentiation, reprogramming and transdifferentiation. *J. R. Soc. Interface* **10**(89), 20130787 (2013).

16. Verd, B., Crombach, A. & Jaeger, J., Classification of transient behaviours in a time-dependent toggle switch model. *BMC Syst. Biol.* **8**, 43 (2014).

17. MacArthur, B. D., Please, C. P. & Oreffo, R. O. C., Stochasticity and the molecular mechanisms of induced pluripotency. *PLoS ONE* **3**(8), e3086 (2008).

18. Parker, G. A. & Maynard Smith, J., Optimality theory in evolutionary biology. *Nature* **348**, 27-33 (1990).

19. Khibnik, A. I. & Kondrashov, A. S., Three mechanisms of Red Queen dynamics. *Proc. R. Soc. Lond. B* **264**(1384), 1049-1056 (1997).

20. Banerji, C. R. S. et al., Cellular network entropy as the energy potential in Waddington's differentiation landscape. *Sci. Rep.* **3**, 3039 (2013).

21. Elowitz, M. B. & Leibler, S., A synthetic oscillatory network of transcriptional regulators. *Nature* **403**, 335-338 (2000).

22. Rabajante, J. F., Equilibrium switching in nonlinear interaction networks with concurrent antagonism. *PeerJ PrePrints* **2**, e382v1, DOI: 10.7287/peerj.preprints.382v1 (2014).

23. Feng, H. & Wang, J., A new mechanism of stem cell differentiation through slow binding/unbinding of regulators to genes. *Sci. Rep.* **2**, 550 (2012).

24. Nené, N. R. & Zaikin, A., Interplay between path and speed in decision making by high-dimensional stochastic gene regulatory networks. *PLoS ONE* **7**(7), e40085 (2012).

25. Alagha, A. & Zaikin, A., Asymmetry in erythroid-myeloid differentiation switch and the role of timing in a binary cell-fate decision. *Front. Immunol.* **4**, 426 (2013).

26. Sasai, M., Kawabata, Y., Makishi, K., Itoh, K. & Terada, T. P., Time scales in epigenetic dynamics and phenotypic heterogeneity of embryonic stem cells. *PLoS Comput. Biol*. **9**(12), e1003380 (2013).

27. Yamanaka, S., Elite and stochastic models for induced pluripotent stem cell generation. *Nature* **460**, 49-52 (2009).

28. Weake, V. M. & Workman, J. L., Inducible gene expression: diverse regulatory mechanisms. *Nat. Rev. Genet.* **11**(6), 426-437 (2010).

29. Crespo, I. & Del Sol, A., A general strategy for cellular reprogramming: the importance of transcription factor cross-repression. *Stem Cells* **31**(10), 2127-2135 (2013).

30. Wang, P. et al., Epigenetic state network approach for describing cell phenotypic transitions. *Interface Focus* **4**(3), 20130068 (2014).

31. Shimojo, H., Ohtsuka, T. & Kageyama, R., Oscillations in notch signaling regulate maintenance of neural progenitors. *Neuron*. **58**(1), 52-64 (2008).

32. Suzuki, N., Furusawa, C. & Kaneko, K., Oscillatory protein expression dynamics endows stem cells with robust differentiation potential. *PLoS ONE*. **6**(11), e27232 (2011).

33. Bonev, B., Stanley, P. & Papalopulu, N., MicroRNA-9 modulates *Hes1* ultradian oscillations by forming a double-negative feedback loop. *Cell Rep.* **2**(1), 10-18 (2012).

34. Purcell, O., Savery, N. J., Grierson, C. S. & di Bernardo, M., A comparative analysis of synthetic genetic oscillators. *J. R. Soc. Interface* **7**(52), 1503-1524 (2010).

35. Kamino, K., Fujimoto, K. & Sawai, S., Collective oscillations in developing cells: Insights from simple systems. *Dev. Growth Differ.* **53**(4), 503-517 (2011).

36. Radde, N., The role of feedback mechanisms in biological network models. *Asian J. Control*. **13**(5), 597-610 (2011).

37. Süel, G. M., Kulkarni, R. P., Dworkin, J., Garcia-Ojalvo, J. & Elowitz, M. B., Tunability and noise dependence in differentiation dynamics. *Science* **315**(5819), 1716-1719 (2007).

38. Chalancon, G. et al., Interplay between gene expression noise and regulatory network architecture. *Trends Genet.* **28**(5), 221-232 (2012).

39. Schultz, D., Lu, M., Onuchic, T. S. J. & Ben-Jacob, E., Turning oscillations into opportunities: lessons from a bacterial decision gate. *Sci. Rep.* **3**, 1668 (2013).

40. Takahashi, K., Cellular reprogramming – lowering gravity on Waddington's epigenetic landscape. *J. Cell Sci*. **125**(11), 2553-2560 (2012).

41. Bielas, J. H., Loeb, K. R., Rubin, B. P., True, L. D. & Loeb, L. A., Human cancers express a mutator phenotype. *PNAS* **103**(48), 18238-18242 (2006).

42. Brock, A., Chang, H. & Huang, S., Non-genetic heterogeneity – a mutation-independent driving force for the somatic evolution of tumours. *Nat. Rev. Genet.* **10**(5), 336-342 (2009).

43. Salk, J. J., Fox, E. J. & Loeb, L. A., Mutational heterogeneity in human cancers: origin and consequences. *Annu. Rev. Pathol. Mech. Dis.* **5**, 51-75 (2010).

44. Pujadas, E. & Feinberg, A. P., Regulated noise in the epigenetic landscape of development and disease. *Cell* **148**(6), 1123-1131 (2012).

45. Huang, S., Genetic and non-genetic instability in tumor progression: link between the fitness landscape and the epigenetic landscape of cancer cells. *Cancer Metast. Rev*. **32**, 423-448 (2013).

46. Huang, S., Ernberg, I. & Kauffman, S., Cancer attractors: A systems view of tumors from a gene network dynamics and developmental perspective. *Semin. Cell Dev. Biol*. **20**(7), 869-876 (2009).

47. Csermely, P. et al., Cancer stem cells display extremely large evolvability: alternating plastic and rigid networks as a potential Mechanism Network models, novel therapeutic target strategies, and the contributions of hypoxia, inflammation and cellular senescence. *Semin. Cancer Biol*. DOI: 10.1016/j.semcancer.2013.12.004 (2014).

48. Daxinger, L. & Whitelaw, E., Understanding transgenerational epigenetic inheritance via the gametes in mammals. *Nat. Rev. Genet*. **13**(3), 153-162 (2012).

49. Wu, M. et al., Engineering of regulated stochastic cell fate determination. *PNAS* **110**(26), 10610-10615 (2013).

50. Danino, T., Mondragón-Palomino, O., Tsimring, L. & Hasty, J., A synchronized quorum of genetic clocks. *Nature* **463**, 326-330 (2010).

51. Moreno-Risueno, M. A. & Benfey, P. N., Time-based patterning in development: the role of oscillating gene expression. *Transcription* **2**(3), 124-129 (2011).

52. Vollmers, C. et al., Circadian oscillations of protein-coding and regulatory RNAs in a highly dynamic mammalian liver epigenome. *Cell Metab*. **16**(6), 833-845 (2012).

53. Levine, J. H., Lin, Y. & Elowitz, M. B., Functional roles of pulsing in genetic circuits. *Science* **342**(6163), 1193-1200 (2013).

54. Prindle, A. et al., Rapid and tunable post-translational coupling of genetic circuits. *Nature* **508**, 387-391 (2014).

55. Creixell, P., Schoof, E. M., Erler, J. T. & Linding, R., Navigating cancer network attractors for tumor-specific therapy. *Nat. Biotechnol.* **30**, 842-848 (2012).

56. Li, Y., Yi, M. & Zou, X., The linear interplay of intrinsic and extrinsic noises ensures a high accuracy of cell fate selection in budding yeast. *Sci. Rep.* **4**, 5764 (2014).

Acknowledgment

We would like to thank B.D. Aguda and C.O. Talaue for discussion about models of cellular regulation, and the organizers and participants of the 2014 Hands-on Research in Complex Systems School at ICTP for feedback and suggestions. JRF is supported by Japanese government (Monbukagakusho: MEXT) fellowship and the University of the Philippines. ALB is supported by Philippine Commission on Higher Education (CHED) fellowship and the University of the Philippines.

Author contributions

JFR and ALB conceived the study. JFR built the model and ran the simulations. JFR and ALB wrote the manuscript.

Competing financial interests

The authors declare no competing financial interests.

Figure legends

Figure 1. (**a**) The epigenetic landscape (adapted from Waddington's illustration[1]). The branching canals depict the various cell lineages towards different fates (cell types/phenotypes). The cell fates are illustrated as valleys and traditionally represented as mathematical attractors (see Supplementary Figs. 1 and 2). The differentiating cells, illustrated as balls, trail a chosen canal towards a specific valley. The canal is chosen based on the landscape's potential (similar to gravitational potential) such that the steeper pathway and deeper valley are preferred[8]. The canals are separated by ridges that restrain cells to switch lineages. The blue pegs (GRFs) and blue strings (GRF-GRF interaction) alter the height of ridges and depth of valleys. The height of ridges and depth of valleys vary through time and affect the route of the differentiating cells. (**b**) A minimal gene regulatory network (GRN) that characterizes decision switches in cell-fate determination[10,24,29]. The nodes represent GRFs, such as master switch genes, transcription factors, or coarse-grained modules of a larger GRN that can be simplified as one node. Each node X_i has auto-activation as represented by the arrows; while the interaction links among GRFs is of repressive behavior (represented by the bars). Strength of repression is not necessarily reciprocal and a one-way repression is possible. The auto-activation and repression can be direct or indirect (e.g., $X_2 \to A \to B \to X_2$).

Figure 2. Illustrations of cells trailing the dynamic pathways in the epigenetic landscape (only the deterministic path is shown; see Supplementary Figs. 3, 5, 6, 8 and 10 for stochastic simulations and for the initial condition and parameter values used). Note that the branch endpoints are the coordinates of an equilibrium point. The different numbers of endpoints are due to the different time scale factors used. (**a**) Five different endpoints. (**b**) No branching and only one endpoint (graphs for GRF 1-5 are superimposed to each other). The state with

equal cell type probabilities represent multipotent or pluripotent cells (undifferentiated state). (**c-e**) Two, three and four endpoints, respectively. (**f**) Time series of the entropy levels for the system in Figs. 2a to 2e. The entropy decreases as cells differentiate.

Figure 3. (**a**) A repressilator-type network with a strong negative feedback loop, $n=3$ (this can be extended for any n). The strengths of repression in one loop (solid black bars) are stronger than the reverse loop (broken bars). The red bars represent inhibition of repression. Note that in our model, only three GRFs are needed to generate sustained oscillations. (**b-c**) Examples of oscillating pathways in the epigenetic landscape generated by repressilator-type networks. There are no deep valleys only continuous zig-zag canals (see Supplementary Figs. 13-15 and 17 for the stochastic simulations and for the parameter values used). (**b**) $n=5$; X_5 is initially silenced. (**c**) $n=4$; X_4 is initially silenced. (**d**) Damped oscillations towards multipotency or pluripotency. The rates of decline of the timescale factors are all equal to 0.001. (**e**) Damped oscillations resulting in partial differentiation and reversal of dominant GRF. The rates of decline of the timescale factors are not all equal. The initial dominant regulatory factor is GRF 5 but eventually becomes inferior as oscillations dampen.

Supplementary Materials

Supplementary Figs. 1 to 20

Figure 1

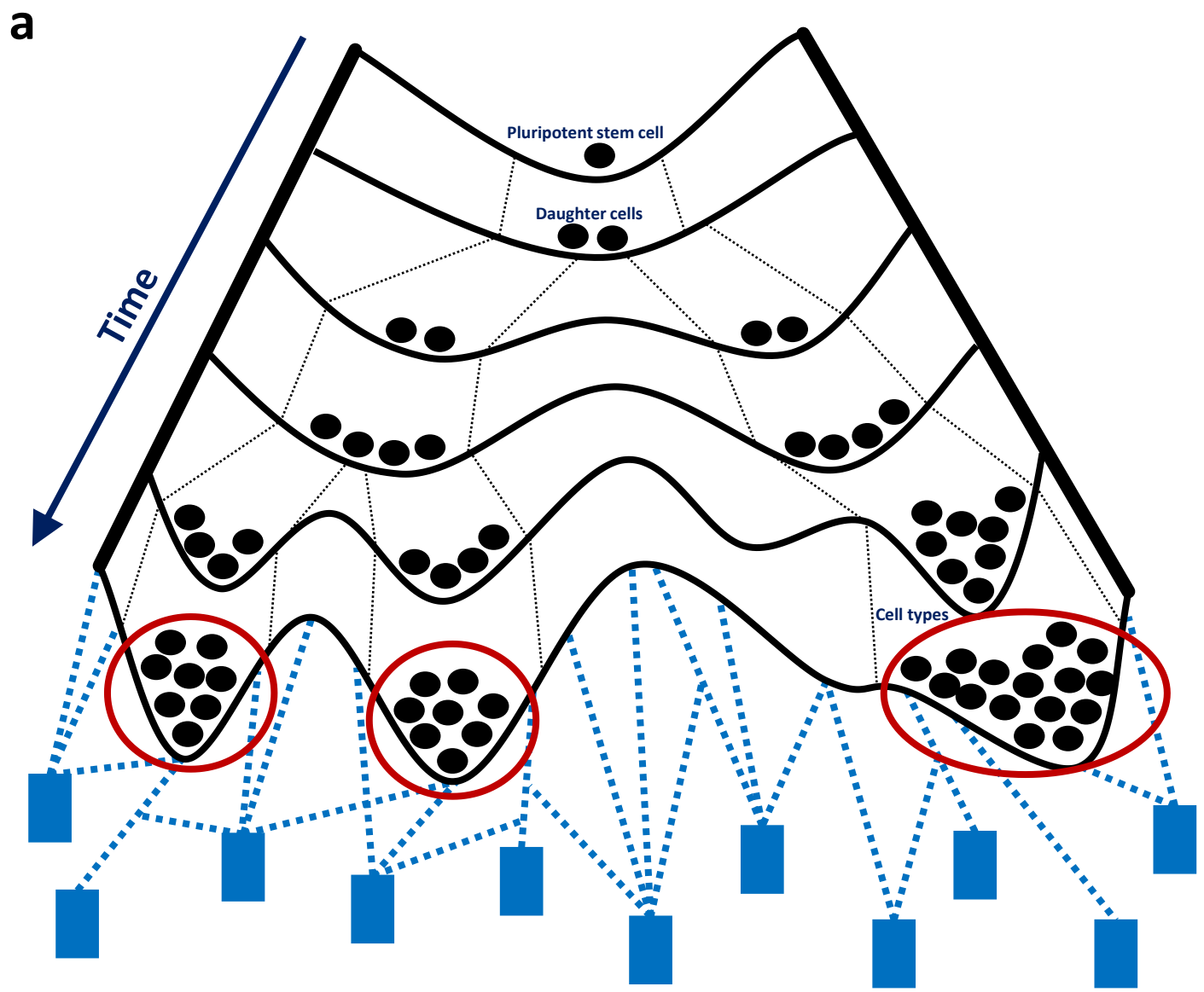


Figure 1

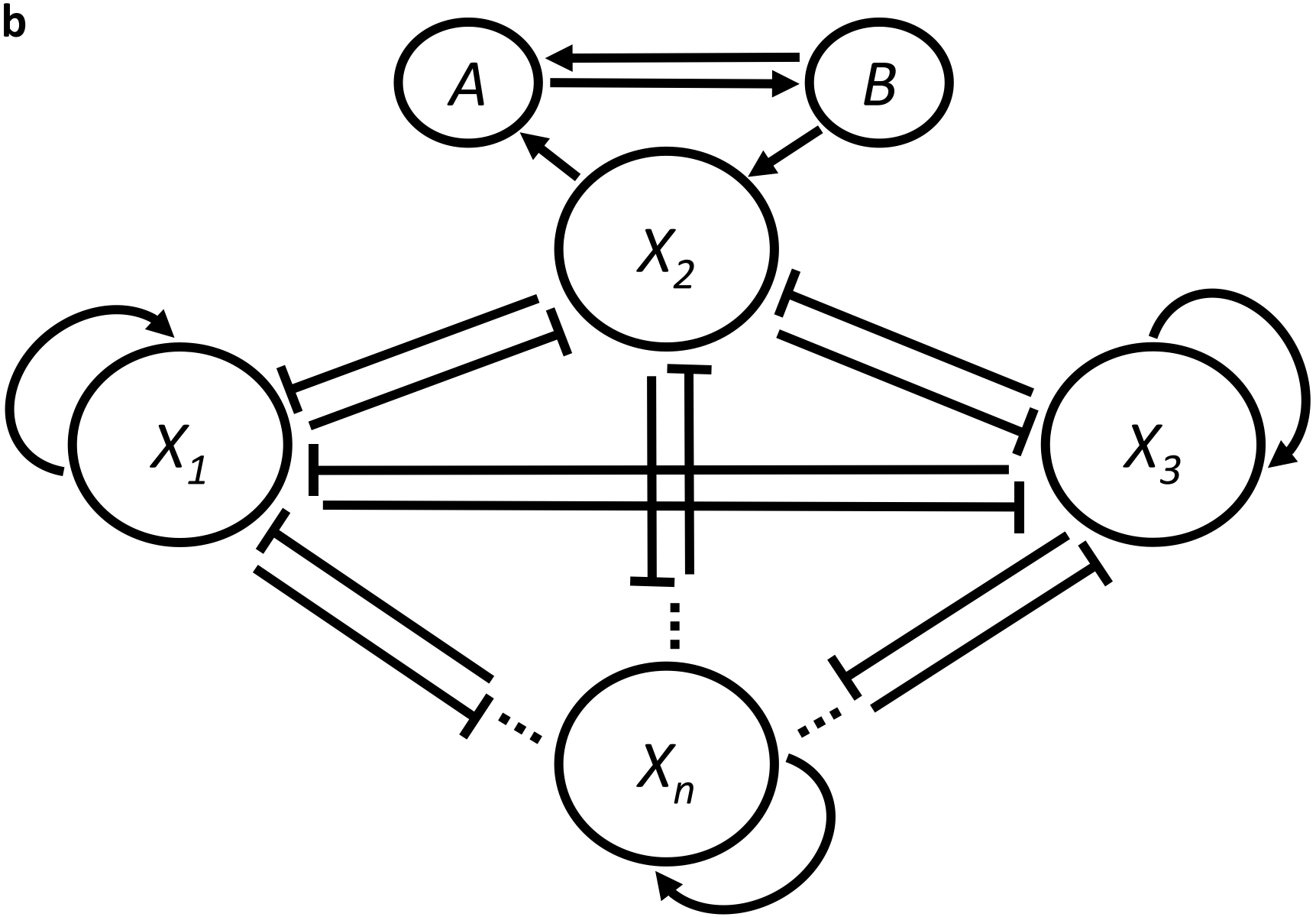


Figure 2

a

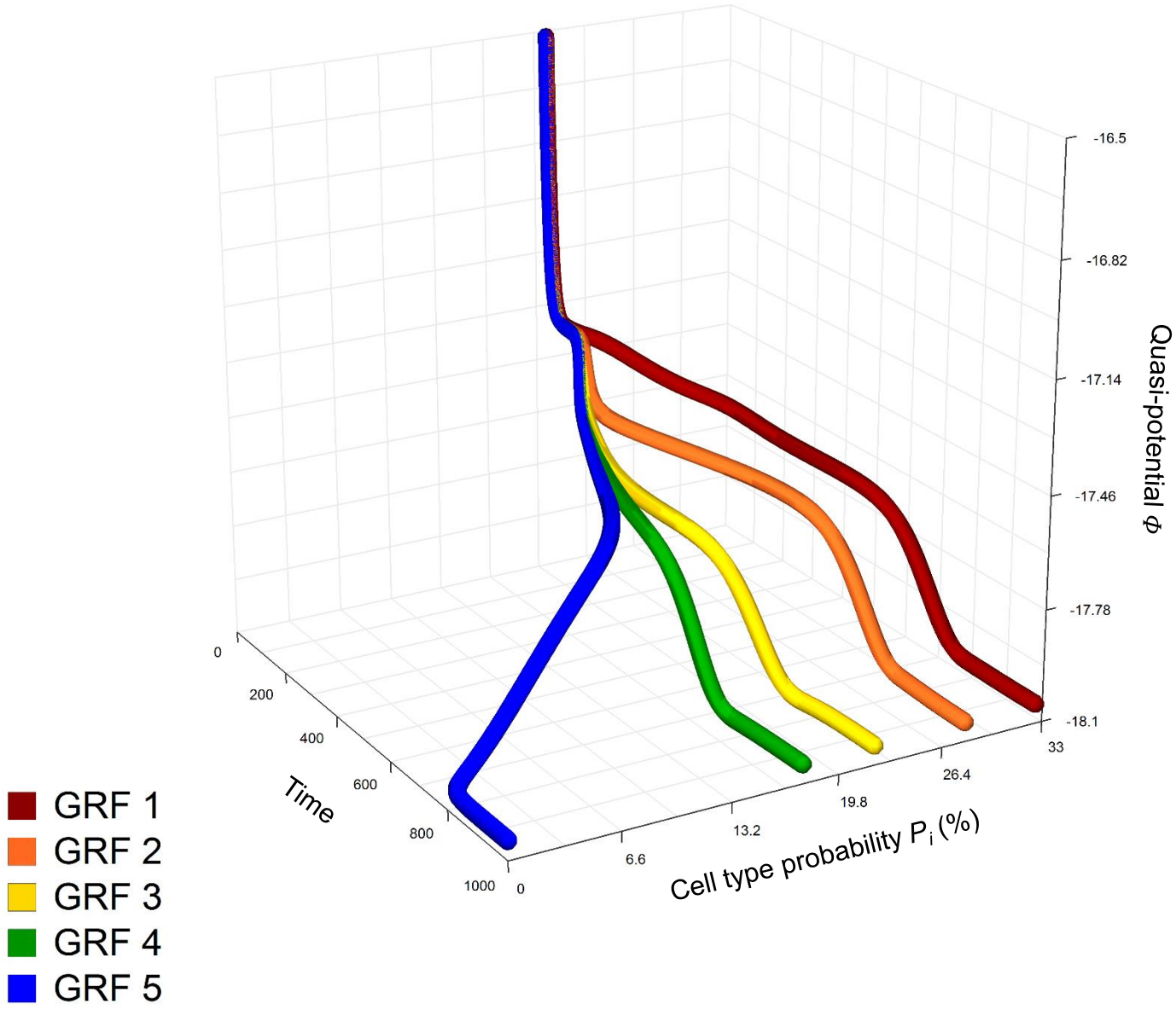

GRF 1
GRF 2
GRF 3
GRF 4
GRF 5
Time
0
200
400
600
800
1000
Cell type probability P_i (%)
0
6.6
13.2
19.8
26.4
33
Quasi-potential Φ
-16.5
-16.82
-17.14
-17.46
-17.78
-18.1

Figure 2

b

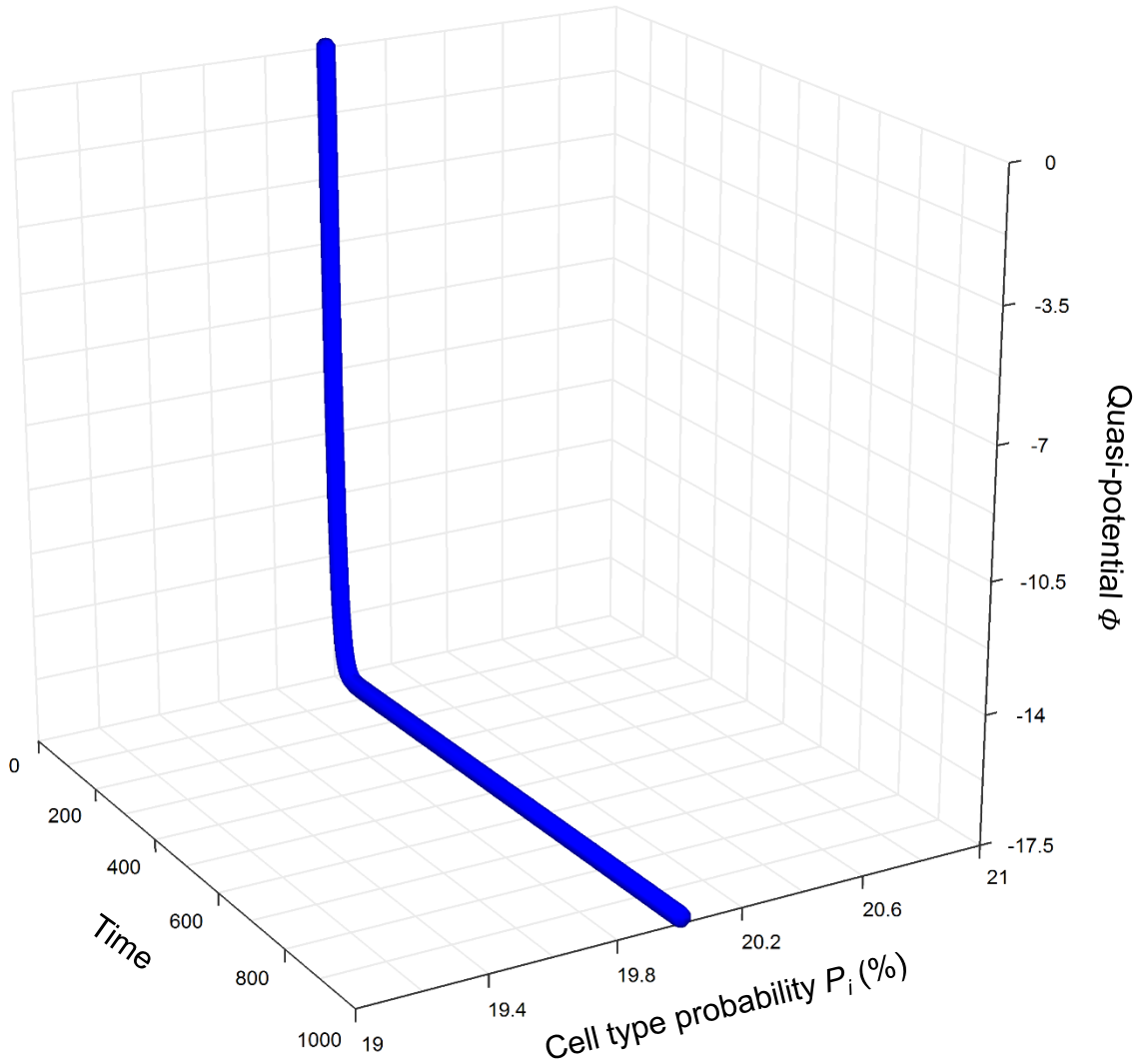

0
-3.5
-7
-10.5
-14
-17.5
Quasi-potential Φ
0
200
400
600
800
1000
Time
19
19.4
19.8
20.2
20.6
21
Cell type probability P_i (%)

Figure 2

c

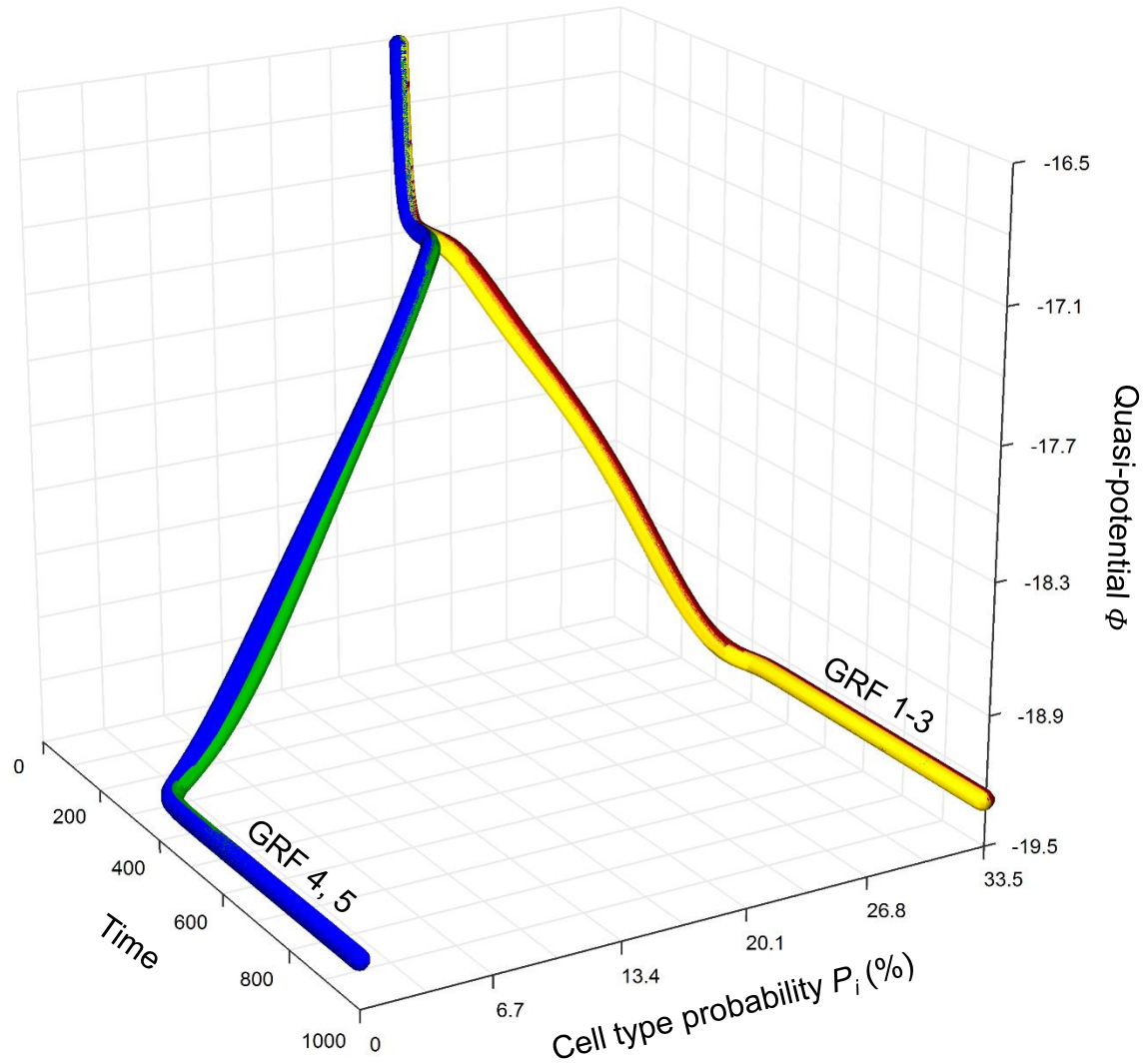

-16.5
-17.1
-17.7
-18.3
-18.9
-19.5
Quasi-potential Φ
GRF 1-3
GRF 4, 5
0
200
400
600
800
1000
Time
0
6.7
13.4
20.1
26.8
33.5
Cell type probability P_i (%)

Figure 2

d

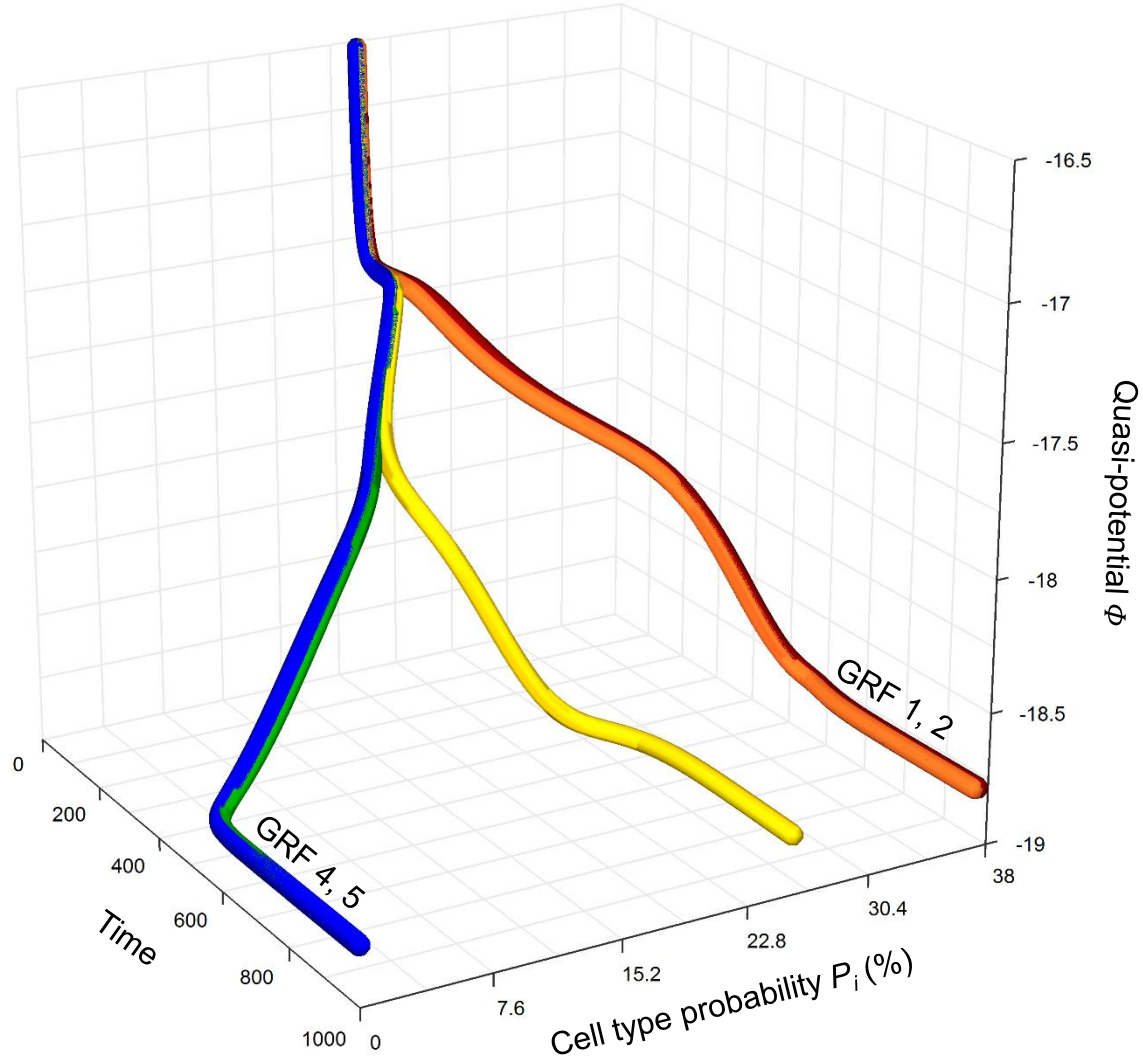

-16.5
-17
-17.5
-18
-18.5
-19
Quasi-potential Φ
GRF 1, 2
GRF 4, 5
0
200
400
600
800
1000
Time
0
7.6
15.2
22.8
30.4
38
Cell type probability P_i (%)

Figure 2

e

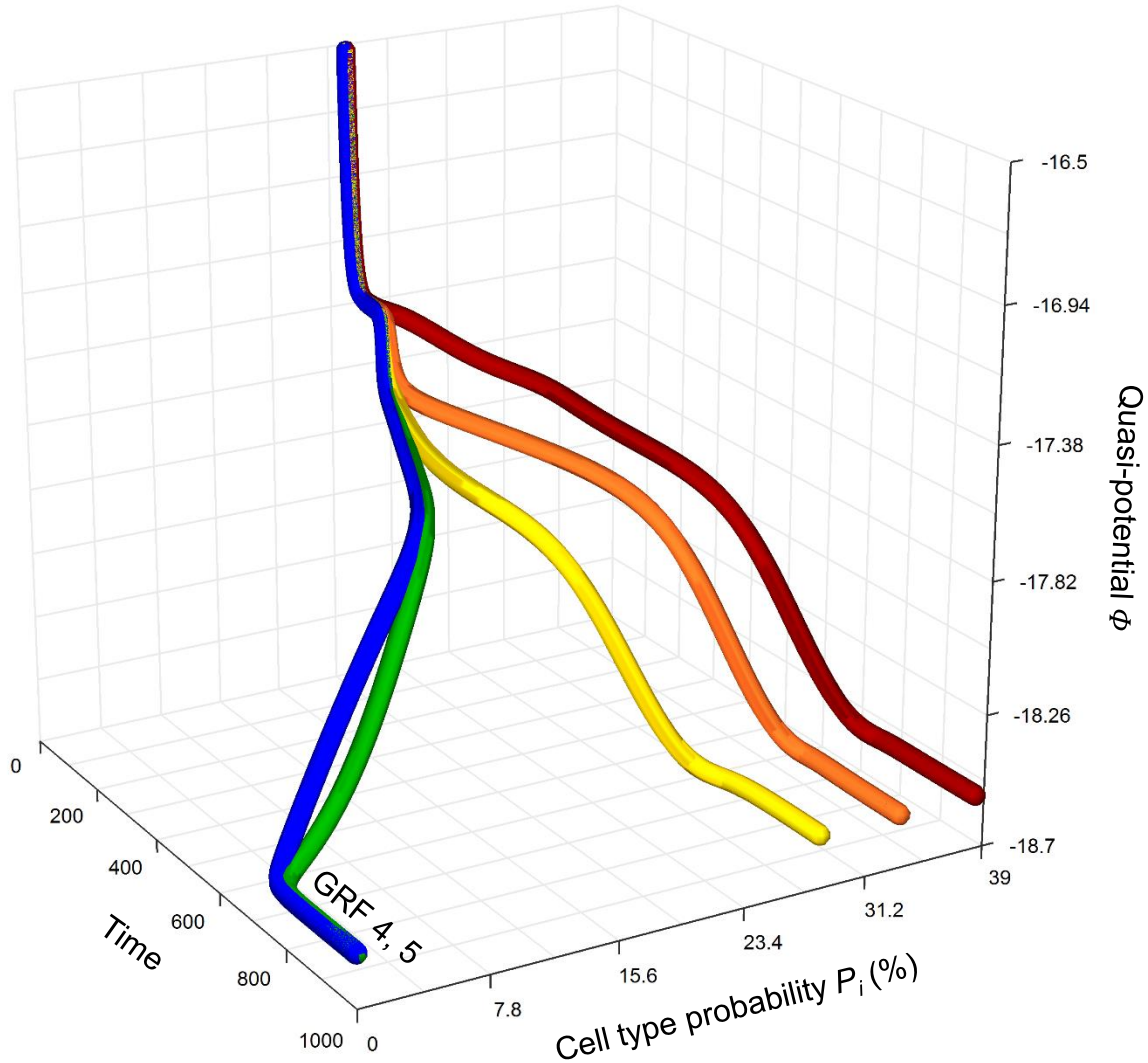

-16.5
-16.94
-17.38
-17.82
-18.26
-18.7
Quasi-potential Φ
0
200
400
600
800
1000
Time
GRF 4, 5
0
7.8
15.6
23.4
31.2
39
Cell type probability P_i (%)

Figure 2

f

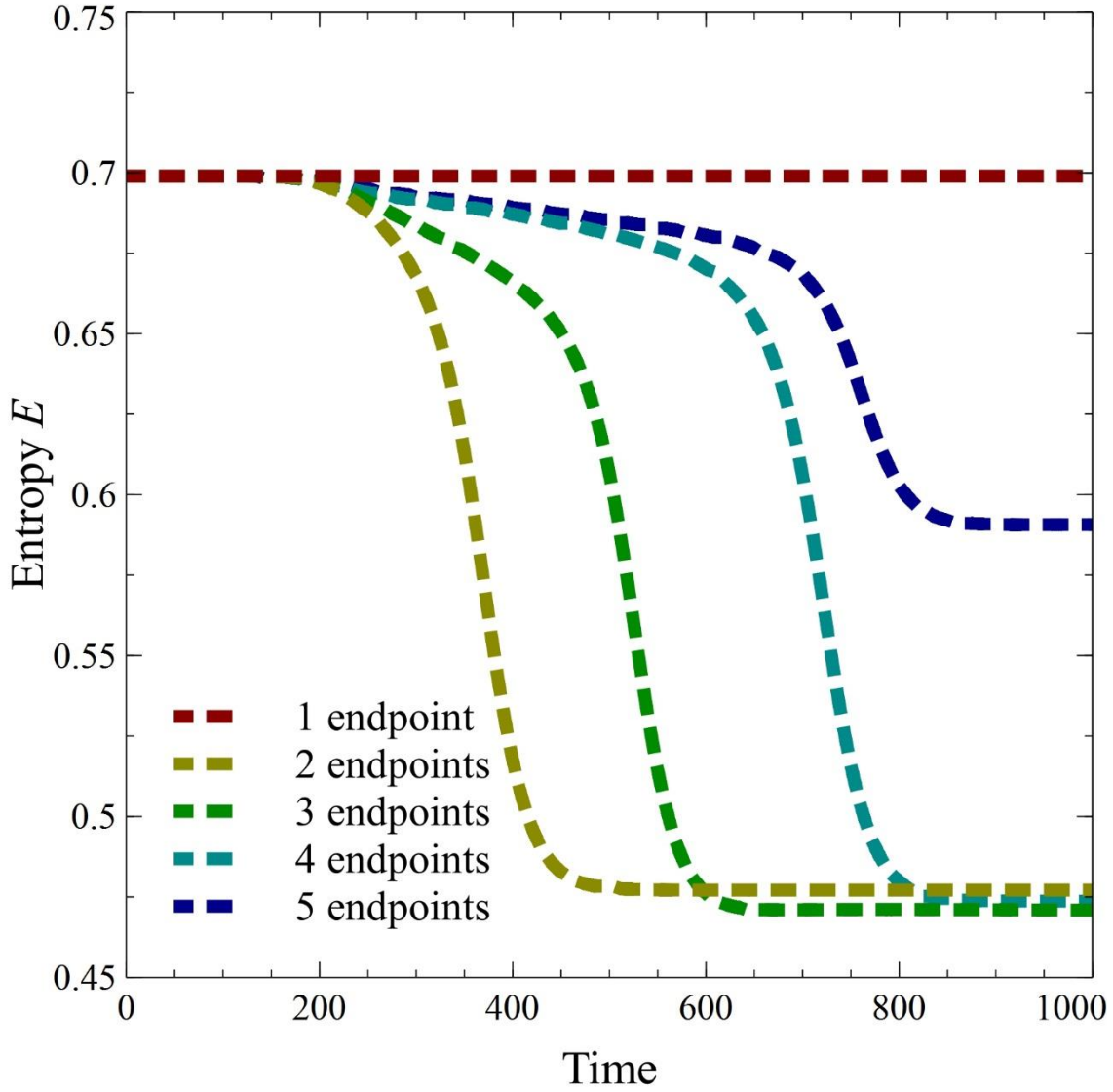

0.75
0.7
0.65
0.6
0.55
0.5
0.45
Entropy E
0
200
400
600
800
1000
Time
1 endpoint
2 endpoints
3 endpoints
4 endpoints
5 endpoints

Figure 3

a

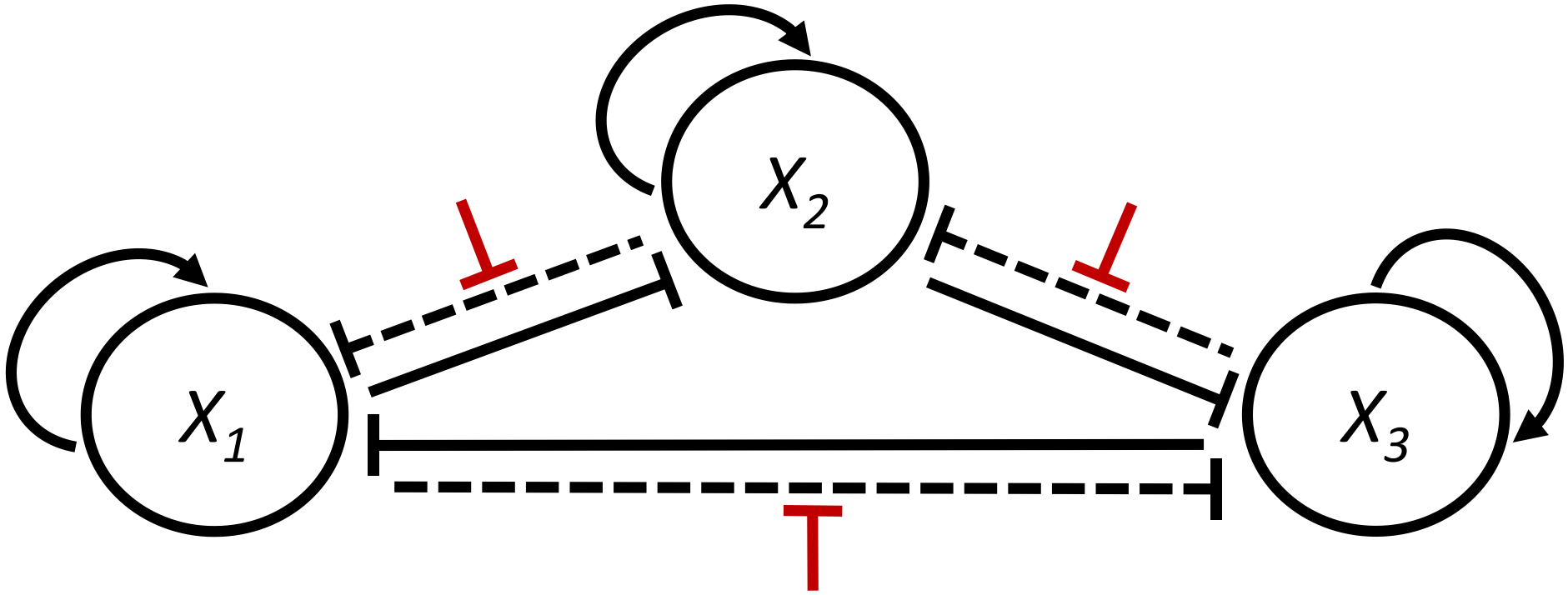


Figure 3

b

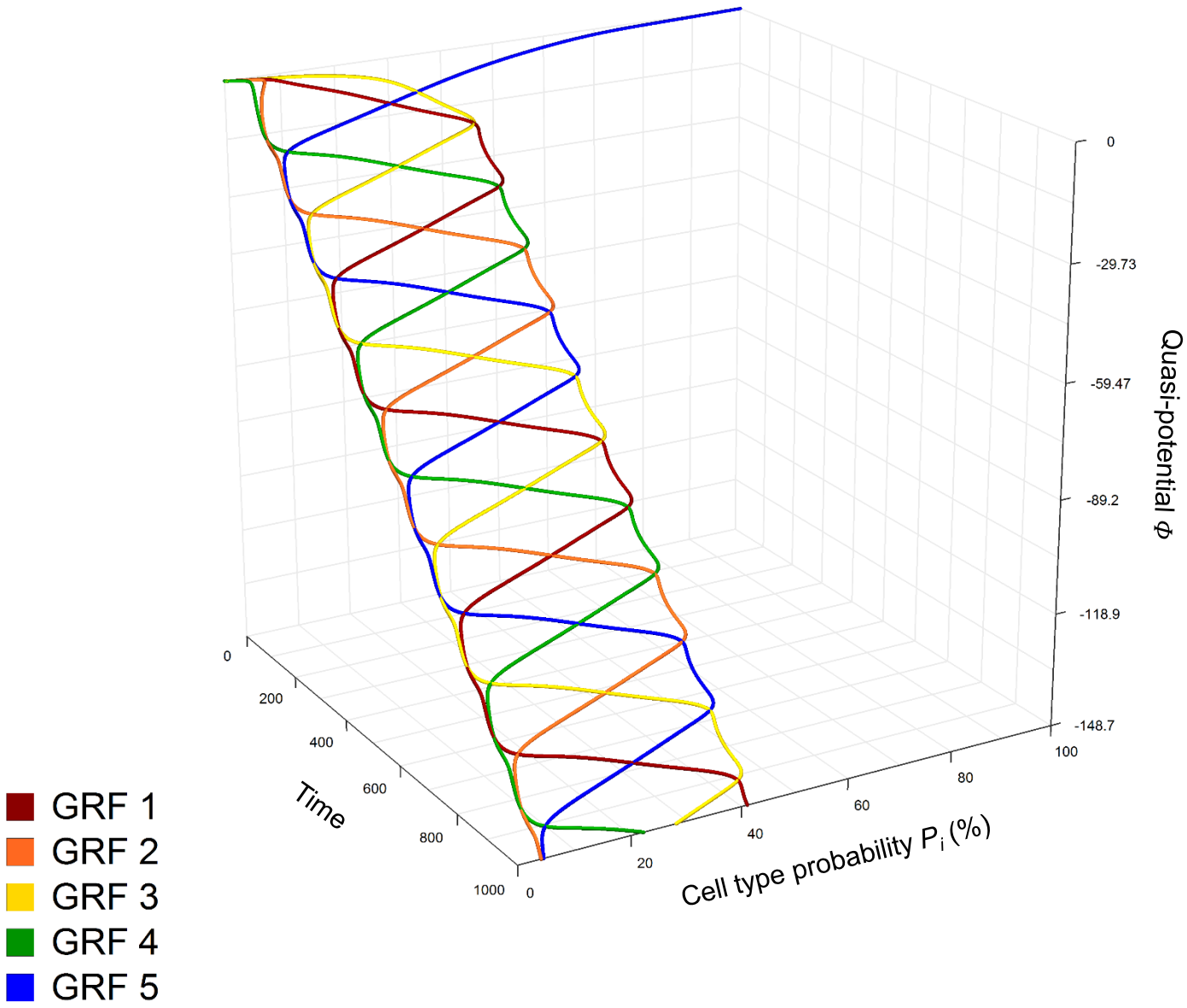

GRF 1
GRF 2
GRF 3
GRF 4
GRF 5
Time
0
200
400
600
800
1000
Cell type probability P_i (%)
0
20
40
60
80
100
Quasi-potential Φ
0
-29.73
-59.47
-89.2
-118.9
-148.7

Figure 3

c

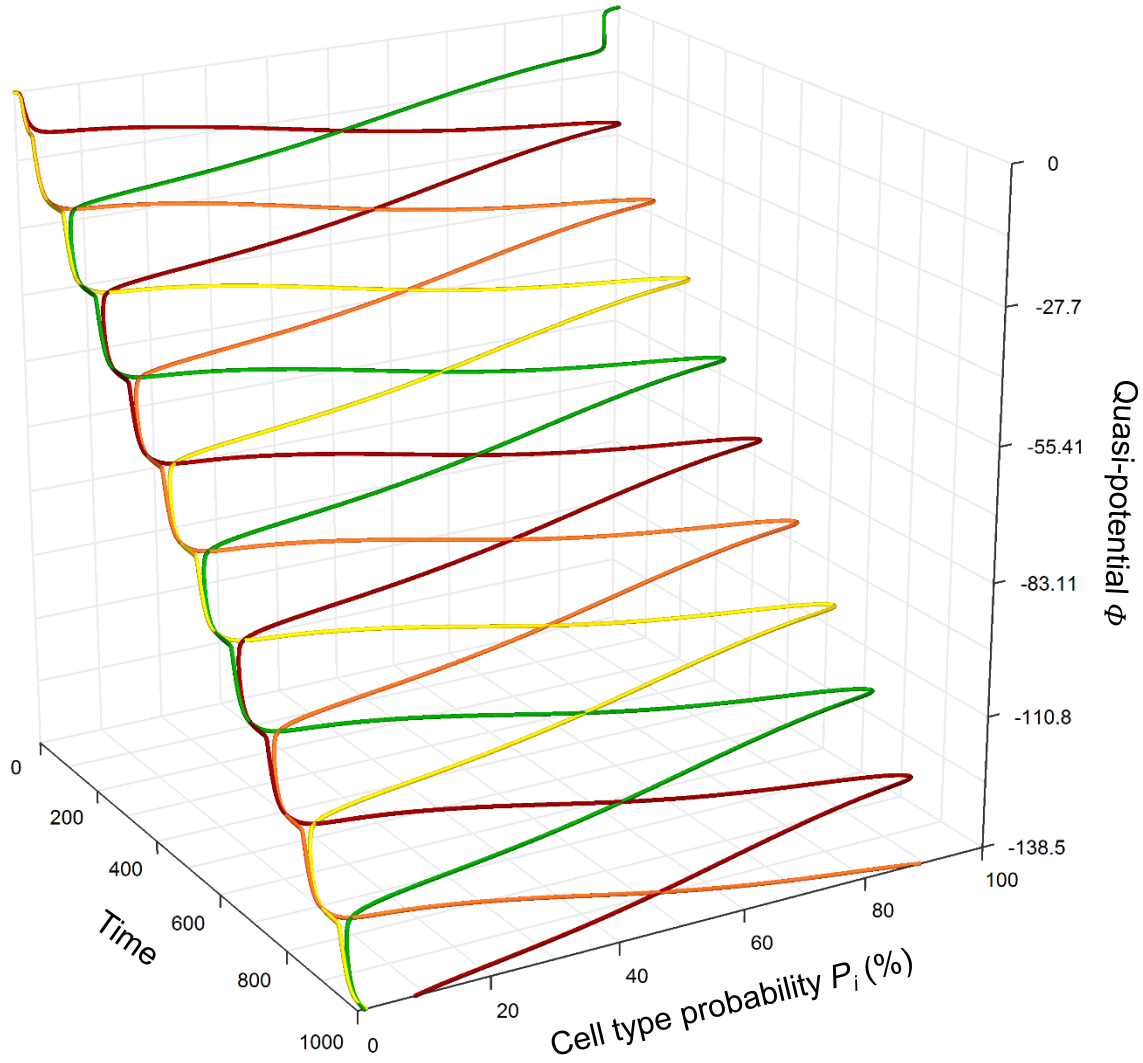

0
-27.7
-55.41
-83.11
-110.8
-138.5
Quasi-potential Φ
0
200
400
600
800
1000
Time
0
20
40
60
80
100
Cell type probability P_i (%)

Figure 3

d

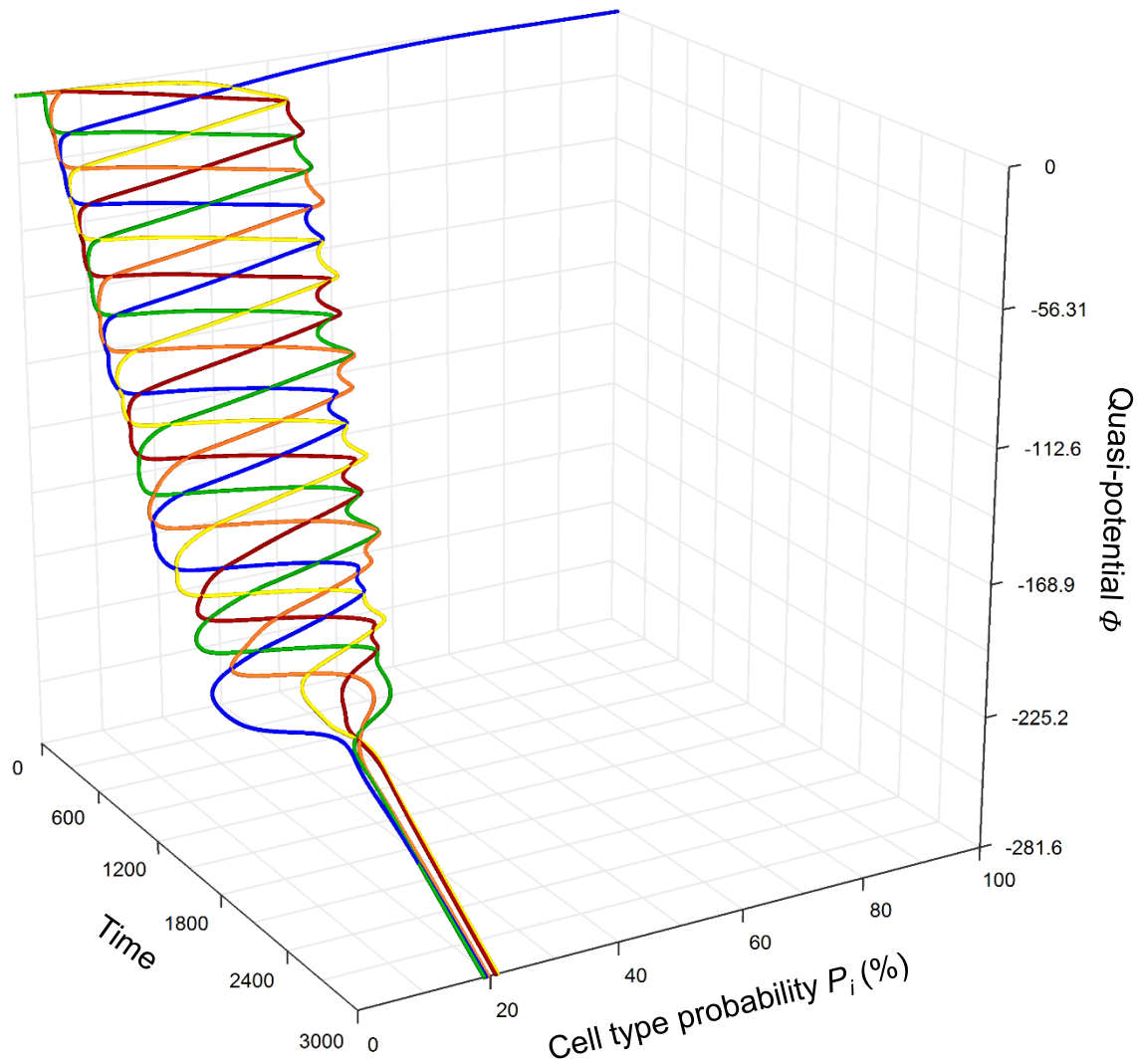
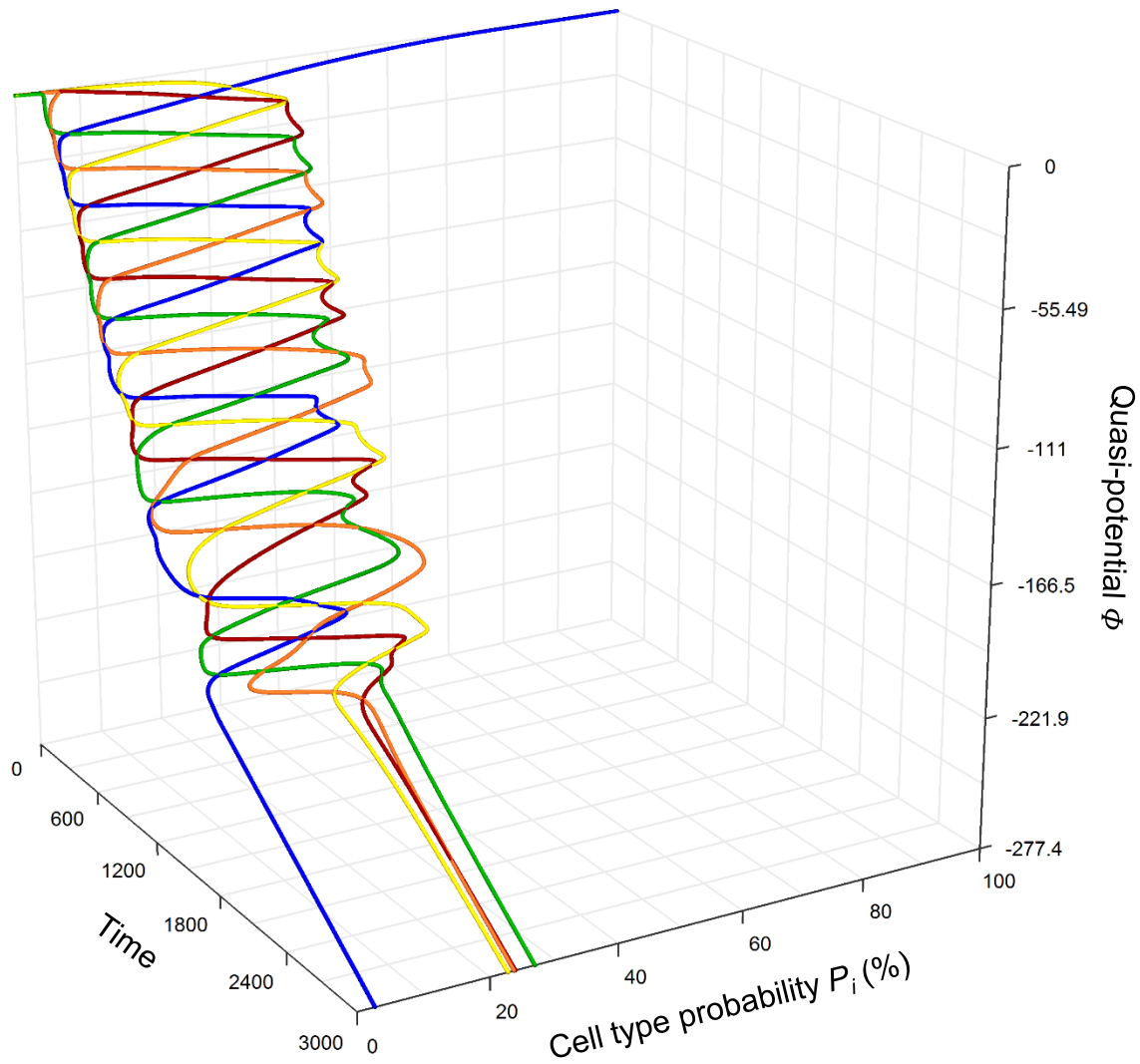

0
-56.31
-112.6
-168.9
-225.2
-281.6
Quasi-potential Φ
0
600
1200
1800
2400
3000
Time
0
20
40
60
80
100
Cell type probability P_i (%)

Figure 3

e


0
-55.49
-111
-166.5
-221.9
-277.4
Quasi-potential Φ
0
600
1200
1800
2400
3000
Time
0
20
40
60
80
100
Cell type probability P_i (%)

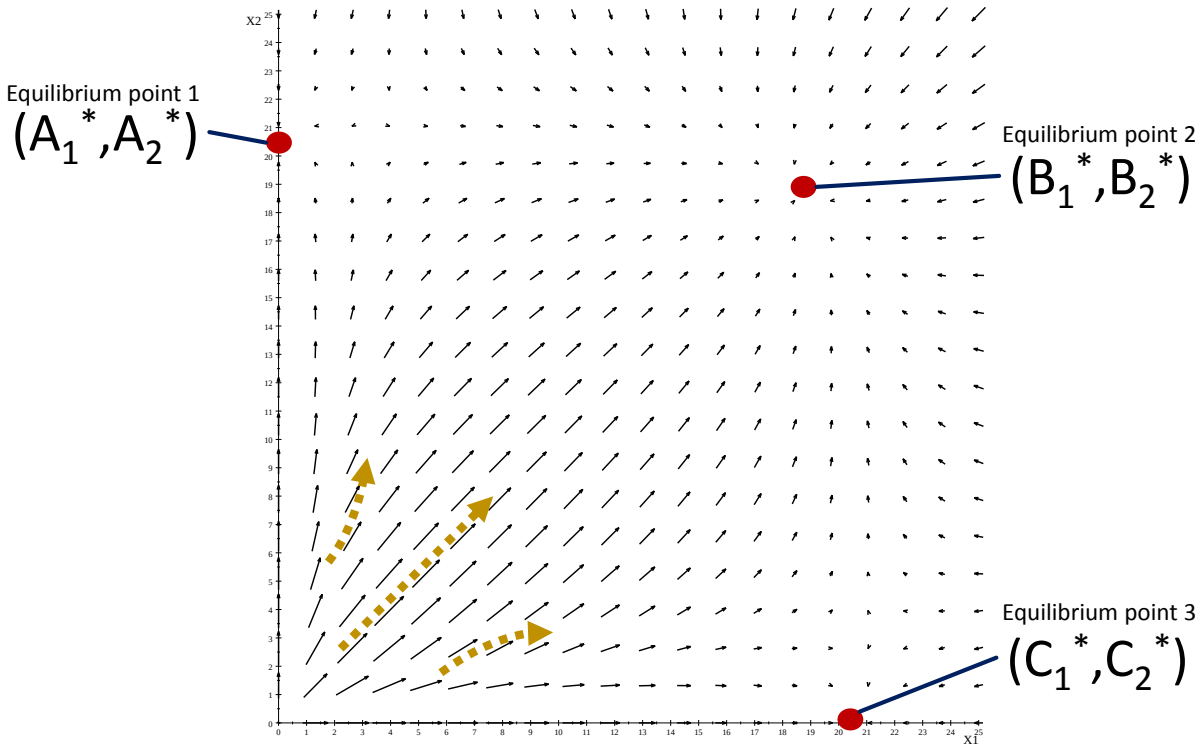
Supplementary Materials for
Branching and oscillations in the epigenetic landscape of cell-fate determination

Jomar Fajardo Rabajante, Ariel Lagdameo Babierra
correspondence to: Jomar Fajardo Rabajante (jfrabajante@up.edu.ph)

This PDF file includes

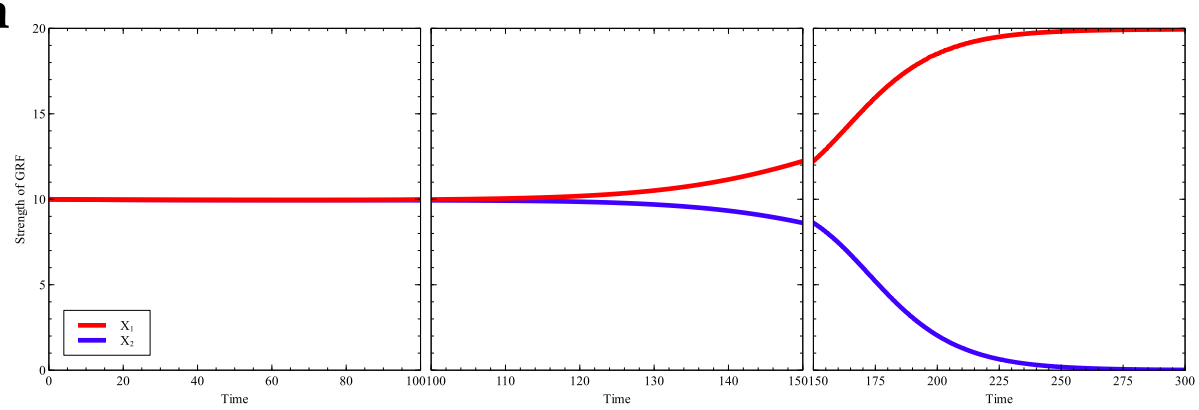
Supplementary Figures 1 to 20

Note: see Methods for the default parameter values.

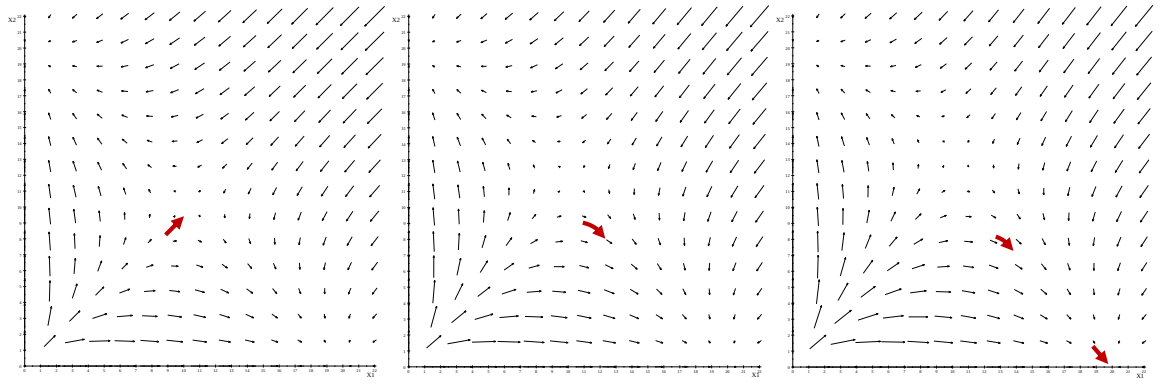


Supplementary Figure 1. Example of phase portrait of the ordinary differential equation model (Box Eq. 1 in the main text) with more than one stable equilibrium point (multistable). The convergence to an equilibrium point depends on the initial condition. The coordinates of an equilibrium point (e.g., A_1^* and A_2^*) are the branch endpoints of the pathways.

a

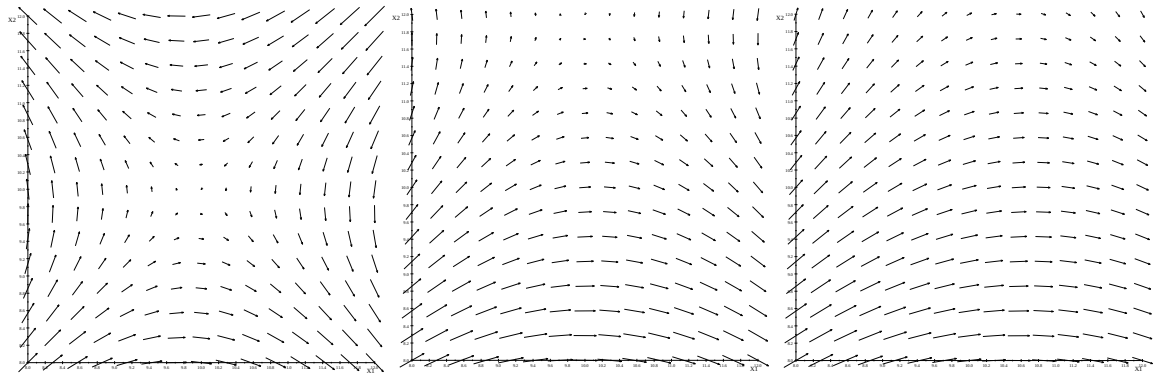


b



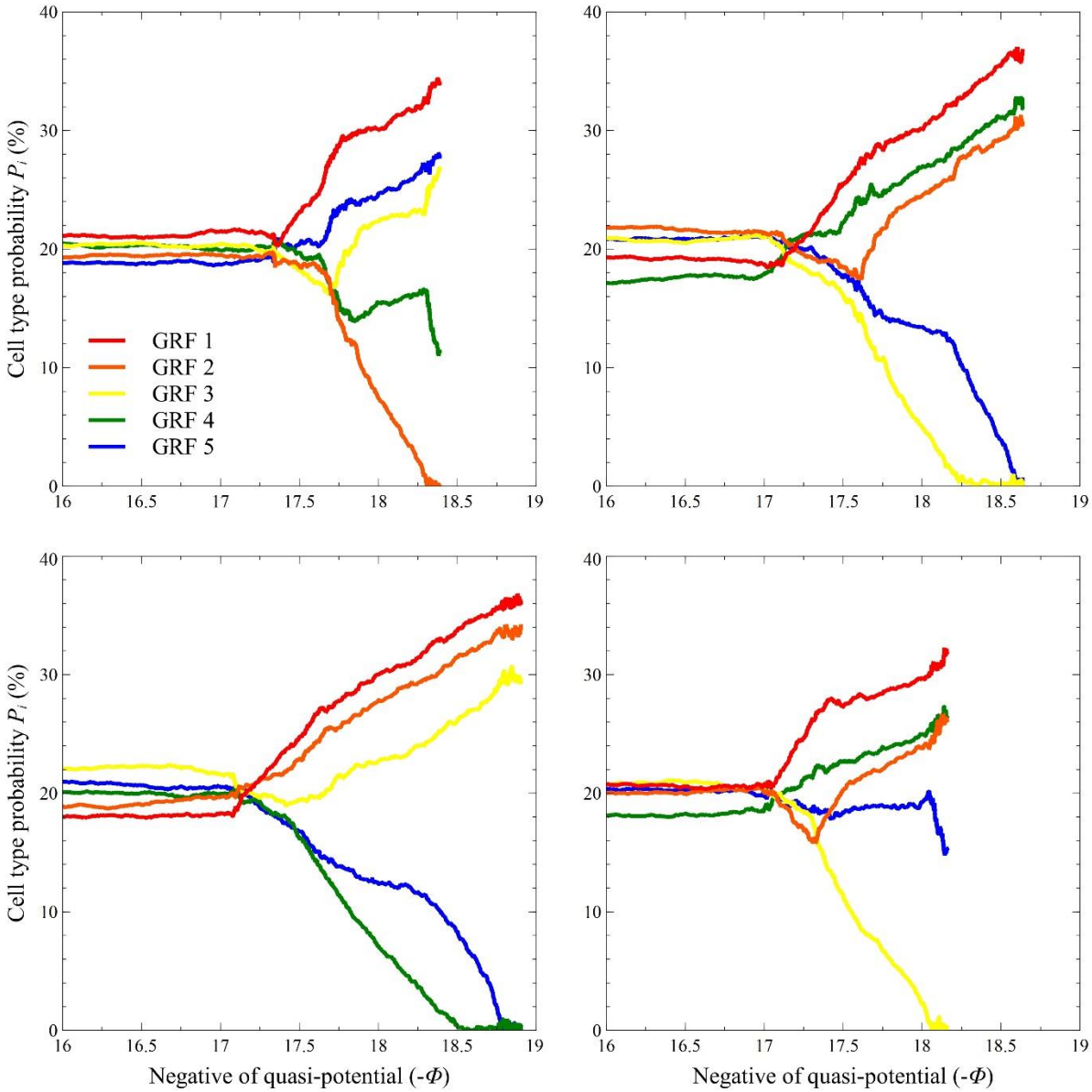
Time=0 Time=150 Time=300

c Zooming in:

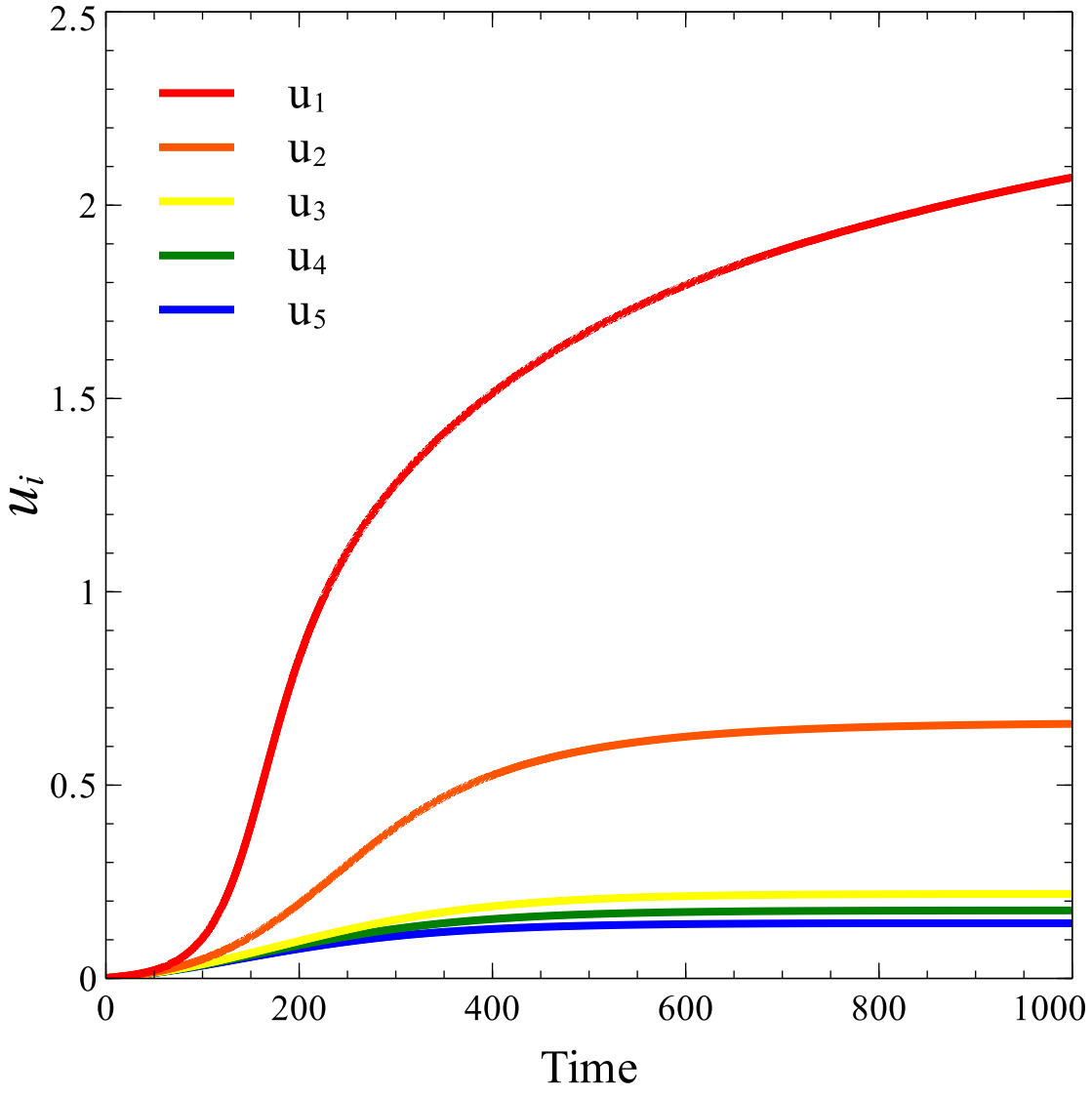


Time=0 Time=150 Time=300

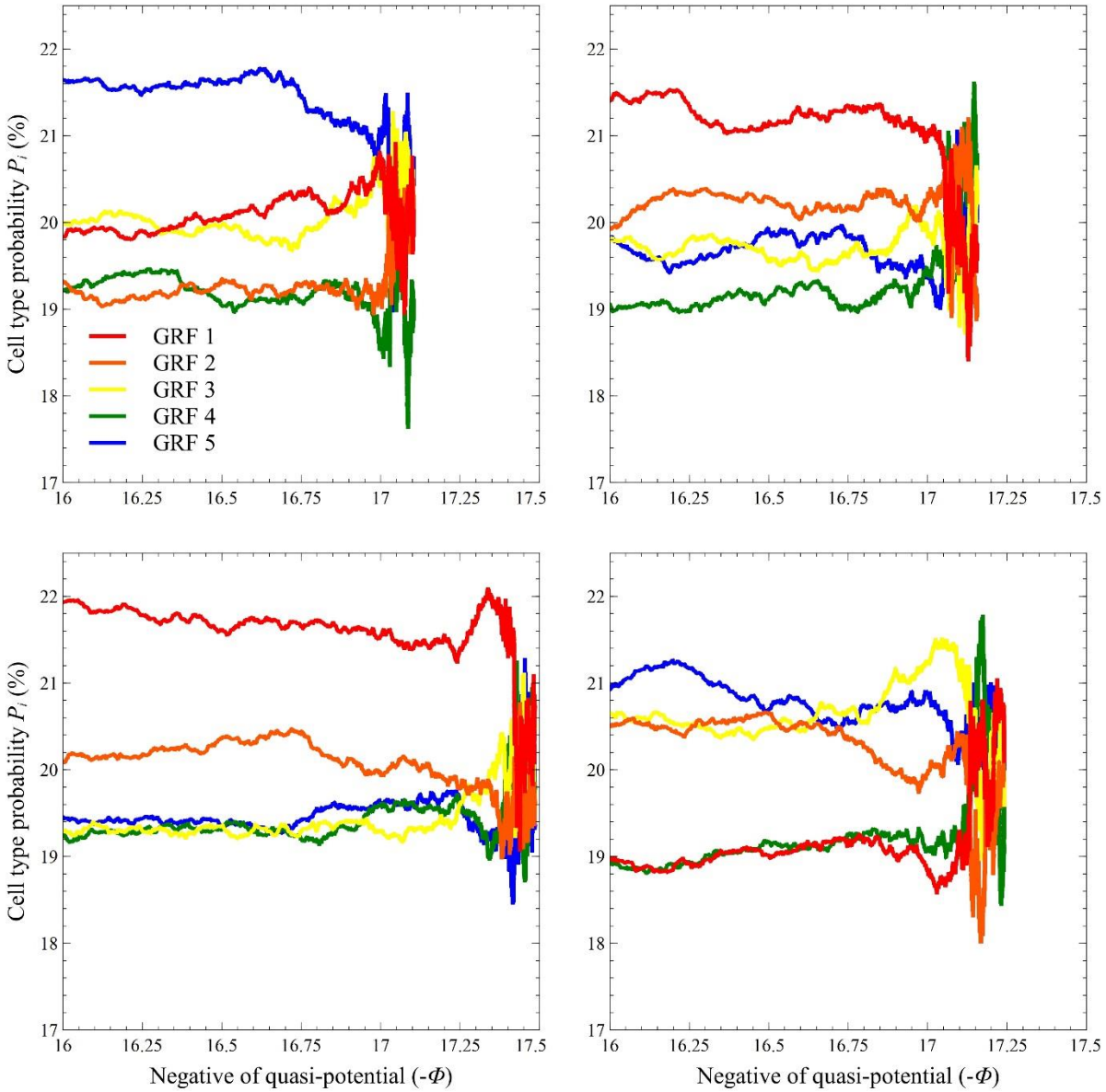
Supplementary Figure 2. The changes in the value of interaction coefficient γ_{ij} drive the transformation of the epigenetic landscape affecting the fate of the differentiating cell. (**a-b**) The initial trajectory of the solution to the differential equation model (Box Eq. 1 in the main text) converges to a state with equal X_1=X_2. However, a slight modification in the topography of the landscape steers the trajectory towards a state with X_1>X_2, hence branching arises. (**c**) Zooming in the phase portrait in Fig. 2b to visualize the transformation of the landscape (horizontal axis is X_1; vertical axis is X_2). Parameters are a_{12}= a_{21}=1, ε_1=0.001 and ε_2=0.005.



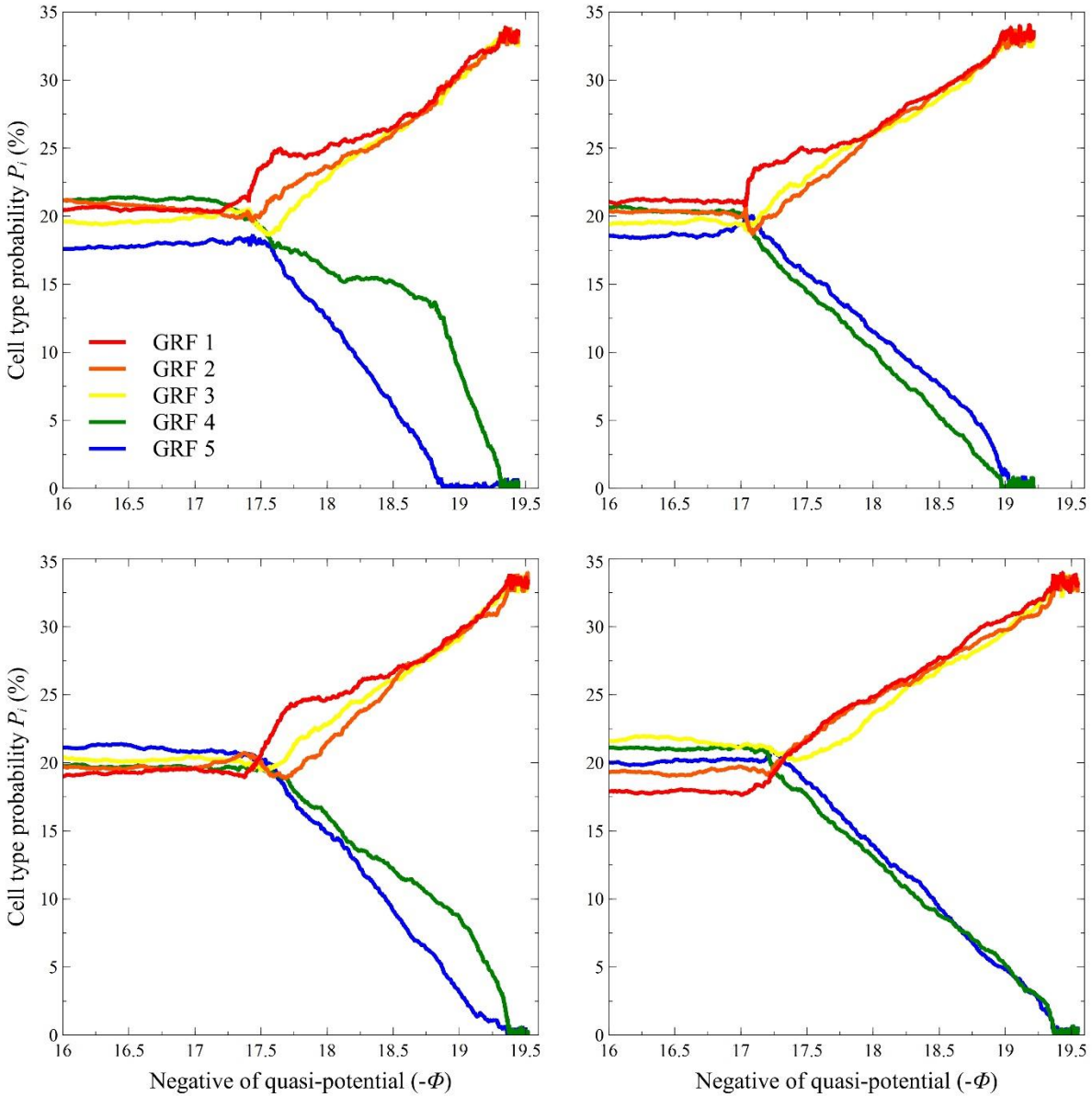
Supplementary Figure 3. Sample stochastic paths of the system in main text's Fig. 2a (five endpoints when deterministic). The initial condition is $X_i=1$ for all i. The parameter values are $a_{ij}=1/8$ for all i,j, $\varepsilon_1=0.0010$, $\varepsilon_2=0.0060$, $\varepsilon_3=0.0085$, $\varepsilon_4=0.0090$ and $\varepsilon_5=0.0095$.



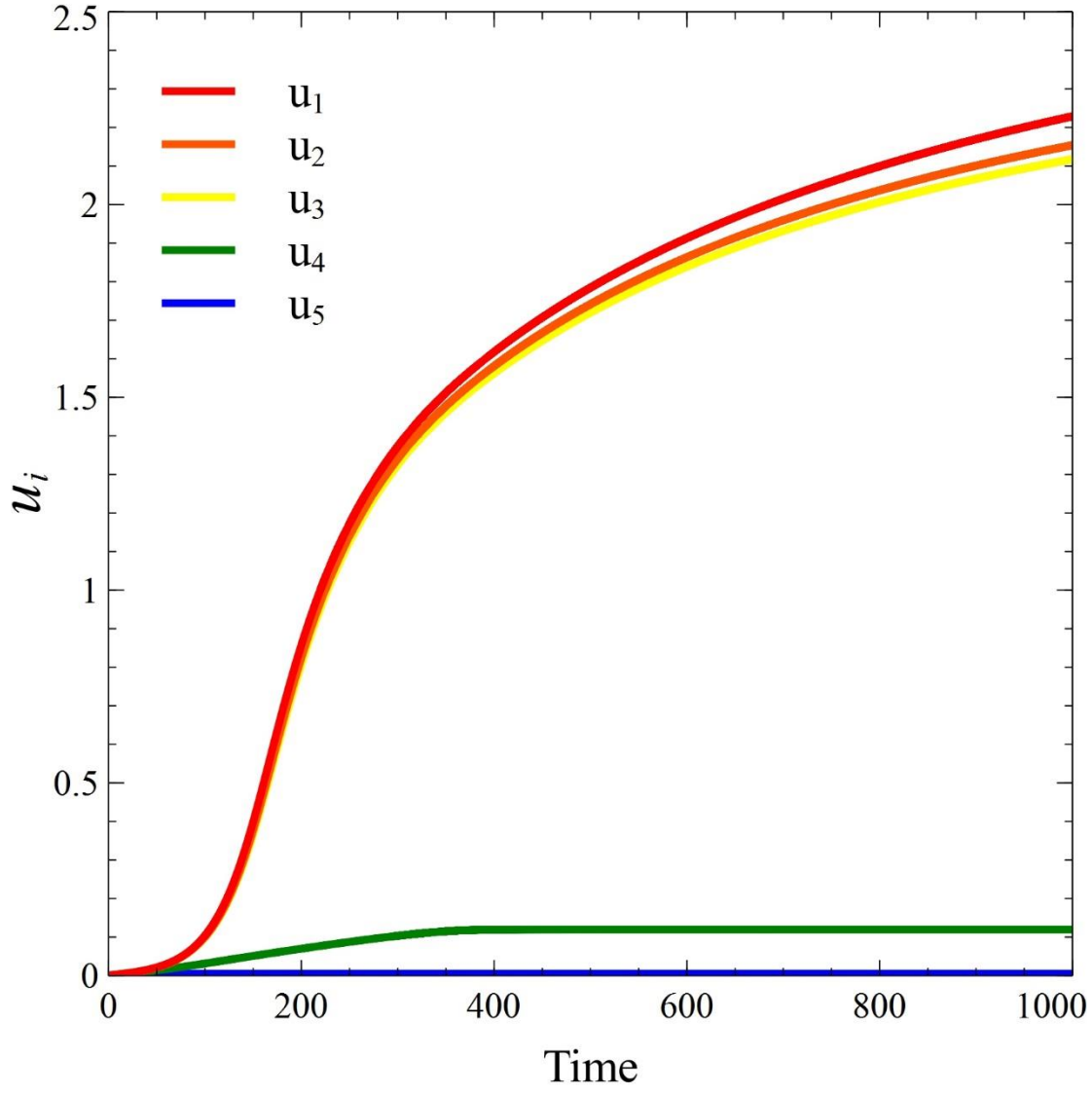
Supplementary Figure 4. Time evolution of u_i for the deterministic system in main text's Fig. 2a. Timescale factor decline rates are ε_1=*0.0010*, ε_2=*0.0060*, ε_3=*0.0085*, ε_4=*0.0090* and ε_5=*0.0095*.



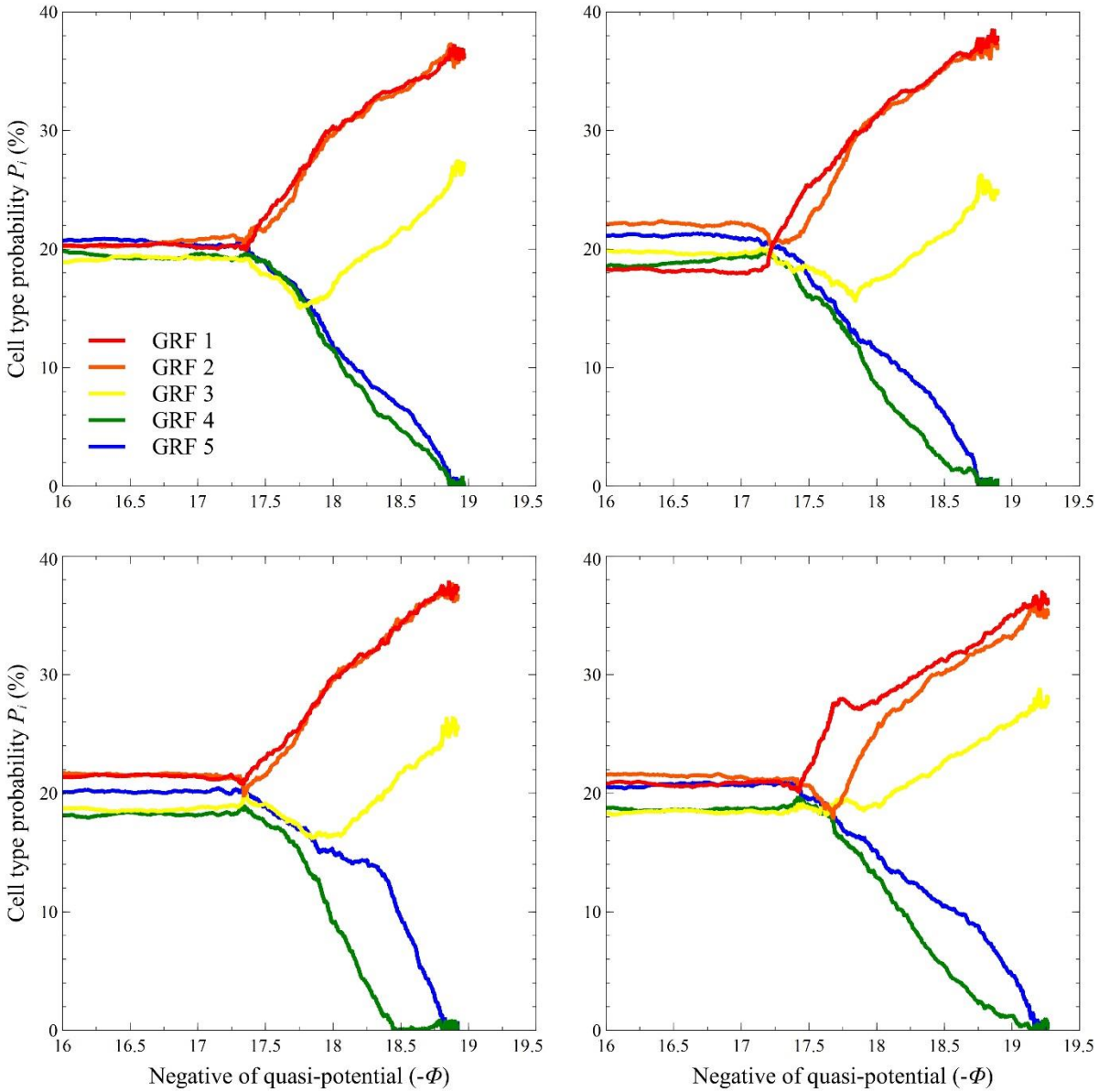
Supplementary Figure 5. Sample stochastic paths of the system in main text's Fig. 2b (one endpoint, undifferentiated state). The initial condition is $X_i=1$ for all i. The parameter values are $a_{ij}=1/8$ and u_{ij} does not evolve for all i,j.



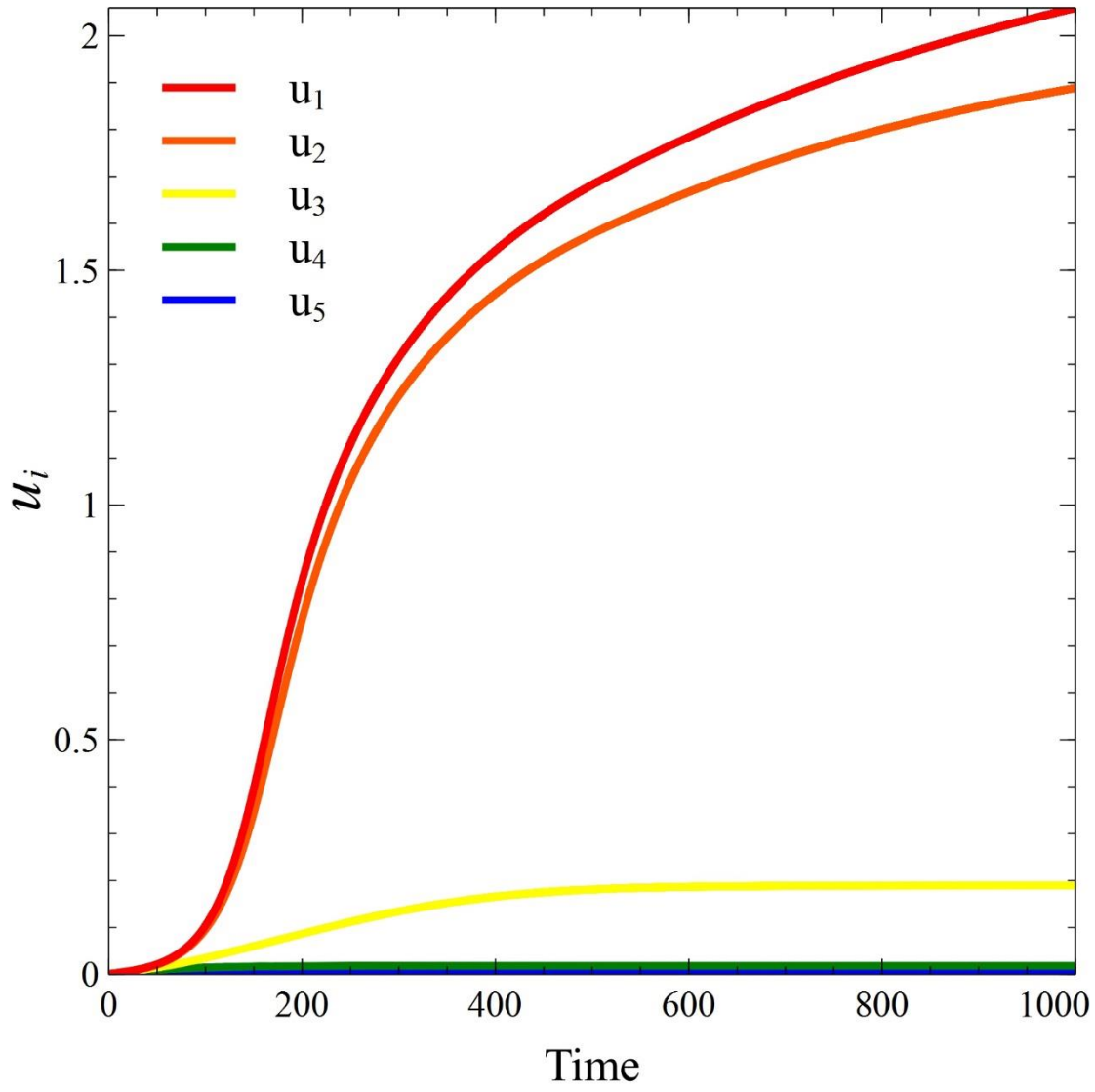
Supplementary Figure 6. Sample stochastic paths of the system in main text's Fig. 2c (two endpoints). The initial condition is $X_i=1$ for all i. The parameter values are $a_{ij}=1/8$ for all i,j, $\varepsilon_1=0.0010$, $\varepsilon_2=0.0012$, $\varepsilon_3=0.0013$, $\varepsilon_4=0.0100$ and $\varepsilon_5=0.0500$.



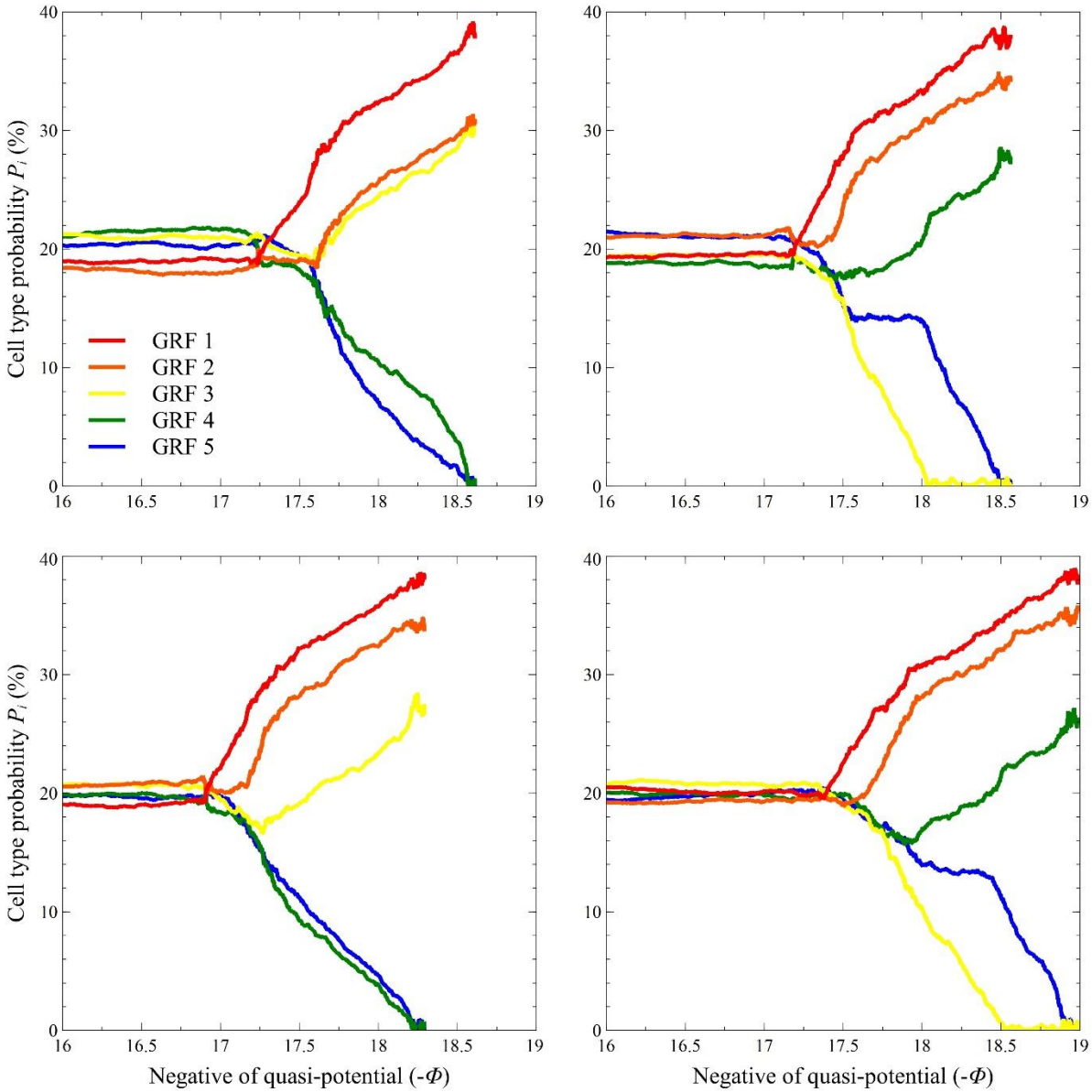
Supplementary Figure 7. Time evolution of u_i for the deterministic system in main text's Fig. 2c. Timescale factor decline rates are ε_1=*0.0010*, ε_2=*0.0012*, ε_3=*0.0013*, ε_4=*0.0100* and ε_5=*0.0500*.



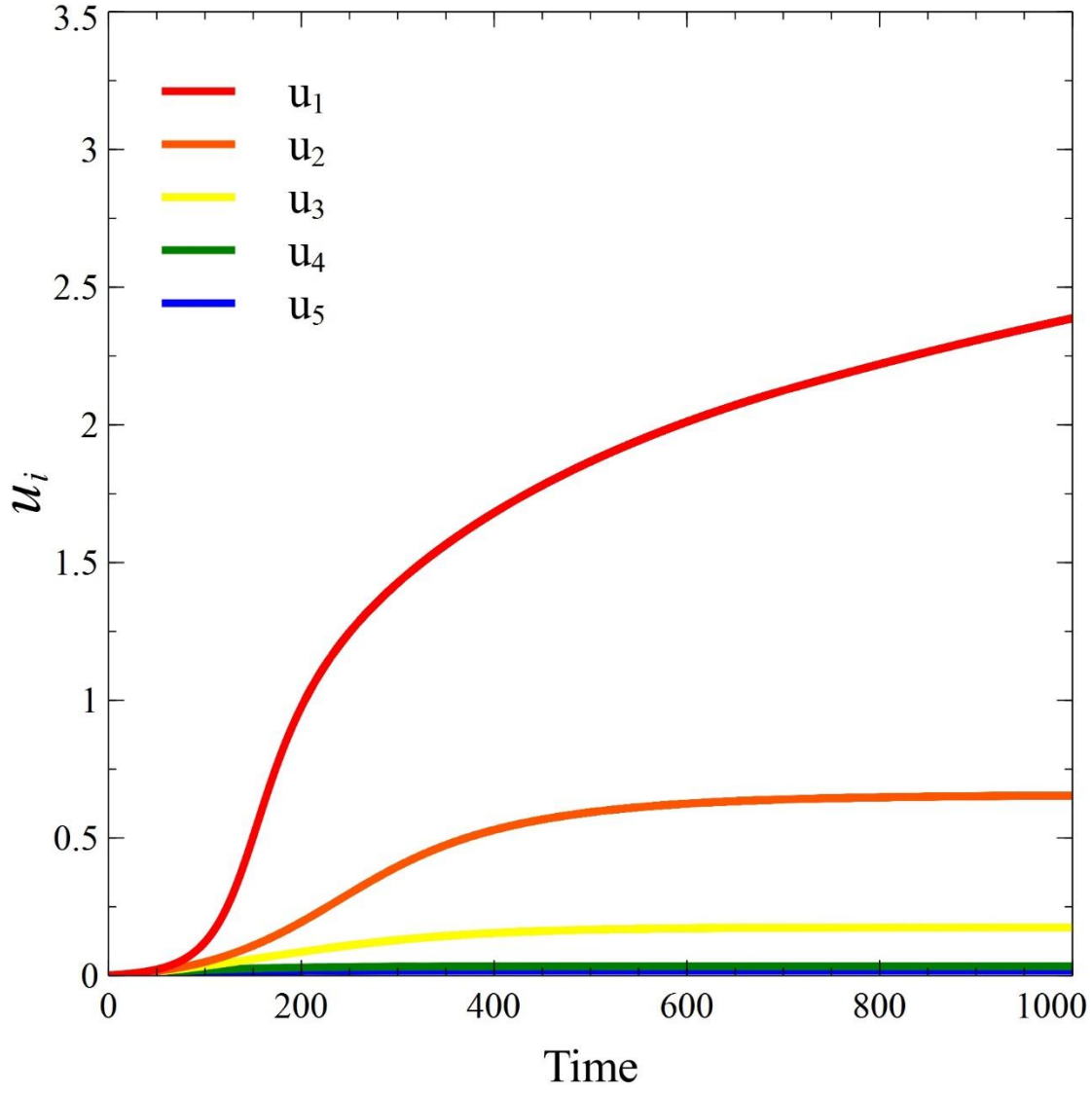
Supplementary Figure 8. Sample stochastic paths of the system in main text's Fig. 2d (three endpoints). The initial condition is $X_i=1$ for all i. The parameter values are $a_{ij}=1/8$ for all i,j, $\varepsilon_1=0.0010$, $\varepsilon_2=0.0015$, $\varepsilon_3=0.0090$, $\varepsilon_4=0.0200$ and $\varepsilon_5=0.0600$.



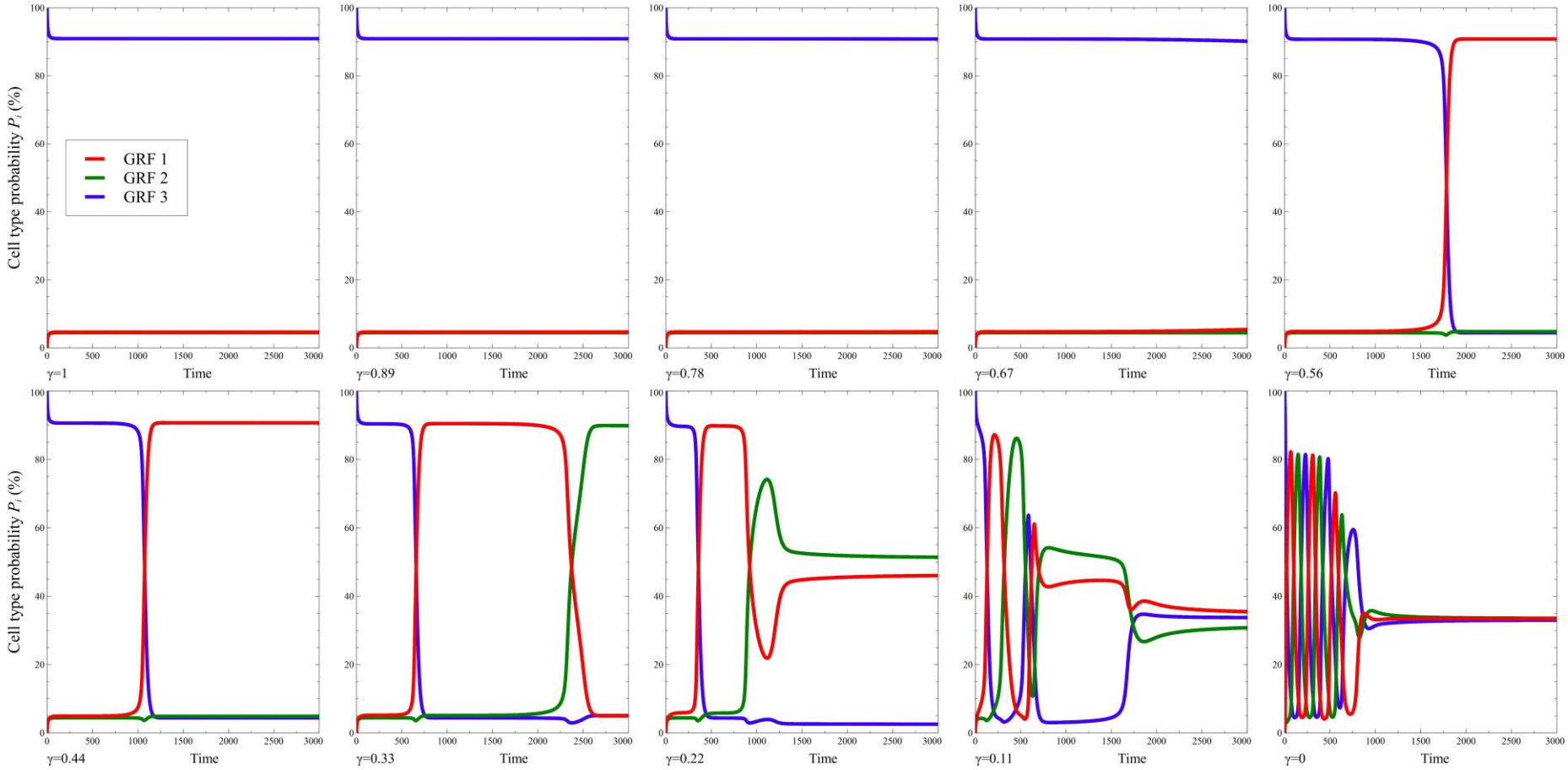
Supplementary Figure 9. Time evolution of u_i for the deterministic system in main text's Fig. 2d. Timescale factor decline rates are ε_1=*0.0010*, ε_2=*0.0015*, ε_3=*0.0090*, ε_4=*0.0200* and ε_5=*0.0600*.



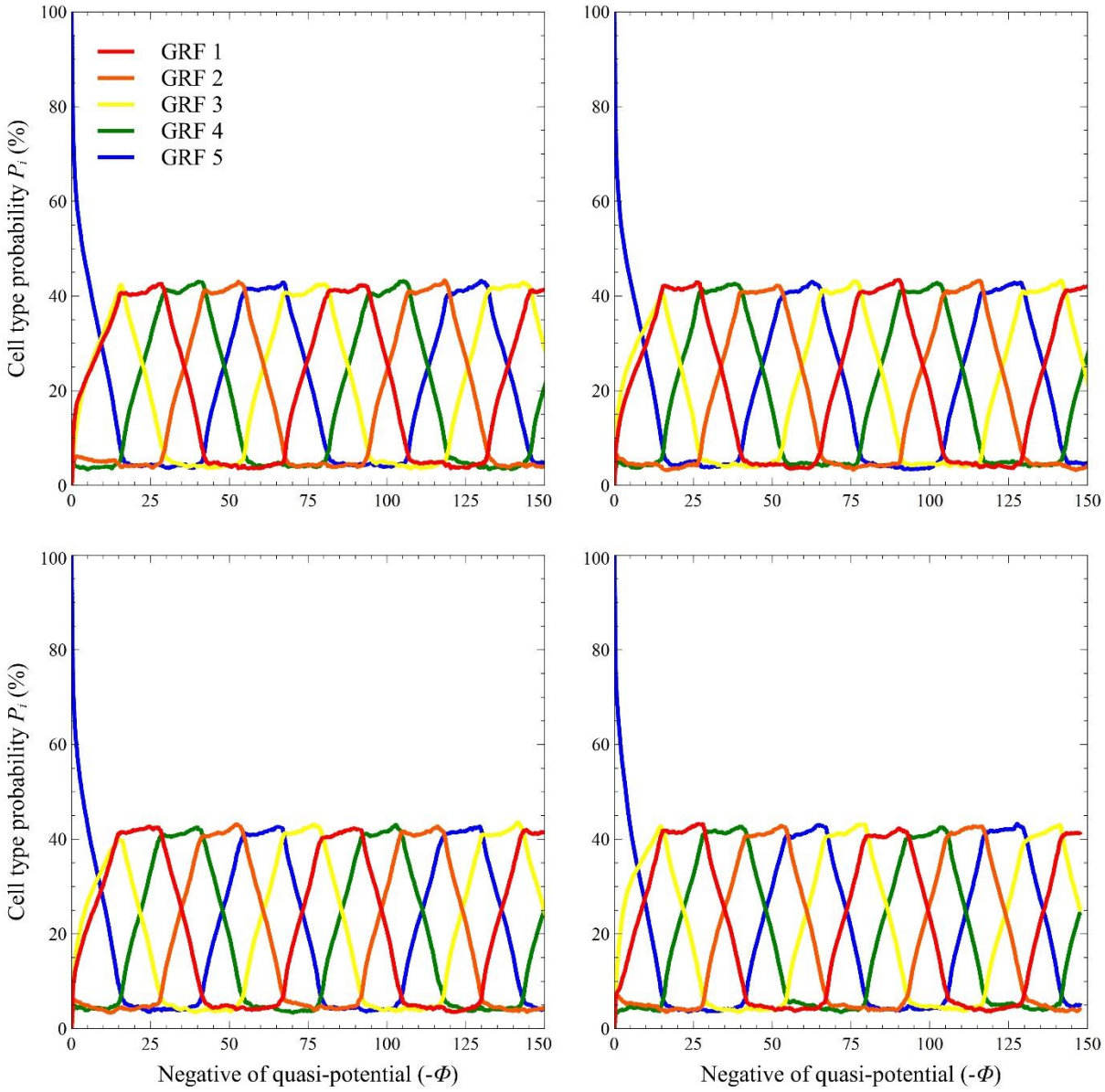
Supplementary Figure 10. Sample stochastic paths of the system in main text's Fig. 2e (four endpoints when deterministic). The initial condition is $X_i=1$ for all i. The parameter values are $a_{ij}=1/8$ for all i,j, $\varepsilon_1=0.0010$, $\varepsilon_2=0.0060$, $\varepsilon_3=0.0090$, $\varepsilon_4=0.0150$ and $\varepsilon_5=0.0200$.



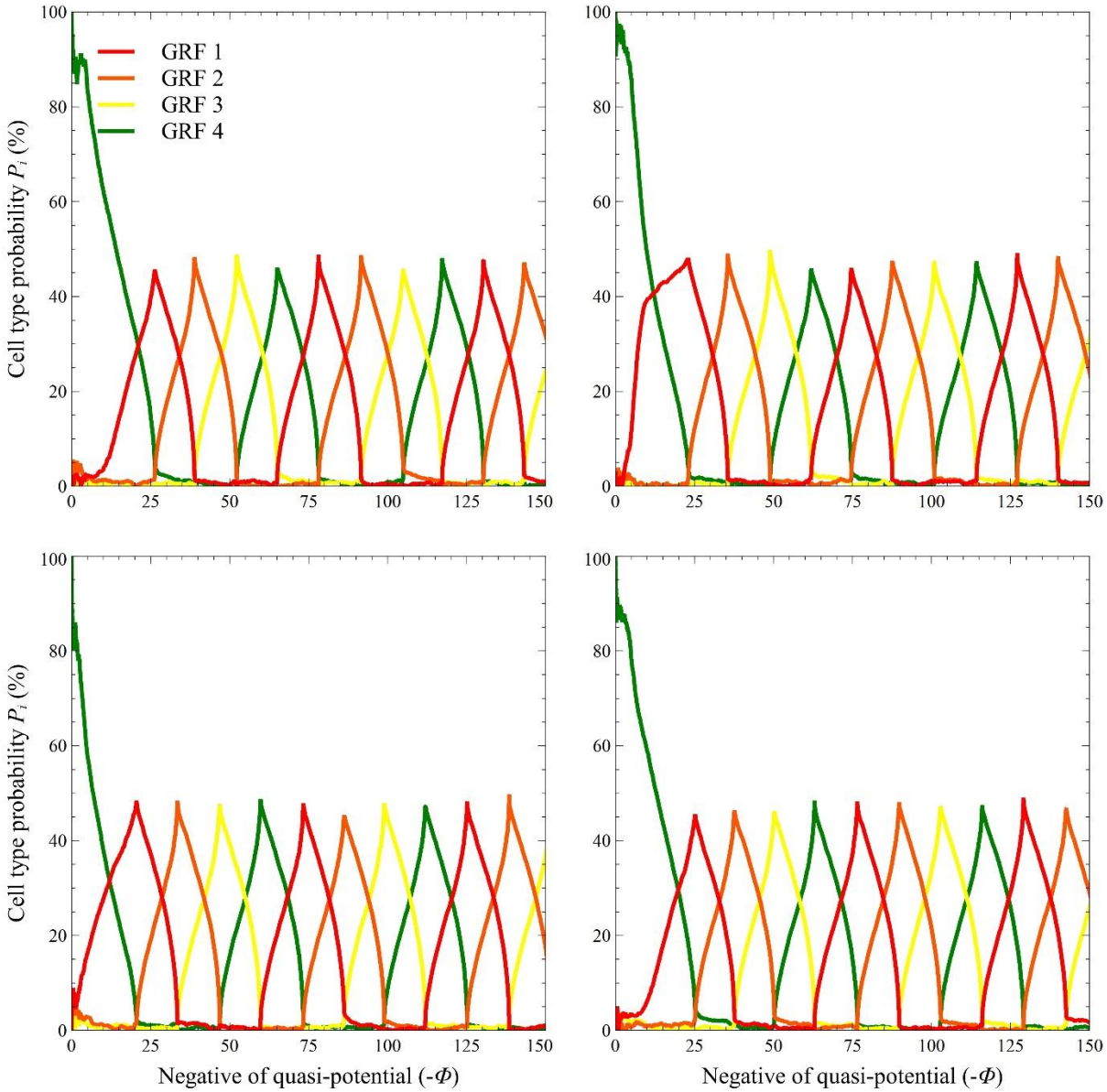
Supplementary Figure 11. Time evolution of u_i for the deterministic system in main text's Fig. 2e. Timescale factor decline rates are ε_1=*0.0010*, ε_2=*0.0060*, ε_3=*0.0090*, ε_4=*0.0150* and ε_5=*0.0200*.



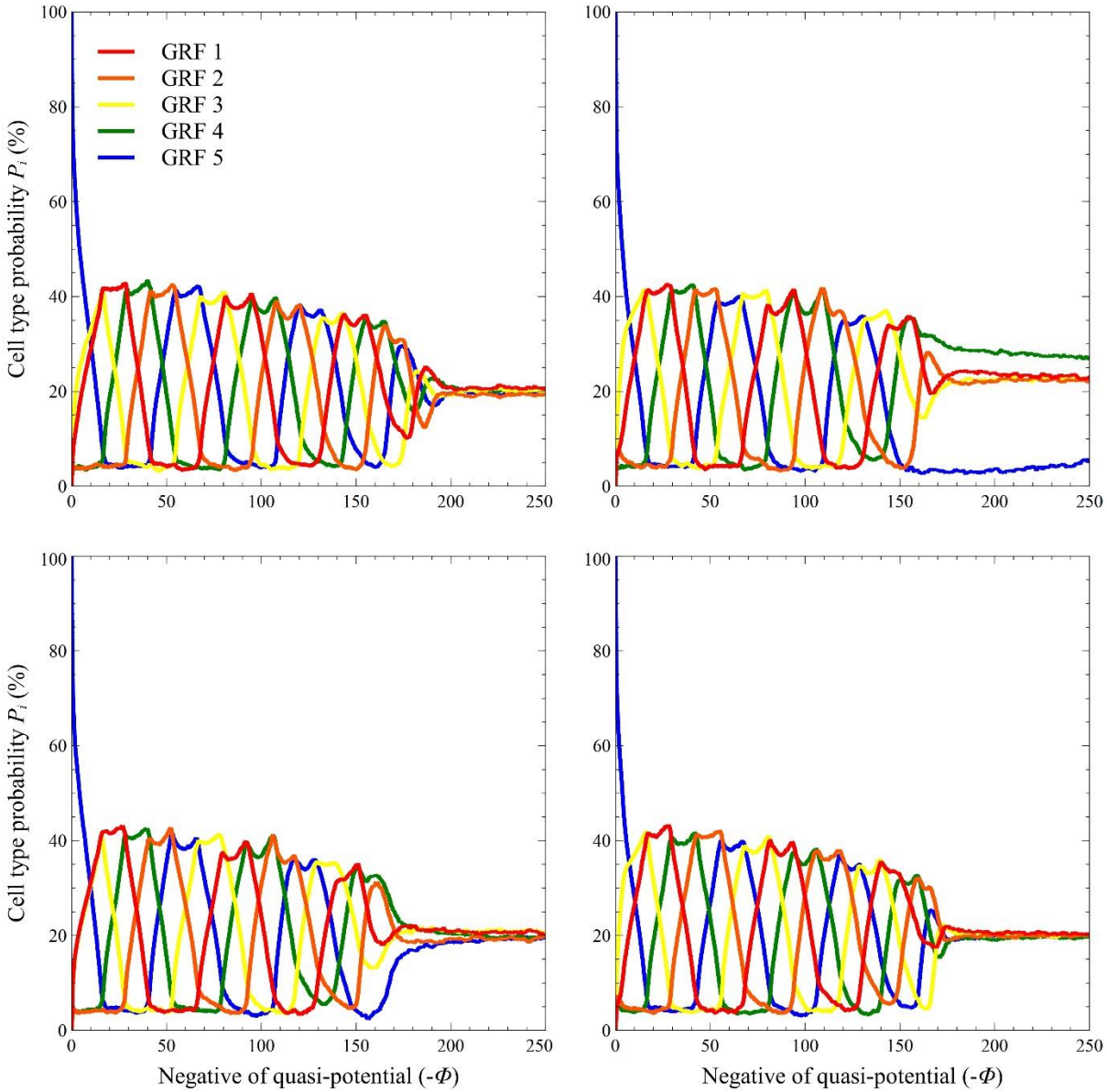
Supplementary Figure 12. Suppose the parameter values are $a_{12}=a_{23}=a_{31}=3$, $a_{13}=a_{21}=a_{32}=\gamma$, $\varepsilon_1=0.001$, $\varepsilon_2=0.001$ and $\varepsilon_3=0.001$. The initial condition is $X_1= X_2=0$ and $X_3=5$. Decreasing the value of γ creates a repressilator network such that one repression loop is stronger than the reverse loop. This results in an attracting oscillatory behavior that activates suppressed genes.



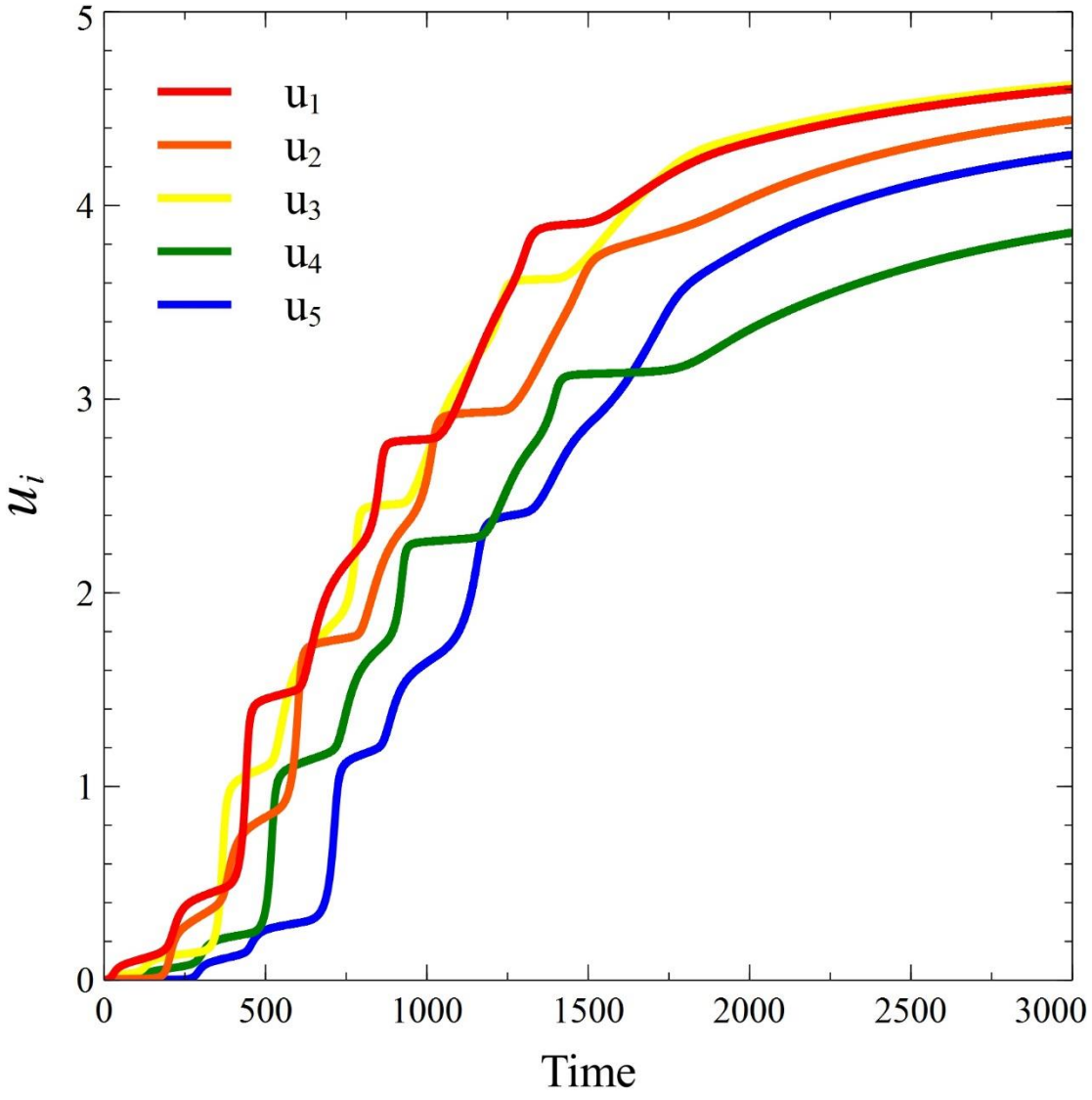
Supplementary Figure 13. Sample stochastic paths of the system in main text's Fig. 3b, $n=5$. The initial condition is $X_i=0$ for $i=1,2,3,4$ and $X_5=15$. The parameter values are $a_{12}= a_{23}= a_{34}= a_{45}= a_{51}=5$ and $a_{ij}=0.01$ for the other i,j. The parameter u_{ij} does not evolve for all i,j.



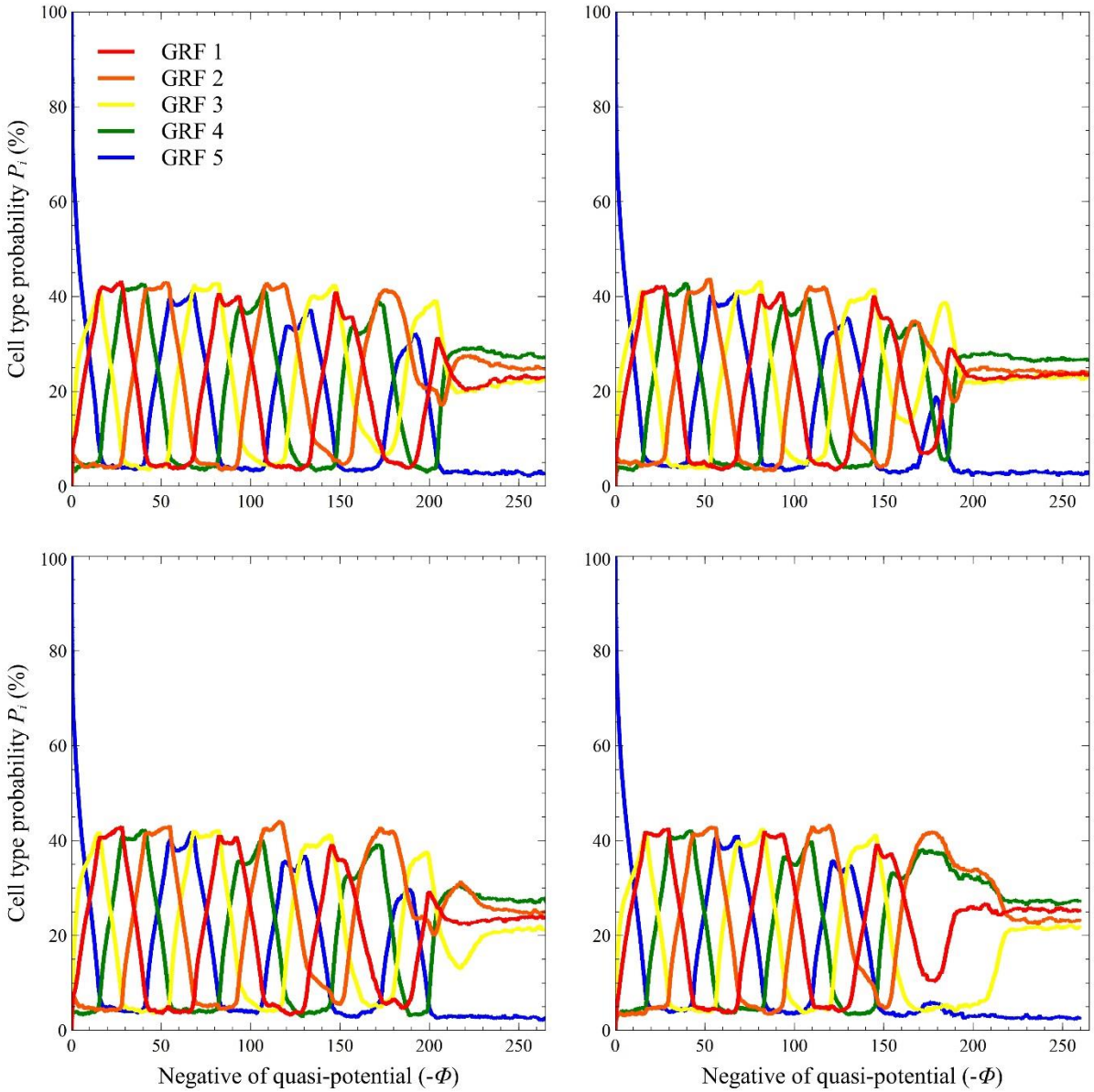
Supplementary Figure 14. Sample stochastic paths of the system in main text's Fig. 3c, $n=4$. The initial condition is $X_i=0$ for $i=1,2,3$ and $X_4=1$. The parameter values are $a_{12}= a_{23}= a_{34}= a_{41}=5$, $a_{13}= a_{24}= a_{31}= a_{42}=1$ and $a_{ij}=0.01$ for the other i,j. The parameter u_{ij} does not evolve for all i,j.



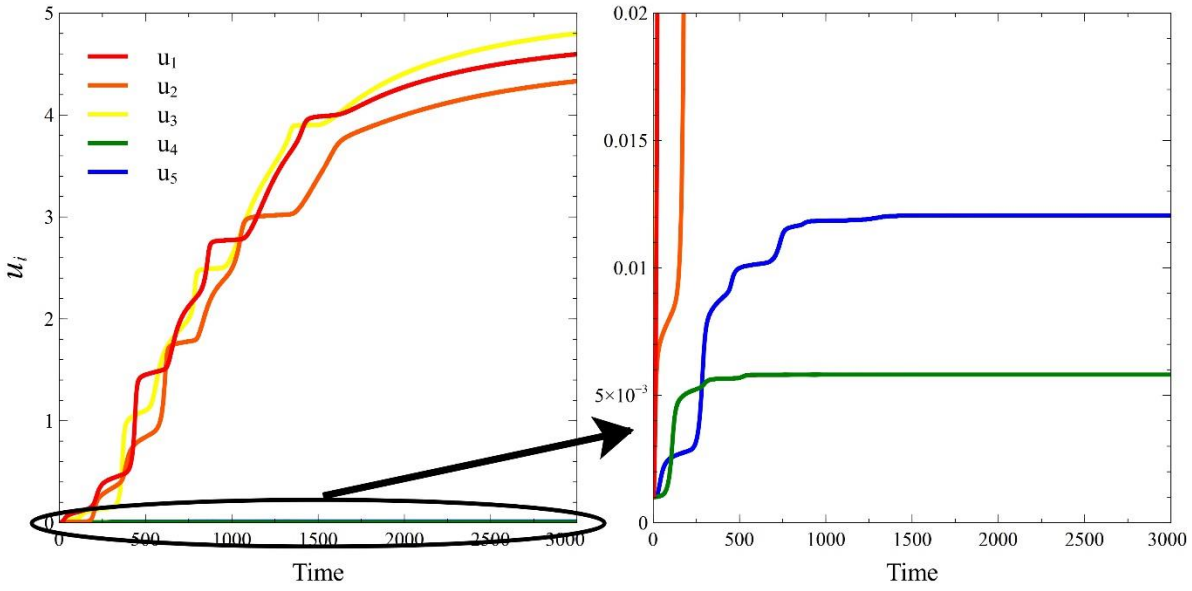
Supplementary Figure 15. Sample stochastic paths of the system in main text's Fig. 3d. The initial condition is $X_i=0$ for $i=1,2,3,4$ and $X_5=15$. The parameter values are $a_{12}= a_{23}= a_{34}= a_{45}= a_{51}=5$, $a_{ij}=0.01$ for the other i,j, $\varepsilon_1=0.001$, $\varepsilon_2=0.\ 0.001$, $\varepsilon_3=0.\ 0.001$, $\varepsilon_4=0.001$ and $\varepsilon_5=0.001$.



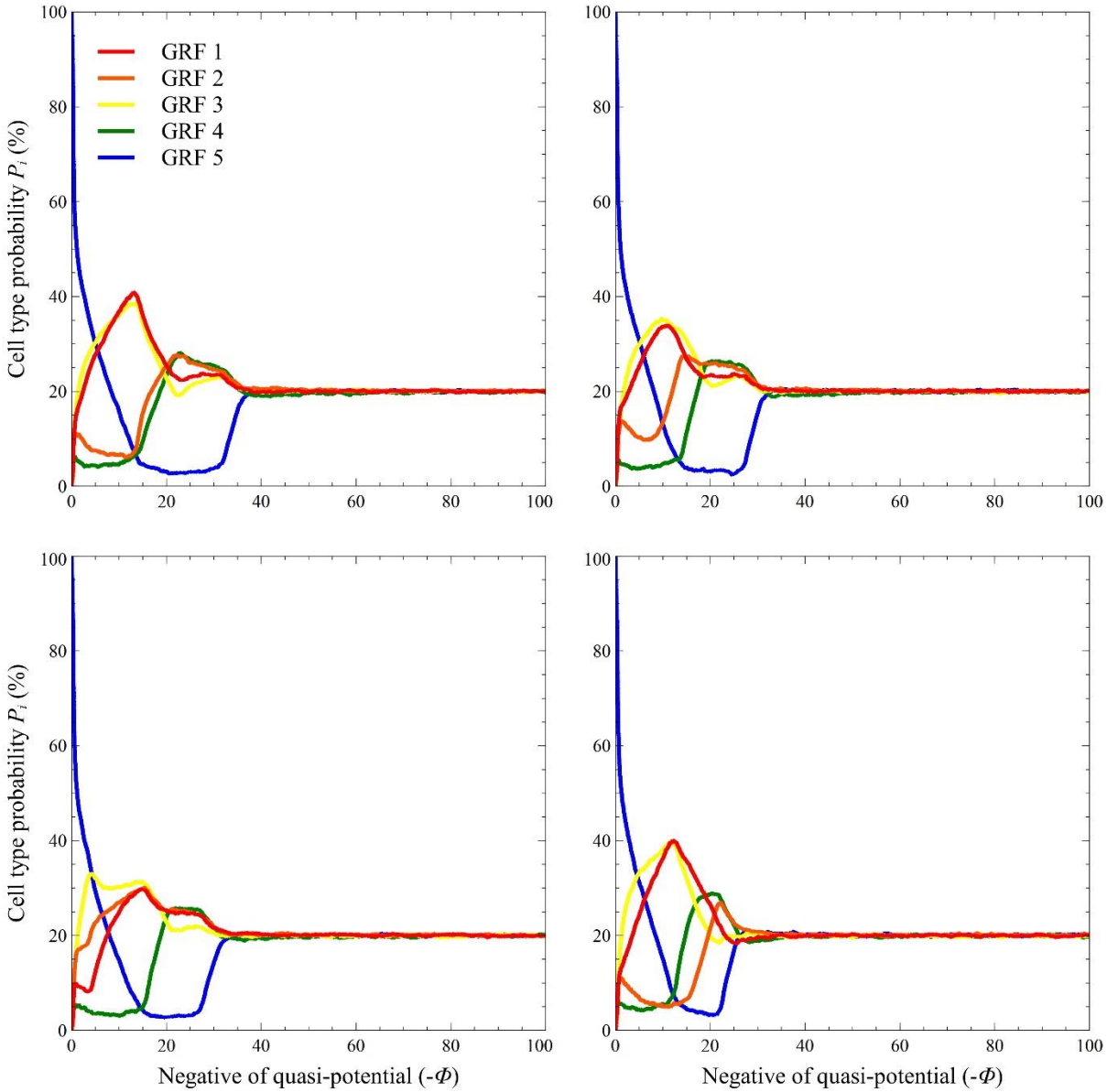
Supplementary Figure 16. Time evolution of u_i for the system in main text's Fig. 3d. Timescale factor decline rates are $\varepsilon_{ij}=0.001$ for all i,j.



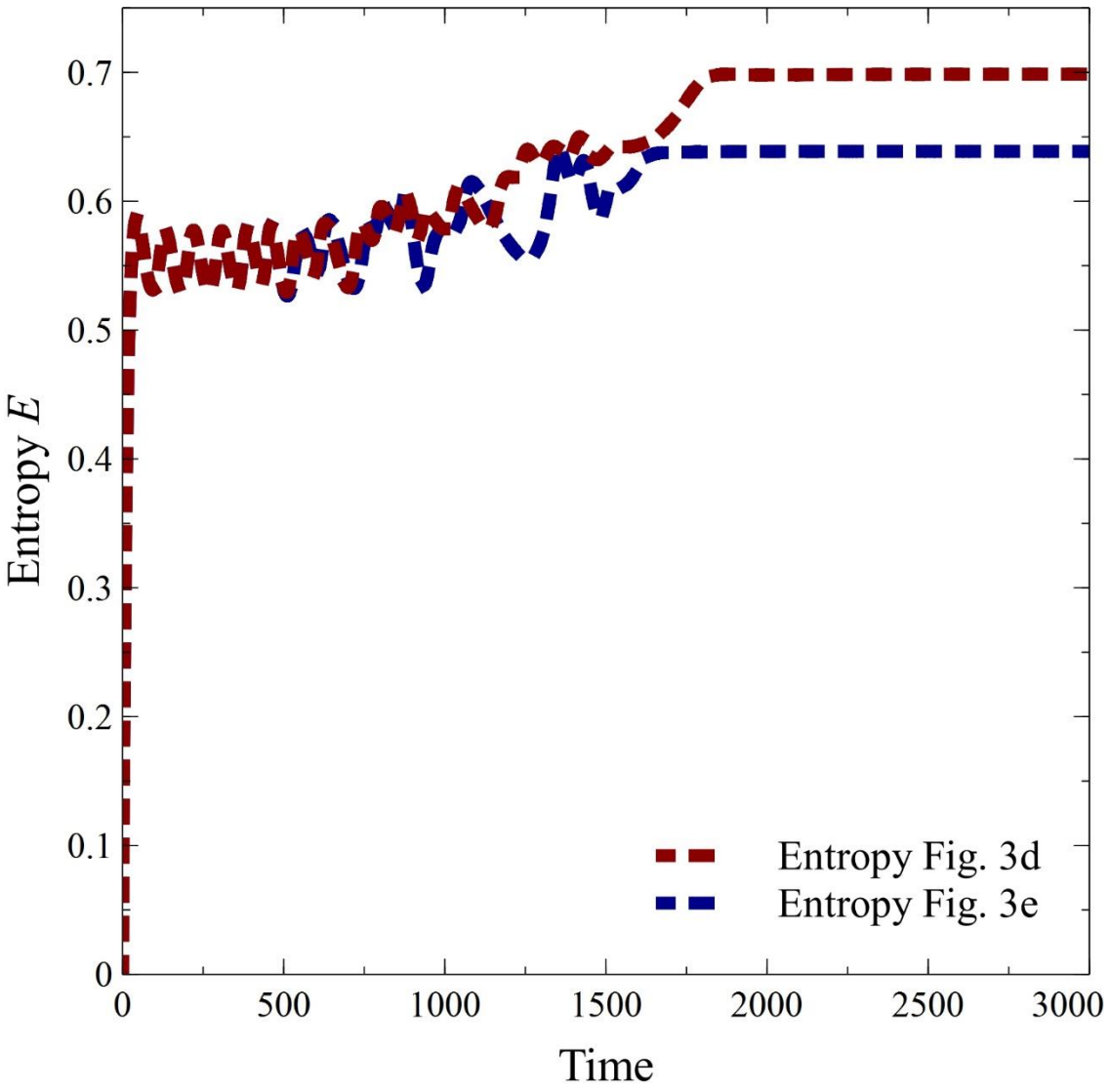
Supplementary Figure 17. Sample stochastic paths of the system in main text's Fig. 3e. The initial condition is $X_i=0$ for $i=1,2,3,4$ and $X_5=15$. The parameter values are $a_{12}= a_{23}= a_{34}= a_{45}= a_{51}=5$, $a_{ij}=0.01$ for the other i,j, $\varepsilon_1=0.001$, $\varepsilon_2=0.\ 0.001$, $\varepsilon_3=0.\ 0.001$, $\varepsilon_4=0.010$ and $\varepsilon_5=0.005$.



Supplementary Figure 18. Time evolution of u_i for the system in main text's Fig. 3e. Timescale factor decline rates are ε_1=*0.001,* ε_2=*0. 0.001,* ε_3=*0. 0.001,* ε_4=*0.010* and ε_5=*0.005*.



Supplementary Figure 19. Sample stochastic paths of the system with slow timescale factor decline rate, $\varepsilon_i=0.0001$ for all i. Here dampening of the oscillations is fast and the initial oscillations are unnoticeable; however, repressilator-type network is still able to activate suppressed genes. The initial condition is $X_i=0$ for $i=1,2,3,4$ and $X_5=15$. The parameter values are $a_{12}= a_{23}= a_{34}= a_{45}= a_{51}=1$ and $a_{ij}=0.1$ for the other i,j.



Supplementary Figure 20. Time series of entropy levels for the systems in the main text's Fig. 3d and 3e.