Homeorhesis in Waddington's landscape by epigenetic feedback regulation

Yuuki Matsushita and Kunihiko Kaneko[*]

Department of Basic Science, Graduate School of Arts and Sciences, University of Tokyo, 3-8-1 Komaba, Meguro-ku, Tokyo 153-8902, Japan

In multicellular organisms, cells differentiate into several distinct types during early development. Determination of each cellular state, along with the ratio of each cell type, as well as the developmental course during cell differentiation are highly regulated processes that are robust to noise and environmental perturbations throughout development. Waddington metaphorically depicted this robustness as the epigenetic landscape in which the robustness of each cellular state is represented by each valley in the landscape. This robustness is now conceptualized as an approach toward an attractor in a gene-expression dynamical system. However, there is still an incomplete understanding of the origin of landscape change, which is accompanied by branching of valleys that corresponds to the differentiation process. Recent progress in developmental biology has unveiled the molecular processes involved in epigenetic modification, which will be a key to understanding the nature of slow landscape change. Nevertheless, the contribution of the interplay between gene expression and epigenetic modification to robust landscape changes, known as homeorhesis, remains elusive. Here, we introduce a theoretical model that combines epigenetic modification with gene expression dynamics driven by a regulatory network. In this model, epigenetic modification changes the feasibility of expression, i.e., the threshold for expression dynamics, and a slow positive-feedback process from expression to the threshold level is introduced. Under such epigenetic feedback, several fixed-point attractors with distinct expression patterns are generated hierarchically shaping the epigenetic landscape with successive branching of valleys. This theory provides a quantitative framework for explaining homeorhesis in development as postulated by Waddington, based on dynamical-system theory with slow feedback reinforcement.

I. INTRODUCTION

In most multicellular organisms, cells differentiate into several types in the course of development, which show distinct gene expression patterns that are robust to external perturbations and internal noise. As a theoretical explanation for this robustness, Waddington introduced the concept of the epigenetic landscape more than 60 years ago, as shown in Fig. 1(a). In this concept, a ball falling along the landscape represents the cell differentiation process over time, and each valley corresponds to a differentiated cell type [1]. Although presented visually as a metaphor, Waddington also proposed that this differentiation process can be understood in terms of the dynamical systems of gene expression. Following his insight, each valley is now interpreted as an attractor of an intracellular dynamical system for gene (protein) expression. The cellular state is represented by a set of gene (protein) expression levels reached as a result of such dynamical system, as represented by an attractor of such dynamical system.

Then, the cellular state remains in the vicinity of the attractor even under internal noise or external perturbation. In fact, several dynamical-systems models with mutual activation and inhibition of protein expression demonstrated the coexistence of multiple attractors that correspond to distinct cell types [2–5], and supporting experiments have been carried out [6,7].

According to this dynamical-systems approach, the *X* axis characterizing the cellular state in Fig. 1(a) is represented by the gene expression pattern. However, since there are thousands of genes (or components) in a cell, the state may not be accurately represented by a one-dimensional variable *X*. Nevertheless, the cellular state can potentially be represented by only a few variables extracted from data reduction of the expression levels of a huge number of components, such as principal component analysis (PCA) [8].

Moreover, the height of the landscape (*Z* axis) represents changeability of the state. Cellular states are attracted to the bottom of the valley, which, in terms of dynamical systems, are fixed-point attractors at which point no more change will occur.

Along with the dynamics falling onto the bottom of the valley, as represented by motion along the *X* axis in Fig. 1(a), the landscape itself is shaped along the other (*Y*) axis representing the developmental course, in which the valleys are shaped successively and are deepened, in a process known as canalization. Therefore, a fundamental question remains: given that the attraction to each valley along the *X* axis is represented by gene-expression dynamical systems, what does

---

[*]kaneko@complex.c.u-tokyo.ac.jp; also at Research Center for Complex Systems Biology, Universal Biology Institute, University of Tokyo.
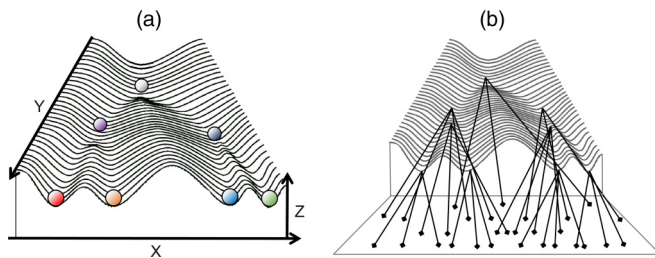
FIG. 1. (a) Waddington's epigenetic landscape. The cell differentiation process is conceptually explained as motion of a ball along the landscape in which valleys correspond to differentiated cell types. Here, the horizontal axis (X) represents a cellular state, the height (Z axis) represents the inverse of the frequency (probability) that a cell takes state X, and the Y axis represents slow developmental change. Adapted from Ref. [1]. (b) Shaping the landscape by genes, represented by strings (also adapted from Ref. [1]). Waddington entitled the figure as the "complex system interaction underlying the epigenetic landscape," where interactions among genes control the landscape as metaphorically represented by strings.

the Y axis representing (slower) landscape change represent? To address this fundamental question, there are three basic questions to resolve with respect to the postulates of Waddington's landscape itself.

First, there is the issue of hierarchical branching. That is, since the valleys are successively generated over developmental time [Fig. 1(a)], many valleys (attractors) are not generated independently, but rather the shallower valleys are generated first and are then branched, and these branching processes are repeated [9,10].

Second, Waddington argued that the developmental process itself, i.e., the motion along the shaping of valleys, is also robust to perturbation, and coined the term homeorhesis to represent such path stability [1,11]. However, the mechanism contributing to the robustness of this shaping process, including successive branchings, remains elusive [4].

Finally, the number ratio of each cell type is also rather robust to perturbations or initial conditions. If we assume that a deeper valley attracts more cells, this robustness implies overall robustness of the landscape, in particular, the depth of each valley [12–18].

Considering these three postulates of the landscape, let us now come back to the fundamental question of the nature of the Y axis representing slow developmental change. Waddington drew the diagram Fig. 1(b) to show schematically how genes control the epigenetic landscape, in particular the change of valleys along the developmental time. After 60 years since proposal of the epigenetic landscape, we have now possible candidates that could cause such a slow landscape change. One candidate is the cell-cell interactions [16–18]. As development progresses and the cell number increases, the influence of cell-cell interactions on the intracellular dynamics for each cell type is stronger. Slow modifications of intracellular expression dynamics can lead to novel attractors or the increase in their robustness.

Another potential source for slow landscape change is change in chromatin structure with epigenetic modification, which is currently one of the hottest topics in cell and devel-

opmental biology [19–25]. Epigenetics is a field derived from a term "epi(above)-genetic" coined by Waddington, 60 years ago [1,11]. At that moment it would cover all possible changes beyond the genetic one. Now, epigenetics are mainly studied as a variety of molecular mechanisms that affect feasibility of the expression of each gene at a given time and place without change in DNA sequences, such as the methylation or some other molecular modifications in DNA [26–28]. These modifications change, for example, the openness of chromatin around a promoter of gene, which affects the flows of gene regulation, i.e., the efficiency of transcription binding promoter changes depending on the degree of openness [29,30]. If the chromatin is more open around a gene, it is more feasible to be expressed by the actions of other genes on its promoter. As for modifications to the DNA influencing on the openness over a long time span, novel molecular mechanisms are still being uncovered, with the ongoing development in the field of epigenetics [31].

Note that epigenetic modification generally depends on a given cellular state, i.e., the expression levels of proteins, whereas the epigenetic changes, e.g., openness of chromatin structures, influence the expression levels. In general, the epigenetic process is slower than the expression dynamics, and the epigenetic change lasts over the time span in the change in protein concentrations [29,32–35]. The epigenetic change leads to stabilization of cellular states, which can be corresponded to deepening the valleys as schematically represented in Fig. 1(b). This stabilization may imply the existence of a positive-feedback process between the expression and epigenetic change as suggested experimentally [30,36], and also theoretically [32–34,37].

Despite these suggestions and growing reports on epigenetic changes, however, the detailed interplay between epigenetic modification and gene expression dynamics has not been fully explored. In contrast to the extensive body of theoretical and empirical literature on expression dynamics or epigenetic modifications, there is little experimental elucidation of the underlying molecular mechanisms nor theoretical model for the interplay proposed to date. Therefore, at this stage, a simple phenomenological model is needed to investigate how such slow epigenetic change can introduce a novel expression pattern or stabilize the existing expression patterns. Such a model would provide a bridge between epigenetic modification and the epigenetic landscape as Waddington conceptualized.

To formulate the epigenetic process in terms of dynamical systems, we here introduce an epigenetic variable for each expressed gene, represented as a threshold level of the input needed for the gene of concern to be expressed. Using the simplest feedback process, we elucidate the possible conditions for the epigenetic landscape and its properties. Rather than seeking detailed models extracted from realistic expression dynamics, we instead consider a minimal conceptual model that captures the interplay between the relatively faster gene expression dynamics and slower epigenetic dynamics to address how an epigenetic landscape satisfying the requisites of (i) hierarchical branching, (ii) homeorhesis, and (iii) robustness in the cell-number ratio is generated. Instead of the simplicity in the model, we have simulated thousands of networks to extract a universal mechanism and draw a general

conclusion, which will hold true in a complicated system with biological reality.

II. MODEL

We consider a cell model with a gene regulatory network (GRN) and epigenetic modification. The cell has N genes and the cellular state is represented by the expression level (or concentration) x_i of each gene i. Here, the GRN represents the mutual control of genes via synthesized proteins. Gene expression typically shows an on-off-type response to the input: a gene is expressed (suppressed) when its input value is above (below) a certain threshold, whereas its expression level is saturated as the input value increases. By normalizing the maximal (i.e., saturated) expression level to unity, we adopt the following gene expression dynamics for simplicity [38–42]:

$$\frac{dx_i}{dt} = F\left(\sum_j^N \frac{J_{ij}}{\sqrt{N}} x_j + \theta_i + c_i \right) - x_i, \tag{1}$$

where J_{ij} is the regulatory matrix. If J_{ij} is positive (negative), gene j activates (represses) the expression of gene i, whereas J_{ij} is set to 0 if no regulation exists. To represent the on-off-type expression of genes, we adopt $F(z) = \tanh(\beta z)$ ($\beta = 40$). $x_i = 1$ indicates the full expression of the ith gene and $x_i = -1$ indicates no expression of the ith gene. c_i is a constant input value, interpreted as an input outside of the N genes (for example, upstream genes) or the natural trend for expression. For most examples, however, c_i is set to 0 unless otherwise noted.

Here, $-\theta_i$ represents the threshold for the input, beyond which the expression is activated. As θ_i is increased (decreased), the ith gene tends to be expressed (repressed). In the standard GRN model, this threshold is fixed. By contrast, we regard it as a variable by assuming that θ_i represents the epigenetic modification level for each gene i, such as histone modification or openness of chromatin structure.

Further, the epigenetic change depends on the expression level of gene i. In accordance with some theoretical and experimental reports [30,32,33,36,44,45], we adopt a positive-feedback process from gene expression to epigenetic modifications: when a gene is expressed (repressed), it tends to be expressed (repressed), as is consistent with the stabilization by the epigenetic changes. Thus the epigenetic change (openness in the chromatin structure) is given by:

$$\frac{d\theta_i}{dt} = v(a x_i - \theta_i). \tag{2}$$

This assumption is rather natural considering the concept of canalization, i.e., each valley becomes deeper (or more stable) as development. In Eq. (2), the parameter $a(> 0)$ represents the strength of the positive-feedback mechanism, and $v(> 0)$ gives the rate of change in the epigenetic modification.

In the model, both θ_i that characterizes the degree of openness and $\sum_j J_{ij} x_j$ that represents the actions from other proteins synergetically determine the expression. In the equation, only if $\sum_j J_{ij} x_j + \theta_i$ is larger than the threshold c_i, i.e., if the input is large and the chromatin is sufficiently open, x_i is expressed. As the term $\sum_j J_{ij} x_j$ is bounded by unity, if θ

decreases goes to a negative value (towards $-\alpha$), the gene is not expressed. Hence if the chromatin is closed, the gene is not expressed. Further, because θ_i increases as a result of positive feedback from x_i, the gene turns to be expressed only when the chromatin is open and activators bind to the promoter.

Remark. Here we remark on the simplification adopted in our modeling and its limitation. Biologically plausible models often adopt the Hill function such as $F^{Hill}(z) = z^\alpha / (z^\alpha + K^\alpha)$ [or $F_s^{Hill}(z) = 1/(z^\alpha + K^\alpha)$] for the expression dynamics, instead of $F(z)$, where combination of $x_m^\alpha / (x_m^\alpha + K_m^\alpha)$ (for activation) and $1/(x_m^\alpha + K_m^\alpha$ (for suppression) provides the expression dynamics [46,47]. The qualitatively same gene-expression dynamics (e.g., on/off expression patterns, multi-stability, oscillation), however, is observed both for the models with $F(z)$ [38–41] and the Hill-type model.

In the Hill function, the value K corresponds to the threshold for the on/off -type expression dynamics. With the increase in K, the threshold for expression increases. Hence, K and θ in our model play the same role. Now K's change with the affinity with the prompter and proteins, or in other words, with the chromatin openness. Considering that epigenetic change to stabilize each expression pattern, as discussed above, the feedback from x to K in the form of Eq. (2) was introduced in a previous research [46] for a Hill-type model extracted from the real data. Indeed, Eq. (2) is obtained by replacing K by θ.

Here, as long as the threshold-type expression dynamics and slow positive epigenetic feedback process in the type of Eq. (2) is adopted, qualitatively similar behaviors are obtained with regard to the classes of the expression dynamics and the fixation from oscillatory dynamics, to be discussed below.[1]

Note that the two basic assumptions (i) threshold-type gene expressions via synthesized proteins and (ii) slow epigenetic modification (chromatin openness) that stabilizes the expression pattern by the feedback are essential to draw the conclusion in the present paper. By our approach, however, one cannot make a direct, quantitative prediction on the epigenetic landscape for a specific example, as it requires quantitative information on the specific form of gene expression dynamics and epigenetic feedback. This is a limitation in our modeling.

III. FIXED-POINT ANALYSIS

The fixed-point solutions of (1) and (2) are obtained by setting each term to zero. From the latter, we get $\theta_i = a x_i$ and from the former we obtain

$$\tanh \beta \left(\sum_i^N \frac{J_{ij}}{\sqrt{N}} x_j^* + a x_i^* \right) - x_i^* = 0 \tag{3}$$

(note that the case with $c_i = 0$ is considered here). In the large β limit, the tanh function is approximated by the step function,

---

[1]The reason for such agreement comes from the dynamical-systems theory. As shown in the later sections, types of attractors (which of the genes are expressed, multistability, oscillation) are determined by how nullclines are crossed, whereas a consequence of the epigenetic change is determined by how the nullclines are shifted. Indeed, our model and the Hill-type model have common behaviors as for the crossings and shift of nullclines.

so that the fixed point x_i^* is given by a sequence of $\{-1, 1\}$ that satisfies (3). The number of fixed points of (3) then increases monotonically with the value of a (Fig. S1 in Supplemental Material [43]). If it is large enough (that is, the second term in the brackets in (3) is sufficiently larger than the first term), all of the 2^N patterns with any combination of $x_i^* = \pm 1$ (with $\theta_i^* = a x_i^*$) satisfy (3). All of these are fixed-point attractors, which are reached by choosing initial conditions close to each $\{-1, 1\}^N$ state. However, for $a = 0$, the number of fixed points satisfying $x_i^* = \tanh \beta (\sum J_{ij} x_j^* / \sqrt{N})$ is much smaller.

Here, we focus on a case with sufficiently strong epigenetic feedback, i.e., sufficiently large a, in which all of the possible 2^N states could exist if any value of x_i and θ_i is initially chosen. However, for studying the canalization dynamics, we restrict the initial condition of θ_i as follows: At the initial stages of development, epigenetic modification is not yet introduced [48–52], so that all of θ_is are set to 0. Under this restriction, we investigate which of the 2^N fixed points with $x_i^* = \pm 1$ and $\theta_i^* = a x_i^*$ is reachable through developmental change of the epigenetic modification. As we limit the dynamics to the state of $\theta_i = 0$, we refer to only the final states reached from such initial conditions as attractors throughout the paper (whereas the initial conditions of x_is cover all possible $\{-1, 1\}$ states).

IV. ATTRACTOR GENERATION AND PRUNING

First, we set $N = 10$ and prepare the initial conditions for all gene expression patterns with null epigenetic modification (i.e., 2^N candidates with $x_i = \pm 1$, $\theta_i = 0$). In the context of the epigenetic landscape, these initial conditions correspond to the balls on the top of the landscape, whereas the valleys are shaped with the change in θ_i and the balls are trapped at the bottom of the landscape that correspond to the attractor. We then examine which and how many attractors are selected depending on the parameter v.

At the limit of $v \rightarrow 0$, i.e., the adiabatic limit in terms of physics, the time scales of the dynamics for x_i and θ_i are well separated. Only after the expression level x_i reaches one of the original attractors with θ_i originally fixed at 0, θ_i begins to show gradual variation. Hence, the number of attractors will be bounded by the expression dynamics when fixing $\theta_i = 0$. At the limit of $v \rightarrow \infty$, θ_i reaches $\theta_i = a x_i$ faster, so that all of the 2^N states $x_i^* = \tanh(\beta a x_i^*)$ are attracted depending on the initial x_i values, as long as a is sufficiently large. By considering these two extreme limits, v generally functions as a parameter that limits the final state from all of the possible 2^N states. Now, from naive expectation based on the above two limits, it might be expected that the number of attractors will monotonically increase with v. Indeed, such monotonic increase could be observed for 80% of randomly chosen networks J_{ij} for $N = 10$.

For $v \sim 0$, the approach to the attractor is completed before epigenetic modification and then θ_i is fixed accordingly. With the introduction of v, θ_i increases or decreases depending on the initial value of x_i. If this process for x_i is fast, x_i is fixed to ± 1 depending on the initial condition; that is, before the approach to the original attractor. Hence, the original basin of attraction is partitioned. With the increase in v, more partitions progress; accordingly, the few attractors that exist at $v = 0$ are successively partitioned toward 2^N states with the increase
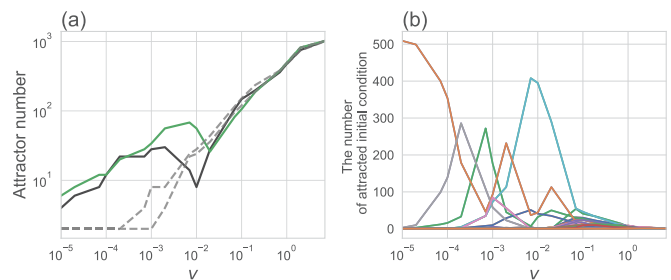

FIG. 2. (a) Dependence of the number of attractors (reached states from $\theta_i = 0$) upon v. Gray dotted lines show the case with a monotonic increase of the attractor number against v. The black and green solid lines are examples with nonmonotonic dependence on v. Attractors are pruned at $2 \times 10^{-3} < v < 1 \times 10^{-2}$. $N = 10$. (b) Dependence of the basin volume of each attractor upon v, for the example of nonmonotonic dependence of attractor number shown as the black line in (a). Basin volume is computed by taking 2^N initial conditions of $\{x_i = \pm 1\}$ and setting $\theta_i = 0$ initially, and then counting the number of initial conditions reaching each fixed-point state. Each line with a different color shows the basin volume for each different attractor.

in v. In this case, for a given v, fixation simply occurs from the neighborhood of each on/off-pattern attractor provided by the initial condition. There exists no hierarchical branching to each attractor over developmental time. Moreover, since only the attractor from the neighborhood of the initial expression state is reached, the final state crucially depends on the initial condition. Neither homeorhesis nor robustness in the cell-number ratio is expected, as will be confirmed later.

However, in the case of $N = 10$, approximately 20% of the randomly chosen matrix J_{ij} shows nonmonotonic dependency of the attractor number on v. Here, different attractors are generated and pruned successively with v in the intermediate range of v. This implies that states separated at smaller v converge again with the increase in v, even though the epigenetic feedback tend to separate each x_i to ± 1. With mutual interference between the fast dynamics of x_i and slower dynamics of θ_i, both the convergence of initial states and divergence to fixed states coexist, as will be discussed below. Further, as will be shown, such convergence of orbits in the initial regime can allow for creation of an epigenetic landscape that satisfies the three postulates of hierarchical branching, homeorhesis, and robustness in the cell-number ratio.

In this nonmonotonic case, the basin volume of each attractor, i.e., the fraction of initial conditions from which each attractor is reached, also changes with v. In particular, dominant attractors successively change with v as shown in Fig. 2(b). This scenario is in stark contrast with the case of a monotonic increase in attractor number, where each basin of attraction is simply partitioned to 2^N successively with the increase in v (Fig. S2).

V. TRAJECTORY SEPARATION BY EPIGENETIC MODIFICATION: SIMPLEST EXAMPLE

To understand how mutual feedback between gene expression and epigenetic modification can lead to the generation and pruning of attractors, we first consider the minimal case
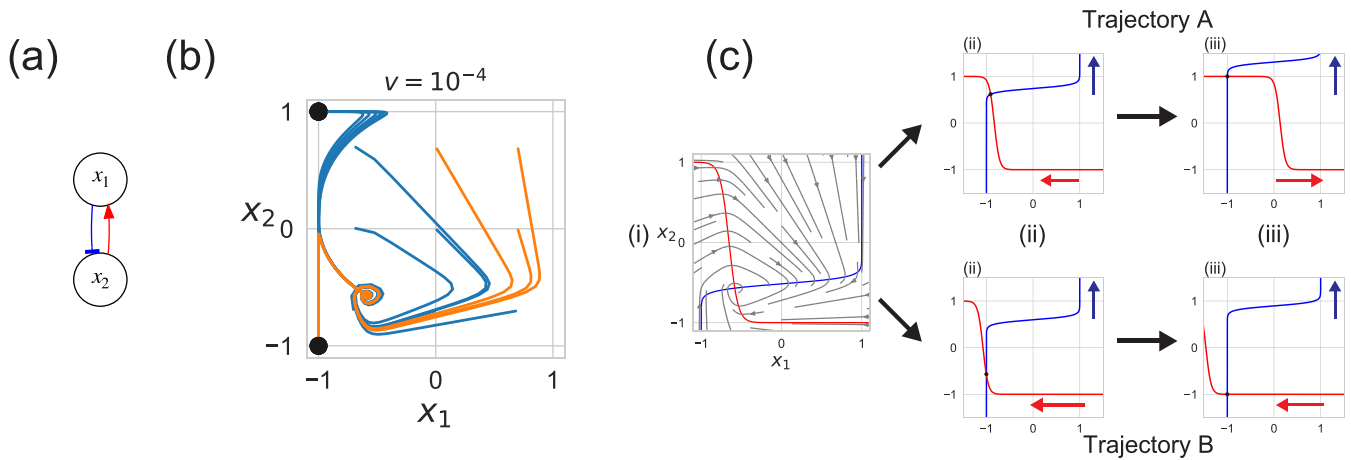
FIG. 3. (a) Two-gene system with activation (red arrow) and inhibition (blue arrow): x_2 activates x_1 and x_1 inhibits x_2. The parameter values are chosen as $J_{12} = 0.44$, $J_{21} = -0.33$, $J_{11} = J_{22} = 0$, $c_1 = 0.16$, $c_2 = -0.15$. (b) Trajectories in x_1, x_2 for $v = 10^{-4}$. Trajectory A reaches the fixed point $(-1, 1)$, whereas Trajectory B reaches $(-1, -1)$. These two types of trajectories coexist (two fixed points as two black dots), depending on the initial condition, for the intermediate value of v. Initial conditions are chosen at even intervals per 0.5 in the phase space of (x_1, x_2). (c) Analyses of the two types of trajectories according to the motion of two nullclines: blue, corresponding to $dx_1/dt = 0$; red, $dx_2/dt = 0$. (i) (x_1, x_2) approaches the cross point of the nullclines if θ_i is fixed, whereas the change in θ_i results in a shift of the nullclines. (ii) For both trajectories, x_1 nullcline (blue line) goes up and x_2 nullcline (red line) goes left first, because (x_1, x_2) first approaches the fixed point at $\theta_i = 0$ starting from any initial condition. Top: As x_2 exceeds 0, the motion of the x_1 nullcline changes its direction, and (x_1, x_2) reaches the fixed point $(-1, 1)$. This gives Trajectory A. Bottom: Before x_2 reaches 0, the x_2 nullcline crosses the x_1 nullcline vertically in Trajectory B so that x_2 remains negative and the motion of the nullclines do not change their direction; thus, (x_1, x_2) reaches $(-1, -1)$.

with only two genes ($N = 2$). In addition, $c_1, c_2 \neq 0$, which may be also regarded as the inputs from genes other than $i = 1, 2$. We consider the case $J_{11} = J_{22} = 0$, $J_{12} > 0 > J_{21}$; i.e., one gene activates the other, which then inhibits the first, as shown in Fig. 3(a).

In this simple case, the number of attractors changes as $1 \rightarrow 2 \rightarrow 1$ with the increase in v over a certain range of parameters c_1, c_2 [Fig. S3(a)]. For $v < 5.7 \times 10^{-5}$, only trajectories reaching $(-1, 1)$ are realized (Trajectory A) [Fig. S3(b)]. By increasing v further, trajectories reaching $(-1, -1)$ then appear (Trajectory B) where $5.7 \times 10^{-5} < v < 7.0 \times 10^{-4}$, and two attractors $(-1, 1)$, $(-1, -1)$ coexist [Fig. 3(b)]. For larger v ($7.0 \times 10^{-4} < v < 9.1 \times 10^{-3}$), the attractor $(-1, 1)$ disappears completely [Fig. S3(c)]. The time course in the development of the two types of trajectories and the change in the basin for each attractor are shown in Figs. S4 and Fig. S5, respectively.

The above v dependency of attractors is explained as follows. When v is small, the dynamics are approximated by the means of adiabatic elimination; i.e., x_i reaches the fixed point for a given θ_i, whereas θ_i changes slowly. For given θ_i, the $\{x_i\}$ dynamics are analyzed by the two nullclines, given by

$$dx_1/dt = 0 \;\rightarrow\; x_1 = \tanh \beta \left(\frac{J_{12}}{\sqrt{N}} x_2 + \theta_1 + c_1 \right), \quad (4)$$

$$dx_2/dt = 0 \;\rightarrow\; x_2 = \tanh \beta \left(\frac{J_{21}}{\sqrt{N}} x_1 + \theta_2 + c_2 \right). \quad (5)$$

When v is small, while x_i moves towards the cross point of the two nullclines, as θ_i slowly changes according to (2), the nullclines are slowly shifted.

When this adiabaticity condition is satisfied, only Trajectory A is realized [Fig. S3(b)]: at $\theta_i = 0$ (null epigenetic

modification), there is a stable fixed point as the cross point of the two nullclines at $x_1 < 0$ and $x_2 < 0$ [Fig. 3(c i)]. Then, according to (2), each nullcline is shifted as follows: the x_1 nullcline (i.e., $dx_1/dt = 0$ nullcline) goes up, whereas the x_2 nullcline (i.e., $dx_2/dt = 0$ nullcline) goes left. As a result, the cross point of the two nullclines itself moves up and left, thus reaching above $x_2 = 0$. Consequently, the shift of the x_2 nullcline changes its direction (as the sign of $d\theta_2/dt$ is approximately given by the sign of x_2). Accordingly, the cross point of the nullclines continues to move up, reaches $(-1, 1)$, and then stops.

However, by increasing v, the faster move of the nullclines generates another trajectory, Trajectory B. First, the cross point of the two nullclines moves to the left and up, in the same way as observed for Trajectory A. However, owing to the faster change in θ, the x_2 nullcline shifts to the left so quickly that the two nullclines cross vertically [see Fig. 3(c ii) Trajectory B], and the cross point does not go above $x_2 = 0$. As a result, the cross point moves to the left and down to $(-1, -1)$, where (x_1, x_2) is fixed for some initial conditions. Here, (x_1, x_2) first approaches the fixed point at $\theta_i = 0$ for both Trajectories A and B, and then owing to slight difference in the initial conditions, (x_1, x_2) is directed either to $(-1, 1)$ or $(-1, -1)$.

By increasing v beyond 9.1×10^{-3}, the shift in the nullclines is accelerated, so that the two nullclines cross vertically for all of the initial conditions. In this case, Trajectory A is not realized for any initial condition, and all of the initial conditions are instead attracted to $(-1, -1)$ [Fig. S3(c)].

Hence, the attractor number increases due to the divergence in the motion of the nullclines depending on the initial conditions of $\{x(i)\}$. With a further increase in v, the attractor is pruned because nullclines move faster and no longer
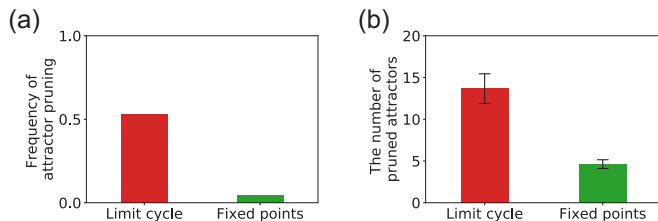
FIG. 4. (a) Fraction of the regulatory matrices J_{ij} that exhibit attractor pruning with the increase in v (here defined as a decrease in attractor number of more than 4). The cases with the initial limit-cycle attractors at $\theta_i = 0$ (red) and those with fixed points (green) are sampled separately. (b) Average number of pruned attractors, defined as the difference between the local maximum and local minimum of attractor number against the change in v. Case with a limit cycle (red) and fixed points (green). See Fig. S6 for more details.

split into two directions of motions due to the faster change of θ_i.

VI. GENERATION AND PRUNING OF ATTRACTORS FROM AN OSCILLATORY STATE

The two-gene minimal model described above suggests how the interplay between fast x dynamics and a slow null-cline shift leads to divergence in trajectories, thereby resulting in nonmonotonic change in the attractor number. By contrast, for $N = 10$, the nonmonotonic behavior of attractor number against v mostly adopts a limit-cycle attractor at $\theta_i = 0$. The frequency of networks showing such behavior is much larger for the limit-cycle case, along with the number of generated and pruned attractors in the intermediate range of v (see Fig. 4).

This relevance of the limit cycle to the generation and pruning of multiple attractors is explained as follows. First, as the limit cycle travels over a larger portion in the phase space of $\{x_i\}$, the variation in the change in $\{\theta_i\}$ is enhanced so that more attractors can be generated with the increase in v. These attractors are generated hierarchically by branching trajectories successively, stemming from the original limit-cycle orbit. However, with the increase in v, the initial oscillation is destroyed due to the faster change in θ (shift of nullclines), so that the top of the hierarchy in branching trajectories is destroyed, leading to a drastic decrease in the attractor number.

This hierarchical attractor generation from limit-cycle (HAGL) is illustrated in a simple three-gene system with a limit-cycle attractor [Fig. 5(a)]. In this three-gene system, only one attractor is reached for small v where the adiabatic condition is satisfied [Fig. S7(a)]. With a further increase in v, however, three attractors are reached ($4 \times 10^{-4} < v < 9 \times 10^{-3}$). The trajectories reaching these attractors initially show oscillation around the original limit cycle at $\theta_i = 0$, and then separate into two groups, as shown in Fig. 5(b): two fixed-point attractors are generated from one group, whereas one fixed-point attractor is generated from the other group. Thus, the attractors are generated hierarchically.

In this v region, the branching in gene expression patterns occur in accordance with that in θ_i as shown in Fig. 5(c), where the first branching occurs in θ_2 at $t \sim 25$, then in θ_1

at $t \sim 50$ and finally at θ_3 at $t \sim 100$. The change in θ_i is triggered by the change in x_i, but it then supports and fixes each of the branching of gene expression patterns. In this sense, the deepening of valleys is guided by θ_i, as is metaphorically represented by strings underlying the Waddington's diagram in Fig. 1(b). With the increase in v, the initial limit-cycle orbit is destroyed before the separation into two groups, so that the number of attractors is reduced from three to one ($v \sim 9 \times 10^{-3}$) (see Fig. S7 for more details).

Most of the generation and pruning of multiple attractors can be understood as HAGL. Note that for much larger N, limit-cycle attractors (or sometimes chaotic attractors) exist more often in the model (1) with $\theta_i = 0$, as previously investigated in neural network models [53,54]. Therefore, the generation and pruning of multiple attractors are expected to be ubiquitous. For confirmation, we simulated the model with $N = 100$. Although sampling all 2^N initial conditions $\{x_i = \pm 1; i = 1, \ldots, N\}$ is numerically difficult, simulations with partial sampling showed that nonmonotonic change in the attractor number occurred for most of the randomly chosen J_{ij} matrices (Fig. S8) where HAGL is commonly observed.

VII. EPIGENETIC LANDSCAPE AND HOMEORHESIS

Thus, HAGL satisfies the first postulate of Waddington's landscape: hierarchical branching. Now, we consider the other two postulates of homeorhesis and robustness in the cell-number ratio. For this purpose, we first need to determine the axes X and Z in the landscape.

As discussed above, the X axis represents the cellular state, which can be extracted from $\{x_i\}$ using PCA. Here, we adopt the first PCA mode of $\{x_i\}$ as X. Each valley corresponds to an attractor stabilized by the slow epigenetic change. To explore robustness in the developmental course and generated epigenetic landscape, we introduce noise in (1) and (2). We adopt the Langevin equation by adding Gaussian white noise $\eta_i(t)$ with $\langle \eta_i(t)\eta_j(t')\rangle = \sigma \delta_{ij}\delta(t - t')$, with δ_{ij} as Kronecker δ and $\delta(x)$ as a δ function. In general, the specific attractor that is reached depends on the initial condition and perturbation by internal noise. By taking the number of cells under noise, each cell reaches one of the attractors (and stays in its vicinity even under noise). Then, one can compute the number distribution of $P(X)$. As Z is lower, the state with X is more frequently reached. By analogy with the relationship between free energy and probability in thermodynamics, one can adopt $Z = \log[1/P(X)]$. Then, the epigenetic landscape can be depicted using the height Z as a function of X.

To compute $P(X)$, we first choose an initial condition of cells (or distribution around a given initial pattern of X). For each initial value, X is computed as a result of time evolution. By starting with a sufficient number of cells, the distribution $P(X)$ is obtained, which may depend on the initial condition of cells. Then, to examine the robustness of the landscape, we explore whether the time evolution of the distribution $P(X)$ is robust against the change in the initial condition of cells.

First, when v is large, any of the 2^N states is approached from the vicinity of each of the initial expression pattern $\{x_i = \pm 1\}$. In this case, the specific state that is attracted as well as the number distribution of cells for each state crucially depend
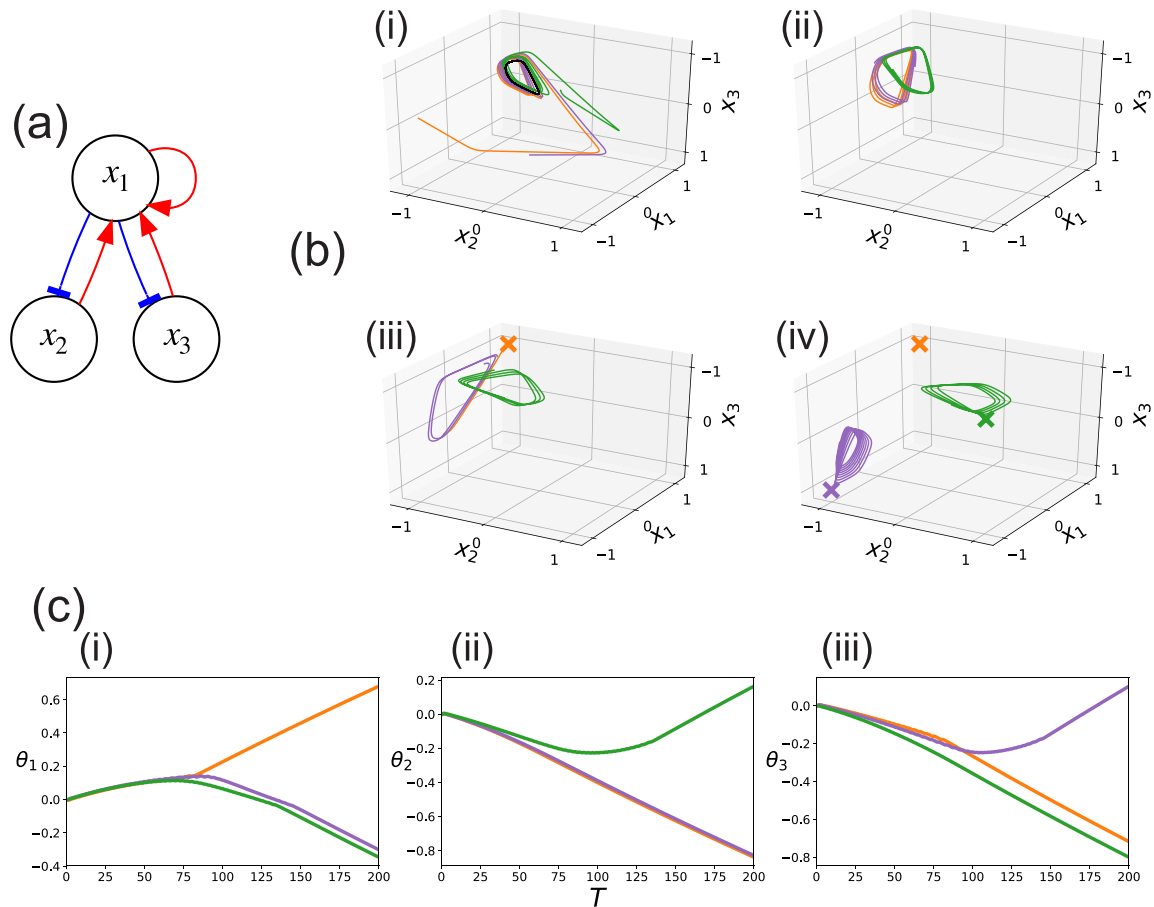
FIG. 5. (a) Three-gene system with $J_{11} = 0.26$, $J_{12} = J_{13} = 0.35$, $J_{21} = -0.4$, $J_{31} = -0.36$, $c_1 = 0.26$, $c_2 = c_3 = 0.17$. (b) Hierarchical attractor generation from limit cycle (HAGL). $v = 10^{-3}$. (i) Trajectories from different initial conditions, plotted by different colors, approach a limit-cycle attractor at $\theta_i = 0$. $0 < t < 10$. (ii) Trajectories are separated into two groups (green line shows one group, and orange and purple lines show another group), depending on the initial initial condition. $40 < t < 50$. (iii) Further separation of the group of trajectories shown in orange and purple. $80 < t < 90$. (iv) Three trajectories reaching distinct fixed points. $120 < t < 150$. (c) Time development of θ_1(i), θ_2(ii), and θ_3(iii). Each line color corresponding to trajectory in (b). HAGL in x space is supported by θ_i, similarity as the strings beyond the epigenetic landscape.

on the choice of initial conditions. Hence, $P(X)$ is not robust to the change in the initial conditions.

Next, we consider the case with monotonic dependence of attractor number upon v. In this case, if v is not so large, the number of attractors n_A is much smaller. Nevertheless, the specific attractor the cell state reaches is still predetermined by how close the expression state at $\theta_i = 0$ is to the final expression state. The initial x_i state is partitioned into n_A basins, from each of which only one attractor (valley) is generated. Hence, $P(X)$ crucially depends on the initial distribution of x_i's (see Fig. S9, S10).

In contrast, for HAGL, the postulated robustness is achieved if v is in the intermediate region in which multiple attractors are generated, as shown in Fig. S9, S10. The obtained $P(X)$ is almost completely independent of the initial conditions of cells (see Fig. S10). For most initial conditions, all of the attractors are reached, and the fraction of cells reaching each attractor under noise is quite stable against the change in the initial distribution of $\{x_i\}$ (Fig. S9, S10). In this case, from any initial conditions, the limit-cycle attractor (at

$\theta_i = 0$) is first reached. With the epigenetic feedback, the cells are then distributed to each attractor depending on the phase of oscillation. Hence, the time course of differentiation to each attractor (cell type), as well as the fraction of each attractor (the number ratio of each cell type) are both independent of the initial distribution of cells.

We can then depict the epigenetic landscape according to the time evolution of $P(X)$. Here, X (in Fig. 1) is given by the first PCA mode from $\{x_i\}$ obtained from a distribution of initial conditions. The landscape is depicted by $Z = -\log P(X)$, so that the bottom of the lower valley has a higher population density. The landscape thus depicted is given in Fig. 6, which shows both the hierarchical branching and robustness to the initial expression or noise. This landscape, as in the case of Fig. 5, is shaped with the change in θ_i. Time development of distribution Θ (corresponding to θ_i in the PCA 1 space) is shown in Fig. S11(a), whereas the landscape thus depicted is given in Fig. S11(b), S11(c). The figures represent how Θ supports the branching trajectories as in the strings in the epigenetic landscape in Fig. 1(b).
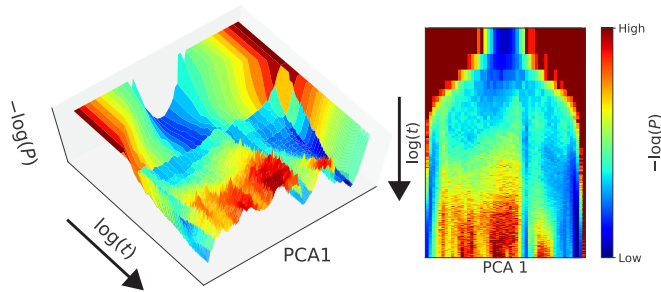
FIG. 6. Epigenetic landscape generated from the temporal evolution of cellular states for given J_{ij} that exhibits HAGL. $N = 40$ and c_i is set to a random value sampled from the normal distribution with average 0, variance 0.1. We adopt one-mode PCA to represent a one-dimensional scalar variable X and $Z = -\log[P(X)]$ indicating the depth of valley, where $P(X)$ is the distribution of X over cells developed under noise, and is plotted against time given by $\log(t)$. Red indicates large (i.e., low frequency) and blue indicates small values. The amplitude of Gaussian white noise $\sigma = 0.1$. The right figure shows a one-dimensional representation with the horizontal axis as X and vertical axis as scaled time (from top to bottom), whereas the color represent Z.

Finally, we quantitatively characterize the robustness of the final distributions of cellular states reached from different initial distributions. Let us define $P^{\mu}(X)$ as the distribution of X reached from a given initial condition of x_i, indexed by μ [e.g., $x_i(t = 0) = \eta_i^{\mu}$, where $\eta_i^{\mu}(i = 1, \ldots, N)$ is one random sequence in $[-1, 1]$, whereas $\nu \neq \mu$ denotes a different random sequence]. As the measure for the distance between two distributions $P^{\mu}(X)$, $P^{\nu}(X)$ generated from different initial distributions, we adopt the KL divergence $D_{KL} = \sum_X P^{\mu}(X) \ln\{P^{\mu}(X)/P^{\nu}(X)\}$ for a pair of two distributions $P^{\mu}(X)$, $P^{\nu}(X)$ obtained from two samples μ and ν starting from different initial conditions. If D_{KL} is small, a similar distribution $P(X)$ (i.e., a similar landscape) is obtained, independent of the initial condition, thereby implying robustness at the distribution level. $\overline{D}_{KL}$ is computed by averaging over the samples μ and ν, which is plotted in Fig. 7 for the case of monotonic attractor number dependency on ν and the nonmonotonic HAGL case. As shown in Fig. 7, the $\overline{D}_{KL}$ value is kept small up to a large value in ν (e.g., $\nu \leqslant 10^{-2}$) for the HAGL case. This quantitatively demonstrates that differentiation from the oscillatory state through epigenetic fixation shows higher robustness in the distribution of cellular states.

VIII. DISCUSSION

We have introduced a model involving mutual interactions between the expression dynamics controlled by a GRN and epigenetic modification. With more efficient execution of the epigenetic feedback regulation, more attractors with different expression patterns, i.e., more cell types, are generated. In some networks, the initial expression levels are simply embedded into epigenetic modifications, whereas for other networks, mutual feedback between expression levels and modifications bring about hierarchically ordered attractors from an oscillatory state. In such a case, the attractor number shows
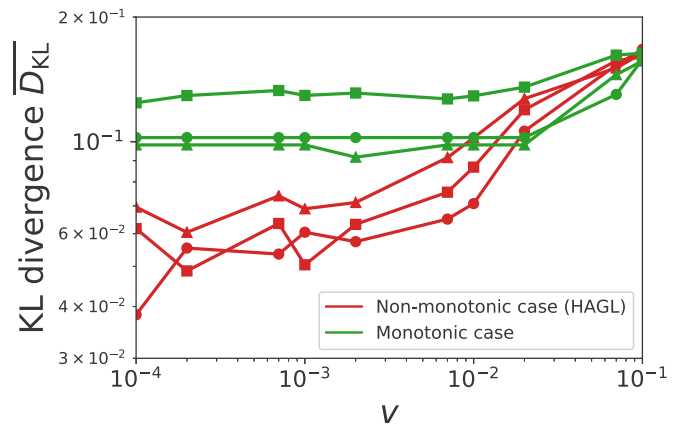

FIG. 7. $\overline{D}_{KL}$ representing the averaged Kullback-Leibler divergence between two distributions of cellular states developed under noise ($\sigma = 0.1$). First, $P(X)$ is computed from 500 cells developed from a given initial condition and J_{ij}. The distribution $P^{\mu}(X)$ is computed over $\mu = 1, 2, \ldots, 10$ starting from different initial conditions. The Kullback-Leibler divergences are then computed over all pairs of 90 distributions and averaged to get $\overline{D}_{KL}$ (see also Fig. S11 for each distribution form). For HAGL, $\overline{D}_{KL}$ remains low up to large $\nu \sim 10^{-3}$ (red lines), whereas in the monotonic case (green lines), it takes on a large value over the full range of ν.

nonmonotonic change against the rate of epigenetic feedback regulation ν. The mechanism of nonmonotonic dependency on ν, i.e., the attractor generation and pruning, is explained in terms of dynamical systems theory.

By using the change in expression dynamics under the slow epigenetic modification process, Waddington's epigenetic landscape is explicitly depicted, in which the landscape axis (X axis in Fig. 1) is given by the principal component of the expression pattern; the depth, Y axis, is given by the developmental time with slow epigenetic modification; and the height is given by $-\log[P(X)]$ with $P(X)$ as the cell-number distribution of X. In particular, when the original attractor in the absence of epigenetic modification is a limit cycle, the timing of branching to different cell types, number of differentiated cell types, and number fraction of each cell type are all robust to perturbations during the course of development and to the variation of initial conditions. Hence, the generated landscape satisfies the three postulates implicitly assumed in Waddington's landscape: (i) hierarchical branching is supported by the hierarchical attractor generation from the limit cycle; (ii) homeorhesis is supported since this branching process is independent of initial conditions and robust to noise; and (iii) the cell-number robustness is demonstrated since $P(X)$ is also independent of initial conditions and robust to noise. This robustness in the path and in the cell-number distribution to perturbation is an essential requirement for the development of multicellular organisms [15].

Our theoretical model assumes epigenetic feedback regulation. Although the transient modification in epigenetic factors has been experimentally confirmed [55,56], the extent to which this modification depends on gene expression is not yet clearly elucidated. Considering that epigenetic change stabilizes the cellular states, it is rather natural to assume positive feedback from the expression level to modification,

i.e., if expressed (repressed), it is easier (harder) to be expressed, whereas some molecular mechanisms for such positive feedback have been suggested [32,33]. However, direct evidence as well as quantitative estimates for the time scale of epigenetic change require further experimental elucidation in the future.

The significance of oscillation in the cellular state for the differentiation process was previously discussed [18]. Indeed, the cell state is not fixed but rather involves several oscillatory modes, including circadian and cell-division cycles. Furthermore, oscillatory expression has recently been uncovered for embryonic stem cells [57–60], which is ultimately lost in cells committed to differentiation. Note that the relevance of an oscillatory state to pluripotency was previously discussed in the context of an alternative approach to the epigenetic landscape with respect to inclusion of cell-cell interactions [46]. In this case, the initial oscillation in expression levels is lost with an increase in the cell number and resulting amplification of cell-cell interactions accordingly. Hence, the two approaches, i.e., cell-cell interactions and epigenetic modification, are compatible. Indeed, a model that includes both approaches was previously investigated, in which epigenetic modification of several genes such as Oct4 and Nanog leads to the commitment of cells from an undifferentiated state, which is consistent with experimental observations [61–63].

The canalization in Waddington's landscape is valid for the normal developmental process. However, through certain external operations, the path of committed cells can be reversed to an undifferentiated state in a process known as reprogramming [64–67]. In the present model, by externally overexpressing some genes for a given time span, the threshold $-\theta_i$ that was initially increased can be decreased so that the expression level recovers, which matches the experimental procedure used to create induced pluripotent stem cells. In the future, it will be important to elucidate the condition required for such reprogramming by identifying the specific genes in the network that need to be overexpressed so as to climb up to the most upstream location in the landscape under the present theoretical framework.

The generation and pruning of attractors that depend on the epigenetic feedback rate is itself an interesting phenomenon in terms of dynamical systems of both fast and slow elements, which requires an analysis beyond the breadth of adiabatic elimination [54]. That is, if the time scales are clearly separated, the change in fast expression would be represented as an attractor change against the slow epigenetic state as a control parameter. In contrast, mutual feedback between the two is important, as shown in the present study with regard to the interaction between the nullclines and the variables. Therefore, an appropriate analytical method that is capable of capturing such feedback dynamics needs to be developed.

Homeostasis, robustness of a steady state in biological systems has gathered much attention over decades. This, for instance, has been discussed as the stability of the final state (attractor) against perturbations. On the other hand, homeorhesis is concerned with the stability of the time course of a state, against the change in the initial conditions or perturbations. So far, studies on homeorhesis are rather limited: a few examples include relaxation dynamics in signal transduction process independent of the initial condition [68], robust developmental process with cell-cell interaction [16,18], and robust ecological dynamics in an experiment consisting of algae and ciliates [69]. For homeorhesis to work, existence of slower time scale and buffering of initial variation will be needed. The hierarchial attractor generation by slower epigenetic feedback after attraction to a limit cycle will provide one general mechanism for the homeorhesis.

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

[2] S. A. Kauffman, J. Theor. Biol. **22**, 437 (1969).

[3] J. Wang, L. Xu, E. Wang, and S. Huang, Biophys. J. **99**, 29 (2010).

[4] C. Furusawa and K. Kaneko, Science **338**, 215 (2012).

[5] G. Forgacs and S. A. Newman, *Biological Physics of the Developing Embryo* (Cambridge University Press, Cambridge, 2005).

[6] J. Wang, K. Zhang, L. Xu, and E. Wang, Proc. Natl. Acad. Sci. USA **108**, 8257 (2011).

[7] M. Mojtahedi, A. Skupin, J. Zhou, I. G. Castaño, R. Y. Leong-Quong, H. Chang, K. Trachana, A. Giuliani, and S. Huang, PLoS Biology **14**, e2000640 (2016).

[8] S. Huang, G. Eichler, Y. Bar-Yam, and D. E. Ingber, Phys. Rev. Lett. **94**, 128701 (2005).

[9] D. E. Wagner, C. Weinreb, Z. M. Collins, J. A. Briggs, S. G. Megason, and A. M. Klein, Science **360**, 981 (2018).

[10] J. A. Briggs, C. Weinreb, D. E. Wagner, S. Megason, L. Peshkin, M. W. Kirschner, and A. M. Klein, Science **360**, eaar5780 (2018).

[11] C. H. Waddington, Endeavour **1**, 18 (1942).

[12] J. Gurdon, P. Lemaire, and K. Kato, Cell **75**, 831 (1993).

[13] N. J. Oviedo, P. A. Newmark, and A. Sánchez Alvarado, Dev. Dynamics **226**, 326 (2003).

[14] I. Rafols, Y. Sawada, A. Amagai, Y. Maeda, and H. K. MacWilliams, Differentiation **67**, 107 (2001).

[15] A. K. Alsing and K. Sneppen, Nucleic Acids Res. **41**, 4755 (2013).

[16] K. Kaneko and T. Yomo, J. Theor. Biol. **199**, 243 (1999).

[17] C. Furusawa and K. Kaneko, J. Theor. Biol. **209**, 395 (2001).

[18] N. Suzuki, C. Furusawa, and K. Kaneko, PLoS ONE **6**, e27232 (2011).

[19] A. Bird, Nature (London) **447**, 396 (2007).

[20] A. D. Goldberg, C. D. Allis, and E. Bernstein, Cell **128**, 635 (2007).

[21] R. Jaenisch and A. Bird, Nature Genet. **33**, 245 (2003).

[22] M. A. Surani, K. Hayashi, and P. Hajkova, Cell **128**, 747 (2007).

[23] S. Huang, Development **136**, 3853 (2009).

[24] R. Cortini, M. Barbi, B. R. Caré, C. Lavelle, A. Lesne, J. Mozziconacci, and J.-M. Victor, Rev. Mod. Phys. **88**, 025002 (2016).

[25] Z. S. Singer, J. Yong, J. Tischler, J. A. Hackett, A. Altinok, M. A. Surani, L. Cai, and M. B. Elowitz, Mol. Cell **55**, 319 (2014).

[26] Y. B. Schwartz and V. Pirrotta, Nature Rev. Genetics **8**, 9 (2007).

[27] T. Rohlf, L. Steiner, J. Przybilla, S. Prohaska, H. Binder, and J. Galle, Epigenomics **4**, 205 (2012).

[28] R. Karlić, H.-R. Chung, J. Lasserre, K. Vlahoviček, and M. Vingron, Proc. Natl. Acad. Sci. USA **107**, 2926 (2010).

[29] O. J. Rando and H. Y. Chang, Annu. Rev. Biochem. **78**, 245 (2009).

[30] S. Hihara, C.-G. Pack, K. Kaizu, T. Tani, T. Hanafusa, T. Nozaki, S. Takemoto, T. Yoshimi, H. Yokota, N. Imamoto *et al.*, Cell Rep. **2**, 1645 (2012).

[31] D. Zhang, Z. Tang, H. Huang, G. Zhou, C. Cui, Y. Weng, W. Liu, S. Kim, S. Lee, M. Perez-Neut *et al.*, Nature (London) **574**, 575 (2019).

[32] I. B. Dodd, M. A. Micheelsen, K. Sneppen, and G. Thon, Cell **129**, 813 (2007).

[33] K. Sneppen, M. A. Micheelsen, and I. B. Dodd, Mol Syst. Biol. **4** 182 (2008).

[34] M. Sasai, Y. Kawabata, K. Makishi, K. Itoh, and T. P. Terada, PLoS Comput. Biol. **9**, e1003380 (2013).

[35] K. Maeshima, K. Kaizu, S. Tamura, T. Nozaki, T. Kokubo, and K. Takahashi, J. Phys. Condens. Matter **27**, 064116 (2015).

[36] M. Grunstein, Cell **93**, 325 (1998).

[37] S. Gombar, T. MacCarthy, and A. Bergman, PLoS Comput. Biol. **10**, e1003450 (2014).

[38] E. Mjolsness, D. H. Sharp, and J. Reinitz, J. Theor. Biol. **152**, 429 (1991).

[39] I. Salazar-Ciudad, J. Garcia-Fernández, and R. V. Solé, J. Theor. Biol. **205**, 587 (2000).

[40] I. Salazar-Ciudad, S. Newman, and R. Solé, Evol. Dev. **3**, 84 (2001).

[41] K. Kaneko, PLoS ONE **2**, e434 (2007).

[42] W. Jia, A. Deshmukh, S. A. Mani, M. K. Jolly, and H. Levine, Phys. Biol. **16**, 066004 (2019).

[43] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevResearch.2.023083 for the monotonic increase of the number of attractors with α.

[44] S. L. Schreiber and B. E. Bernstein, Cell **111**, 771 (2002).

[45] J. C. G. Spainhour, H. S. Lim, S. V. Yi, and P. Qiu, Cancer Inform. **18**, 1 (2019).

[46] T. Miyamoto, C. Furusawa, and K. Kaneko, PLoS Comput. Biol. **11**, e1004476 (2015).

[47] K. Fujimoto, S. Ishihara, and K. Kaneko, PLoS ONE **3**, e2772 (2008).

[48] N. Feldman, A. Gerson, J. Fang, E. Li, Y. Zhang, Y. Shinkai, H. Cedar, and Y. Bergman, Nat. Cell Biol. **8**, 188 (2006).

[49] G. A. Challen, D. Sun, M. Jeong, M. Luo, J. Jelinek, J. S. Berg, C. Bock, A. Vasanthakumar, H. Gu, Y. Xi *et al.*, Nature Genet. **44**, 23 (2012).

[50] W. Reik, W. Dean, and J. Walter, Science **293**, 1089 (2001).

[51] R. D. Hawkins, G. C. Hon, L. K. Lee, Q. Ngo, R. Lister, M. Pelizzola, L. E. Edsall, S. Kuan, Y. Luu, S. Klugman *et al.*, Cell Stem Cell **6**, 479 (2010).

[52] K. Tripathi and G. I. Menon, Phys. Rev. X **9**, 041020 (2019).

[53] H. Sompolinsky, A. Crisanti, and H.-J. Sommers, Phys. Rev. Lett. **61**, 259 (1988).

[54] H. Aoki and K. Kaneko, Phys. Rev. Lett. **111**, 144102 (2013).

[55] S. Kangaspeska, B. Stride, R. Métivier, M. Polycarpou-Schwarz, D. Ibberson, R. P. Carmouche, V. Benes, F. Gannon, and G. Reid, Nature (London) **452**, 112 (2008).

[56] S. Rulands, H. J. Lee, S. J. Clark, C. Angermueller, S. A. Smallwood, F. Krueger, H. Mohammed, W. Dean, J. Nichols, P. Rugg-Gunn *et al.*, Cell Sys. **7**, 63 (2018).

[57] T. Kobayashi, H. Mizuno, I. Imayoshi, C. Furusawa, K. Shirahige, and R. Kageyama, Genes Dev. **23**, 1870 (2009).

[58] I. Imayoshi, A. Isomura, Y. Harima, K. Kawaguchi, H. Kori, H. Miyachi, T. Fujiwara, F. Ishidate, and R. Kageyama, Science **342**, 1203 (2013).

[59] A. Aulehla and O. Pourquie, Curr. Opin. Cell Biol. **20**, 632 (2008).

[60] M. A. Canham, A. A. Sharov, M. S. Ko, and J. M. Brickman, PLoS Biol. **8**, e1000379 (2010).

[61] H. Niwa, J.-i. Miyazaki, and A. G. Smith, Nature Genet. **24**, 372 (2000).

[62] L. A. Boyer, T. I. Lee, M. F. Cole, S. E. Johnstone, S. S. Levine, J. P. Zucker, M. G. Guenther, R. M. Kumar, H. L. Murray, R. G. Jenner *et al.*, cell **122**, 947 (2005).

[63] S.-J. Dunn, G. Martello, B. Yordanov, S. Emmott, and A. Smith, Science **344**, 1156 (2014).

[64] K. Takahashi and S. Yamanaka, Cell **126**, 663 (2006).

[65] K. Hochedlinger and K. Plath, Development **136**, 509 (2009).

[66] P. Hajkova, K. Ancelin, T. Waldmann, N. Lacoste, U. C. Lange, F. Cesari, C. Lee, G. Almouzni, R. Schneider, and M. A. Surani, Nature (London) **452**, 877 (2008).

[67] S. Huang, Bioessays **31**, 546 (2009).

[68] J. T. Young, T. S. Hatakeyama, and K. Kaneko, PLoS Comput. Biol. **13**, e1005434 (2017).

[69] J. S. Chuang, Z. Frentz, and S. Leibler, Proc. Natl. Acad. Sci. USA **116**, 14852 (2019).