

## RESEARCH ARTICLE

# Delineating the respective impacts of stochastic curl- and grad-forces in a family of idealized core genetic commitment circuits

Marc Turcotte<sup>1,2,\*</sup>

<sup>1</sup> Department of Mathematics, The University of Texas at Arlington, Arlington, TX 76109, USA

<sup>2</sup> Biological Sciences, The University of Texas at Dallas, Richardson, TX 75080, USA

\* Correspondence: marc.turcotte@utdallas.edu, marc.turcotte@uta.edu

Received December 15, 2015; Revised January 19, 2016; Accepted February 23, 2016

**Stochastic dynamics pervades gene regulation. Despite being random, the dynamics displays a kind of innate structure. In fact, two stochastic forces combine driving efforts: one force originates from the gradient of the underlying stochastic potential, and the other originates from the mathematical curl of the probability flux. The curl force gives rise to rotation. The gradient force gives rise to drift. Together they give rise to helical behavior. Here, it is shown that around and about the vicinity of attractive fixed points, the gradient force naturally wanes but the curl force is found to remain high. This leads to a locally noticeably different type of stochastic track near and about attractive fixed points, compared to tracks in regions where drift dominates. The consistency of this observation with the experimental fact that, in biology, fate commitment appears to not be a-priory locked-in, but rather necessitating active maintenance, is discussed. Hence attractive fixed-points are not only fuzzy, but may effectively be, locally, “more free”.**

**Keywords:** systems biology; theoretical biology; gene regulation; nonlinear dynamics; stochasticity

## INTRODUCTION

In previous work, focus was on shedding light on the role of parameter asymmetry and time-scale separation and the relationship to phenotype commitment in a few prototypical core gene regulatory circuits [1]. Here, following the lead of Wang *et al.* [2–14] and by using the set of example circuits from [1]: DS1, DS2, DS3 and DS4, the previous study is extended by focusing on the respective roles of the mathematical curl of the probability flux and the gradient of the stochastic potential. In a series of groundbreaking papers, Wang *et al.* have shown that the stochastic motion of gene regulatory systems can be explained by the combined effect of two main influences: a force originating from the gradient of the underlying stochastic potential function and another force, this one originating in any curl of the probability flux. The split up of influences between curl-free and divergence-free sources is valid in any number of system dimensions,

but in three dimensions (3D) this results in a direct mathematical correspondence with the familiar motion of a charged particle in combined electric and magnetic fields. There exists, in fact, yet a third component driving statistical systems, this one due to the possible systematic spatial variation in the strength of the noise term, but study of this effect is outside the scope of this work.

### The conceptual curl

The formal definition of the curl is the amount of circulation of a vector field around the closed boundary of a surface, divided by the area of that surface. The curl thus points normal to the surface. Intuitively, the curl is thus how much rotation there is in the field. In general, the curl is expressed by anti-symmetric combinations [15]:

$$\text{Curl } \mathbf{a} = \left( \frac{\partial a_2}{\partial x_1} - \frac{\partial a_1}{\partial x_2} \right) \mathbf{k}_{12} + \left( \frac{\partial a_3}{\partial x_1} - \frac{\partial a_1}{\partial x_3} \right) \mathbf{k}_{13} + \dots$$

in which  $\mathbf{a}$  is a general vector field with components  $x_1, x_2, x_3, \dots$  and the  $\mathbf{k}_{ij}$ 's are unit bi-vectors, e.g.,  $\mathbf{k}_{12}$  is a unit surface area lying in the 1, 2-plane with a unit vector pointing in the 3-direction. Thus the curl of a two-dimensional vector field is a one-dimensional function, the curl of a three-dimensional vector field is itself a three-dimensional vector, etc.

### General stochastic processes and curl

As mentioned above, the role of curl in driving statistical processes has been delineated in several seminal works. In a spacially constant noise system, the driving force is split into two components [16]:

$$\mathbf{F} = -D \frac{\partial}{\partial x} U + \mathbf{J}_{ss}(x,t)/P_{ss}.$$

In the above equation,  $\mathbf{F}$  is the driving force of the statistical process,  $D$  the drift tensor,  $U$  the stochastic potential,  $\mathbf{J}$  the probability flux and  $P$  the probability; the subscript “*ss*” stands for the steady-state maintained by a constant entropy increase rate. This equation is nothing but the well-known decomposition of a vector field into a curl-free component (first term) and a divergence-free component (second term). The first term is thus the component of the driving force originating from the gradient of the stochastic potential, and the second term is the component of the driving force due to the curl of the probability flux. Of course, as extensively covered in the works cited previously, in 3D this is completely analogous to the motion of a charged particle under the combined action of an electric field and magnetic field. Generally we can expect a general stochastic process to drift down a potential well along a spiraling trajectory, rotation being sourced into how far the process happens to be from equilibrium. If the process is at equilibrium, there will be no rotation; the curl term will be null.

The work reported upon herein focuses much attention on the role of mathematical curl in several core regulatory gene differentiation circuits. Curl has special interest because it is related to the rate of entropy increase [17] and thus to how far the stochastic systems lie from equilibrium. In fact, in the context of core genetic commitment circuits, it is not *a priori* clear what the specific consequences are for stochastic processes to deploy far from equilibrium. This notion is particularly still obscure as a function of circuit topology and in view of the inherent parameter-asymmetry in the circuits. This study is based on four basic core regulatory topologies whose behavior in terms of parameter-asymmetry and time-scale separation has been extensively reported on in [1]: DS1, DS2, DS3 and DS4. As explained in the reference, DS1 is the basic positively auto-regulated and

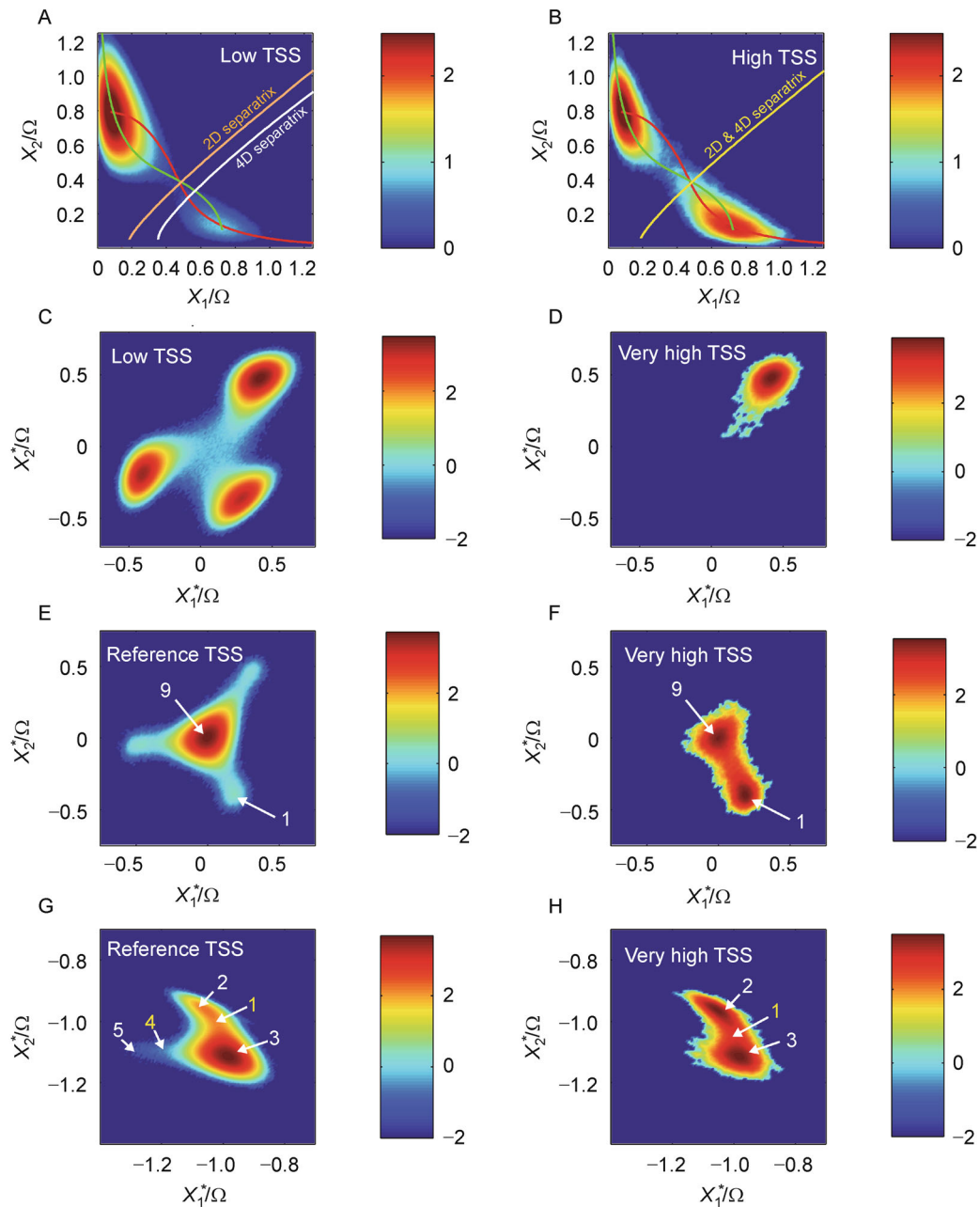
mutually repressive canonical two-gene switch. DS2, DS3 and DS4 are three, four and five gene switches stemming from direct generalization of DS1, multiply connecting all genes in an analogous manner. For more details, the reader is referred to [1].

Time scale separation (TSS) refers to the fact that the messenger RNA sub-manifold in the system reaches steady-state much faster than the protein processes. In mathematical models, this is achieved by imposing much slower protein production and degradation rates relative to those of messenger RNA. In biology, transcription is fast whereas translation is slower. More details can be found in [1].

## RESULTS

### Residency

Because, while highly desirable, an analytical solution of the Chemical Master Equation [18,19] represents a significant challenge in some cases that may be insurmountable in others, a numerical approach based on the well-known Gillespie algorithm [18–21] was adopted. More details can be found in Methods. For each the four core circuits, one hundred stochastically independent tracks were simulated and the results analyzed, as function of  $(X_1, X_2)$  for DS1 and the  $(X_1, X_2$  and  $X_3$  state variables) for DS2, DS3 and DS4. Note, for DS3 and DS4, this represents a choice since additional state variables exist. The behavior of the systems is identical with respect to these additional state variables ( $X_4$  for DS3;  $X_4$  and  $X_5$  for DS4), so they are omitted for brevity. In all cases, for clarity and brevity, the messenger RNA sub-manifold is also not shown. Again, it was verified that this leads to no loss of information (data not shown for brevity). Also, only in the limit of infinite number of particles (so-called thermodynamic limit; infinitely small noise) would the first moment (mean) of the residency distribution originating from a stochastic model agree with the prediction of its associated deterministic version. This thermodynamic limit, in similar contexts, was studied elsewhere [22]. Therefore, here, for brevity, we focus on results from stochastic models examined at finite levels of noise commensurate with biologically realistic conditions. Further, for each of the core circuits, two versions of the stochastic systems were simulated: one with low intrinsic time-scale separation and one with very high intrinsic time-scale separation between the transcription and translation sub-manifolds. The results are shown on Figure 1. For completeness, Figure 1A and 1B of row 1 are for DS1 and are reproduced verbatim from [1]. Panel C through H (DS2, DS3, and DS4) of rows 2, 3 and 4 respectively



**Figure 1. Occupancy diagrams.** Each diagram is the average of 100 statistically independent trajectories. The residency color scales are  $\log_{10}$ . Rows 1, 2, 3 and 4 are for DS1, DS2, DS3 and DS4 respectively. Left and right columns are for low TSS and high TSS respectively. In high TSS, the messenger RNA sub-manifold is essentially at rest. The lines on panels A and B are separatrix for infinite TSS (“2D”) and finite low TSS (“4D”), see text for details. For panels C through H, the original  $X_1X_2X_3$  view is first rotated about the outward pointing eigenvector of the central fixed point in the system to a yield a new coordinate system:  $X_1^*X_2^*X_3^*$  in which  $X_3^*$  is in the direction of this outward pointing eigenvector. The reference is also shifted to the central fixed point. Data are then histogrammed in the  $X_1^*X_2^*$  plane by integrating over the third coordinate  $X_3^*$  to produce a view from above. More details in Methods. The numbers on panels E, F, G and H are the various fixed points in the systems. On panels E and F, 9 and 1 are stable attractive fixed points. On panel G, 1 and 4 are saddle nodes. 2, 3 and 5 are stable attractive fixed points. The location of #5 is approximate because of sparse sampling. On panel H, 1 is saddle node; 2 and 3 are stable attractive fixed points.

present new simulation results. The DS2, DS3 and DS4 results are shown in convenient rotated view (see Methods). The leftmost column of results shows systems in low time-scale separation. The rightmost column shows results for very high TSS. Based on the bifurcations diagrams in [1], DS2, DS3 and DS4 results are shown for parameter asymmetry away from rotation. Thus there is no orbit possible in any of the systems shown here. The reason for this choice of asymmetry is obvious: the focus of the work herein is to determine the role of mathematical curl in the systems in regimes where this role is non-trivial. Thus, the dynamical regions in lack of explicit closed rotation are of prime interest. For each system, the stochastic tracks are initiated near one fixed point and the systems are allowed to reach steady-state over very long simulation times.

Perhaps the most striking feature of the simulations is that the presence of un-equal depths underlying attractive fixed points of the dynamics in the low TSS regime is broken for DS2 (Figure 1C). Thus on Figure 1A, 1E and 1G, we can clearly see that the stochastic potential underlying one fixed point of the dynamics is much deeper than the other. In case of DS1, Figure 1A, the upper fixed point. In case of DS3, Figure 1E, the central fixed point (#9 on the figure) and in the case of DS4, Figure 1G, the lower fixed point (#3 on the figure). The color in the residency diagrams represents the probability of the system being found at the corresponding values of the state variables (irrespective of any possible other variables, not shown). For DS2, Figure 1C, the steady-state of the system shows that each fixed points of the dynamics can be occupied with similar probability. It is important to mention again that the dynamics DS2 in low TSS (Figure 1C) does not present an explicit orbit. This break from a trend is puzzling and will be addressed more deeply in the next section focusing on curl. Results from the rightmost column are for high TSS and they consistently re-establish the relative probabilistic equality of opposite wells separated by a saddle node (see [1] for more details). Interestingly, DS2 (Figure 1B) appears to be out of trend. High TSS basically restricts the dynamics to one fixed point. However, the consistency of results is re-established when the location of the general separatrix surface is considered and, also, whether additional dynamical variables (not shown on the diagram) affect the system. In DS1, the location of the separatrix curve shifts with TSS (see curves on Figure 1A and 1B) thus allowing the populations of the two wells to equalize for Very High TSS. In DS2 (Figure 1C vs 1D), as was also shown in [1], the general separatrix surface rotates with TSS about a central unstable fixed point (Figure S1). In DS2 Low TSS (Figure 1C), one of the sheets of the general separatrix surface is located asymmetrically between attractive fixed

points so the dynamics can hop from one fixed point to another nearby. But, by the tree-fold symmetry of the system, this possibility is also necessarily created for the other two fixed points, leading to equalization of the dynamics between all three fixed points. In DS2 very high TSS however (Figure 1D), the three-fold symmetric general separatrix surface is centrally located between fixed points resulting in confinement of the dynamics to one fixed point. In DS3 (Figure 1E and 1F) and DS4 (Figure 1G and 1H), the general separatrix surface location also shifts as a function of TSS (Figure S2) but in a very subtle way. In these cases, the location of the general separatrix surface is complicated by the fact that it also depends on the value of any additional variables ( $X_4$  for DS3;  $X_4$  and  $X_5$  for DS4). For DS3, Figure S2 shows that because of asymmetric proximity of the general separatrix surface — one  $X_4$  layer at a time — to one of two fixed points, the dynamics falls into the attractor of the central fixed point where it remains confined. This behavior is relieved for VHTSS as the general separatrix surface is more centrally located, again,  $X_4$ -layer by  $X_4$ -layer. A similar concept is expected to prevail for DS4, although there it would be by  $X_4$  and by  $X_5$  layer.

Overall, because TSS affects the general separatrix surface, this leads to low TSS/high TSS consistency of behavior between DS1, DS3 and DS4: sequestration vs. equalization for low/high TSS. However DS2 appears seemingly to present contrary dynamical behavior (equalization/sequestration for low/high TSS). Yet this behavior is completely understood as explained above. The primary reason for DS2 to appear contrary is that for this system there is three-fold symmetry of the general separatrix surface matching the three attractive fixed points and most importantly, in addition, for DS2 there are no extra state variables to affect the location of the general separatrix surface.

The situation is quite different for DS3 and for DS4 however. For these circuits, the concept of general separatrix surface comes with the caveat that its position is affected by the existence of additional state variables that the dynamics does visit, but whose values cannot be shown on a three-variable only graph. Consider plotting only the three-dimensional sub-manifold of a higher-dimensional general separatrix surface. Given that only three state variables can be rendered on a three-axis plot, one can ask for what value of the extra (unseen) variables should this sub-manifold be drawn? As shown on Figure S2B, the location of the surface sub-manifold does depend heavily on these extra variables. So, in the case of DS3 and DS4, for some values of these extra variables (and not for others), it would theoretically be possible for the dynamics to find a road from one side of a three-dimensional sub-manifold of the multi-dimensional gen-

eral separatrix surface to the other, via travel through any of the unseen (hidden) state-variables. The actual outcome depends on the initial values of the unseen hidden variables. Or more precisely, which sub-manifold surface to consider strongly depends on the values of the hidden (unseen) variables. The extra (unseen) variables strongly matter. Because of these subtleties, strictly speaking, only at fixed value of any additional state variables over the plotted three, does the concept of the three-dimensional sub-manifold of a multi-dimensional general separatrix surface make sense. Hence, this is the approach followed herein (Figure S3).

## Curl

Figure 2 and Figure 3 present the same curl modulus data. However, on Figure 3, the color scale of the curl modulus ranges from the minimum to the maximum of the data plotted on each panel, as opposed to being common to all the panels. This varying finer scale is particularly useful as it reveals subtle low/high TSS systematics. For DS1 (Figure 3A and 3B) and DS2 (Figure 3C and 3D), low/high TSS presents lower/higher relative curl in the saddle region (fixed point #2 for DS1 and fixed points #5, #6 and #7 for DS2), with the extremities of the basins of attractions of the outer fixed points (#3 and #1 for DS1 and #1, #2 and #3 for DS2) presenting a complementary opposite behavior. There is a striking gradation with circular symmetry in curl modulus size resulting in the appearance of characteristic color bands. This symmetry is present in the curl function (data omitted for brevity). However, this dynamical behavior is not carried over to DS4 (Figure 3E and 3F) and DS5 (Figure 3G and 3H). These circuits however do present a common trend. These two circuits, as opposed to DS1 and DS2, have extra dimensions not shown on the plot (not including mRNAs: one ( $X_4$ ) for DS3 and two ( $X_4$  and  $X_5$ ) for DS4). The curl modulus in the vicinity of the saddles (#2 and #1 for DS3 and DS4, respectively) remains high in all cases. For DS3, the 1-2-9, 3-4-9 and 7-8-9 axes are all relatively high in curl modulus, leaving only the perpendicular 5-6-9 axis to present a low-to-high gradation at low TSS. However, low-to-high gradation on the 1-2-9 axis is recovered at high TSS. No data exists for the other axes. For DS4 however, only one axis could be sampled (2-1-3) and it presents a low-high-low curl modulus behavior in both the low and the high TSS. Figure 4 shows partial data from one stochastic track, rather than integrating over a large sample of 100 tracks. Here, the time-sequential information is retained along the propagating track; color indicates the track segment's local curl modulus. These partial data are in accordance with the trend shown on previous figures, but they also testify to the inherent intrinsic variations.

## DISCUSSION

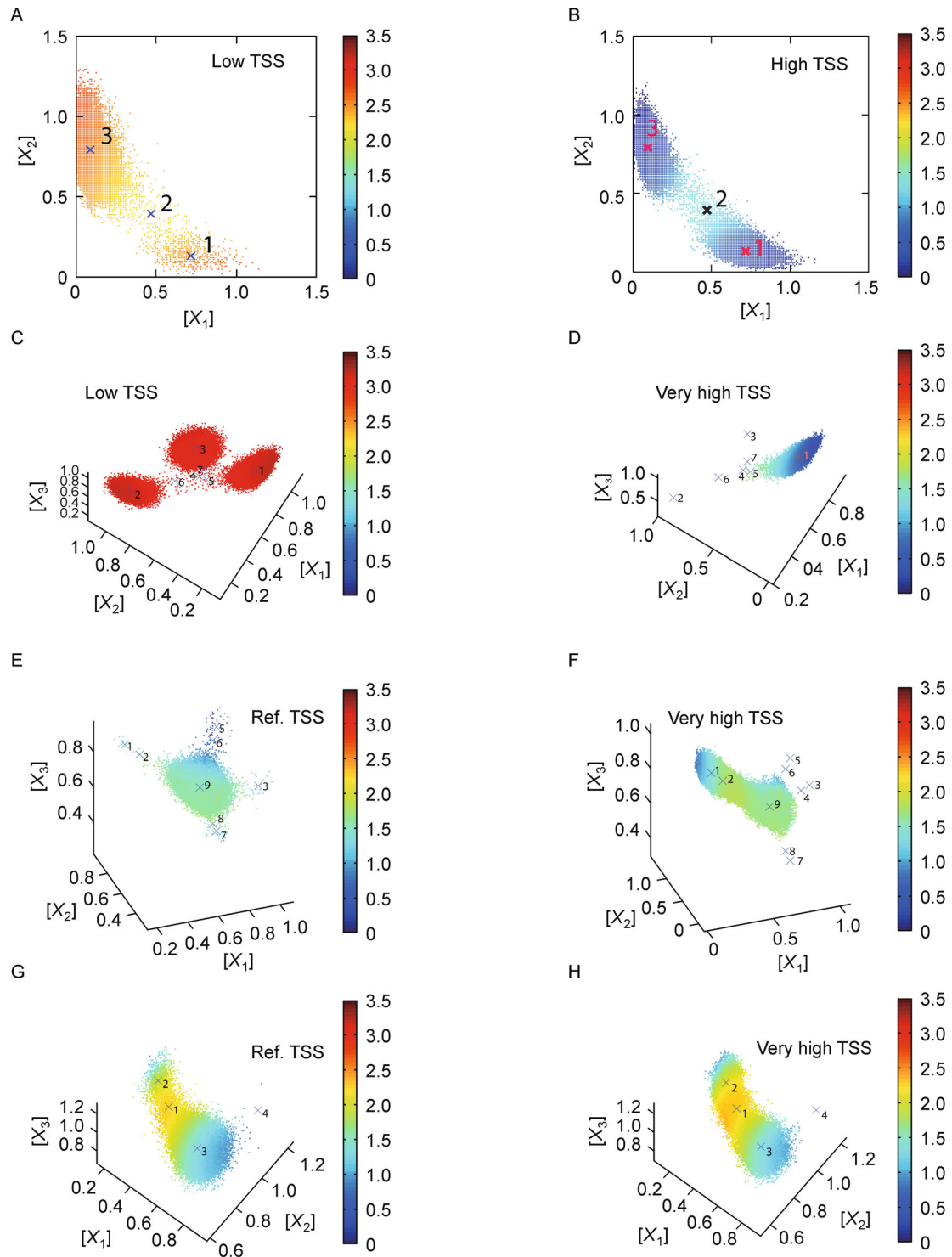
### Curl-forces versus grad-forces

As introduced earlier, stochastic processes with homogeneous noise size are driven by the combination of two forces: a curl-free force and divergence-free force, that is to say: a force arising from the gradient of the stochastic potential, and a force arising from the curl of the probability flux [14]. The role of curl has been studied in varied contexts [23,24]. In this work, the focus is to delineate the various effects in simple canonical core genetic regulation circuits. To this end, a measure of the stochastic potential was obtained as follows

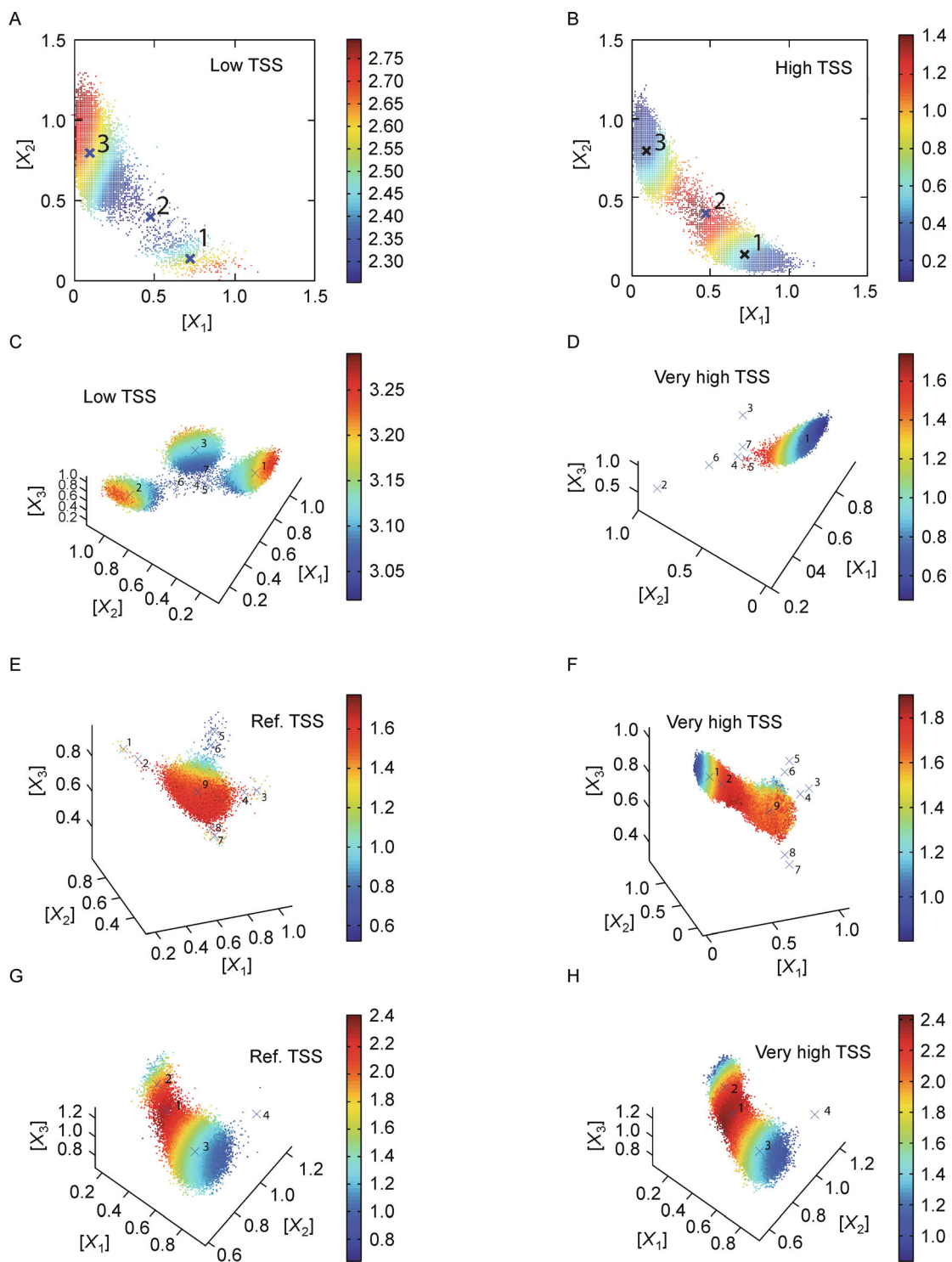
$$U = -\log_{10}R$$

Here,  $R$  is the residency obtained by rotating the systems about one of the outward pointing eigenvectors of one of the fixed points (usually chosen to be the central saddle point) such that the  $X_3^*$  axis (“\*” denotes rotated directions) is now pointing outwards, and the perpendicular  $X_1^*$  and  $X_2^*$  directions conveniently define a plane containing at least three fixed points of the dynamics, in the basins of attraction of which all stochastic tracks of the simulation propagate. For purposes of simplicity, the origin is redefined as the position of the fixed point about which the rotation is applied. The residency is obtained by histogramming the rotated data in the  $X_1^*-X_2^*$  plane (Figure 1). Figure 5 shows the result of numerically computing, bin by bin, the modulus of the gradient of the stochastic potential obtained directly from  $U$ . The analysis reveals that for both low TSS and high TSS, the gradient modulus is peaking in the far periphery of fixed points, and is tapering off to zero in the immediate vicinity of the fixed points. Noteworthy is the fact that the regions relatively free of gradient-force (i.e., near the fixed stable fixed points) are large: the potential is in fact locally relatively flat over an extended dynamical region.

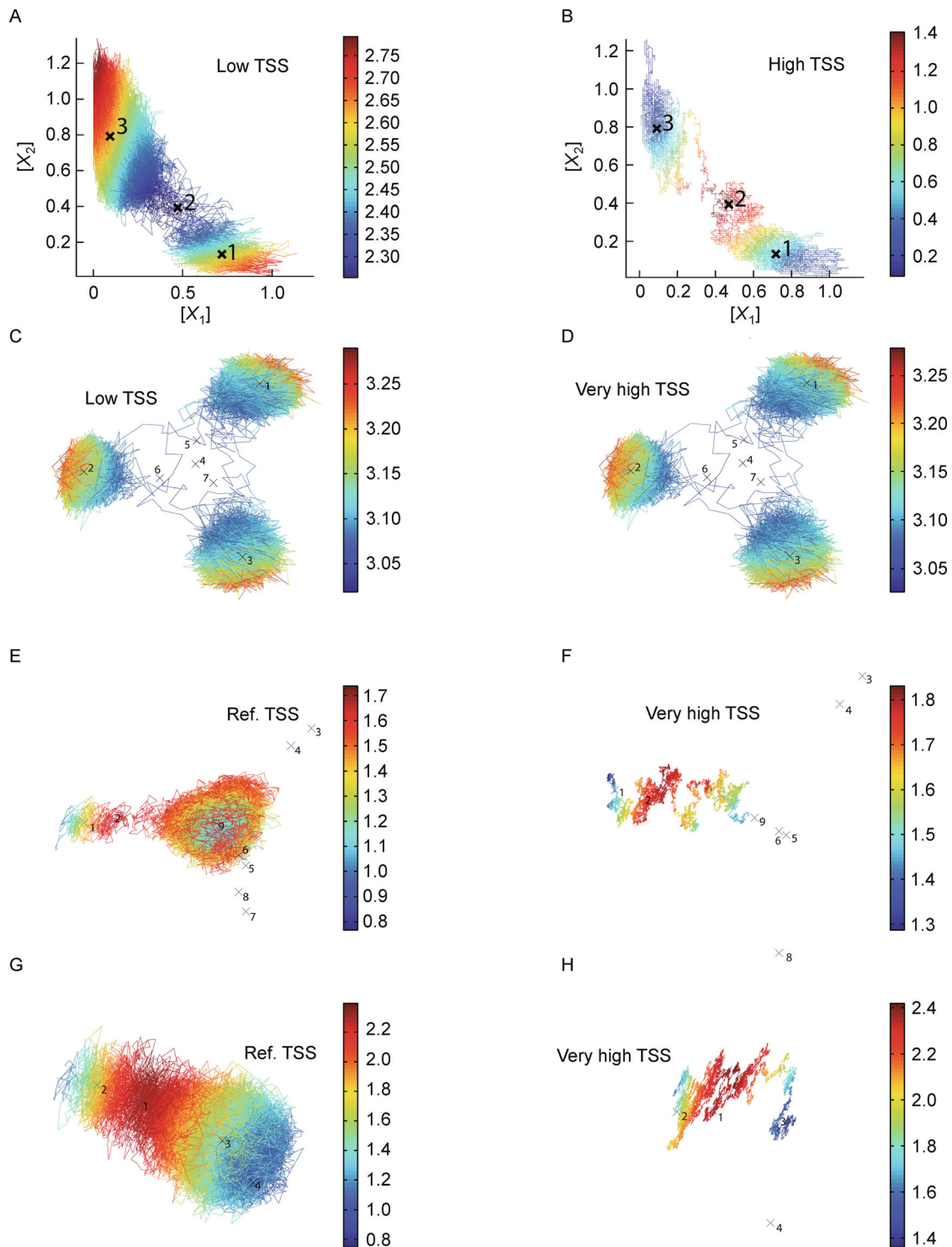
In order to make an equitable comparison between systems, the systems' respective multi-dimensional curls were evaluated along each stochastic track directly relying on the time-ordered sequence of n-dimensional jumps in the original un-rotated phase space (See Figure 3 for all tracks, see Figure 4 for one representative track). For DS1, the  $X_1-X_2$  data was then directly histogrammed. For DS2, DS3 and DS4 which all present  $X_1-X_2-X_3$  data, because of the presence of that  $X_3$  dimension, these systems' data need first be rotated to a suitable view, and then histogrammed. The rotation was performed to a convenient view (as described above, see also in Methods) and then histogrammed in the  $X_1^*-X_2^*$  rotated plane. Note that since the representation of multi-dimensional curl is prohibitively complex, only the modulus of the curl was studied (see Methods for details).



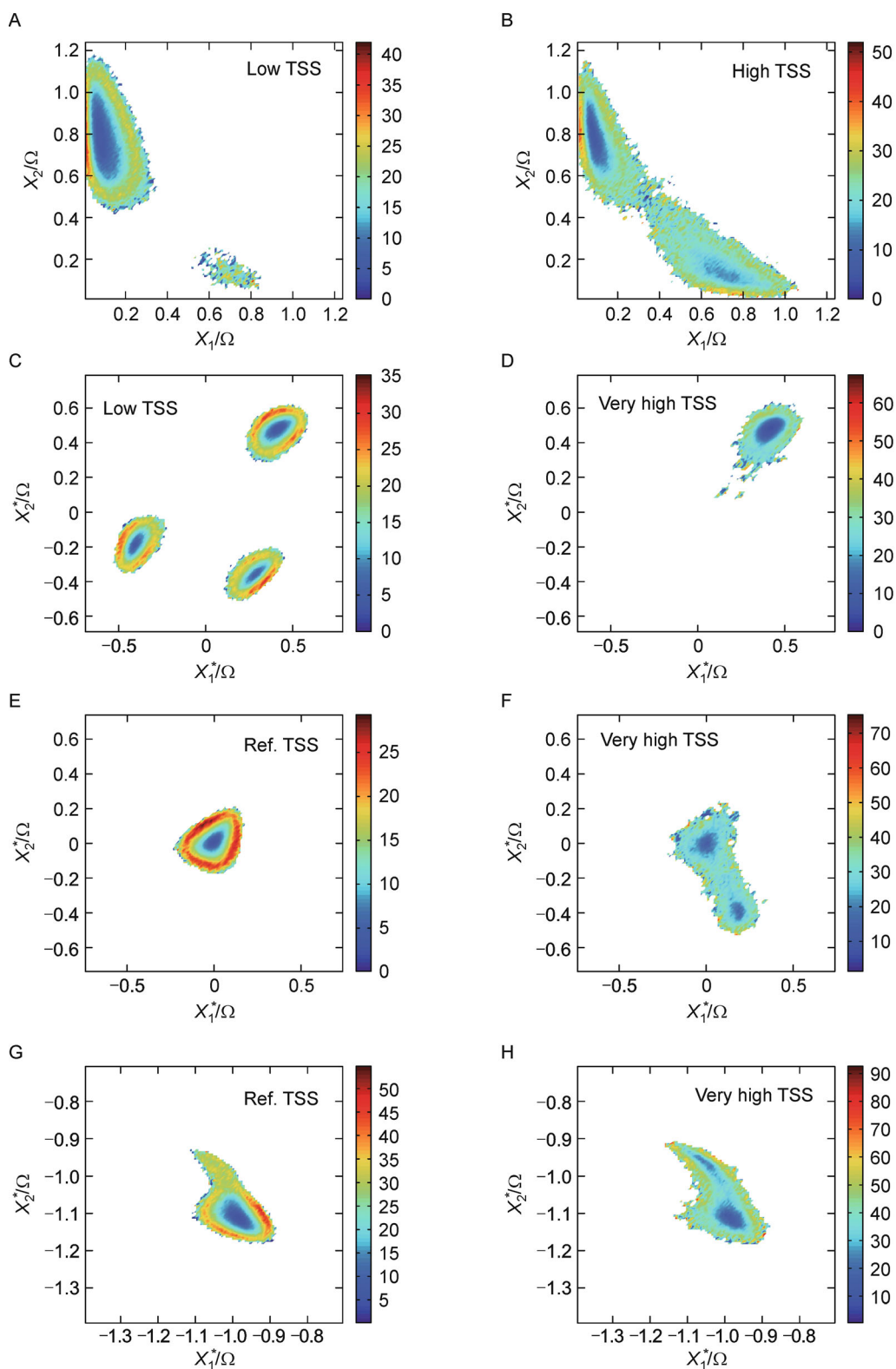
**Figure 2. Diagrams of curl sampled along the stochastic trajectories.** Each diagram includes 100 statistically independent trajectories plotted every 500<sup>th</sup> step. The contribution by components 1, 2 (top row) and 1, 2, 3 (rows 2, 3 and 4) to the local curl modulus is indicated by the color on a linear common scale. Rows 1, 2, 3 and 4 are for DS1, DS2, DS3 and DS4 respectively. Left and right columns are for low TSS and high TSS respectively. In high TSS, the messenger RNA sub-manifold is essentially at rest. All color bars have the same range. This facilitates comparison across conditions.



**Figure 3. Diagrams of curl sampled along the stochastic trajectories.** Each diagram includes 100 statistically independent trajectories plotted every 500<sup>th</sup> step. The contribution by components 1, 2 (top row) and 1, 2, 3 (rows 2, 3 and 4) to the local curl modulus is indicated by the color ranging from each figure’s minimum curl to its maximum; the scale is linear. Rows 1, 2, 3 and 4 are for DS1, DS2, DS3 and DS4 respectively. Left and right columns are for low TSS and high TSS respectively. In high TSS, the messenger RNA sub-manifold is essentially at rest. Contrasting to Figure 2, each color bar is different; each scale ranges from the shown track’s lowest to highest curl modulus. This facilitates focusing on each set of conditions.



**Figure 4. Diagrams of curl sampled along one single stochastic trajectory.** Propagation is displayed by sequential colored line segments joining successive locations along the stochastic track. Only a relevant part of the track is displayed; details in the text. The contribution by components 1, 2 (top row) and 1, 2, 3 (rows 2, 3 and 4) to the local curl modulus is indicated by the color ranging from each figure's minimum curl to its maximum. The curl scale is linear. Rows 1, 2, 3 and 4 are for DS1, DS2, DS3 and DS4 respectively. Left and right columns are for low TSS and high TSS respectively. In high TSS, the messenger RNA sub-manifold is essentially at rest. The  $X_1$ ,  $X_2$  and  $X_3$  axes are omitted on panel C through H because they would force an adversely distant viewpoint making details of the tracks imperceptible. See Table S1.



**Figure 5. Histograms of the modulus of the gradient of the stochastic potential.** The stochastic potential was computed from the corresponding residency histograms. Rows 1 through 4 are for DS1, DS2, DS3 and DS4 respectively. Left and right columns are for low and high TSS, respectively. Each panel carries its own linear scale, from the minimum to the maximum of the respective data.

Further, since the data is examined in the 1-2-3 space, only those 1-, 2- and 3- contributions were included in the computation of the modulus. This procedure allows a consistent comparison of grad moduli (Figure 5) and curl moduli (Figure 6). Note that on Figure 6, each panel carries its own color range. It is appropriate to focus on regions where the corresponding grad-force is null (islands surrounding the stable fixed points). Paying attention to each panel's own curl-color scale, it can immediately be seen from these data that, in all cases, the curl-force is relatively high in these regions. In short, where it matters most—in the dynamical regions near the attractive fixed points—curl-forces can be seen to overwhelm grad-forces. This is a feature that is seen to be common to all circuit topologies studied herein.

### Stochastic motion under combined curl- and grad-forces

Generally speaking, the notion of motion under combined curl and grad-forces is well understood. The grad-forces provide drift, herein towards the attractive fixed points of the dynamics, while the curl-forces provide rotation about this drift. The combined effect is a spiral along the direction of motion. But as the results shown above suggest, the very-close approach to the attractive fixed points is actually under the majority influence of curl-forces, albeit in the constant presence of random biochemical fluctuations. It is therefore interesting to delineate the dynamical meaning of this finding, and to explore any biological consequences. Note that, of course, the stochastic potential must evidently taper off near the stable fixed points because it reaches an extremum there. It is therefore the extent of the region near an attractive fixed point where essentially the gradient forces wane in favor of curl-forces, that is of interest. Below, the potential impact on the biology is explored.

Figure 7A shows one example drawn from many of the passage of a stochastic process through a stable attractive fixed point (DS3, very high TSS). See Figure S4 for several other examples. Near attractive fixed point #9, grad-forces wane; curl-forces dominate. The spiraling stochastic process flies through the fixed point literally unaffected because, locally, the grad-forces are negligible; curl-forces dominate. The process follows a typical spiraling (helical) trajectory along a general locally linear direction of motion. But at the process meanders off, it repeatedly reaches the boundary where grad-forces build up sufficiently again. There, the dynamics is veered back to perpetuate dwelling in the vicinity of the attractive fixed point (Figure 7 B). In short, the region near the attractive fixed point is essentially grad-free and dominated by curl.

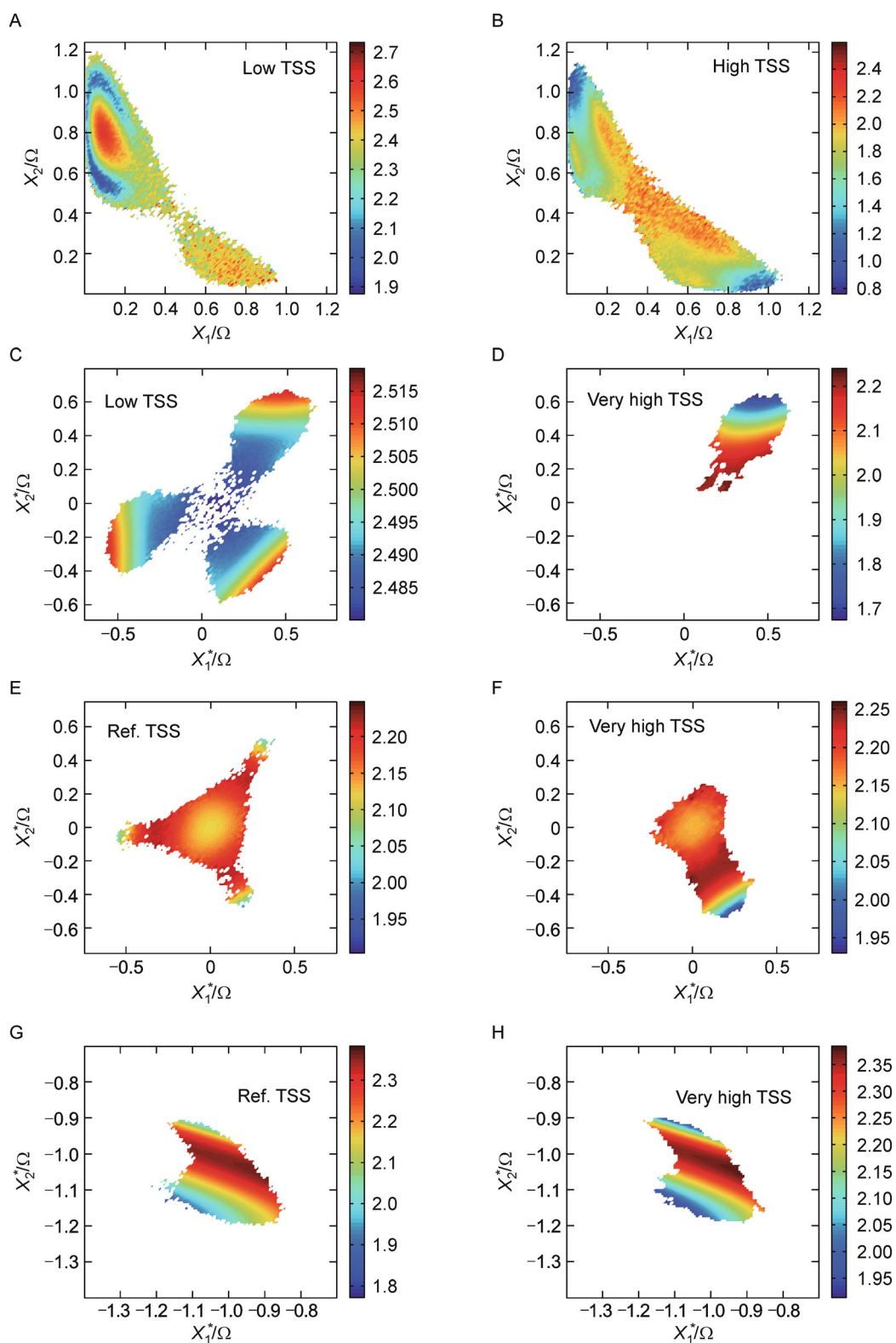
Figure 7C quantitatively shows how much the curl modulus and the gradient modulus respectively dip across the  $X_1$  or  $X_2$  directions while traversing the potential well center. The curl modulus is nearly constant while the gradient modulus, as expected, starting from the periphery plunges to zero at the fixed point. The well bottoms out. At and near the fixed point, the dynamics is mostly helical dominating null to weak drift (Figure 7A and 7B). Moving outwards, the dynamics is progressively less helical and more and more influenced by drift. Very far from the fixed point (Figure 7D and Figure S4O) the dynamics is totally dominated by drift since curl is locally negligible: the track is noisy but basically linear; it does not show any helical behavior.

### The dynamical regions nearest the stable fixed points are grad-forces free random mixers

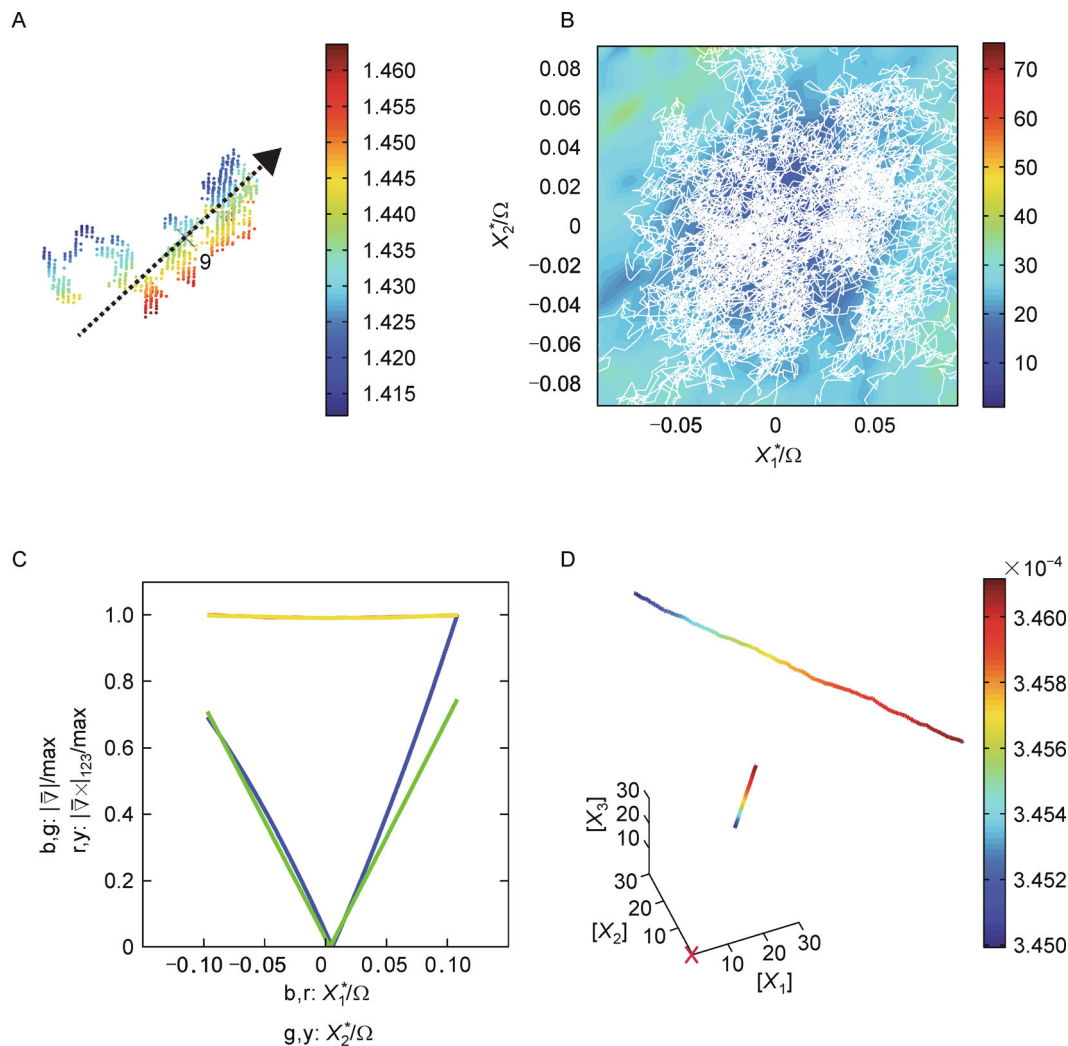
The results obtained here suggest that near attractive fixed points there exist bounded regions where curl-forces dominate, and grad-forces are comparatively negligible. In these regions, curl-forces help propagate the dynamics in characteristic combined directed and spiraling (helical) motion towards the region periphery where grad-forces build sufficiently up again and repeatedly turn the process back. In essence, in these regions, the presence of the nearby attractive fixed point is no longer the dominating dynamical feature of the system. Naturally, due to the finite number of molecules in realistic biological systems, noise sourced in biochemical fluctuations will impart a concomitant blur on this picture. What consequences, if any, can this be to living systems?

To address this question, it is interesting to notice that helical behavior in tracks near the fixed points is only observed in higher time scale separation simulations. Further, the helical behavior near the fixed points (grad forces waning and thus revealing the effect of curl forces) seems to be a prominent feature only for very high TSS versions of DS3 and DS4. One can readily see this on Figure 4: firstly, no data on panels of the left column display tracks with helical behavior. Secondly, data on Figure 4F and 4H from the right column clearly do. Hence, these simulation data suggest that time scale separation and circuit regulation complexity (DS3 and DS4 are 4- and 5- fully-connected gene circuits, respectively compared to DS1 and DS2 that are 2- and 3- fully-connected gene circuits, respectively) may impart more freedom to the dynamics near fixed points (rendering it “more free”) through the mechanism discussed above: curl forces dominating waning grad forces near fixed points.

What is described above is essentially a locally free system. It permits the dynamics to meander potentially far from an attractive fixed point, further away than if the



**Figure 6. Histograms of the modulus of the 1-2-3-curl in the various systems.** Each panel comprises 100 stochastic tracks. The curl was computed along each track. Rows 1 through 4 are for DS1, DS2, DS3 and DS4 respectively. Left and right columns are for low and high TSS, respectively. Each panel carries its own linear scale ranging from the minimum to the maximum of the respective data.



**Figure 7. Stochasticity near and far attractive fixed points.** (A) The part of one typical stochastic track in DS3 at very high TSS, as it passes through the location of attractive fixed point #9. The color linearly indicates the local curl modulus in the 1-2-3 axes. The dotted arrow indicates the general direction of the trajectory. Although fixed point #9 is a stable (attractive) fixed point, the dynamics in this region of phase space is dominated by curl forces. Hence, the motion is spiraling around the general direction of motion drifting through (but not converging to) the fixed point. This is because, locally, the gradient forces effectively vanish. Further from the fixed point, gradient forces relatively increase in strength and coerce the dynamics to dwell in the basin of attraction of fixed point #9. For the purpose of comparison and asserting the commonality of this behavior, several similar tracks are shown on Figure S4. (B) Most of the track is shown in white, rotated to the view discussed in the text, and overlaid on top of the stochastic potential gradient modulus underlying the attractive fixed point #9. Because the curl forces in the neighborhood of the fixed point dominate the dynamics, the track meanders in the neighborhood of the fixed point. But as the track repeatedly approaches the boundary where grad forces become significant again, the process is veered back. (C) Normalized gradient modulus and normalized 123 curl modulus profiles in the  $X_1$  and  $X_2$  directions crossing at the stochastic potential center. The labels b, g, r, and y stand for blue, green, red and yellow. The curl modulus is essentially constant whereas the gradient modulus falls nonlinearly to zero at the fixed point located at the origin. (D) Stochastic track started at  $(X_1, X_2, X_3, X_4) = [30, 30, 30, 30]$  in a far region of negligible curl as indicated by the linear color bar. As shown by the inset, the track is headed for the fixed points (red X at origin) driven by the gradient force. In this very low curl region however, the track despite being noisy, is mostly linear, lacking the helix nature seen on panel A. The scales of panels A and D (main) are similar.

dynamics were constantly restrained to the fixed point's immediate vicinity by grad-forces. At issue, then, is the impact of the local shape of the stochastic potential.

In an interesting scenario (e.g., Figure 5, elevated TSS Panels B, D, F and H), these regions correspond to relatively flattish areas of stochastic potential through

which the dynamics roams relatively freely and away from the fixed point, thus essentially roaming unrestrained to its immediate vicinity. So the regions, in addition to adding relative indeterminacy to the fixed point location, are in effect presenting an effective mechanism for destabilization of the fixed point. This appears to be done via the shape of the stochastic potential in the vicinity of the fixed point, allowing regions where grad-forces are negligible and curl-forces dominate. These aspects of the dynamics, to the author's knowledge, were not appreciated before. In short, the regions are consistent with biology's requirement for the underlying regulatory dynamics to accommodate a certain amount of indistinctness in the location of dynamical features. And the regions are, in addition, consistent with aiding destabilization. This particular aspect, as discussed below, is integral to biology.

### **Phenotype commitment: the long-term stability of differentiated states**

In recent years, experiments have begun to shed much needed light on the topic of how living systems achieve phenotype stability. Complementary to Conrad Waddington's historical metaphor for cellular differentiation [25] that has been recast in sound dynamical systems terms [16,26,27] perhaps the most interesting aspect of the emerging experimental picture is that evolution seems to be implementing a kind of active control system over cell fates. That in fact, in many cases, phenotype commitment is more than an inherent property of the underlying core regulatory genetic systems, but it involves the impact of regulatory transcription factors, or networks thereof, and/or chromatin state level regulation throughout the lifetime of the cell. In addition to the mentioned factors sourced in the cell proper, others are found sourced in the local environment (niche) it lives in. So, in fact, multiple mechanisms may act individually or in concert to lock-in a cellular fate, yet upon disruption they would allow lock-out or even reverting to a less mature differentiation state. In fact, perhaps a more accurate view of all this would be to consider that living systems have both built-in phenotype commitment and de-commitment mechanisms ([28], also see [29] for a theoretical treatment). The exciting picture revealed by careful study of development in living systems becomes even more fascinating considering that diseases such as cancer are apparently subject to intriguingly similar regulatory notions adding increasing weight to the idea that if cancer is a possible state of a dynamical system [11,30–32], it would be actively maintained [33]. This last aspect seems to offer new therapeutic opportunities including multimodal approaches (See multiple references in [33]).

The results presented herein support the notion that, in

many core regulatory systems, near fixed point of the dynamics, stochastic forces sourced in the gradient of the stochastic potential are overwhelmed locally by those forces sourced in the curl of the probability flux. This presents an intriguing local freedom mechanism that, as discussed above, could decrease the stability of a differentiated state, and yet in the more biologically meaningful context of an embedding active commitment-stabilizing network, the stability would indeed be maintained and so would phenotype commitment. So in fact, nonlinear dynamics and stochasticity collaborate to underlie phenotype commitment and de-commitment. Hence we hope to be getting closer to building a deep theoretical view not merely focused on elucidating instances of the "how" in living systems, but also shedding light on the "why".

In a nutshell, previous results from [1] and new results from this work are all consistent with biology in a very interesting and non-trivial way. Earlier results are pointing to core gene regulation circuits favoring being topologically complex rather than simple [1]. This work, however, is showing that the innate structure of the underlying stochastics (the waning of grad forces in favor of curl forces near fixed points, particularly in more complex networks) is conspiring to destabilize them. In [1], based on dynamics, it was shown why evolution selecting more complex networks is desirable from the point of view of fate commitment, whereas here, based on stochasticity, it is shown that more complex networks could be more prone to destabilization. Hence, the yin and the yang (阴阳) of biology.

## **METHODS**

### **Models**

DS1, DS2, DS3 and DS4 are, respectively, two-, three-, four- and five-gene idealized core genetic circuits whose structure, equations, parametrization, and dynamics are described in details in [1].

### **Parameter asymmetries and configurations**

The DS1, DS2, DS3 and DS4 parameter asymmetries were 0.9, 0.85, 1.0, and 1.39 respectively, so chosen to be away from intrinsic oscillations. The reader is referred to [1] for all parameter configurations.

### **Stochastic simulations**

All stochastic simulations were achieved by the exact same methods described in [1]. Briefly, the dynamical systems were decomposed into the set of discrete molecular generation and degradation events. The

probabilistic occurrence of these events was numerically simulated by the Gillespie algorithm [18–21]. A multitude of statistically independent simulations thus amounts to the solution of the Master Equation of the dynamics [18,19]. All simulations were developed, performed and analyzed using the Matlab (The Mathworks, Natick, Massachusetts) computational framework. Production of simulation data required dedicated use of two Dell Precision T5500 multi-core Intel I7 Pentium computers for a total of about 6,000 hours (~8 months) of non-stop integrated computing time. These computers typically ran between 25 to 50 parallel processes. The development of all code and the analysis of all data were performed on a single Intel I7 quad-core Dell XPS15z computer. This computer was also used extensively for parallel computations for a total of about 400 hours.

### Separatrix curve and general separatrix surfaces

A multitude of deterministic tracks are initiated over a regularly spaced grid in phase space. Each track is numerically integrated to its terminal fixed point of the dynamics. Each such track is thus tagged with the terminal fixed point ID. The separatrix line is then numerically estimated to reside between the tagged ID change-over locations. In multi-dimensional space, the general separatrix surface is obtained by straightforward generalization of this procedure.

### Residency

Residency diagrams are two-dimensional  $[X_1^*, X_2^*]$  histograms of the stochastic trajectory rotated to a suitable view (see below):  $[X_1 X_2 X_3] \rightarrow [X_1^* X_2^* X_3^*]$ .

### Linear analysis

All fixed point locations are numerically discovered. Using Mathematica (Wolfram, Champaign, Illinois), the Jacobian matrices for the DS1, DS2, DS3 and DS4 systems are symbolically derived from the sets of differential equations of the respective systems. The matrices are numerically evaluated at the fixed points. The eigenvalues and eigenvectors of the matrices are then computed using Matlab (The Mathworks, Natick, Massachusetts) built-in functions. The nature and stability of fixed points is then deduced accordingly [34].

### Rotations

Rotations are achieved by aligning the rotated  $z$ -axis with a suitably chosen eigenvector, at a convenient fixed point: usually the saddle between two stable fixed points. The origin of the rotated views is rotation point.

### Curl and grad

Using Mathematica (Wolfram, Champaign, Illinois), the curl is symbolically derived from the sets of differential equations of the various systems. Curl equations are lengthy and terse, so they are omitted for brevity. The modulus of curl is computed the usual way by summing the square of components and by taking the square root of the resulting sum. The gradient is obtained by numerically computing the local slopes in the residency. The modulus of the gradient is computed as just described.

### Profiles of the gradient moduli along the $X_1$ and $X_2$ directions

The moduli of the gradients of the stochastic potential shown on Figure 7, panel C, were computed as follows. Residency data shown on Figure 1 was used to compute the stochastic potential  $U = -\log_{10} R$  where  $R$  is the residency. Profiles along the  $X_1$  and  $X_2$  directions at the well center were fitted to a 3<sup>rd</sup> degree polynomial using Matlab (The Mathworks, Natick, Massachusetts). The local slope was computed using the derivative.

### SUPPLEMENTARY MATERIALS

The supplementary materials can be found online with this article at DOI 10.1007/s40484-016-0070-5.

### ACKNOWLEDGEMENTS

Special thanks go to Hongguang Xi for his unfailing assistance in completing the extensive parallel numerical simulations over extended periods of time. Many thanks to Michael Zhang for his continued hospitality at UT Dallas where MT designed the research, wrote the code, performed the analysis, and wrote the paper. The author is indebted to anonymous reviewers for their help in improving the manuscript.

### COMPLIANCE WITH ETHICS GUIDELINES

Marc Turcotte declares that he has no conflict of interest. This article does not contain any studies with human or animal subjects performed by the author.

### REFERENCES

1. Xi, H. and Turcotte, M. (2015) Parameter asymmetry and time-scale separation in core genetic commitment circuits. *Quant. Biol.*, 3, 19–45
2. Wang, J., Xu, L. and Wang, E. K. (2008) Potential landscape and flux framework of nonequilibrium networks: robustness, dissipation, and coherence of biochemical oscillations. *Proc. Natl. Acad. Sci. USA*, 105, 12271–12276
3. Wang, J., Zhang, K., Xu, L. and Wang, E. (2011) Quantifying the Waddington landscape and biological paths for development and differentiation. *Proc. Natl. Acad. Sci. USA*, 108, 8257–8262

4. Li, C., Wang, E. and Wang, J. (2011) Landscape and flux decomposition for exploring global natures of non-equilibrium dynamical systems under intrinsic statistical fluctuations. *Chem. Phys. Lett.*, 505, 75–80
5. Li, C., Wang, E. and Wang, J. (2011) Landscape, flux, correlation, resonance, coherence, stability, and key network wirings of stochastic circadian oscillation. *Biophys. J.*, 101, 1335–1344
6. Zhang, F., Xu, L., Zhang, K., Wang, E., Wang, J. (2012) The potential and flux landscape theory of evolution. *J. Chem. Phys.* 137, 065102
7. Li, C., Wang, E. and Wang, J. (2012) Landscape topography determines global stability and robustness of a metabolic network. *ACS Synth. Biol.*, 1, 229–239
8. Li, C. and Wang, J. (2013) Quantifying Waddington landscapes and paths of non-adiabatic cell fate decisions for differentiation, reprogramming and transdifferentiation. *J. R. Soc. Interface*, 10, 20130787
9. Wu, W. and Wang, J. (2013) Potential and flux field landscape theory. I. Global stability and dynamics of spatially dependent non-equilibrium systems. *J. Chem. Phys.*, 139, 121920
10. Wu, W. and Wang, J. (2013) Landscape framework and global stability for stochastic reaction diffusion and general spatially extended systems with intrinsic fluctuations. *J. Phys. Chem. B*, 117, 12908–12934
11. Li, C. and Wang, J. (2014) Quantifying the underlying landscape and paths of cancer. *J. R. Soc. Interface*, 11, 20140774
12. Li, C. and Wang, J. (2014) Landscape and flux reveal a new global view and physical quantification of mammalian cell cycle. *Proc. Natl. Acad. Sci. USA*, 111, 14130–14135
13. Xu, L., Zhang, F., Zhang, K., Wang, E. and Wang, J. (2014) The potential and flux landscape theory of ecology. *PLoS One*, 9, e86746
14. Wang, J. (2015) Landscape and flux theory of non-equilibrium dynamical systems with application to biology. *Adv. Phys.*, 64, 1–137
15. Gilbert, D. R., Gilber, N. L. (1910) On Four-Dimensional Vector Analysis, and its application in electrical theory. *Proc. Am. Acad. Arts Sci.*, 46, 165–181
16. Wang, K. Zhang, L. Xu, E. Wang, E. (2011) Quantifying the Waddington landscape and biological paths for development and differentiation. *Proc. Natl. Acad. Sci. USA*, 108, 8257–8262
17. Qian, H. (2011) Nonlinear stochastic dynamics of mesoscopic homogeneous biochemical reaction systems — an analytical theory. *Nonlinearity*, 24, R19
18. Gillespie, D. T. (1991) *Markov Processes, An Introduction for Physical Scientists*. Boston : Academic Press
19. Gillespie, D. T. (2007) *Stochastic simulation of chemical kinetics*. Annu. Rev. Phys. Chem., 58, 35–55
20. Gillespie, D. T. (1976) A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J. Comput. Phys.*, 22, 403–434
21. Gillespie, D. T. (1977) Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.*, 81, 2340–2361
22. Xi, H., Duan, L. and Turcotte, M. (2013) Point-cycle bistability and stochasticity in a regulatory circuit for *Bacillus subtilis* competence. *Math. Biosci.*, 244, 135–147
23. Feng, H. and Wang, J. (2011) Potential and flux decomposition for dynamical systems and non-equilibrium thermodynamics: Curvature, gauge field, and generalized fluctuation-dissipation theorem. *J. Chem. Phys.*, 135, 234511
24. Feng, H., Zhang, K. and Wang, J. (2014) Non-equilibrium transition state rate theory. *Chem. Sci.*, 5, 3761–3769
25. Waddington, C. H. and Kacser, H. (1957) *The Strategy of the Genes: A Discussion of Some Aspects of Theoretical Biology*. London: George Allen & Unwin
26. Xu, L., Zhang, K. and Wang, J. (2014) Exploring the mechanisms of differentiation, dedifferentiation, reprogramming and transdifferentiation. *PLoS One*, 9, e105216
27. Ferrell, J. Jr. (2012) Bistability, bifurcations, and Waddington's epigenetic landscape. *Curr. Biol.*, 22, R458–R466
28. Holmberg, J. and Perlmann, T. (2012) Maintaining differentiated cellular identity. *Nat. Rev. Genet.*, 13, 429–439
29. Li, C. and Wang, J. (2013) Quantifying Waddington landscapes and paths of non-adiabatic cell fate decisions for differentiation, reprogramming and transdifferentiation. *J. R. Soc. Interface*, 10, 20130787
30. Ao, P., Galas, D., Hood, L. and Zhu, X. (2008) Cancer as robust intrinsic state of endogenous molecular-cellular network shaped by evolution. *Med. Hypotheses*, 70, 678–684
31. Huang, S. (2013) Genetic and non-genetic instability in tumor progression: link between the fitness landscape and the epigenetic landscape of cancer cells. *Cancer Metastasis Rev.*, 32, 423–448
32. Li, C. and Wang, J. (2015) Quantifying the landscape for development and cancer from a core cancer stem cell circuit. *Cancer Res.*, 75, 2607–2618
33. Lasorella, A., Benezra, R. and Iavarone, A. (2014) The ID proteins: master regulators of cancer stem cells and tumour aggressiveness. *Nat. Rev. Cancer*, 14, 77–91
34. Strogatz, S. H. (1994) *Nonlinear Dynamics and Chaos*, Cambridge: Westview Press