Sloppy Systems

Ka Man (Ambrose) Yim

St John's College

University of Oxford

Dissertation submitted for

Honour School of Mathematical and Theoretical Physics Part C

Trinity 2017

Acknowledgements

It has been an illuminating experience to work on this fascinating project with Professor Ard Louis and Chico Camargo and I am most grateful to them for their advice and guidance. I am also indebted to my friends Joe Kidson, Matias Janvin and Alex Thorne for their counsel and support. None of my work would have been possible if not for the unwavering love of my family.

Abstract

A sloppy system is one which displays a logarithmic hierarchy of sensitivity to certain parameter combinations. To study the effect of sloppiness on evolution in biological systems driven by random mutations in parameter space, measures of evolvability and robustness are derived by seeking inspiration from analogous measures in discrete genotype-phenotype maps. Under this set of definitions, sloppiness is found to decrease state evolvability. However, it is not a sufficient condition for determining state robustness, parameter evolvability and parameter robustness.

Contents

List of Figures

Chapter 1

Introduction

Models in system biology describe the complex biochemical processes in living things. Very often, these complex nonlinear models display a property known as 'sloppiness', which may be a general property of a very wide class of systems described by many parameters [5]. A sloppy system is one that displays a logarithmic hierarchy of sensitivity to certain specific combinations of parameters [2] [8] [9]. The phenomenon of sloppiness is best understood in the context of 'information geometry' which formalises the relationship between parameters and the observable states of the system.

1.1 Information Geometry

Models of dynamical systems describe how a collection of quantities changes over time. In systems biology, the time evolution of a set of quantities $\boldsymbol{y} = (y_1, \ldots, y_m)$ is often modelled by a set of coupled nonlinear ordinary differential equations (ODEs) involving a set of parameters $\boldsymbol{\theta} = (\theta^1, \ldots, \theta^D)$. Given a set of parameters $\boldsymbol{\theta}$, the model predicts the evolution of the system $\boldsymbol{y}(t) = \boldsymbol{F}(t, \boldsymbol{\theta}) = (F_1(t, \boldsymbol{\theta}), \ldots, F_m(t, \boldsymbol{\theta}))$ (see figure 1.1). This defines a trajectory parametrised by t in *phase space* Γ, the m-dimensional space containing all possible values of $\boldsymbol{y}$. In an abstract sense, a model maps parameters from parameter space $\Theta \subseteq \mathbb{R}^D$ to trajectories in phase space: $\Omega : \Theta \mapsto \Gamma$. Yet in reality it is uncommon that the evolution of the system is observed at all times. Experimentalists often sample a trajectory at discrete points in time $(t_1, \ldots, t_n)$. Thus the behaviour observed by an experimentalist is the set of sampled positions on the trajectory $\boldsymbol{\varphi}(F) = (\boldsymbol{F}(t_1), \ldots, \boldsymbol{F}(t_n))$. The act of observation, $\varphi : \Gamma \mapsto Z$, maps the trajectories to an $N = n \times m$ dimensional space of observables $Z \subseteq \mathbb{R}^N$. This process defines a composite map $f = \varphi \circ \Omega : \Theta \mapsto Z$. The image of f in Z, i.e. the collection of all possible observations that are described by the model, is called the 'model manifold' $\mathcal{M}$ [9]. Since $\Theta \subseteq \mathbb{R}^D$ and $Z \subseteq \mathbb{R}^N$ are both differentiable manifolds, assuming $N > D$ and f is a smooth (infinitely differentiable) map, then f is an *immersion* of Θ in Z; if f is also *injective*, f is an *embedding* of Θ in Z. Thus $\mathcal{M}$ can be thought of as a manifold parametrised by coordinates $\boldsymbol{\theta} \in \Theta$.

A natural object to consider on manifolds is its metric which encodes its geometry. The metrics of Z and Θ are often assumed to be Euclidean [4]. Since $\mathcal{M}$ is immersed

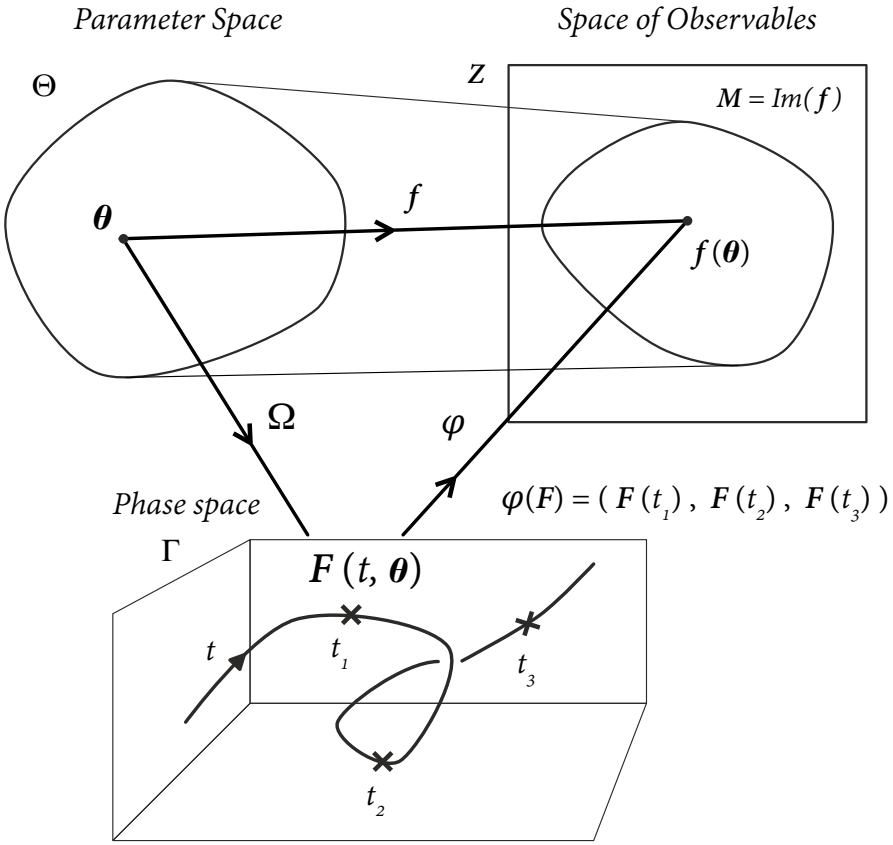


Figure 1.1: Illustration of the relationship between parameter space Θ, a 3 dimensional phase space and the space of observables Z. Ω maps parameters to trajectories $\boldsymbol{F}(t, \boldsymbol{\theta})$ in phase space. φ samples the trajectory at times (t_1, t_2, t_3) and maps the position of $\boldsymbol{F}$ at those times to a space of observables Z. This defines a composite map $f = \varphi \circ \Omega$.

by f in a Euclidean space Z , the metric *induced* on $\mathcal{M}$ is

$$g_{\mu\nu} = \delta_{ab} \frac{\partial f^a}{\partial \theta^\mu} \frac{\partial f^b}{\partial \theta^\nu} \tag{1.1}$$

Thus the infinitesimal distance in $\mathcal{M}$ due to an infinitesimal difference in parameter coordinates $d\boldsymbol{\theta}$ is

$$ds^2 = g_{\mu\nu} d\theta^\mu d\theta^\nu = \delta_{ab} \frac{\partial f^a}{\partial \theta^\mu} \frac{\partial f^b}{\partial \theta^\nu} d\theta^\mu d\theta^\nu \tag{1.2}$$

Unlike Θ, the distance in $\mathcal{M}$ is not Euclidean. The geometry of the model manifold is dependent on the Ω which maps parameters to trajectories in phase space; and φ, the experimental design which samples the trajectory for observable behaviours.

Remark 1.1. The current study limits its scope to models that are at least *locally identifiable* [2] [4]. A *globally* identifiable model f is one that is *injective* everywhere on the domain Θ. An example of a globally non-identifiable model is $y(t) = F(t, \theta^1, \theta^2) = e^{-t/\theta^1} + e^{-t/\theta^2}$, since exchanging the values of (θ^1, θ^2) maps to the same trajectory in phase space. Though $F(t, \theta^1, \theta^2)$ is not globally identifiable, it is *locally* identifiable, in the sense that for almost any point in $\boldsymbol{\theta} = (\theta^1, \theta^2) \in \Theta$ there exists a neighbourhood $U(\boldsymbol{\theta})$ such that $f = \varphi \circ \Omega$ is injective in the domain $U(\boldsymbol{\theta})$.

Remark 1.2. For many models described by ODEs (such as those in systems biology), an exact solution for phase space trajectories cannot be obtained analytically. However, the ODEs can be solved numerically for each parameter; by sampling across parameter space a 'numerical model manifold' can be constructed [9] [4].

1.2 Parameter Sensitivity

Given a variation in parameter $\boldsymbol{\theta} \to \boldsymbol{\theta} + \delta\boldsymbol{\theta}$, what is the magnitude of the change in behaviour? For any sufficiently small $\delta\boldsymbol{\theta}$, this can be well approximated by (1.2):

$$\delta s^2 \approx g_{\mu\nu}(\boldsymbol{\theta}) \delta\theta^\mu \delta\theta^\nu \tag{1.3}$$

Since $g_{\mu\nu}$ is a symmetric matrix, it has real eigenvalues $\{\lambda_a\}$ and a basis of orthonormal eigenvectors $\{\hat{\boldsymbol{n}}_\alpha\}$; moreover since $ds^2 \geq 0 \quad \forall d\boldsymbol{\theta}$, $g_{\mu\nu}$ is positive (semi-)definite and its eigenvalues are non-negative. Because $\{\hat{\boldsymbol{n}}_\alpha\}$ is a basis of Θ, any $\delta\boldsymbol{\theta}$ can be written as

$$\delta\boldsymbol{\theta} = \delta\theta'^\alpha \hat{\boldsymbol{n}}_\alpha \quad \text{i.e.} \quad \delta\boldsymbol{\theta} = P\delta\boldsymbol{\theta}' \tag{1.4}$$

where the columns of P are $\hat{\boldsymbol{n}}_\alpha$. Recognising that $(P^T P)_{ij} = \hat{\boldsymbol{n}}_i \cdot \hat{\boldsymbol{n}}_j = \delta_{ij}$, the coordinate transformation (1.4) between $\delta\boldsymbol{\theta} = (\delta\theta^1, \ldots, \delta\theta^D)$ and $\delta\boldsymbol{\theta}' = (\delta\theta'^1, \ldots, \delta\theta'^D)$ preserves the norm of $\delta\boldsymbol{\theta}$ in Θ:

$$\delta\boldsymbol{\theta}'^T \delta\boldsymbol{\theta}' = \delta\boldsymbol{\theta}^T P P^T \delta\boldsymbol{\theta} = \delta\boldsymbol{\theta}^T \delta\boldsymbol{\theta} \tag{1.5}$$

By spectrally decomposing the metric into $g = P\Lambda P^T$ (where $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_D)$)

$$\delta s^2 = \delta\boldsymbol{\theta}^T P \Lambda P^T \boldsymbol{\theta} = \delta\boldsymbol{\theta}'^T \Lambda \delta\boldsymbol{\theta}' = \lambda_1 (\delta\theta'^1)^2 + \cdots + \lambda_D (\delta\theta'^D)^2 \tag{1.6}$$

One can immediately read off the entries of the metric in the eigen-coordinates:

$$g'_{\mu\nu} = \lambda_\mu \delta_{\mu\nu} \tag{1.7}$$

This admits a geometric interpretation of the eigenvalues: $\sqrt{\lambda_\alpha}$ it is the change in behaviour (distance on $\mathcal{M}$) per unit distance of parameter variation $\delta\theta'^\alpha$ in the direction $\hat{\boldsymbol{n}}_\alpha$ in Θ. The eigenvalue spectrum of the metric characterises the local sensitivity of the model behaviour to variations in parameter space. If $\{\lambda_\alpha\}$ are

all equal and constant everywhere, then $\mathcal{M}$ is flat; if $\{\lambda_\alpha\}$ changes from $\boldsymbol{\theta}$ to $\boldsymbol{\theta}$, $\mathcal{M}$ is 'inhomogeneous' with respect to Θ; if $\{\lambda_\alpha\}$ takes a hierarchy of values, $\mathcal{M}$ is 'anisotropic' with respect to Θ.

Remark 1.3. When data analysts fit models to data, they often quantify the goodness of fit with the 'cost function' or χ^2, which is the Euclidean distance in Z between the model prediction $f(\boldsymbol{\theta})$ and data $\tilde{z}$. The best fit parameters $\boldsymbol{\theta}_*$ are those that minimise χ^2. Practitioners often characterise the parameter sensitivity of the model at $\boldsymbol{\theta}_*$ using the Hessian of χ^2, which is mathematically identical to $g_{\mu\nu}(\boldsymbol{\theta}_*)$ [9].

1.3 Sloppiness

In systems biology models, the dimension of parameter space is often quite large. For example, a model for the growth-factor-signaling network in PC12 cells is dependent on 48 parameters [1]. Due to the complexity of such systems, it is difficult a priori to identify the relevant mechanisms that play a dominant role and distinguish them from the irrelevant mechanisms. To identify the relevant mechanisms, a complex model that incorporates a large variety of mechanisms into account are fitted to data [11]. If a mechanism is relevant, then the system should be sensitive to the parameters which tune the mechanism. For example, if a relevant mechanism of a model is tuned by parameters $(\theta^1, \theta^2, \theta^3)$, $\delta\boldsymbol{\theta}$ in the subspace of $(\theta^1, \theta^2, \theta^3)$ should induce a larger δs^2 in $\mathcal{M}$, compared to variations of parameters which characterise mechanisms of lesser relevance.

However, it is rather uncommon in systems biology models that mechanisms can be distinctively classified as relevant or irrelevant. In their study of PC12 cells Brown et al. found that the system shows great sensitivity to all 48 parameters, suggesting that all mechanisms engage a concerted effort in influencing the gross dynamics of the system [1]. In a study of 17 cell cycle models, Gutenkunst et al. found that very few parameters induce a behavioural response of great significance or insignificance [5] (see figure 1.2). Given the complex nonlinear coupling between mechanisms in systems biology models, it should not be surprising that contributions from individual mechanisms to the overall dynamics of the system cannot be separated from each other.

Yet this is not to say parameter space Θ is isotropic. In all of the cell cycles investigated by Gutenkunst et al., the eigenvalues of the Hessian of these models are found to be approximately evenly spaced in their logarithms, the smallest eigenvalue being several orders of magnitudes smaller than the leading eigenvalue [5] (see figure 1.2). This phenomenon is known as 'sloppiness' [2] [8] [9]. Such a 'sloppy system' exhibits a hierarchy of sensitivity to perturbations in the eigen-directions of the metric $\{\hat{\boldsymbol{n}}_\alpha\}$. Note that sloppiness makes no statement about the absolute size of the eigenvalues, only their relative magnitudes. Why sloppiness occurs is still a matter of active research. While studies show that experimental design φ could play a role in increasing the separation between eigenvalues [2] [8], the fact that sloppiness has only emerged in some nonlinear models suggests that these models possess certain intrinsic properties which provide the necessary conditions for sloppiness [5].

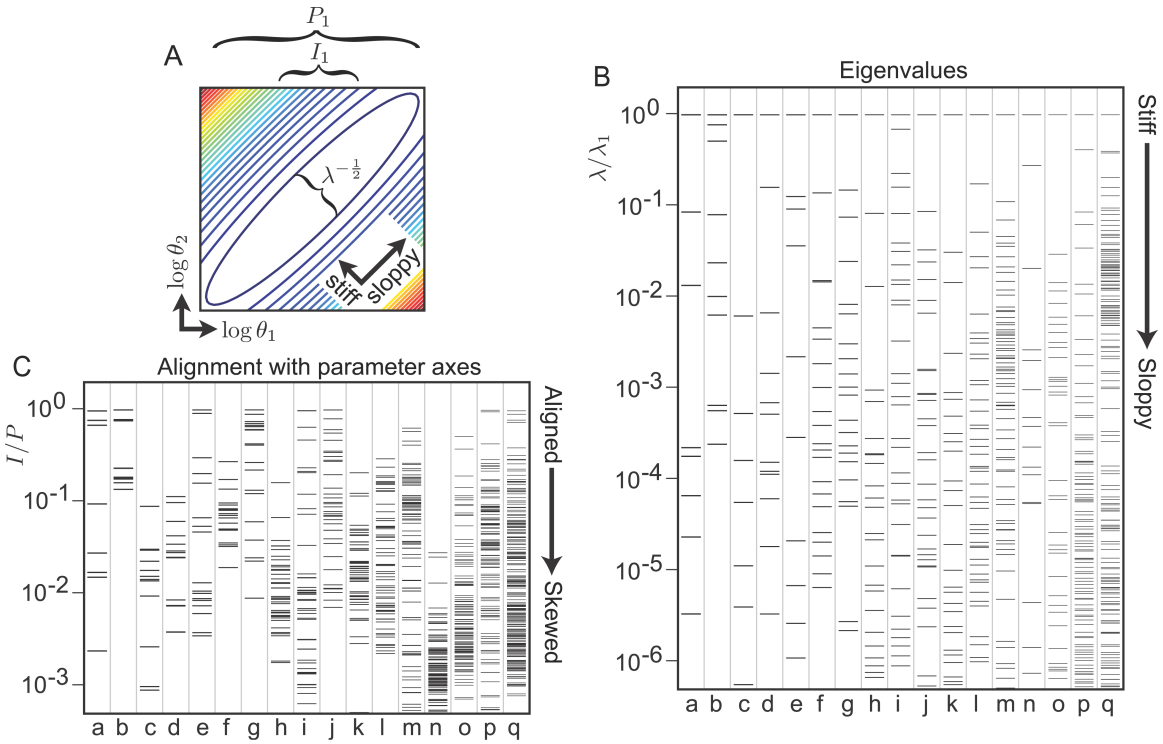


Figure 1.2: Figure from 'Universally sloppy parameter sensitivities in systems biology models' by Gutenkunst et al. [5]. A shows a surface of constant χ^2 (δs^2) projected onto parameter space Θ; such surfaces are ellipsoids whose semi-major axes are characterised by $(\lambda_i)^{-1/2}$. B shows the eigenvalue spectra of various systems biology models describing cell cycles, mitosis, circadian rhythm, growth-factor signaling, regulatory networks, metabolism and etc. In C, I_i/P_i is the ratio between the intersection of the χ^2 ellipsoid with the parameter axes of θ^i and the projection of the ellipsoid onto the axes. The smaller the ratios, the smaller the alignment between the bare parameter axes and the eigenvectors of the Hessian.

1.4 Evolvability and Robustness

The question of how easily organisms encounter new traits is of great importance to understanding the process of adaptation in evolutionary biology. In a study of $L = 15$ RNA sequences and their secondary structures, it was discovered that 50% of all possibles sequences fold into only 6% of the structures that appear [6]. The folding process - an instance of what is called a 'genotype-phenotype' (GP) map - shows a huge bias towards a small number of 'frequent' secondary structures. Thus a 'mutating agent' randomly searching the space of sequences (genotypes) is far more likely to encounter a very small subset of structures (phenotypes) that occur frequently. Such a bias limits the variety of phenotypes that is available for natural selection.

The example of RNA folding illustrates how a map between a space of parameters (genotypes) and a space of behaviours (phenotypes) affects the likelihood of encountering different behaviours under random mutation in parameter space. One would similarly expect parameter space inhomogeneity and anisotropy in continuous models

to influence the likelihood of finding new behaviour in $\mathcal{M}$ under random mutations in Θ. Since the processes that systems biology models describe (e.g. metabolism) must adapt their behaviour to changes in their circumstances, the question of how easily these systems find new traits via parameter perturbations is just as pertinent.

This dissertation focuses two ideas that are relevant to this question. One is *robustness*, which describes the likelihood of a system to change its behaviour under mutation; the other is *evolvability*, which describes the diversity of behaviour accessible under mutation. Both are desirable characteristics of a biological system. A system should be able to resist mutation and retain favourable phenotypes, yet also have the flexibility to adapt to its circumstances. Notions of evolvability and robustness are precisely quantified in discrete GP maps by a set of definitions developed by Wagner [10] and an attempt has been made to generalise the discrete definitions to continuous models by Sethna et al. [3]. Sethna et al. made the extraordinary claim that sloppiness enables mutations in parameter space to explore a diverse range of behaviour. This dissertation sets out to assess the validity of this claim and examine their definitions of evolvability and robustness. In chapter 2, definitions of evolvability and robustness in discrete GP maps proposed by Wagner are reviewed; using the machinery of information geometry, chapters 3 and 4 put forward modified or new measures of evolvability and robustness that are more consistent with Wagner's definitions; working with a fresh set of definitions, the role of sloppiness on evolvability and robustness is investigated in chapter 5.

1.5 The Null Model

It will be useful for subsequent discussions to introduce an object called 'the null model' at this juncture. The null model $\mathcal{M}^*$ is an artificial object constructed for the analysis of a model $\mathcal{M}$. It is nothing more than a flat manifold (homogeneous and isotropic with respect to Θ) that takes up the same volume as $\mathcal{M}$ in Z. Given that it is flat, the metric on $\mathcal{M}^*$ takes the form

$$\eta_{\mu\nu} = \Lambda\delta_{\mu\nu} \tag{1.8}$$

where Λ is a constant scale. The volume of the model manifold over Θ is

$$V = \int d\boldsymbol{\theta}\sqrt{\det(g)} \tag{1.9}$$

The volume of the null manifold is

$$V^* = \int d\boldsymbol{\theta}\sqrt{\det(\eta)} = \Lambda^{D/2}\int d\boldsymbol{\theta} \tag{1.10}$$

The volume of parameter space is simply $V_p = \int d\boldsymbol{\theta}$, so forcing $V = V^*$ fixes Λ to be

$$\Lambda = \left(\frac{V}{V_p}\right)^{\frac{2}{D}} \tag{1.11}$$

Chapter 2

Evolvability and Robustness in Discrete Genotype-Phenotype Maps

Following on from the discussion in section 1.4, it would seem that robustness and evolvability are competing characteristics of a system. Consider figure 2.1 which represents a space of discrete genotypes. Each site on the lattice is a genotype, and the genotypes that can mutate into each other through one mutation are connected by an edge. Comparing genotypes A and B, one can observe how genotypes with more neighbours of the same phenotype (neutral neighbours) are connected to fewer alternative phenotypes that are different from each other, since there are fewer non-neutral neighbouring genotypes left over to support a diverse collection of alternative phenotypes. Hence genotypes that are more robust (more neutral neighbours and less likely to change phenotype after mutation) are less evolvable (fewer alternative phenotypes to mutate into), a correlation that Wagner has discovered in RNA secondary structure GP maps [10]. Wagner argued that evolvability and robustness can also be assessed from the perspective of phenotypes [10]. Consider figure 2.1 as an example again. A more frequent phenotype occupies a larger extent of genotype space and increases the number of alternative phenotypes within the reach of the set of genotypes that map to the phenotype - as such the *phenotype* is more evolvable. Yet Wagner also empirically observed in RNA secondary structure GP maps that genotypes belonging to the same phenotype are connected to each other more often than what one would expect by chance. Hence the robustness and evolvability of secondary structures (phenotypes) are positively correlated with each other. If this correlation holds generally for other GP systems, it would have great implications on how evolutionary biology considers the emergence of evolutionary novelty. The rest of this chapter provides more precise definitions and distinctions between genotype evolvability, genotype robustness, phenotype evolvability and phenotype robustness, and makes initial contact with sloppiness.

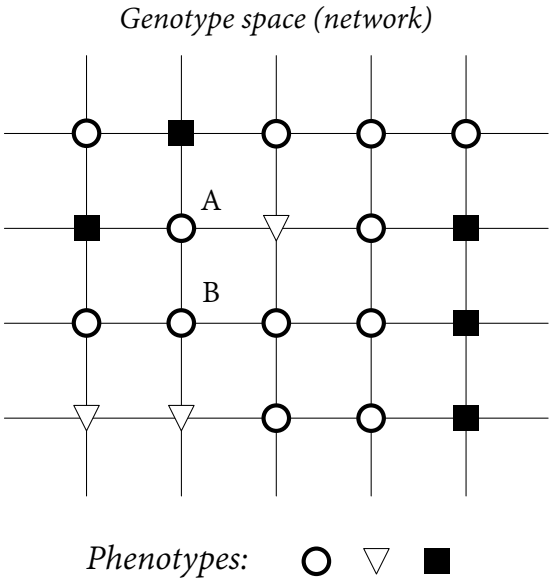


Figure 2.1: Representation of a genotype-phenotype map in genotype space. Genotype A has very few neighbours of the same kind, yet its neighbourhood contains a diverse set of alternative phenotypes. Most of the neighbours of genotype B have the same phenotype, leaving little room for a diverse set of alternative phenotypes in its neighbourhood. Observe how the mapping from genotype to phenotype is many-to-one. The most frequent phenotype, the circle, forms a large neutral network that extends across parameter space, allowing it to mutate into triangles and squares. The triangle and square which are less frequent can only mutate into the circle, but not into each other.

2.1 Evolvability and Robustness of Genotypes

In a discrete GP system, genotypes can mutate into one another by one-step mutations; this mutational relationship between genotypes can be described by a network, in which genotypes are the nodes and edges between nodes represent allowed mutations. An example of a genotype space is the space of RNA sequences of length n. The genotype - an RNA sequence - can be thought of abstractly as a string of n letters where each letter in the sequence can be either A, G, C or U. Changing one letter in the string mutates one sequence into another. It is useful to define the notion of a 1-neighbourhood of a genotype:

Definition 2.1 (1-neighbourhood of a genotype)**.** The 1-neighbourhood of a genotype g are the genotypes that can be accessed by g in one mutation.

Stepping through the network of genotypes, a mutating individual can potentially encounter new phenotypes at each node. It has been observed that the map between

genotypes and phenotypes are very often many-to-one [6]. In other words, there tends to be a redundancy of genotypes for each phenotype. Moreover, as Wagner has observed, the genotypes that map to the same phenotype are often connected to each other by one-step mutations in the genotype network [10]. Given this structure in GP systems, it is convenient to introduce three related concepts:

Definition 2.2 (Neutral neighbour of a genotype)**.** A neutral neighbour of a genotype g is a genotype in the 1-neighbourhood of g which maps to the same phenotype as g.

Definition 2.3 (Neutral set)**.** The neutral set of phenotype p is the maximal subset of genotypes which map to p.

Definition 2.4 (Neutral network)**.** In a genotype network, connected components of genotypes that map to the same phenotype p are the neutral networks of p.

The existence of neutral networks has significant implications on the mutational robustness of genotypes. If the majority of a genotype's 1-neighbourhood are neutral neighbours, it is more likely to mutate into a genotype that maps to the same phenotype. In otherwords, the phenotype is likely to *persist* after mutation. A definition of 'genotype robustness' can be used to quantify the likelihood of persistence [10]:

Definition 2.5 (Genotype robustness)**.** The robustness of a genotype is the fraction of neutral neighbours of a genotype g.

The evolvability of a genotype can be defined in a similar way. A more evolvable genotype should be able to explore a more diverse set of phenotypes after mutation. Wagner quantifies this by enumerating the number of different phenotypes in the 1-neighbourhood of a genotype:

Definition 2.6 (Genotype evolvability)**.** The evolvability of a genotype is the number of different phenotypes that are accessible in the 1-neighbourhood of g.

Remark 2.1. It is important to point out a flaw in Wagner's definition in quantifying genotype evolvability. Suppose the 1-neighbourhood of a genotype g contains nine neighbours mapping to three different phenotypes A, B and C. Consider two cases: (i) A, B and C splits the nine neighbours evenly amongst themselves, i.e. A, B and C each corresponds to three neighbours; (ii) 7 neighbours map to A while B and C only correspond to 1 each. Wagner's definition does not distinguish between cases (i) and (ii) where (i) obviously creates a greater diversity of outcomes for the genotype. In (ii) it is far more likely that the genotype mutates into neighbours which map to A.

2.2 Evolvability and Robustness of Phenotypes

The story could also be told from the perspective of phenotypes. Given the redundancy in GP maps, the likelihood of a phenotype p to persist after mutation should take into account all the genotypes that map to p. Consider a scenario in which a

phenotype p yields the optimum adaptation to selection pressure. Averaging over an ensemble of the system's possible evolutionary histories, the population of individuals with phenotype p should be evenly distributed in the neutral set of p, as evolutionary pressure does not distinguish between the genotypes that map to p. The robustness of the phenotype can be quantified by the fraction of individuals that retain the same phenotype p after mutation. This can be computed by adding up the probabilities of each individual retaining the same phenotype after mutation. The likelihood of an individual with genotype g retaining the same phenotype is simply the robustness of g. Since the population is evenly distributed in the neutral set, the fraction of the population that retains the same phenotype is computed by summing up the genotype robustness of the neutral set. This argument justifies Wagner's definition of phenotype robustness [10]:

Definition 2.7 (Phenotype robustness)**.** The robustness of a phenotype p is the average genotype robustness of its neutral set.

The evolvability of a phenotype can also be motivated by the same argument used for genotype evolvability. A phenotype is able to 'mutate' into another phenotype if and only if there exists at least one genotype in their respective neutral sets that are connected to one another in the genotype network. Wagner quantifies the evolvability of a phenotype by the number of different phenotypes accessible by a population of individuals dispersed in its neutral set [10]; in other words,

Definition 2.8 (Phenotype evolvability)**.** The evolvability of a phenotype p is the number of unique phenotypes that are accessible by genotypes in the neutral set of p.

Remark 2.2. Wagner's definition of phenotype evolvability suffers from the same issue plaguing genotype evolvability which was discussed in remark 2.1: it does not distinguish between the relative likelihood of outcomes after mutations and so fails to give a full measure of the diversity of mutational outcomes.

2.3 Sloppiness and Neutral Spaces

Most of the idioms of GP maps are not immediately suitable for describing continuous models [3]. While there are only a finite number of genotypes and phenotypes in the domain and image of discrete GP maps, parameters and behavioural states live in spaces of real numbers which are uncountable sets; hence the counting schemes employed in the definitions above need to be carefully modified for those ideas to make sense in continuous models. However, an important theme is carried through: the distinction between the evolution of a single *individual* and that of a *population* of individuals remains relevant. Parameter evolvability/robustness, the continuous analogue of genotype evolvability/robustness, characterises the mutational outcome of an individual occupying a single point in parameter space; state (behaviour) evolvability/robustness, the continuous analogue of phenotype evolvability/robustness, describes an ensemble of individuals that belong to the same state on the model manifold. This idea is a very useful starting point for formulating definitions of evolvability and robustness in continuous models.

Sethna et al. makes an analogy between sloppiness in continuous models and neutral sets in GP maps [3]. Since sloppiness reduces a system's sensitivity to perturbations in some directions of parameter space, Sethna et al. argued that sloppiness would create 'neutral subspaces' - parameter insensitive subspaces of Θ - analogous to neutral networks in GP maps. Wagner proposed that large neutral networks in discrete GP systems enable individuals on the neutral network to explore a diverse set of phenotypes [10]. In response to this, Sethna et al. claims that large neutral subspaces, similar to large neutral networks in discrete GP maps, allow mutating agents with similar behaviour to explore a larger extent of parameter space and increase their likelihood of encountering new behaviour in $\mathcal{M}$ [3]. This claim is examined in chapter 5.

Chapter 3

Robustness and Evolvability of Parameters

3.1 Parameter Robustness

Inspired by Wagner's definition of genotype robustness, Sethna et al. defined the robustnesss of a parameter $R_p(\boldsymbol{\theta})$, or what they refer to as the 'chemotype', as the fraction of mutations $\delta\boldsymbol{\theta}$ in Θ satisfying $|\delta\boldsymbol{\theta}|^2 < \Delta$ that do not change the distance in $\mathcal{M}$ beyond a tolerance $\delta s^2 < \epsilon^2$, where $\Delta, \epsilon \to 0$ with finite Δ/ϵ [3]. This definition is best interpreted geometrically (see figure 3.1). Consider an agent at $\boldsymbol{\theta}_0$ under the influence of a random isotropic perturbation of size less than Δ. The agent's possible positions in Θ after the perturbation is the volume enclosed by a D-dimensional hypersphere of radius Δ. Expressing this in the norm-preserving coordinates $\delta\theta^i$ that (locally) diagonalise the metric of the model manifold $g_{\mu\nu}(\boldsymbol{\theta}_0)$ (equation 1.4)

$$\Delta^2 = (\delta\theta^1)^2 + \cdots + (\delta\theta^D)^2 \tag{3.1}$$

This is the equation of a sphere in Θ.

Now change perspective to $\mathcal{M}$ and consider the distance in $\mathcal{M}$ in $\delta\theta^i$ coordinates:

$$\delta s^2 = \lambda_\mu \delta_{\mu\nu} \delta\theta^\mu \delta\theta^\nu \tag{3.2}$$

Define a new set of local 'normal' coordinates

$$\delta\hat{\theta}_i = \delta\theta_i \sqrt{\lambda_i} \tag{3.3}$$

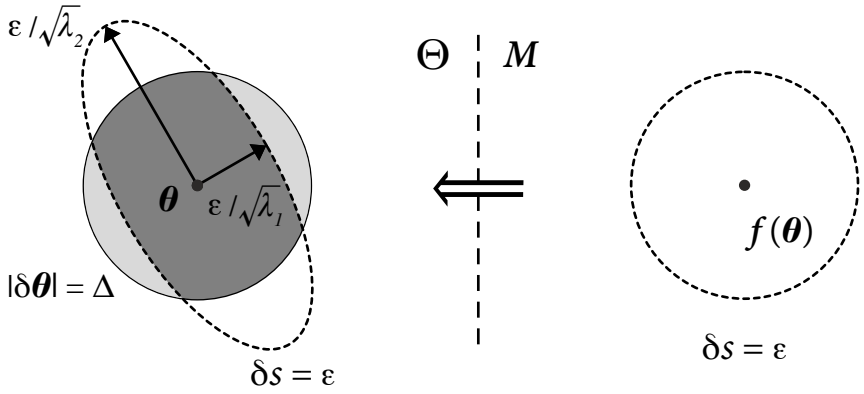
such that in these coordinates the metric is flat:

$$\delta s^2 = g_{\mu\nu}(\hat{\theta}) \delta\hat{\theta}^\mu \delta\hat{\theta}^\nu = \delta_{\mu\nu} \delta\hat{\theta}^\mu \delta\hat{\theta}^\nu \tag{3.4}$$

For a perturbation satisfying the tolerance $\delta s^2 < \epsilon^2$, it must lie within a D-dimensional hypersphere of radius ϵ on $\mathcal{M}$; in $\delta\hat{\theta}^i$ coordinates,

$$\epsilon^2 = (\delta\hat{\theta}^1)^2 + \ldots + (\delta\hat{\theta}^D)^2 \tag{3.5}$$

(a) Parameter robustness by volume subtraction



(b) Parameter Evolvability

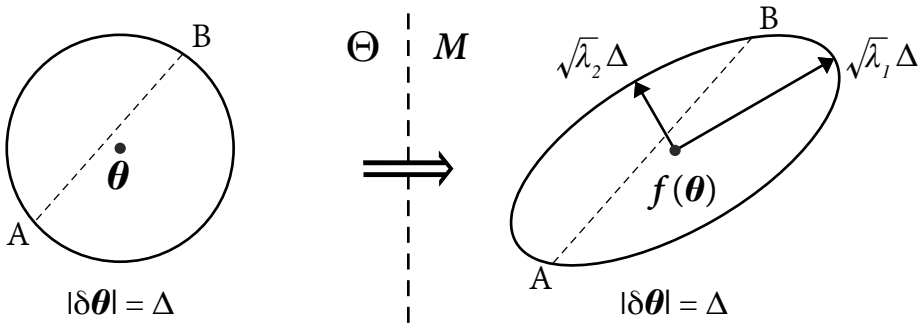


Figure 3.1: (a) Projection of ϵ-tolerance sphere from $\mathcal{M}$ onto Θ. Parameter robustness (by volume subtraction) can be understood as the fraction of perturbations $|\delta\boldsymbol{\theta}| < \Delta$ that satisfy $\delta s < \epsilon$, indicated by the volume shaded dark grey. (b) Projection of Δ-perturbation sphere from Θ onto $\mathcal{M}$. Evolvability is the average distance between two points A and B on the surface $|\delta\boldsymbol{\theta}| = \Delta$ in $\mathcal{M}$.

This ellipse can be mapped onto Θ by recasting the equation in $\delta\theta$ coordinates,

$$\epsilon^2 = \lambda_1(\delta\theta^1)^2 + \cdots + \lambda_D(\delta\theta^D)^2 \tag{3.6}$$

The semi-major axes of the ellipse in the θ^i direction is $\epsilon/\sqrt{\lambda_i}$. Returning to the definition of $R_p(\boldsymbol{\theta})$, finding the fraction of perturbations satisfying $|\delta\boldsymbol{\theta}|^2 < \Delta$ and $\delta s^2 < \epsilon^2$ is equivalent to calculating the intersecting volume between the sphere and ellipse defined by equations (3.1) and (3.6) respectively as a fraction of the volume of the sphere. In other words, robustness is the fraction of the perturbation sphere

remaining once the points that lie outside the tolerance ellipse is subtracted. However, as Sethna et al. points out, this is a difficult calculation [3]!

To get around this difficulty, Sethna et al. softened the hard boundaries of the spheres and ellipses and represented them with probability distributions that decay to infinity [3]. Suppose the agent mutating away from $\boldsymbol{\theta}_0$ travels a distance $|\delta\boldsymbol{\theta}|$ that is normally distributed with a width scale Δ

$$P(|\delta\boldsymbol{\theta}|, \Delta) = \frac{1}{(\sqrt{2\pi}\Delta)^D} \exp\left(-\frac{|\delta\boldsymbol{\theta}|^2}{2\Delta^2}\right) \tag{3.7}$$

$$= \frac{1}{(\sqrt{2\pi}\Delta)^D} \exp\left(-\frac{1}{2\Delta^2}\sum_{i=1}^{D}(\delta\theta^i)^2\right) \tag{3.8}$$

To take away the perturbations that lie outside the ellipse, imagine a scenario where the agent is removed with some probability that increases with distance δs. Sethna et al. chose the probability of survival Q to be a Gaussian with a typical length scale $\delta s \sim \epsilon$:

$$Q(\delta s, \epsilon) = \exp\left(-\frac{\delta s^2}{2\epsilon^2}\right) \tag{3.9}$$

$$= \exp\left(-\frac{1}{2\epsilon^2}\sum_{i=1}^{D}\lambda_i(\delta\theta^i)^2\right) \tag{3.10}$$

where the normalisation of $Q(\delta s, \epsilon)$ ensures the probability of survival for not mutating is $Q(\delta s = 0, \epsilon) = 1$. If an ensemble of mutating agents start off at $\boldsymbol{\theta}_0$ and diffuse away from $\boldsymbol{\theta}_0$ with a scale $|\delta\boldsymbol{\theta}| \sim \Delta$ a la equation (3.7), the fraction of agents that survive removal can be used as a measure of the robustness of $\boldsymbol{\theta}_0$. In terms of $P(|\delta\boldsymbol{\theta}|, \Delta)$ and $Q(\delta s, \epsilon)$,

$$R_p(\boldsymbol{\theta}, \Delta, \epsilon) = \int d\delta\boldsymbol{\theta} P(|\delta\boldsymbol{\theta}|, \Delta) Q(\delta s, \epsilon) \tag{3.11}$$

$$= \frac{1}{(\sqrt{2\pi}\Delta)^D} \int d\delta\boldsymbol{\theta} \exp\left(-\frac{1}{2\Delta^2}\sum_{i=1}^{D}(\delta\theta^i)^2\left(1 + \lambda_i\frac{\Delta^2}{\epsilon^2}\right)\right)$$

$$= \prod_{i=1}^{D} \frac{1}{\sqrt{1 + \frac{\lambda_i}{\lambda_c}}} \tag{3.12}$$

where $\lambda_c = \frac{\epsilon^2}{\Delta^2}$.

This definition of parameter robustness has a few undesirable properties. First of all, it is dependent on an arbitrary scale λ_c whose value is left to whoever's applying the definition to decide. In their analysis of a model of an EGF/NGF signaling pathway, Sethna et al. chose λ_c to be the fourth-largest eigenvalue of the local Hessian, which changes from $\boldsymbol{\theta}$ to $\boldsymbol{\theta}$ [3]. This is a problematic choice when the robustness of different parameters are being compared.

Fortunately, the model manifold $\mathcal{M}$ has a natural scale in relation to parameter space Θ, Λ. Insight into why this is an appropriate scale for λ_c can be gained by going back to the geometric interpretation of hard Δ perturbation spheres and ϵ tolerance ellipses. In the null model, a perturbation $|\delta\boldsymbol{\theta}| = \Delta$ in Θ corresponds to $\delta s = \sqrt{\Lambda}\Delta$. For the Δ perturbation sphere to be fully enclosed by the ϵ tolerance ellipse (a sphere in the null model), ϵ must be chosen such that $\epsilon \geq \sqrt{\Lambda}\Delta$. If $\epsilon < \sqrt{\Lambda}\Delta$, the ϵ tolerance ellipse is smaller than the Δ perturbation sphere for any parameter in the null model. Hence any parameter will be designated as *not* robust under this scheme, which is not desriable for the null model. A choice of $\epsilon = \sqrt{\Lambda}\Delta$, i.e. $\lambda_c = \Lambda$ would ensure that all parameters in the null model have a neutral robustness. Given this choice,

$$R_p(\boldsymbol{\theta}) = \prod_{i=1}^{D} \frac{1}{\sqrt{1 + \frac{\lambda_i}{\Lambda}}} \tag{3.13}$$

However if this definition of robustness is computed for the null model $\mathcal{M}^*$ with $\lambda_i = \Lambda$,

$$R_p^* = \frac{1}{2^{D/2}} \tag{3.14}$$

One would expect the measure of robustness for the null model to evaluate to unity, since the choice of $\lambda_c = \Lambda$ ensures neutral robustness in the null model, i.e. the Δ-sphere and ϵ-ellipse overlap exactly. The D-dependence is also somewhat troubling. The incongruity of R_p^* is due to the infinite extents of the Gaussian distributions and the agent removal process. Any agent that is perturbed ever so slightly has a finite probability of being removed, and there are always agents perturbed sufficiently far away that are definitely removed. Since there is always a loss of agents, R_p must be smaller than 1, even in the null model. A more appropriate measure of robustness would be to compare the survival rate of agents in a particular model to that of the null model. Hence one can define a new robustness measure $\rho_p(\boldsymbol{\theta})$:

$$\rho_p(\boldsymbol{\theta}) = \frac{R_p(\boldsymbol{\theta})}{R_p^*} = \prod_{i=1}^{D} \sqrt{\frac{2}{1 + \frac{\lambda_i}{\Lambda}}} \tag{3.15}$$

3.2 Parameter evolvability

The evolvability of a genotype g is the number of different phenotypes found in the one-neighbourhood of g (definition 2.6). While differences between phenotypes are discontinuous and easy to enumerate in discrete GP systems, the difference between behaviour in a continuous model is measured by the distance on the model manifold $\mathcal{M}$. How do we 'enumerate' the number of 'different' behaviours that are accessible in the mutational neighbourhood of $\boldsymbol{\theta}_0$?

The response of Sethna et al. to this challenge is to bypass this argument completely. They defined the evolvability of a 'chemotype' (parameter) to be the maximum change in 'fitness' averaged over a spherically distributed 'force' $\boldsymbol{F}$ in Z, where fitness is defined to be the inner product between the change in behaviour $\delta\boldsymbol{f}$ and $\boldsymbol{F}$

on $\mathcal{M}$ [3]. This is a rather unsatisfactory definition since it introduces the unnecessary complication of selection pressure ($\boldsymbol{F}$) and fitness which is intrinsically independent of a system's innate ability to discover new behaviour via parameter space fluctuations. Moreover, it is unclear how this relates to the diversity of behaviour in the neighbourhood of a parameter.

While the alternative definition proposed here does not follow Wagners definition to the letter, it respects the idea behind Wagner's definition. The evolvability of a parameter $\boldsymbol{\theta}_0$ should be a measure of the potential of a mutating agent at $\boldsymbol{\theta}_0$ to explore a 'diverse' set of outcomes. Since the difference between behaviours is measured by their distance on $\mathcal{M}$, the diversity of the mutational neighbourhood can be quantified by picking two random points on the Δ sphere and computing their mean squared distance $\langle \delta s^2 \rangle$ on $\mathcal{M}$ (see figure 3.1). If $\langle \delta s^2 \rangle$ is 'large', the behavioural outcomes in the mutational neighbourhood are well separated in $\mathcal{M}$; conversely if $\langle \delta s^2 \rangle$ is 'small', there is little diversity to be found in behaviour.

The first step in computing $\langle \delta s^2 \rangle$ is to project the Δ perturbation sphere in Θ onto $\mathcal{M}$. Adopting spherical coordinates in θ, the Δ perturbation sphere is parametised by angular coordinates $(\phi^1, \ldots, \phi^{D-1})$:

$$\begin{aligned}
\delta\theta^1 &= \Delta \cos(\phi^1) \\
\delta\theta^2 &= \Delta \sin(\phi^1)\cos(\phi^2) \\
\delta\theta^3 &= \Delta \sin(\phi^1)\sin(\phi^2)\cos(\phi^3) \\
&\vdots \\
\delta\theta^{D-1} &= \Delta \sin(\phi^1)\ldots\sin(\phi^{D-2})\cos(\phi^{D-1}) \\
\delta\theta^{D} &= \Delta \sin(\phi^1)\ldots\sin(\phi^{D-2})\sin(\phi^{D-1})
\end{aligned} \tag{3.16}$$

Using the normal coordinates defined in equation (3.3), the Δ sphere in $\hat{\theta}$ is

$$\begin{aligned}
\delta\hat{\theta}^1 &= \sqrt{\lambda_1}\Delta \cos(\phi^1) \\
\delta\hat{\theta}^2 &= \sqrt{\lambda_2}\Delta \sin(\phi^1)\cos(\phi^2) \\
&\vdots \quad \text{etc.}
\end{aligned} \tag{3.17}$$

Consider $\boldsymbol{\theta}_A = (\phi^1_A, \ldots, \phi^{D-1}_A)$ and $\boldsymbol{\theta}_B = (\phi^1_B, \ldots, \phi^{D-1}_B)$ on the Δ sphere. The distance between $\boldsymbol{\theta}_A$ and $\boldsymbol{\theta}_B$ in $\mathcal{M}$ is simply

$$\delta s^2_{AB} = (\delta\hat{\theta}^1_A - \delta\hat{\theta}^1_B)^2 + \ldots + (\delta\hat{\theta}^D_A - \delta\hat{\theta}^D_B)^2 \tag{3.18}$$

Averaging the expression over the solid angles $d\Omega_A$ and $d\Omega_B$ of $\boldsymbol{\theta}_A$ and $\boldsymbol{\theta}_B$ respectively,

$$\langle \delta s^2 \rangle = \int \frac{d\Omega_A}{\Omega}\frac{d\Omega_B}{\Omega}\delta s^2_{AB} \tag{3.19}$$

$$= \lambda_1\Delta^2 \int \frac{d\Omega_A}{\Omega}\frac{d\Omega_B}{\Omega}(\cos(\phi^1_A) - \cos(\phi^1_B))^2 \tag{3.20}$$

$$+ \lambda_2\Delta^2 \int \frac{d\Omega_A}{\Omega}\frac{d\Omega_B}{\Omega}(\sin(\phi^1_A)\cos(\phi^2_A) - \sin(\phi^1_B)\cos(\phi^2_B))^2$$

$$+ \quad \text{etc.}$$

While the integrals multiplying λ_i in the expression appear different, they in fact evaluate to the same value. $\langle \delta s^2 \rangle$ should be invariant to permutations of $(\delta\theta^1, \ldots, \delta\theta^D)$ on the left hand side of equations (3.16), since the permutation of $(\delta\theta^1, \ldots, \delta\theta^D)$ chosen in equations (3.16) was arbitrary in the first place (ignoring the handedness of the coordinate system). Such permutations would replace the integral multiplying λ_i with another one, yet the contribution of λ_i to the sum cannot change, so the two integrals that have exchanged places must evaluate to the same value. Thus after going through all such permutations for all λ_i, one is forced to conclude that all the integrals in this sum evaluate to the same value. Hence

$$\langle \delta s^2 \rangle = \mathcal{N} \Delta^2 \sum_{i=1}^{D} \lambda_i \tag{3.21}$$

where $\mathcal{N}$ is the value of the integrals. For the null model,

$$\langle \delta s^2 \rangle^* = \mathcal{N} \Delta^2 D \Lambda \tag{3.22}$$

$\langle \delta s^2 \rangle$ is a measure of behavioural diversity in the mutational neighbourhood and hence provides a measure of evolvability. Taking the null model as the baseline reference, the evolvability of a parameter is thus defined as

$$\eta_p = \frac{\langle \delta s^2 \rangle}{\langle \delta s^2 \rangle^*} = \frac{1}{D} \sum_{i=1}^{D} \frac{\lambda_i}{\Lambda} \tag{3.23}$$

If $\eta_p > 1$, it is more evolvable than the null model. Coincidentally, the definition proposed by Sethna et al. is actually proportional to $\sqrt{\eta_p}$; yet not only is this derivation closer to Wagner's intentions, it also does away with the need of appealing to selection pressure. It is noteworthy that the mean squared distance of a mutating agent subject to a random isotropic perturbation in Θ is also proportional to η_p. Hence η_p can not only be interpreted as a measure of diversity in a parameter's mutational neighbourhood, it can also be understood as a measure of the average change in behaviour in response to perturbations in Θ.

Chapter 4

Robustness and Evolvability of States

4.1 States

A straightforward phenotype analogue in continuous models is the behaviour $v = f(\boldsymbol{\theta})$. While a behaviour is strictly a point on $\mathcal{M}$, it is more useful to consider an infinitesimally small volume of behaviours in the neighbourhood of v. The set of behaviours enclosed by this volume is referred to as the state $\Psi(v)$ (a.k.a 'dynatype' in [3]). By relaxing a point into an infinitesimally small volume, definitions of robustness and evolvability could be more readily conceived for the phenotype analogue.

While Sethna et al. pictures the state (dynatype) as a D-dimensional hypersphere in $\mathcal{M}$, it is far more convenient to consider an infinitesimally small D-dimensonal hypercube with sides of length ϵ as measured in $\mathcal{M}$ (see figure 4.1). Thus, in this construction, $\Psi(v)$ is the set of points enclosed in the hypercube with v at the cube's centre. Each hypercube touches 2D other neighbouring hypercubes. The hypercube can be chosen to be oriented along the normal coordinates $\delta\hat{\theta}^i$ as defined in equation (3.3). Using equation (3.3), $\Psi(v)$ can be transformed into $\delta\theta^i$ coordinates and projected back to Θ. The projection $\tilde{\Psi}(\boldsymbol{\theta})$ is a hypercuboid with sides $|\delta\theta^i| = \frac{\epsilon}{\sqrt{\lambda_i}}$.

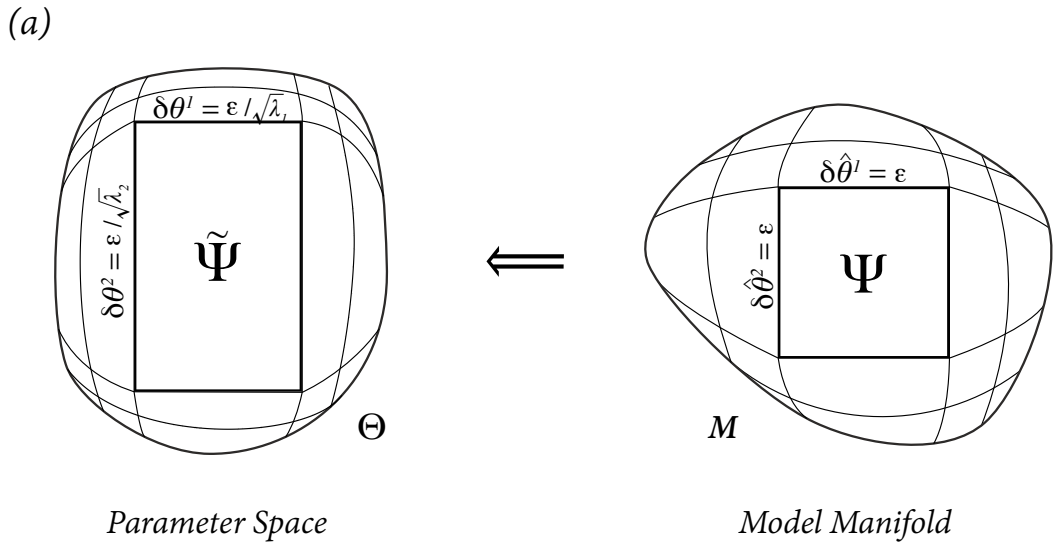
Remark 4.1. The process of relating $\Psi(v)$ to $\tilde{\Psi}(\boldsymbol{\theta})$ has implicitly assumed that the model f is injective on some subset of Θ i.e. f is locally identifiable (see remark 1.1). If f is locally identifiable but not globally identifiable, this one-to-one correspondence between parameters and states can only make sense by restricting Θ to a subset on which f is injective.

Remark 4.2. The volume of $\tilde{\Psi}(\boldsymbol{\theta})$ is

$$V = \epsilon^D \prod_{i=1}^{D} \frac{1}{\sqrt{\lambda_i}} \tag{4.1}$$

and the surface area is

$$A = 2\epsilon^{D-1} \prod_{i=1}^{D} \frac{1}{\sqrt{\lambda_i}} \sum_{i=1}^{D} \sqrt{\lambda_i} = 2\frac{V}{\epsilon} \sum_{i=1}^{D} \sqrt{\lambda_i} \tag{4.2}$$



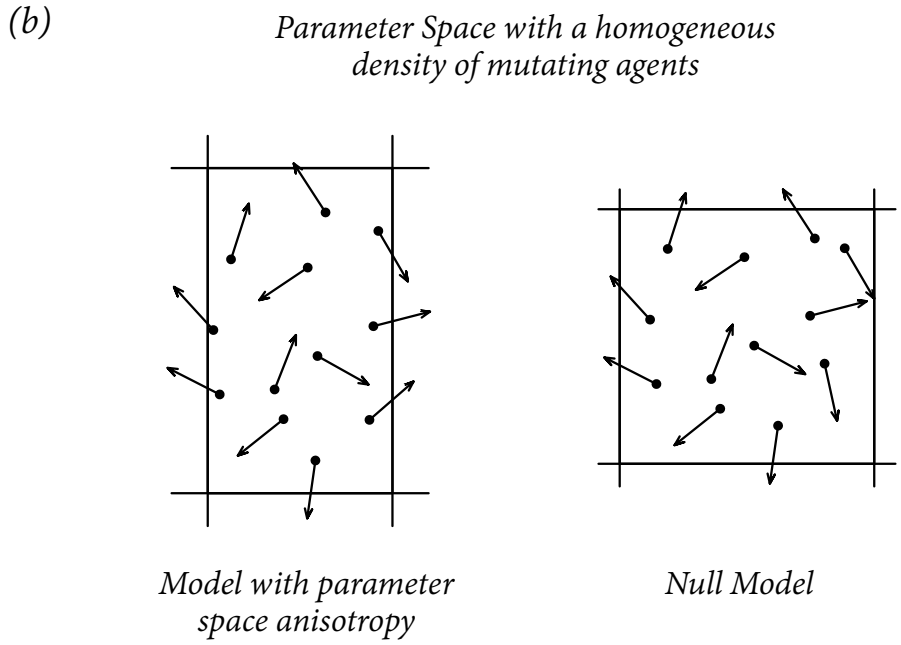


Figure 4.1: (a): State Ψ in $\mathcal{M}$ and its corresponding projection $\tilde{\Psi}$ in Θ. Θ and $\mathcal{M}$ is tiled by Ψ and $\tilde{\Psi}$ respectively. (b) Left: diffusion of mutating agents out of $\tilde{\Psi}$ in a model with parameter space anisotropy; Right: diffusion out of $\tilde{\Psi}^*$ in the null model. Observe how mutating agents leave each side of $\tilde{\Psi}^*$ with equal probability, yet they are biased towards leaving the longer sides of $\tilde{\Psi}$.

Remark 4.3. Since the null model is isotropic, Ψ^* with sides $|\delta\hat{\theta}^i| = \epsilon$ in $\mathcal{M}$ corresponds to a hypercube $\tilde{\Psi}^*$ in Θ with sides $|\delta\theta^i| = \epsilon/\sqrt{\Lambda}$. The volume of a null state is

$$V^* = \epsilon^D/\sqrt{\Lambda}^D \tag{4.3}$$

The surface area of a null state is

$$A^* = 2D\epsilon^{D-1}/\sqrt{\Lambda}^{D-1} \tag{4.4}$$

4.2 State Robustness

Sethna et al. did not attempt to define the robustness of a 'dynatype' [3]. Since the robustness and evolvability of a phenotype p describe the evolvability and robustness of a population of individuals with phenotype p, it is meaningful to consider the evolvability and robustness of a state Ψ in terms of the outcome of an *ensemble* of agents that are enclosed in $\tilde{\Psi}$ at a particular time. Consider a thought experiment in which Θ is filled with a homogeneous density of agents in brownian motion. These agents walk around Θ randomly and explore different states $\tilde{\Psi}$ as they do so. At a particular time t_0, the permeable boundary of $\tilde{\Psi}$ is suddenly made impermeable to agents exterior to $\tilde{\Psi}$. However, agents in $\tilde{\Psi}$ are allowed to cross the boundary and leave $\tilde{\Psi}$.

Conjecture: the robustness of the state is quantified by the length of time taken for the ensemble of agents to escape $\tilde{\Psi}(\boldsymbol{\theta})$. Such a time scale τ is

$$\tau = \frac{N}{\Phi} \tag{4.5}$$

where N is the number of agents in $\tilde{\Psi}(\boldsymbol{\theta})$ and Φ is the flux of agents out of $\tilde{\Psi}(\boldsymbol{\theta})$ averaged over time. Since Θ is filled initially with a constant density of agents, the number of agents trapped in $\tilde{\Psi}(\boldsymbol{\theta})$ is proportional to the volume V of $\tilde{\Psi}(\boldsymbol{\theta})$. If the total flux Φ across the surface of $\tilde{\Psi}(\boldsymbol{\theta})$ is assumed to be proportional to the surface area A,

$$\tau(\boldsymbol{\theta}) = \chi\frac{V}{A} = \chi\frac{\epsilon}{2\sum_{i=1}^{D}\sqrt{\lambda_i}} \tag{4.6}$$

where the factor χ accounts for the conditions such as the mobility and density of agents which is unrelated to the model's geometry. The time scale for a state in the null model is

$$\tau^* = \chi\frac{V^*}{A^*} = \chi\frac{\epsilon}{2D\sqrt{\Lambda}} \tag{4.7}$$

The robustness of a state Ψ is defined as the time scale of escape relative to that of the null model, therefore

$$\rho_\Psi(\boldsymbol{\theta}) = \frac{\tau(\boldsymbol{\theta})}{\tau^*} = \left(\frac{1}{D}\sum_{i=1}^{D}\sqrt{\frac{\lambda_i}{\Lambda}}\right)^{-1} \tag{4.8}$$

In loose terms, a state is in its most robust form if its surface area to volume ratio is minimised.

4.3 State Evolvability

Sethna et al. defined the evolvability of a 'dynatype' (state) as the optimum response within a population of agents at $\boldsymbol{\theta}$ to a force $\boldsymbol{F}$ in Z [3]. Just like their definition of 'chemotype' (parameter) evolvability, selection forces are involved for no good reason and it is not clear how it relates to Wagner's definition. Wagner quantifies the evolvability of a phenotype p by the number of unique phenotypes that are accessible by genotypes in the neutral set of p (definition 2.8) [10]. For states in continuous models, a D-dimensional box always has $2D$ number of faces, so the number of neighbouring boxes of $\tilde{\Psi}$ is always $2D$. If Wagner's definition for phenotypes is naively applied to states, all states would have the same evolvability, regardless of thir geometries. However this way of counting is clearly problematic. Consider an ensemble of agents diffusing out of a rectangle in figure 4.1 (b). Such agents are more likely to cross the longer sides of the rectangles. Hence most of the agents end up in the two states that share the longer side with $\tilde{\Psi}$. The ensemble of agents access an 'effective' number of neighbours that is closer to two than four.

The thought experiment in subsection 4.2 can be employed to formalise this argument. If the flux through the k^{th} face of the box, Φ_k, is proportional the face's area, A_k, the probability that any agent escapes through the k^{th} face of the box p_k is given by

$$p_k = \frac{\Phi_k}{\sum_k \Phi_k} = \frac{A_k}{\sum_k A_k} \tag{4.9}$$

The expected value of the surface area of a face crossed by any randomly chosen agent escaping $\Psi(\boldsymbol{\theta})$ is

$$\langle A \rangle = \sum_k p_k A_k = \frac{\sum_k A_k^2}{\sum_k A_k} \tag{4.10}$$

We define the effective number of faces, or effective number of neighbours, $\langle N \rangle$ to be the total area divided by $\langle A \rangle$

$$\langle N \rangle = \frac{\sum_k A}{\langle A \rangle} = \frac{\left(\sum_k A_k\right)^2}{\sum_k A_k^2} \tag{4.11}$$

In a null model, the state is a simple cube in Θ. With all faces equal in area the effective number of neighbours is simply $\langle N \rangle^* = 2D$. Defining the evolvability of a state η_Ψ as the effective number of neighbours relative to that of a null state,

$$\eta_\Psi = \frac{\langle N \rangle}{\langle N \rangle^*} = \frac{1}{2D} \frac{\left(\sum_k A_k\right)^2}{\sum_{k=1}^{2D} A_k^2} \tag{4.12}$$

This could be rewritten as

$$\eta_\Psi = \frac{(\sum_{k=1}^{2D} A_k/2D)^2}{\sum_{k=1}^{2D} A_k^2/2D} = \frac{\overline{A}^2}{\overline{A^2}} \tag{4.13}$$

The explicit expression of $\eta_\Psi(\boldsymbol{\theta})$ in terms of the eigenvalues of the metric tensor at $\boldsymbol{\theta}$ can be computed. Consider the face of the cube in the plane perpendicular to the eigenvector with eigenvalue $\sqrt{\lambda_i}$. The area of that face is

$$A_j = \epsilon^{D-1} \left(\prod_{i=1}^{D} \frac{1}{\sqrt{\lambda_i}} \right) / \left(\frac{1}{\sqrt{\lambda_i}} \right) = \sqrt{\lambda_j} V / \epsilon \tag{4.14}$$

Each cube has two such faces for each axes. Substituting (4.14) into equations (4.11) and (4.12),

$$\eta_\Psi(\boldsymbol{\theta}) = \frac{(\sum_{i=1}^{D} \sqrt{\lambda_i}/D)^2}{\sum_{i=1}^{D} \lambda_i / D} \tag{4.15}$$

Remarkably, ρ_Ψ, η_Ψ and η_p can be summarised into a rather concise relation

$$\rho_\Psi^2 = \frac{1}{\eta_\Psi \eta_p} \tag{4.16}$$

Remark 4.4. Loosely speaking, the two evolvabilities capture different aspects of the local geometry at $\boldsymbol{\theta}$: η_p is a measure of the local length scale and η_Ψ is a function of the angular distribution around $\boldsymbol{\theta}$. ρ_Ψ depends on both aspects of the geometry that are encapsulated in the two evolvabilities respectively.

4.4 Globally Non-identifiable but Locally Identifiable Models

It was noted in remark 4.1 that the geometric construction above relies on a one-to-one correspondence between $\tilde{\Psi}(\boldsymbol{\theta})$ and $\Psi(v)$ on a subset of Θ, i.e. the model f being *locally* identifiable. If the model f *globally* non-identifiable (see remark 1.1), a state Ψ on $\mathcal{M}$ corresponds to more than one point in Θ. Fortunately, restricting f to be locally identifiable, one can find disjoint neighbourhoods for each of these points. Consider the maximal set of points $S = \{\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_m\} \in \Theta$ that map to the same point v in $\mathcal{M}$ by a locally identifiable model f: $v = f(\boldsymbol{\theta}_1) = f(\boldsymbol{\theta}_2) = \cdots = f(\boldsymbol{\theta}_m)$. For each $\boldsymbol{\theta}_k \in S$, one can compute the eigenvalues of the metric $g_{\mu\nu}(\boldsymbol{\theta}_k)$. The state $\Psi(v)$ can be projected onto the individual disjoint neighbourhoods U_k of $\boldsymbol{\theta}_k$. Let the projections be $\{\tilde{\Psi}_k\}$. Casting this in the context of the thought experiment in subsection 4.2, the state Ψ now encloses populations of mutating agents in disjoint regions $\tilde{\Psi}_k \subset U_k$ of Θ. Recall that ρ_Ψ is simply a normalised escape time-scale of agents out of Ψ; averaging over all agents escaping out of every $\tilde{\Psi}_k$, the mean escape time $\overline{\tau}$ out of $\Psi(v)$ can be computed by

$$\frac{1}{\overline{\tau}} = \frac{\sum_k \Phi_k}{\sum_k N_k} = \sum_k \left(\frac{N_k}{\sum_k N_k}\right) \frac{\Phi_k}{N_k} = \sum_k \frac{\nu_k}{\tau_k} \tag{4.17}$$

where $\nu_k = N_k / \sum_k N_k = V_k / \sum_k V_k$ and $\tau_k = N_k / \Phi_k$ is the escape time-scale out of $\tilde{\Psi}_k$. Normalising the lifetime with τ^* (4.7), the robustness of the state $\Psi(v)$, $\rho_\Psi(v)$ is

$$\frac{1}{\rho_\Psi(v)} = \frac{\tau^*}{\overline{\tau}} = \sum_{\boldsymbol{\theta} \in S} \frac{\nu(\boldsymbol{\theta})}{\rho_\Psi(\boldsymbol{\theta})} \tag{4.18}$$

All the quantities on the right hand side can be computed explicitly in terms of the eigenvalues at $\boldsymbol{\theta}_k$ using equations (4.1) and (4.8).

Since the evolvability of a state is simply the normalised effective number of neighbours, one can simply sum over all the individual evolvabilities $\eta_\Psi(\boldsymbol{\theta_k})$ of $\tilde{\Psi}_k$ to obtain the normalised total number of effective neighbours for $\Psi(v)$:

$$\eta_\Psi(v) = \sum_{\boldsymbol{\theta} \in S} \eta_\Psi(\boldsymbol{\theta}) \tag{4.19}$$

Chapter 5

Sloppiness, Robustness and Evolvability

In their paper 'Sloppiness, Evolvability and Robustness in Systems Biology', Sethna et al. argued that sloppiness - parameter indeterminacy in certain dimensions of parameter space - induces 'neutral subspaces' in Θ. Such neutral subspaces allow mutating agents of similar behaviour to explore larger extents of parameter space and increase their likelihood of encountering new behaviour (section 2.3) [3]. This argument is rooted in Wagner's demonstration that larger neutral networks in genotype space (analogous to neutral subspaces in Θ) increases the evolvability of phenotypes [10]. If the argument proposed by Sethna et al. holds water, it should be reflected in the evolvability η_Ψ of states, the analogue of phenotypes in continuous models. If the eigenvalue spectrum of $g_{\mu\nu}(\boldsymbol{\theta})$ is sloppy, $\Psi(f(\boldsymbol{\theta}))$ should be an evolvable state (assuming global identifiability), i.e. $\eta_\Psi > 1$.

As it turns out, parameter space anisotropy is a necessary and sufficient condition for $\eta_\Psi < 1$. Consider

$$\eta_\Psi = \frac{\langle N \rangle}{\langle N \rangle^*} = \frac{\overline{A}^2}{\overline{A^2}} \qquad \text{(4.13 revisited)}$$

where A is the area of faces of $\tilde{\Psi}$. Recognising

$$\overline{A^2} - \overline{A}^2 = \sum_k (A_k - \overline{A})^2 \geq 0 \tag{5.1}$$

It is apparent that

$$\eta_\Psi = \frac{\langle N \rangle}{\langle N \rangle^*} \leq 1 \tag{5.2}$$

In other words, the effective number of neighbours of any state cannot be greater than that of the null model, where there is no parameter space anisotropy at all. More insight can be gained by applying η_Ψ to a phenomenological toy 'meta-model' of a sloppy eigenvalue spectrum. A sloppy system is characterised by a roughly even spread of eigenvalues over a logarithmic scale. It would be natural to consider the spectrum

$$\lambda_n = \lambda_0 e^{-2n\mu} \tag{5.3}$$

where $n = 0, \ldots, D-1$ and $2\mu = \log(\lambda_n) - \log(\lambda_{n+1})$ is the constant log-separation between eigenvalues. This model reduces a sloppy spectrum to two 'meta-parameters': the scaled leading order eigenvalue $\beta = \lambda_0/\Lambda$ and the log-separation μ. Substituting this into equation (4.15),

$$\eta_\Psi = \frac{1}{D}\frac{\tanh(D\mu/2)}{\tanh(\mu/2)} \tag{5.4}$$

η_Ψ is always smaller than $\eta_\Psi(\mu = 0) = 1$ and is a monotonically decreasing function of μ for $\mu > 0$. As the separation between the eigenvalues, μ, increases, the local parameter space becomes more anisotropic. This shows that, to first approximation, sloppiness decreases the evolability of a state. It seems that under a strict adherence to Wagner's original definitions, Sethna's claim that neutral subspaces allow individuals in it to reach a broader range of behavioural changes [3] [9] is thoroughly debunked.

This dissertation makes two conjectures as to why the neutral subspace argument fails. Firstly, the diversity of behaviour is limited by the hypercube geometry of Ψ in $\mathcal{M}$: it has a fixed number of neighbours, a fact that is independent of parametrisation. No amount of deformation in parameter space can increase that. In contrast, this limit is not imposed on phenotypes in discrete GP maps as the mutational relationships between phenotypes are established by the topology of the genotype network rather than any a priori measures of similarity between phenotypes. Hypothetically, a neutral set can grow in size to increase the number of accessible phenotypes up to the number of phenotypes allowed in the system. Secondly, sloppiness discourages individual agents within the neutral subspace to explore certain directions in Θ. As a result, mutating agents can only explore a low dimensional subspace of Θ and is unable to access a maximally diverse set of behaviours.

It is manifest from 4.15 that η_Ψ is invariant under rescaling of the eigenvalues:

$$\eta_\Psi = \frac{(\sum_{i=1}^{D}\sqrt{\lambda_i}/D)^2}{\sum_{i=1}^{D}\lambda_i/D} \tag{4.15 revisited}$$

Hence η_Ψ is only dependent on the sizes of the eigenvalues relative to each other. The invariance under rescaling is not true of η_p, ρ_p and ρ_Ψ; in their case the absolute value of λ (in units of Λ) matters. Hence one is cautioned against developing reasonings about η_p, ρ_p and ρ_Ψ on the basis of sloppiness only - sloppiness is merely a description of the relative sizes of eigenvalues, not their absolute magnitude. Sethna et al. noted a negative correlation between parameter evolvability and robustness in a sloppy model of an EGF/NGF in PC12 cells [3][1], which mirrors the negative correlation between genotype evolvability and robustness in RNA second structures in Wagner's findings [10]. Though Sethna et al. did not make an explicit connection between this phenomenon and sloppiness, it is worth emphasising that there is insufficient evidence to determine the significance of sloppiness in correlating these two quantities. If this correlation appears in other sloppy systems biology models, it is prudent to examine other possible reasons for its occurence rather than hastily attributing it to sloppiness.

[1]This correlation should hold under the definitions of parameter evolvability and robustness in this dissertation as they are only slightly different to Sethna's definitions.

Chapter 6

Summary and Outlook

Evolvability and robustness are desirable qualities in biological systems: favourable traits need to persist and resist random mutation, and a diverse set of behaviour accessible by mutation helps the system adapt to changes in circumstances. What is the role of sloppiness in influencing the evolvability and robustness of a system? This dissertation has demonstrated that parameter space anisotropy is the only factor in determining the evolvability of a state. In particular, increasing sloppiness decreases a state's evolvability. Yet sloppiness is not a sufficient condition to determine state robustness, parameter evolvability and paramter robustness. While sloppiness is a relevant to the discussion on adaptation, it cannot provide an elegant and unified account of evolvability and robustness in system biology models.

This dissertation has developed a set of measures of evolvability and robustness which can be applied to any continuous model which is locally identifiable. Local metric eigenvalues of models can be numerically computed efficiently using tools such as Sloppy Cells [7]. Since the measures developed are purely functions of metric eigenvalues (in fact, apart from ρ_p, all of them can be computed from the trace of the metric or metric square rooted), η_p vs ρ_p and η_Ψ vs ρ_Ψ correlations can be evaluated for continuous models. This provides an apparatus for further research to investigate whether systems biology models can be simultaneously evolvable and robust.

Bibliography

[1] Kevin S Brown, Colin C Hill, Guillermo A Calero, Christopher R Myers, Kelvin H Lee, James P Sethna, and Richard A Cerione. The statistical mechanics of complex signaling networks: nerve growth factor signaling. *Physical biology*, 1(3):184, 2004.

[2] Oana-Teodora Chis, Alejandro F Villaverde, Julio R Banga, and Eva Balsa-Canto. On the relationship between sloppiness and identifiability. *Mathematical Biosciences*, 282:147–161, 2016.

[3] Bryan C Daniels, Yan-Jiun Chen, James P Sethna, Ryan N Gutenkunst, and Christopher R Myers. Sloppiness, robustness, and evolvability in systems biology. *Current opinion in biotechnology*, 19(4):389–395, 2008.

[4] Emilie Dufresne, Heather A Harrington, and Dhruva V Raman. The geometry of sloppiness. *arXiv preprint arXiv:1608.05679*, 2016.

[5] Ryan N Gutenkunst, Joshua J Waterfall, Fergal P Casey, Kevin S Brown, Christopher R Myers, and James P Sethna. Universally sloppy parameter sensitivities in systems biology models. *PLoS Comput Biol*, 3(10):e189, 2007.

[6] Ard A Louis. Contingency, convergence and hyper-astronomical numbers in biological evolution. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 58:107–116, 2016.

[7] Christopher R Myers, Ryan N Gutenkunst, and James P Sethna. Python unleashed on systems biology. *Computing in Science & Engineering*, 9(3):34–37, 2007.

[8] Christian Tönsing, Jens Timmer, and Clemens Kreutz. Cause and cure of sloppiness in ordinary differential equation models. *Physical Review E*, 90(2):023303, 2014.

[9] Mark K Transtrum, Benjamin B Machta, Kevin S Brown, Bryan C Daniels, Christopher R Myers, and James P Sethna. Perspective: Sloppiness and emergent theories in physics, biology, and beyond. *The Journal of chemical physics*, 143(1):07B201_1, 2015.

[10] Andreas Wagner. Robustness and evolvability: a paradox resolved. *Proceedings of the Royal Society of London B: Biological Sciences*, 275(1630):91–100, 2008.

[11] Andrew White and Mark Transtrum. Limitations of model-based experimental design in systems biology. *Bulletin of the American Physical Society*, 59, 2014.