

Methods for Measuring Viability and Evaluating Viability Indicators

Matthew D. Egbert^{*,**}

Harvard University
University of Auckland

Juan Pérez-Mercader[†]

Harvard University
Santa Fe Institute

Abstract Life and other dissipative structures involve nonlinear dynamics that are not amenable to conventional analysis. Advances are being made in theory, modeling, and simulation techniques, but we do not have general principles for designing, controlling, stabilizing, or eliminating these systems. There is thus a need for tools that can transform high-level descriptions of these systems into useful guidance for their modification and design. In this article we introduce new methods for quantifying the *viability* of dissipative structures. We then present an information-theoretical approach for evaluating the quality of *viability indicators*, measurable quantities that covary with, and thus can be used to predict or influence, a system's viability.

Keywords

Viability, complex systems, analysis, computational tools, dissipative structures

1 Introduction

The second law of thermodynamics states that a closed system will tend toward disorder. Nevertheless, a wide variety of highly ordered *dissipative structures* are found in nature. These are systems that maintain their local ordered state by creating disorder elsewhere, or equivalently, by dissipating free energy [20]. They are, in this sense, *necessarily open* systems in that they rely upon their interaction with their environment to persist.

Bénard convection cells, vortices, and life are all frequently referenced as examples of dissipative structures. When conditions are right, these systems can persist for extended periods of time despite their intrinsic instability. But their precariousness draws attention to a number of interesting and potentially important questions that we can ask about their viability. What are the minimum conditions necessary to cultivate a colony of bacteria? How long is the Great Red Spot on Jupiter likely to persist? Is Dorothy growing more healthy as time passes, or is she getting increasingly ill? Even in the simpler non-biological cases, the high dimensionality and nonlinearity of these systems makes it difficult to describe or predict their dynamics. Despite advances being made in theory, modeling methods, and computational simulation techniques, we do not yet have general principles for describing or measuring the viability of these systems.

A common approach when investigating an entity's viability is to systematically vary a parameter of interest so as to evaluate its effect upon viability—see, for instance, investigations into the viability of

* Contact author.

** Department of Earth and Planetary Sciences, Harvard University; and Department of Computer Science, University of Auckland. mde@matthewegbert.com

† Department of Earth and Planetary Sciences, Harvard University; and Santa Fe Institute. jperezmercader@fas.harvard.edu

bacteria [13, 15], bee colonies [7], and biologically inspired artificial systems such as self-reproducing software [17]. This technique requires first establishing a way to quantify viability, and there are many ways to do so. Perhaps the simplest is the binary concept of *essentiality* (see, e.g., [14]), where the system of interest is directly observed as being either *alive* or *dead*, and ranges of parameter values are identified for which the system lives. This concept of viability has seen widespread use in the study of biomolecular networks, where proteins or genes are seen as either absolutely necessary for the survival of an organism (i.e., *essential*) or not—see [16] for a review of some of this work.

Of course, the reality is often more complicated than this. Although some proteins are essential for an organism's survival, other proteins, which are not essential for survival, may nevertheless help quite a bit. In other words, many of the factors that contribute to a system's viability do so in concert with other factors and in a way that is not all-or-nothing. Combinatorics make it difficult to evaluate interaction between these factors, and so, even when multiple studies occur on the same system, such as the response of *Escherichia coli* to limited availability of consumable energy [13] and temperature variation [15], the results are usually evaluated independently, despite the interaction that one might expect.

The temporally extended nature of viability further complicates its quantification. First, there is the question of how long you have to observe a system before you can say that it has or has not survived a particular event or condition. In many cases the answer to this question comes from informal intuitions that emerge from a scientist's experience with the system being studied. Second, distinguishing the set of initial conditions where the system will survive from those where the system will die is often difficult, even in mathematical models or simulations, where complete knowledge of the system's state is available and/or where there exist equations that explicitly describe the system's instantaneous dynamics. As a case in point, we can consider a deterministic dynamical system, where every state implies a single trajectory and thus a single final outcome. If one knows the regions in state space that are associated with the system's death, it is theoretically possible to categorize any single state as being either *viable* (it avoids those regions associated with death) or *nonviable* (it does not). But delimiting these sets of initial conditions and the *viability boundary* between them is equivalent to identifying the boundaries of the basins of attraction, which in almost all cases is analytically intractable [24]. In these cases, one can use heuristic-based sampling of initial conditions, where for each sampled initial condition, an initial-value problem is solved to determine whether that state results in survival or death; but when the model is numerically stiff or involves a high-dimensional state space, this kind of initial-condition survey can be computationally expensive to the point of being impossible. In more forgiving systems it is possible to perform this kind of analysis, and a number of investigations have done so, evaluating the *viability space* of modeled protocells or other precarious systems. See for instance the investigation of artificial-chemistry-based metabolisms [19], simulated metabolism-boundary co-construction systems [1, 2], coupled metabolism-behavior systems [4, 11], and dissipative reaction-diffusion structures [28]. One particularly thorough investigation along these lines examines the viability space of dynamic structures called *gliders* observed in the discrete cellular automaton of Conway's Game of Life [5, 6].

There have also been a number of studies of viability in nondeterministic systems. For example, [21, 22] and [23] have investigated the stochastic dynamics of simulated protocells and how these relate to the system's viability. The nondeterminism of these systems means that statistical rather than categorical measures of viability are more appropriate. Empirical studies also often report on statistical rather than (or in addition to) categorical effects of various influences upon viability. For example, [13] and [15] report on the proportion of bacteria that survive various environmental conditions. The statistical approach means that instead of being simply classified as either viable or nonviable, each state is associated with a distribution of possible viable and nonviable final states. By comparing these distributions, it is possible to rank states in terms of their viability, and this opens up a number of interesting lines of inquiry. For instance, it becomes possible to evaluate how external perturbations impact viability, even if they are not so strong as to cause the system to become nonviable. This means that instead of being limited to describing proteins as either essential or non-essential, it becomes possible to compare their relative contributions, even if they

are both non-essential. It also becomes possible to evaluate the effect of *intrinsic* dynamics upon viability—for example, to quantify the extent to which the movement of a vortex or the unperturbed behavior or metabolic dynamics of an organism helps or hinders its persistence—see [3, 4] for examples of this kind of study and further discussion. Finally, it also becomes possible to evaluate viability *dynamics*. For instance, one can compare the viability of the current state of a system with that of an earlier state to evaluate whether the system is becoming more or less viable.

We have just described some benefits of a quantitative rather than categorical measure of viability. While statistical approaches can provide a quantitative measure, it is often difficult to acquire the necessary data. This article presents a new method for generating a model of viability when the data available are more limited. In its simplest form, the method requires only the results of a number of initial-condition experiments, where associations are made between initial conditions and their *viability class*, that is, whether each results in survival or death. We present a method for extrapolating from such data so as to be able to make a quantitative prediction of the viability of other states. In subsequent steps we explain how additional data, such as knowledge of the system’s dynamical tendencies, can be used to produce additional, complementary measures of viability. The techniques we propose make it possible to (i) predict which of a set of system configurations are more or less likely to persist; (ii) predict which conditions must be maintained for a given system to persist; (iii) predict the typical impact of a given perturbation upon the system’s viability; (iv) identify methods for influencing system viability; and (v) identify and evaluate the quality of *viability indicators*, defined as measurable quantities that covary with system viability and that can thus be used to predict and in some cases influence viability.

In the next section, we present the new measures. We see these new techniques as first steps towards a general method for generating the best possible model of viability from the various data that are available. As additional data are acquired, the model can be iteratively refined, but even an incomplete model can still provide information that is often useful in (i) understanding the target system’s viability and (ii) guiding future data acquisition in the investigation of the target system’s viability. It should be pointed out that although we began our introduction with a discussion of the relationship between organisms and dissipative structures, the methods presented are independent of this view, and likewise of other theories about what life is or how it works (maximum entropy production, Friston’s free-energy principle, autopoiesis, etc.). All that is required to apply the basic presented method is the ability to associate a variety of initial conditions with whether the system survives or dies.

One final clarification must be made before we present our methods for evaluating viability. We opened this section by drawing a comparison between dissipative structures and life. In a closed system, the presence of a dissipative structure accelerates approach to thermodynamic equilibrium, where dissipative structures cannot persist. It might thus seem paradoxical to associate life and viability with a structure that essentially accelerates its own demise. This apparent paradox disappears, however, once one recognizes that dissipative structures can persist indefinitely when they are part of an open system that provides a steady source of free energy. As a case in point, life involves highly ordered structures that according to the second law of thermodynamics are intrinsically unstable, yet has managed to persist on Earth for approximately four billion years thanks to the energetic input of the Sun, geothermal radioactive decay, and so on. Thus life and dissipative structures share in common a precarious form of existence, where both are essentially constantly “dying” but, when the conditions are right, can avoid (or at least delay) death. This common organization makes dissipative structures interesting models for living systems and their viability.

2 Method

2.1 Example System

To illustrate the new methods for quantifying viability, we will use a minimal and abstract example dynamical model that is loosely inspired by the idea of a vesicle-based bioreactor, or *protocell* [18].

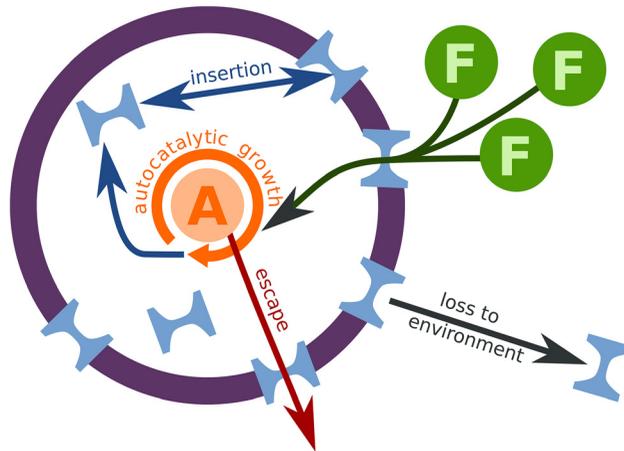


Figure 1. Schematic diagram of the example model. This figure indicates the processes that govern the dynamics of the concentration of the compartmentalized metabolism (A) and the bioreactor’s membrane permeability (X).

The model (see Figure 1) has two state variables: X , representing the permeability of the bioreactor’s membrane, and A , representing the intravesicular concentration (micromolar) of a metabolic network. The metabolism requires resources from outside the membrane to grow. It also produces transmembrane proteins that can become embedded in the vesicle membrane and increase its permeability. There is, therefore, a nonlinear feedback relationship between these variables, which can be described by the following abstraction:

$$\frac{dX}{dt} = \underbrace{2(0.2A - X)}_{\text{insertion}} - \underbrace{X/4}_{\text{loss to environment}}, \tag{1}$$

$$\frac{dA}{dt} = \underbrace{10A\sigma_{0.4}(X)}_{\text{autocatalytic growth}} - \underbrace{25A\sigma_{0.5}(X)}_{\text{escape}} - \underbrace{4A}_{\text{degradation}}, \tag{2}$$

$$\sigma_{x^*}(x) = 1 - \left(1 + e^{100(x-x^*)}\right)^{-1}, \tag{3}$$

where X and A represent the variables mentioned above, and each term in these equations describes a process that contributes to the evolution of the vesicles and its contents. The first term in Equation 1 represents the insertion of transmembrane proteins from the interior of the vesicle into the membrane. The second term represents loss of these particles from the membrane to the extravesicular solution, where they have no further influence on the system. The first term in Equation 2 represents the autocatalytic growth of the metabolic network. To derive this term, we approximate the autocatalytic growth of the network as a single autocatalytic reaction, $A + F \rightarrow 2A$, where F is a precursor *food* molecule, and we assume this autocatalysis to occur much more quickly than the diffusion of F across the membrane. This allows us to approximate the concentration of F as being proportional to its rate of diffusion across the membrane, which is given by $\sigma_{0.4}(X)$. The second term in this equation represents the direct escape of the metabolic-network constituents through the membrane, which occurs at a rate proportional to the concentration of the network

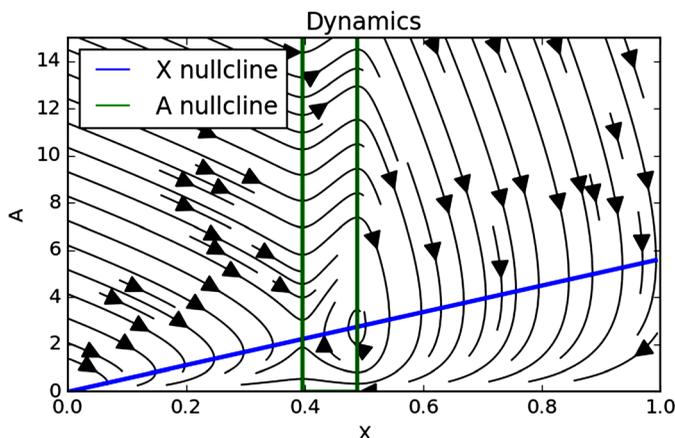


Figure 2. Flow and nullclines of the minimal example model.

and to a function of the permeability of the membrane, $\sigma_{0.5}(X)$. In these equations, the parameterized sigmoidal function σ_{X^*} (Equation 3) represents a nonlinear threshold behavior where above a certain permeability (i.e., for $X \gg X^*$) the reactant crosses the membrane. The final term in Equation 2 represents the degradation of A into products that have no further influence on the system's dynamics. The coefficients were chosen to represent a situation where the bioreactor would function given certain initial conditions, but fail given others.

Figure 2 indicates the nullclines and dynamic flow of this model. In it, we can see that for certain initial conditions, the system reaches a steady state where $(A, X) \approx (5, 0.5)$, and that there is an unstable equilibrium at $(A, X) \approx (4, 0.4)$. For all of the other initial conditions in the domain $X \in [0, 1]$, $A \in [0, 15]$, we have $\lim_{t \rightarrow \infty} A = 0$, that is, the concentration of autocatalyst, the membrane permeability, and the rate of production of autocatalyst all go to 0, so that the bioreactor ceases to be active.

2.2 Quantifying Viability

2.2.1 Viability Class

The most simple description of viability that we discuss is the viability class. This quantification of viability, already introduced above, describes long-term survival, distinguishing between initial conditions for which the system of interest persists for an extended period of time and those for which it does not. The former belong to the set of *viable* initial conditions (V^+), and the latter belong to its complement, the set of *nonviable* initial conditions (V^-). We can write the viability class as a function that maps initial conditions \mathbf{x} into either V^+ or V^- ; thus: $S(\mathbf{x}) \mapsto \{V^+, V^-\}$.

To evaluate $S(\mathbf{x})$ in our example model, we first observe that when there is no autocatalyst, the system is dead, and we use this to define the collection of dead states, $\omega = \{(X, A) : A = 0\}$. The initial conditions that approach ω belong to V^- , and the initial conditions that (at least for some finite but distant future) do not enter ω belong to V^+ . Figure 3 provides a visualization of V^- and V^+ in $X \in (0, 1)$, $A \in (0, 15)$, generated by numerically integrating 2^{15} randomly selected initial conditions for a long period of time and assigning it to V^- if A approaches 0, and V^+ otherwise. In this simple model, it is possible to thoroughly sample the viability space, but when evaluating the viability class for a given initial condition is more costly (e.g., in experimental scenarios or in more computationally expensive simulations), interpolation and extrapolation methods could be used to make informed predictions about the viability region of untested initial conditions. Similarly, when the viability of an initial condition is nondeterministic, additional methods or approximations must be employed to evaluate $S(\mathbf{x})$, but this is left as future work.

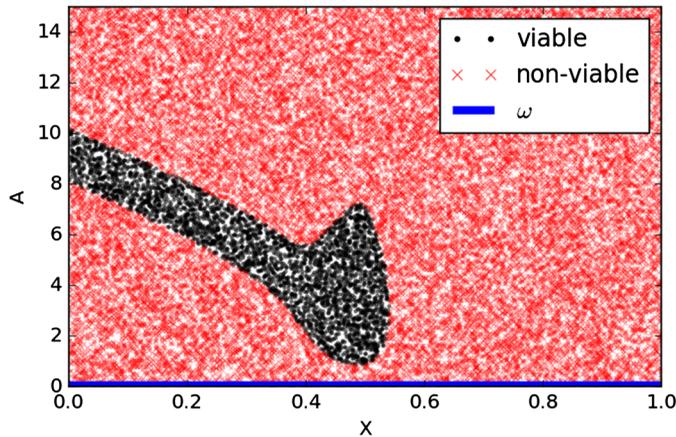


Figure 3. Viable (V^+), nonviable (V^-), and dead (ω) regions for the minimal example system. Plotted points indicate randomly selected initial conditions and their associated survival value $S(\mathbf{x})$. The blue line across the bottom of the plot indicates ω , those states for which the system of interest does not exist.

2.2.2 Vulnerability

The classification of $S(\mathbf{x})$ is a useful, but limited, description of viability in that it only provides information about which side of the viability limits the system is on. Essentially, states are categorized as either “good” (viable) or “bad,” (nonviable), and there is no way to compare the viability of two states in the same viability class. With $S(\mathbf{x})$ alone, it is impossible to determine if the viability is getting worse or better, or if a perturbation was beneficial or detrimental (except in the specific and perhaps not very useful case of the system crossing the viability boundary from V^+ into V^- or vice versa). Put another way, $S(\mathbf{x})$ does not provide any information about the distance from or relative velocity with respect to the viability boundary.

As a first approximation, the farther away a V^+ state is from V^- , the more robust it is, as a larger perturbation is required for it to become nonviable. We can use this property as the basis of a metric that allows us to compare states that belong to the same viability class. However, the dimensions of the viability space can be in entirely different units (e.g., concentration and temperature) and/or have different scales (do we measure A in μM or nM ?), making the measurement of distances nontrivial. To resolve this, we can use the shape of V^+ to normalize the viability space (see Figure 4). This is accomplished

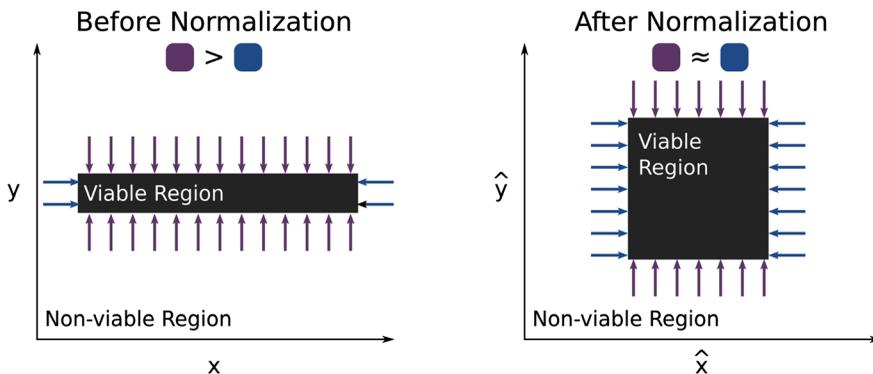


Figure 4. Normalizing viability space. Before normalization (left), a perturbation in the y dimension is more likely to cross the viability boundary than a perturbation of the same magnitude in the x dimension. After rescaling (right), the likelihood of a perturbation to cross the viability boundary is determined by the shape of V^+ , not the choice of units used to describe the space.

by calculating the extent to which the viability interface is orthogonal to each dimension and then scaling the values in that dimension by this amount. More formally, for each dimension of the viability space X we identify I_X , the average magnitude of the viability-interface surface normals in that dimension:

$$I_X = \frac{\iint_I \|\hat{\mathbf{n}} \cdot \hat{\mathbf{e}}_X\| dI}{I}, \quad (4)$$

and use this value to rescale values in that dimension into normalized units, thus: $\hat{X} = X \cdot I_X$. In Equation 4, $\hat{\mathbf{e}}_X$ is the basis unit vector for the dimension X , and $\hat{\mathbf{n}}$ is the surface normal to I , the viability interface.

This normalization method allows us to remove the effects of the arbitrary scales and units used to describe the system, and thus to evaluate the variables in terms of their effect upon viability. To demonstrate this, we evaluated 10^6 small random perturbations in the example model. These were generated by selecting a random initial condition and a random displacement vector, $(\Delta X, \Delta A) = (m \cdot \cos(\alpha), m \cdot \sin(\alpha))$, where $\alpha \in [0, 2\pi)$ and $m = 0.01$. When the viability class of the initial condition is different from the displaced initial condition $(X + \Delta X, A + \Delta A)$, we note the angle of the vector (α). The distribution of angles is plotted for both the non-normalized and the normalized systems in Figure 5. When the perturbations are in non-normalized space (left), the viable region appears much taller than it is wide and a perturbation with a strong X component is therefore more likely to cross the viability interface than a perturbation of equal magnitude but with a strong A component. After normalization (right), an asymmetrical distribution of angles remains, but this is due to the shape of the viable region, which involves pronounced faces that are roughly parallel to $\hat{A} = -\hat{X}$ (see Figure 6), and corresponds with the higher proportion of perturbations at $\approx 45^\circ$ and $\approx 225^\circ$ in the right plot of Figure 5). Instead of being distracted by artefacts imposed by the choice of variable scales, the normalized coordinates allow us to see that the perturbations that are most likely to cross the viability threshold are those that are perpendicular to that face, that is, perturbations where $\Delta \hat{X} \approx \Delta \hat{A}$.

After normalizing viability space, we can define ρ as the minimum distance from a given state to the viability interface in the rescaled coordinates. For points in V^- , the smaller this value, the better, but for points in V^+ , the reverse is true. Taking this into account, we can define our second viability metric, *vulnerability*, as a continuous scalar field over the state space:

$$\gamma(\mathbf{x}) \equiv \begin{cases} -\rho & : \mathbf{x} \in V^+, \\ +\rho & : \mathbf{x} \in V^-, \end{cases} \quad (5)$$

where $\gamma(\mathbf{x}) = 0$ at the viability interface, and more negative values of $\gamma(\mathbf{x})$ correspond with states that are more likely to persist.

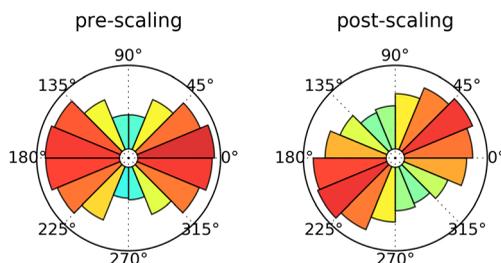


Figure 5. Distribution of viability-interface-crossing perturbations for different perturbation angles in the example system. Monte Carlo methods were used to randomly sample small perturbations of equivalent magnitude in the non-normalized and normalized viability spaces. Before normalization, the difference in scales between the axes suggests that perturbations in X are the most likely to cause the system to cross the viability boundary, an artefact of the choice of scales used to describe the system. After normalization, the artefact disappears and we see results that arise from the shape of V^+ , specifically that perturbations with angles of 45° or 225° (i.e., where $\Delta \hat{A} \approx \Delta \hat{X}$) are the most likely to cause the system to cross the viability boundary.

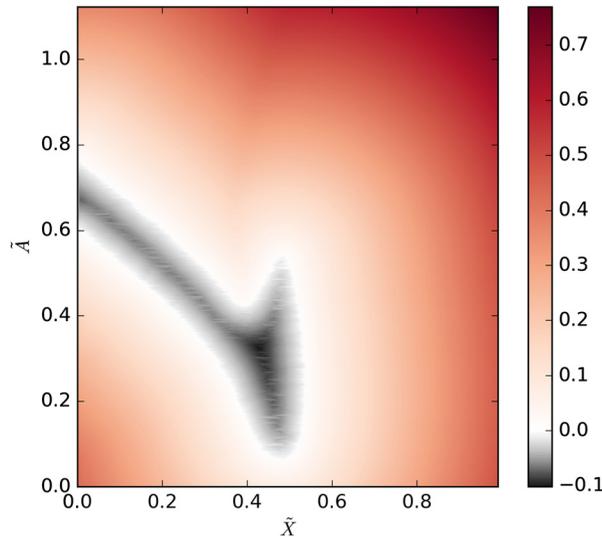


Figure 6. Vulnerability, that is, signed minimum viability-interface distance in normalized viability space. Red values indicate distances from V^- states to the nearest known V^+ state. Black values indicate the negative of the distance from V^+ states to the nearest known V^- state. According to this metric, the closer to $-\infty$ the indicated value, the more viable it is.

The vulnerability measure corresponds with the intuition that the farther away a viable state is from the viability boundary, the more viable it is, as a larger perturbation would be necessary to cause the system to die. The measure comes with some assumptions, in particular that the distribution of likely perturbations is independent of the system’s state and evenly distributed in all directions. If data become available describing the distribution of perturbations, they could be used to augment the model, but in the absence of such data, our measure enables the comparison of two states in terms of their relative viability. For example, if we examine Figure 6, which presents this scalar field for the bioreactor model now plotted in rescaled coordinates, where $\hat{X} = 0.995 \cdot X$ and $\hat{A} = 0.0462 \cdot A$, then we can see that a state of $(\hat{X}, \hat{A}) = (0.6, 1.0)$ is more viable than $(0.9, 1.0)$, but not as viable as $(0.6, 0.25)$, even though all three of these initial conditions fall into V^- . This figure was generated by calculating, for each of the 2^{15} initial conditions used to generate Figure 3, the distance in normalized viability space to the nearest initial condition to an initial condition of the other viability class (V^+/V^-). These values were then used to generate an interpolated rectangular lattice mesh.

2.2.3 Vulnerability Velocity

Evaluating the distance from the viability interface allows for the comparison of two states, but it does not take into account the inherent system dynamics. For example, certain initial conditions may start deep inside the viable region, but after a short period of time move quite close to the viability limits, where a small perturbation could make the system nonviable. By combining our knowledge of how the state of the system changes over time ($\frac{d\mathbf{x}}{dt}$) with how vulnerability changes with changes in viability space, ($\frac{\partial\gamma(\mathbf{x})}{\partial\mathbf{x}}$), it is possible to calculate a third quantification of viability: the instantaneous *vulnerability velocity*, that is, the rate and direction at which the vulnerability is changing for a particular state:

$$\frac{d\gamma(\mathbf{x})}{dt} = \frac{\partial\gamma(\mathbf{x})}{\partial\mathbf{x}} \cdot \frac{d\mathbf{x}}{dt}. \quad (6)$$

Figure 7 was generated by convolving the data plotted in Figure 6 with a Gaussian filter to remove noise due to the sampling process, and then calculating in a regular lattice the value of $\frac{d\gamma(\mathbf{x})}{dt}$ as described in Equation 6. We can see in this figure how the natural dynamics of the system

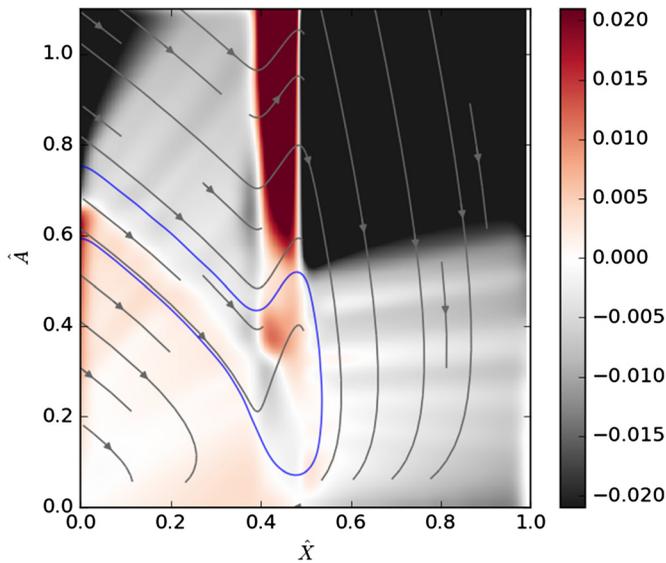


Figure 7. Instantaneous vulnerability velocity in the example system. A positive velocity indicates motion that is increasing in vulnerability. Dynamical flow and the viability boundary are approximated in gray and blue, respectively.

become more or less viable over time. As an example, the red area of V^+ states around $(\hat{X}, \hat{A}) \approx (0.45, 0.4)$ is viable, but moving closer to the viability interface (becoming increasingly vulnerable). It is currently less likely to be destroyed by perturbation than it will be in the near future. The opposite is the case for the gray regions within V^+ . To preempt any confusion: Some readers might be surprised to see nonzero vulnerability velocities found at the viability boundary. This boundary is the separatrix between the V^+ and V^- basins of attraction, and so no trajectories cross this manifold. Nevertheless, it is possible for states at the boundary to have nonzero *instantaneous* vulnerability velocities. An example of this is found in the trajectories near the top viability boundary of the “stem” of V^+ . In this area trajectories in V^+ and V^- are both becoming less vulnerable, but subsequent dynamics mean that those above the boundary never reach V^+ . This observation highlights the latent nature of the viability boundary in the sense that it need not be co-located with any local dynamical structure. Put another way, there may be no obvious indication in a system’s dynamics that it is approaching the viability boundary. We can see in this example that the vulnerability velocity only captures the current instantaneous change in vulnerability. We will now look at one measure that takes into account the long-term change in system vulnerability.

2.2.4 Minimum Future Vulnerability

Figure 8 shows the *minimum future vulnerability* for the example system. These values were calculated by solving initial-value problems for 2^{15} randomly selected initial conditions, $\mathbf{x} \in X \in [0, 1]$, $\mathcal{A} \in [0, 15]$, and identifying the minimum value of $\gamma(\mathbf{x})$ encountered by each trajectory. When the system state is not inside the interpolated field, as occurs, for instance, soon after the system starts with initial conditions $X = 1$, $\mathcal{A} = 15$ (see Figure 2), the vulnerability is assumed to be higher than any encountered, and these states have no influence upon the minimum future vulnerability. This measure is an example of how the viability metrics mentioned above could be extended as additional information is acquired concerning the system’s dynamics to address specific questions concerning a system’s viability, and how one could maximally influence it. The minimum future vulnerability indicates the minimum perturbation in the system’s future that would cross the viability boundary. This information could be useful in planning, for example, an intervention to bring a dying protocell back into V^+ , or vice versa, to identify when in the future the smallest perturbation would transform a viable system into one that is nonviable.

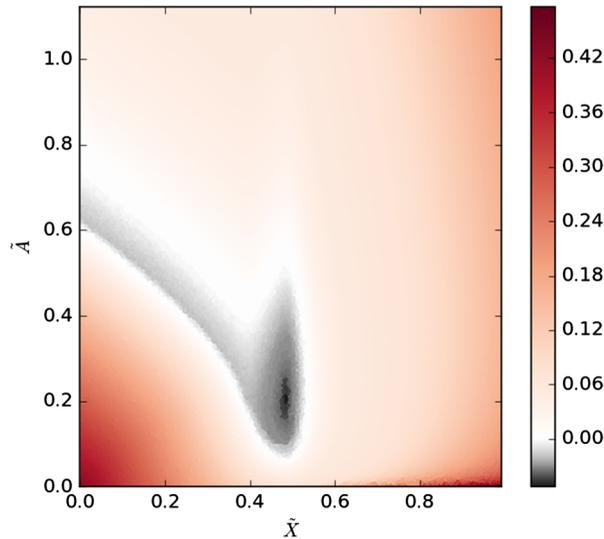


Figure 8. Minimum future vulnerability. The colors here indicate the minimum future value of $\gamma(\mathbf{x})$ encountered for various initial conditions.

The minimum future vulnerability captures the long-term dynamics, but in a nondeterministic system, noise or externally driven events or perturbations limit the value of this metric. In fact, none of the four proposed viability metrics capture *all* of the relevant aspects of system viability, and some combination is necessary if the goal is a complete quantification of viability. It is also interesting to note that none of these viability metrics are quantities that can be directly measured. They are representations of how a system state compares to other states of the same system.

2.3 Evaluating Viability Indicators

In general, no single measurable state variable will fully capture a system’s viability. By calculating the mutual information shared by candidate viability indicators and the viability measures described above, one can establish which of the candidate variables provide more information about a system’s viability.

Figure 9 indicates the mutual information between candidate variables \hat{X} and \hat{A} and each of the four presented measures of viability. For most of the measures, \hat{X} is a better indicator of viability,

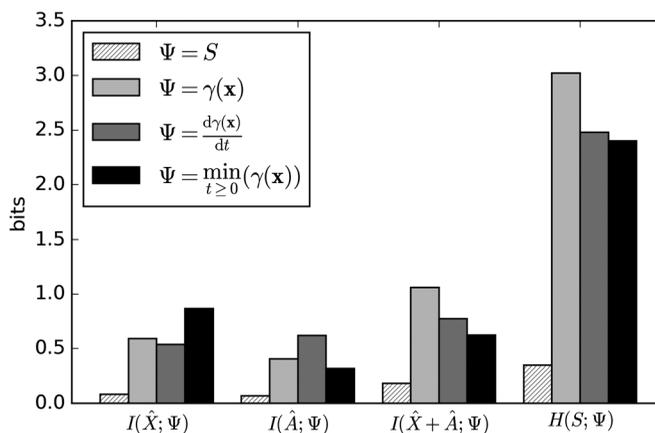


Figure 9. Mutual information shared between candidate variables and viability metrics, and the entropy (H) of each metric.

as it provides more information than \hat{A} about all of the viability metrics. It is noteworthy that one variable (\hat{X}) provides better information about viability class, vulnerability, and minimum future vulnerability, while \hat{A} is better at reducing uncertainty in the vulnerability velocity.

Similar methods can be used to evaluate linear or nonlinear combinations of variables to identify improved viability indicators. For instance, in the example system a simple linear combination of \hat{A} and \hat{X} provides more information about the protocell's viability than either variable alone.

It is important to recognize that these mutual-information statistics are influenced by the distributions of variables studied. Here we have assumed even distributions across the studied portion of viability space, but, as with the measures presented above, as additional data about the system and its environment are acquired, such as states the system tends to be in, the information measures would be improved by taking this information into account.

3 Discussion

We have introduced four measures for quantifying a system's viability. The viability class, $S(\mathbf{x})$, indicates whether or not a given initial state will survive in the absence of noise or perturbation. The vulnerability, $\gamma(\mathbf{x})$, indicates a signed minimum distance to the viability boundary, where increasingly negative values are considered to have a greater chance of survival. The main assumption underlying this measure is that the closer V^+ states are to the viability boundary, the less viable they are, as it takes less of a perturbation to push them into V^- . The vulnerability velocity, $\frac{d\gamma(\mathbf{x})}{dt}$, provides an indication of the effects of the system's intrinsic dynamics upon vulnerability, and finally, the minimum future vulnerability, $\min_{t>0} \gamma(\mathbf{x})$, identifies the magnitude of the minimum future perturbation that will cause the system to cross the viability boundary.

The newly presented viability measures depend upon a normalization of viability space to avoid artefacts introduced by the arbitrary selection of units for different variables. We described one way to do so, where the shape of V^+ , specifically the extent to which it faces each essential variable, is used to normalize viability space. This makes possible a more meaningful measurement of distances in viability space and, more broadly, the comparison of different variables in terms of their influence upon viability. This normalization of variables with respect to their effect upon viability could further be useful in identifying which environmental conditions and variations are the most threatening to a system's viability.

We then presented a simple application of information-theoretical analysis that quantifies the covariance of state variables with the measures of viability. In the minimal model, the information-theoretical analysis showed that \hat{X} was more strongly correlated with three of the four investigated viability metrics, and can thus be thought of as a better viability indicator than \hat{A} . In other words, if you had to gamble on whether or not the bioreactor was going to persist, it would be better to have information about \hat{X} than about \hat{A} . For similar reasons, if this had been a model of a real synthetic protocell or bioreactor, and we were looking for feedback-based mechanisms for increasing its robustness, it would be more useful to have these mechanisms respond to the state of \hat{X} than to the state of \hat{A} . That said, \hat{A} is a better indicator of vulnerability velocity than \hat{X} . How to prioritize these metrics in this kind of contradictory case depends upon factors specific to the problem at hand. We also gave a brief example of how the quality of linear or nonlinear combinations of variables could be evaluated in terms of their ability to predict a system's viability.

We deliberately chose a simple, two-dimensional model to demonstrate our viability measures, so as to allow visualization. Increasing the dimensionality of the model presents no direct challenge to applying the presented measures, but there are other factors that might make the measures more difficult to apply in more complex systems. In particular, the toy model used in this article is deterministic, with no hidden or stochastic variables. This means that each initial condition clearly lies either in V^+ or in V^- . If the data used to generate the initial map of the viable region come from experimental data or a stochastic model, it is likely that the line between V^+ and V^- will not be so distinct. This would complicate the analysis, as it would no longer be so straightforward to identify the viability boundary, nor measures of distance to that boundary. This problem could be overcome

by using thresholding and a convex-hull algorithm to approximate the viability boundary. In future work we will evaluate the utility of our measures in more complex (higher-dimensional) systems and in systems for which the initial V^+ and V^- data come from systematic experimentation with a real-world system.

Viability indicators are also worth considering in the context of *viability-based* behaviors where an organism's behavior is a response to one of its essential variables (see [10, 12]). The aerotaxis of *E. coli* is driven by a response to the state of the organism's electron-transport system [9]. Presumably, the ETS provides an excellent indicator of the state of the organism's metabolism and thus health. Has evolution selected this viability indicator because it is the best available indicator of the organism's health, or are there other state variables that more strongly covary with the system's viability?

In this context it is also worth repeating that the viability boundary is not necessarily co-located with any dynamical structure. In the example system, states just above and below the “stem” of V^+ have almost identical derivatives, making it very difficult or impossible to evaluate viability from the current instantaneous change in the model's variables. The latent nature of viability means that behavioral mechanisms cannot respond directly to viability, but must instead respond to state variables (or combinations thereof) that correlate with, and thus can be used to predict and/or influence, viability—that is, viability indicators.

Developing a better understanding of viability dynamics could help us to evaluate the impact of design decisions or interventions. It also could help advance theory pertaining to the adaptive behavior of living systems. In particular, the enactive approach to understanding life and mind [25, 26] is built largely upon the concept of viability. In this framework, living systems are seen as precarious entities (comparable to dissipative structures) that manage to remain viable thanks to processes of self-maintenance, and it is this organizational property rather than spatial contiguity that distinguishes them as autonomous entities [27]. Building on this framework, adaptivity has been defined in terms of the transformation of trajectories that tend to approach the viability boundary (i.e., become more at risk of dying) into trajectories that do not [8]. But as discussed above, measuring distance and velocities in viability space is nontrivial. The method proposed herein for normalizing viability space may be a way to overcome this challenge. This is one example of how developing measures of viability allows us to make these kinds of theoretical definitions more formal, improving our ability to connect them with our experiments on natural and artificial systems.

Acknowledgments

Thanks are due to Eran Agmon, James Dyke, Nathaniel Virgo, and Inman Harvey for their useful comments on various versions of this manuscript.

References

1. Agmon, E., Gates, A., Churavy, V., & Beer, R. (2014). Quantifying robustness in a spatial model of metabolism-boundary co-construction. In H. Sayama, J. Rieffel, S. Risi, R. Doursat, & H. Lipson (Eds.), *Artificial Life 14, Proceedings of the Fourteenth International Conference on the Synthesis and Simulation of Living Systems* (pp. 514–521). Cambridge, MA: MIT Press.
2. Agmon, E., Gates, A. J., Churavy, V., & Beer, R. D. (2016). Exploring the space of viable configurations in a model of metabolism-boundary co-construction. *Artificial Life*, 22(2), 153–171.
3. Aubin, J.-P., Bayen, A. M., & Saint-Pierre, P. (2011). *Viability theory: New directions*. Springer.
4. Barandiaran, X. E., & Egbert, M. D. (2013). Norm-establishing and norm-following in autonomous agency. *Artificial Life*, 20(1), 5–28.
5. Beer, R. D. (2004). Autopoiesis and cognition in the Game of Life. *Artificial Life*, 10(3), 309–326.
6. Beer, R. D. (2014). The cognitive domain of a glider in the Game of Life. *Artificial Life*, 20(2), 183–206.
7. Bryden, J., Gill, R. J., Mitton, R. A. A., Raine, N. E., Jansen, V. A. A., & Hodgson, D. (2013). Chronic sublethal stress causes bee colony failure. *Ecology Letters*, 16(12), 1463–1469.

8. Di Paolo, E. A. (2005). Autopoiesis, adaptivity, teleology, agency. *Phenomenology and the Cognitive Sciences*, 4(4), 429–452.
9. Edwards, J. C., Johnson, M. S., & Taylor, B. L. (2006). Differentiation between electron transport sensing and proton motive force sensing by the Aer and Tsr receptors for aerotaxis. *Molecular Microbiology*, 62(3), 823–837.
10. Egbert, M. D., Barandiaran, X. E., & Di Paolo, E. A. (2010). A minimal model of metabolism-based chemotaxis. *PLoS Computational Biology*, 6(12), e1001004.
11. Egbert, M. D., Di Paolo, E. A., & Barandiaran, X. E. (2009). Chemo-ethology of an adaptive protocell: Sensorless sensitivity to implicit viability conditions. In G. Kampis, I. Karsai, & E. Szathmáry (Eds.), *Advances in Artificial Life, Proceedings of the 10th European Conference on Artificial Life, ECAL* (pp. 242–250). Berlin: Springer.
12. Egbert, M. D., & Pérez-Mercader, J. (2016). Adapting to adaptations: Behavioural strategies that are robust to mutations and other organisational-transformations. *Scientific Reports*, 6, 18963.
13. Hoehler, T. M., & Jørgensen, B. B. (2013). Microbial life under extreme energy limitation. *Nature Reviews Microbiology*, 11(2), 83–94.
14. Jeong, H., Mason, S. P., Barabási, A.-L., & Oltvai, Z. N. (2001). Lethality and centrality in protein networks. *Nature*, 411(6833), 41–42.
15. Lee, J., & Kaletunç, G. (2002). Evaluation of the heat inactivation of *Escherichia coli* and *Lactobacillus plantarum* by differential scanning calorimetry. *Applied and Environmental Microbiology*, 68(11), 5379–5386.
16. Mason, O., & Verwoerd, M. (2007). Graph theory and networks in biology. *IET Systems Biology*, 1(2), 89–119.
17. Meyer, T., Schreckling, D., Tschudin, C., & Yamamoto, L. (2008). Robustness to code and data deletion in autocatalytic quines. In C. Priami, F. Dressler, O. B. Akan, & A. Ngom (Eds.), *Transactions on Computational Systems Biology X* (pp. 20–40). Berlin, Heidelberg: Springer.
18. Noireaux, V., & Libchaber, A. (2004). A vesicle bioreactor as a step toward an artificial cell assembly. *Proceedings of the National Academy of Sciences of the U.S.A.*, 101(51), 17669–17674.
19. Piedrafitra, G., Montero, F., Morán, F., Cárdenas, M. L., & Cornish-Bowden, A. (2010). A simple self-maintaining metabolic system: Robustness, autocatalysis, bistability. *PLoS Computational Biology*, 6(8), e1000872.
20. Prigogine, I. (1967). *Introduction to thermodynamics of irreversible processes*. New York: Wiley.
21. Ruiz-Mirazo, K., & Mavelli, F. (2007). Simulation model for functionalized vesicles: Lipid-peptide integration in minimal protocells. In F. Almeida e Costa, L. M. Rocha, E. Costa, I. Harvey, & A. Coutinho (Eds.), *Advances in Artificial Life, 9th European Conference, ECAL 2007, Proceedings* (pp. 32–41). Berlin, Heidelberg: Springer.
22. Ruiz-Mirazo, K., & Mavelli, F. (2008). On the way towards ‘basic autonomous agents’: Stochastic simulations of minimal lipid-peptide cells. *Biosystems*, 91(2), 374–387.
23. Shirt-Ediss, B., Solé, R., & Ruiz-Mirazo, K. (2013). Steady state analysis of a vesicle bioreactor with mechanosensitive channels. In P. Lió, O. Miglino, G. Nicosia, S. Nolfi, & M. Pavone (Eds.), *Advances in Artificial Life, ECAL 2013, Proceedings of the Twelfth European Conference on the Synthesis and Simulation of Living Systems* (pp. 1162–1169). Cambridge, MA: MIT Press.
24. Sprott, J. C., & Xiong, A. (2015). Classifying and quantifying basins of attraction. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 25(8), 083101.
25. Stewart, J., Gapenne, O., & Di Paolo, E. A. (Eds.) (2010). *Enaction: Toward a new paradigm for cognitive science*. Cambridge, MA: MIT Press.
26. Thompson, E. (2010). *Mind in life*. London: Harvard University Press.
27. Varela, F. J. (1979). *Principles of biological autonomy*. New York: North Holland.
28. Virgo, N. (2011). *Thermodynamics and the structure of living systems*. Doctoral thesis, University of Sussex, UK.