Biophysical Journal
Article

Biophysical Society

Numerical Parameter Space Compression and Its Application to Biophysical Models

Chieh-Ting (Jimmy) Hsu,[1] Gary J. Brouhard,[1,2,*] and Paul François[1,2,*]
[1]Department of Physics, McGill University, Montréal, Quebec, Canada and [2]Department of Biology, McGill University, Montréal, Quebec, Canada

ABSTRACT Physical models of biological systems can become difficult to interpret when they have a large number of parameters. But the models themselves actually depend on (i.e., are sensitive to) only a subset of those parameters. This phenomenon is due to parameter space compression (PSC), in which a subset of parameters emerges as "stiff" as a function of time or space. PSC has only been used to explain analytically solvable physics models. We have generalized this result by developing a numerical approach to PSC that can be applied to any computational model. We validated our method against analytically solvable models of a random walk with drift and protein production and degradation. We then applied our method to a simple computational model of microtubule dynamic instability. We propose that numerical PSC has the potential to identify the low-dimensional structure of many computational models in biophysics. The low-dimensional structure of a model is easier to interpret and identifies the mechanisms and experiments that best characterize the system.

SIGNIFICANCE Computational models are integral to many domains of science. But are these models overly complex? Von Neumann quipped that only four parameters can fit an elephant, and five can make its trunk wiggle. Here, we show how to compress the parameter space of computational models, which allows us to discover their underlying structure and to extract key parameters. We validate our method against two analytically solvable models. We then compress a well-known computational model of microtubule dynamic instability, which is the nonequilibrium switching of tubulin polymers between phases of growth and shrinkage. We show that only two effective parameters are sufficient to describe dynamic instability. Our work opens the door to the rigorous analysis of any computational model in biophysics.

INTRODUCTION

A central goal of biophysics is to develop mathematical and computational models that describe biological systems. These models can operate at different temporal and spatial scales. In the case of the microtubule cytoskeleton, models range from molecular dynamics simulations of $\alpha\beta$-tubulin heterodimers (1) to Monte Carlo simulations of microtubule dynamic instability (2–4) to analytical theories that treat the mitotic spindle as a nematic liquid crystal (5). These models vary in their degree of complexity, i.e., in the number of parameters they use.

A central problem in biophysical modeling is defining the "right" number of parameters to explain and predict experimental data, which we refer to as observables. We prefer simple models; in the well-known quip from Von Neumann, four parameters are sufficient to fit an elephant, and five can make its trunk wiggle (6), as was indeed later demonstrated (7). More parameters can sometimes improve a model's performance—namely, its ability to reproduce observables—but too many can be a problem. Unnecessary parameters obfuscate those that determine the model's output and render the model less interpretable—all without any gain in predictive power (8). In other words, complex models can be black boxes. Thus, we need a rigorous way to define which parameters determine the performance of any model.

The behavior of a model can be described within a so-called parameter space, which has as many dimensions as there are parameters. Moving within this parameter space (by changing the values of parameters) should change a model's output of observables. But usually a given observable significantly changes along only a few directions in parameter space (9). In other words, most directions in parameter space are irrelevant. To make sense of complex models, an important scientific problem is to reliably extract relevant directions in parameter space, defining the true,

Submitted August 13, 2019, and accepted for publication January 14, 2020.
*Correspondence: gary.brouhard@mcgill.ca or paul.francois2@mcgill.ca
Editor: David Sept.
https://doi.org/10.1016/j.bpj.2020.01.023

lower-order "dimensionality" of the model. There are several ways to solve this problem. In the 1980s, classical principal-component analysis was proposed as a method to reduce ordinary differential equations (ODE)-based models of biochemical systems (10). More recently, the manifold boundary approximation method has been developed to fit data while minimizing dimensionality (11); similarly, fitness-based asymptotic parameter reduction can extract the "core working module" of a model (12). Other machine-learning approaches can develop realistic models with a minimal number of parameters, e.g., using Bayesian information criterion (13). These methods are focused on ODE-based models, however, so there is an acute need for universal methods that are applicable to stochastic computational models as well.

Recently, parameter space compression (PSC) (14) has been proposed as the reason why fundamental models in physics operate successfully with simple parameter sets (15). PSC is related to the properties of the Fisher information matrix (FIM), which quantifies the relative significance of a model's parameters. More specifically, for a generic dynamical physical system, the eigenvalues of the FIM change over time to identify combinations of parameters that become "stiff" (viz, those with strong effects on model outputs) versus others that become "sloppy" (those with very weak effects) (16). The sloppy parameters or parameter combinations are thus "compressed away" to reveal the simpler dimensionality that underlies a model's performance (17). In other words, most parameters are sloppy and thus irrelevant; this sloppiness is the main reason why coarse-grained models in physics provide such satisfying descriptions of the natural world (15).

PSC explains how the predictive, low-dimensional structure of a model emerges (14), but currently PSC has been applied to a very limited number of analytically solvable physics models because of the difficulty of generalizing PSC to other contexts. To study the phenomenon of dimensional reduction via PSC in more general contexts, we developed a numerical PSC method, allowing us to explicitly study stochastic biophysical systems. To validate our method, we recovered the analytical results of a simple one-dimensional random walk model (14) and its perturbations. We further tested our method on an analytically solvable model of protein production and degradation. To test our method on a bona fide computational model, we applied numerical PSC to a classic Monte Carlo model of microtubule dynamic instability. In all three test cases, we show that the eigenvalues of the FIM provide critical insights into the behavior of a model and the importance of its parameters. Thus, our numerical PSC method opens the door to an analysis of computational models in biophysics that reveals the minimal yet predictive descriptions of living systems.

MATERIALS AND METHODS

Mathematical formulation

Our approach implements numerically the theoretical ideas of Machta and colleagues on the computation of the FIM to describe parameter space compression (14,16,17). We thus study the distribution of an observable x and its sensitivity to a vector of parameters $\vec{\theta} = \{\theta_\mu\}$. The central idea is that changes in an important parameter will result in larger changes in the probability distributions $y(\vec{\theta}, x)$ relative to changes in a less important one (Fig. 1 *A*). Thus, the important parameters will dominate the eigenvalues and eigenvectors of the FIM. Dominating eigenvalues define directions in parameter space where observables vary significantly. We call these directions "effective parameters." In general, the effective parameters of a model are not the original parameters but rather combinations of them (see below). Thus, the goal of PSC is to identify these dominating eigenvalues and eigenvectors, which will define the most important directions in parameter space and the effective parameters defining the distribution $y(\vec{\theta}, x)$ of observable x (16,17). In particular, for a dynamical system, we expect that a hierarchy of eigenvalues will appear for the FIM of $y(\vec{\theta}, x, t)$ as time t progresses, such that only a few effective parameters define the

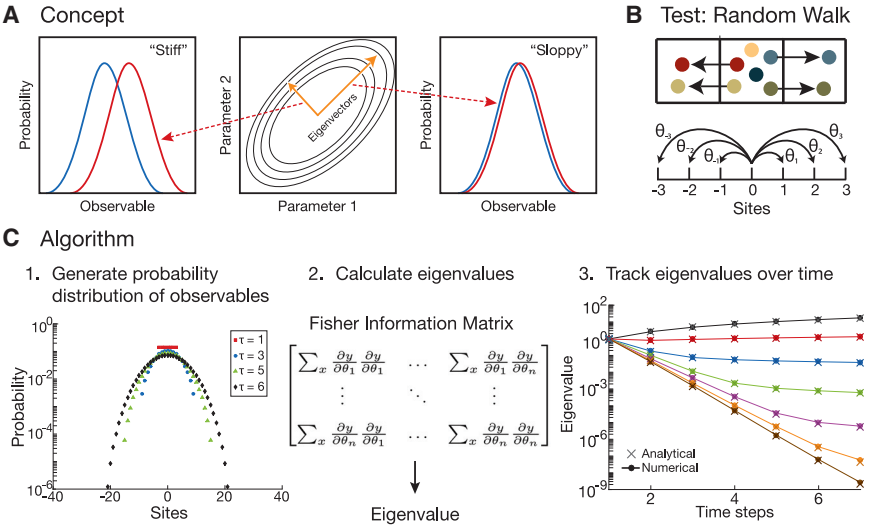

FIGURE 1 Numerical PSC. (*A*) Shown is a schematic of a stiff parameter versus a sloppy parameter in parameter space. The stiff parameter changes the observable more significantly than a sloppy one. Note: the red and blue curves are probability distributions when shifting different parameters by the same amount. (*B*) Shown is a schematic of one-dimensional random walk with parameters θ_i, the probability of jumping to neighboring sites. (*C*) Shown are the three steps of our numerical PSC method: 1) generate the probability distributions of the observables needed (see the plot of particle density at different time steps), 2) calculate the finite derivatives for the Fisher information matrix and its corresponding eigenvalues, and 3) repeat at each time step of the simulation and track the eigenvalues over time. Our numerical PSC reproduces the analytical result from (14). Note: each color corresponds to the same eigenvalue tracked over time. To see this figure in color, go online.

observed dynamics at any given time. These few effective parameters are sufficient to completely describe the system.

As shown in (14), with $y(\vec{\theta},x)$ being the probability distribution of observable x with parameters $\vec{\theta}$, the FIM at any given time t can be rewritten with standard assumptions as a simple "metric":

$$g_{\mu,\nu}(t) = \sum_x \frac{\partial y(\vec{\theta},x,t)}{\partial \theta_\mu} \frac{\partial y(\vec{\theta},x,t)}{\partial \theta_\nu}, \tag{1}$$

where y can be evaluated as a function of time t. Notice that $g_{\mu,\nu}(t)$ then becomes a simple function of the Jacobian with respect to its parameters. A detailed derivation adapted from (14) is provided in Appendix 1 in the Supporting Material. We independently consider several observables x and, for each x, compute the FIM of distribution $y(\vec{\theta},x,t)$ and its eigenvalues as a function of time t by summing over the entire observable landscape.

Scaling and algorithm

A challenge in analyzing computational models, especially in biology, is that the parameters have different units and scales. Some parameters are energies (e.g., the ΔG^o of bond formation), and some are kinetic rate constants (e.g., the rate constant of a GTP hydrolysis reaction). Because rate constants are exponentially distributed to thermal energy k_BT, we choose to rescale the parameters to express them all in terms of energies when calculating the FIM. Energies are more fundamental quantities, and their variations are easier to interpret physically. We thus define newly rescaled parameters, $\tilde{\theta}_\mu$, so that $\tilde{\theta}_\mu = \theta_\mu$ for parameters that are energies and $\tilde{\theta}_\mu = \log\theta_\mu$ for rate constants (with an implicit conversion factor to remove units, which normalizes the different parameters, equivalent to the common procedure of taking the derivatives with respect to the log of the parameter, as done previously (9)). Therefore, Eq. 1 becomes

$$g_{\mu,\nu} = \sum_x \frac{\partial y}{\partial \tilde{\theta}_\mu} \frac{\partial y}{\partial \tilde{\theta}_\nu} = \sum_x \frac{\partial y_i}{\partial \theta_\mu} \frac{\partial y_i}{\partial \theta_\nu} \theta_\mu^{\alpha_\mu} \theta_\nu^{\alpha_\nu}, \tag{2}$$

where $\alpha_\mu = 0$ if θ_μ is an energy and $\alpha_\mu = 1$ if θ_μ is a rate constant.

To calculate the FIM numerically using Eq. 2, we developed a three-step algorithm, shown in Fig. 1 *C*. First, we numerically compute the probability distributions $y(\vec{\theta},x,t)$ of each observable x at any time t for incremental variations of parameters θ_μ. We then compute finite derivatives to evaluate $(\partial y/\partial\tilde{\theta})$ (corresponding to the Jacobian), which implies that we need to generate $2N+1$ probability distributions $y(\vec{\theta}\pm\Delta\vec{\theta},x,t)$ for each observable x for a model with N parameters. Second, we evaluate each element of the FIM using the $2N+1$ probability distributions that we generated. Therefore, Eq. 2 becomes

$$g_{\mu,\nu} = \sum_x \frac{y(\theta_\mu+\Delta\theta_\mu,x,t) - y(\theta_\mu-\Delta\theta_\mu,x,t)}{2\Delta\theta_\mu} \times \frac{y(\theta_\nu+\Delta\theta_\nu,x,t) - y(\theta_\nu-\Delta\theta_\nu,x,t)}{2\Delta\theta_\nu} \tag{3}$$

Using Eq. 3, we sum over the entire observable landscape for each element of the FIM and then calculate the eigenvalue of the FIM at a given time. Third, we track the eigenvalues of the FIM over time. In general, the eigenvalues of the FIM are logarithmically distributed (16). The important feature of the eigenvalues is not their absolute values but rather their relative values, which is to say that the largest eigenvalue points to the most important direction in parameter space.

When evaluating the finite derivatives in Eq. 3, the choice of $\Delta\theta$ is arbitrary. It is clear that very large changes in parameter values will cause numerical instability when calculating the finite derivatives. Rather, the issue is whether small changes in parameter values are nevertheless large enough to cause meaningful changes in the distributions of observables. In our experience, very small changes cause the eigenvalues and eigenvector components of the FIM to become noisy. Therefore, we recommend $\Delta\theta = 0.05$ k_BT (leading to a change of 5% for corresponding rate constants; see Appendix 2 in the Supporting Material) as a robust choice to avoid numerical instability while keeping significant changes in model output. The best value of $\Delta\theta$ may be be model dependent. The consequences of small changes in the parameters (e.g., 0.01 k_BT) will be demonstrated below.

Computational costs mostly come from running enough simulations to generate good probability distributions. For the microtubule study, we ran $\geq 40,000$ independent simulations of microtubule growth per parameter set, which represents a bit less than 1 h of computation on a single 3.6-GHz core. As mentioned above, for N parameters, we need $2N+1$ simulations to compute derivatives from Eq. 3. Thus, for five parameters, roughly 10 core hours of computation are needed. Computation can be easily parallelized because simulations of individual microtubules are independent, and distributions of all observables can be obtained from the same set of simulations.

RESULTS

Test case: one-dimensional random walk

To test our numerical PSC method, we benchmarked our algorithm by simulating a model for which an analytical solution is available. We chose the one-dimensional random walk model introduced in Machta et al. (14), which is the model used to develop the concept of PSC. The parameters of the model are the probabilities of a particle jumping to one of six neighboring sites (Fig. 1 *B*); the observable x of the model is the position of a particle, and $y(\vec{\theta},x,t)$ is the distribution of particle positions as a function of time (viz, the particle density in a mean-field approximation when there are many particles). We simulated the random walk and plotted the eigenvalues of the FIM over time (Fig. 1 *C*), and our results precisely match those derived from the analytical expression (see Appendix 2 in the Supporting Material). In particular, the eigenvalues start at unity; as time progresses, the distribution of eigenvalues expands, establishing a clear hierarchy of eigenvalues at later times.

As pointed out in Machta et al. (14), the first two eigenvalues can be interpreted as a drift term and a diffusion coefficient, respectively; the spreads of the eigenvalues are enough to reproduce most of the data in an effective theory (as discussed in (14)). We further tested the correspondence between our numerical results and the analytical theory by introducing drift into the random walk, which was not done previously. The particle density over time is shown in Fig. 2 *A*. The eigenvalues of the FIM over time for the perturbed random walk are shown in Fig. 2 *B*. The result is similar to uniform diffusion in the sense that a hierarchy of eigenvalues appears as time progresses. We are able to show that the eigenvalues are proportionally defined by the probabilities of particles jumping to neighboring sites (the derivation is shown in Appendix 3 in the Supporting Material). Most importantly, our numerical results precisely

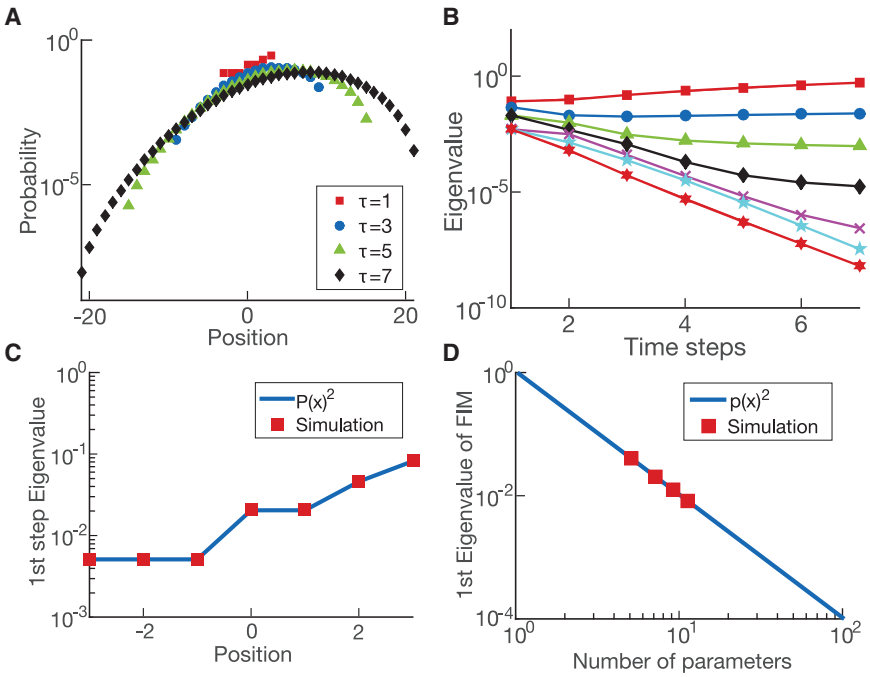

FIGURE 2 Random walk and random walk with drift. (*A*) Shown is the particle density for a nonuniform one-dimensional random walk over time (with probability higher to the right side of the space). (*B*) Shown are the eigenvalues for the nonuniform one-dimensional random walk. The eigenvalues at the first step of the simulation are not at unity contrary to the uniform random walk. (*C*) At the first step of the simulation for the one-dimensional random walk, the eigenvalue is equal to the squared rate given for the nonuniform simulation. This result is universal for any rates given. (*D*) The eigenvalues at the first step of a uniform simulation are also equal to the squared rate of the simulation, i.e., squared of one over the number of parameters. To see this figure in color, go online.

match the analytical solutions we derived for a random walk with drift and a uniform random walk with different numbers of parameters (see Fig. 2, *C* and *D*, respectively). Thus, our numerical PSC method successfully compressed this classical system and its variations.

Test case: a simple protein production and degradation system

Having benchmarked our algorithm against the random walk model, we next wondered how our numerical PSC method would handle a model in biophysics, in which the distributions of observables are often complex. Therefore, we applied our numerical PSC method to a textbook biophysical model of protein production and degradation. The model has only two parameters, the production rate ρ and the degradation rate δ (see Fig. 3 *A*). The observable x of the model is the number of proteins in the system at any given time. Importantly, the stationary distribution $y(\vec{\theta}, x)$ of protein number is a Poisson distribution of the parameter combination ρ/δ (representing the expectation value for the number of proteins) (18). Using this stationary distribution, we can analytically solve for the dominating eigenvalue of the corresponding FIM in the continuous limit:

$$\lambda_1 \simeq \frac{1}{2}\sqrt{\frac{\rho}{\delta\pi}} \tag{4}$$

The derivation of the eigenvalues and the expression for Eq. 4 can be found in Appendix 4 in the Supporting Material.

Starting from an initial condition with no proteins, we simulated this process using the Gillespie algorithm (19) and computed the eigenvalues of the system over time (Fig. 3 *B*). One eigenvalue is always over two orders of magnitude larger than the other, indicating that the system

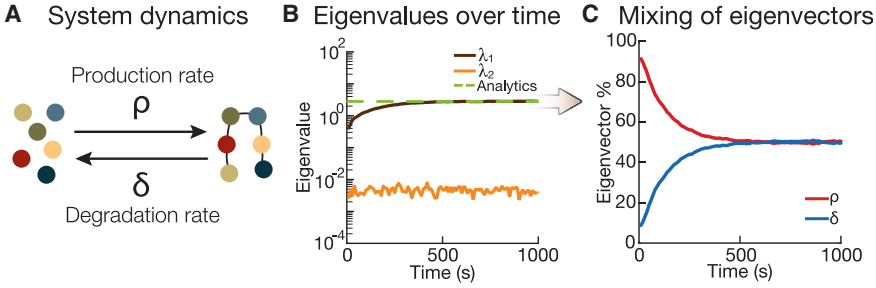

FIGURE 3 Protein production and degradation. (*A*) Shown is a schematic of a simple protein production-degradation system with production rate ρ and degradation rate δ. (*B*) Shown is a plot of eigenvalues over time for the protein production-degradation system. There is one dominating eigenvalue, and it matches the analytical result. (*C*) Shown is the plot of eigenvector percent from the dominating eigenvalue of (*B*). The production rate ρ dominates at early time points, but at stationarity, the production rate and degradation rate contribute equally. Note: the eigenvector percent is the absolute value of the parameter component. To see this figure in color, go online.

is governed by one effective parameter, which is to say that there is only one relevant direction in parameter space that determines the model's output. Looking at the relative contribution of the eigenvector components of the dominating eigenvalue in Fig. 3 *C*, we can see that during the early stages of the simulation, the production rate ρ dominates, corresponding to the net production of proteins from the initial condition. The system then reaches stationarity, at which point the eigenvector components of the dominating eigenvalue are an equal mix of the production rate ρ and degradation rate δ, as expected from our derivation (Fig. 3 *C*). We checked that our method recovers the analytical result of Eq. 4 (in the asymptotic limit) for different ratios of production rate over degradation rate (see Fig. 4 *A*). Thus, our numerical PSC method is able to compress out irrelevant directions and extract the effective parameter defining the distribution of protein number (here, a Poisson distribution).

In the continuous limit, the second eigenvalue of the system goes to 0. In a discrete simulation, however, it is impossible for the second eigenvalue to reach 0 because of the limitations of numerical precision and the physical definition of the system. For example, for a production rate $\rho = 1$ and a degradation rate $\delta = 0.01$, we know that the steady-state solution will have a peak at $N = 100$ proteins. However, this means that when calculating the finite derivatives, any shift smaller than 1% will result in a change of less than one protein, which is nonphysical. We show that even using analytical values of the Poisson distribution, we will not be able to reach 0 when calculating finite derivatives (see Fig. 4 *B*).

Microtubule dynamics: a complex biophysical system

Having fully characterized our method, we applied it to a biophysical system that cannot be solved analytically—namely, the dynamic instability of microtubules (20). Microtubules are polymers of $\alpha\beta$-tubulin, and dynamic instability is the nonequilibrium behavior in which the polymers stochastically switch between periods of growth and shrinkage. This complex, nonequilibrium phenomenon was first simulated numerically in the 1980s (21,22) and has remained a subject of considerable interest for computational biologists, who have developed increasingly sophisticated models (3,4,23,24). The long-term goal of these collective efforts is to develop a powerfully predictive yet minimal model that can be used to explain microtubule physiology. Our numerical PSC method has the power to determine whether existing models have an underlying low-dimensional structure.

Our model is based on VanBuren et al. (2) (see Fig. 5 *A*); a similar model is used by Ayaz et al. (25). We chose this model because it is a classic and because understanding its underlying dimensionality will inform ongoing modeling work on microtubules. Briefly, tubulin subunits associate head-to-tail to create protofilaments (pfs), forming longitudinal bonds described by an energy parameter ΔG^o_{long}. In our model, 13 pfs are connected by lateral bonds between adjacent subunits with an energy parameter ΔG^o_{lat} (26). The rate at which tubulin binds to pf ends is described by an association rate constant, k^+. Because tubulin is a GTPase, these incoming tubulin subunits contain GTP in the tubulin nucleotide pocket. This GTP becomes hydrolyzed after 1) the subunit incorporates into the polymer and 2) another GTP-tubulin binds on top of it, contributing catalytic residues that complete the nucleotide pocket (27). The rate of GTP hydrolysis is described by a rate constant parameter k_H. GTP hydrolysis and phosphate release converts GTP-tubulin to GDP-tubulin and weakens the bonds between tubulin subunits in the polymer (28,29). After VanBuren, this weakening of energies is described by an energy parameter, $\Delta\Delta G^o_{lat}$, which is assigned to the lateral bonds of the new GDP-tubulin subunit.

We performed a parameter sweep and arrived at parameter values similar to Castle et al. ((4), see Fig. 5 *A*). Using these parameters, our simulation produces microtubule growth curves that correspond reasonably with measurements from multiple labs using 8 μM brain tubulin (e.g., (30), see Fig. 5 *B*). More specifically, microtubules grow as long as their ends are protected by a "cap" of GTP-

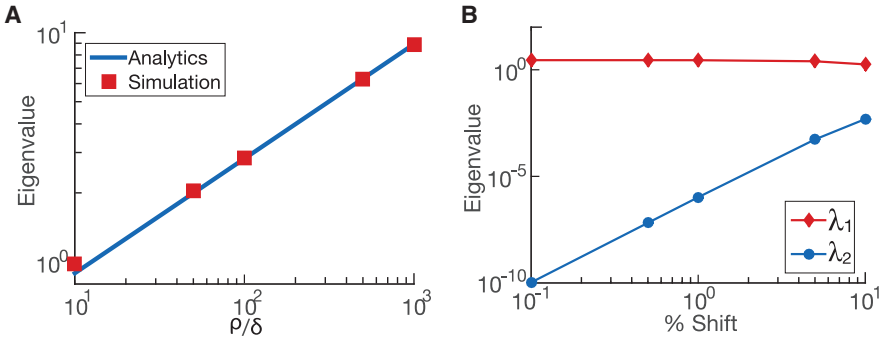

FIGURE 4 Generalization of protein system and understand numerical limitations. (*A*) The dominating eigenvalue for the protein production-degradation system is shown to be a ratio of the production rate over the degradation rate times some constant. The simulations of different ratios match the analytic solution. (*B*) The second eigenvalues for the protein production and degradation rate are nonzero during simulation because of the limitation of the physical system itself. At steady state, the average number of proteins is 100 proteins, which means that the smallest shift for probability to calculate the finite derivatives is one protein (1% of the average number), which gives a nonzero eigenvalue. To see this figure in color, go online.

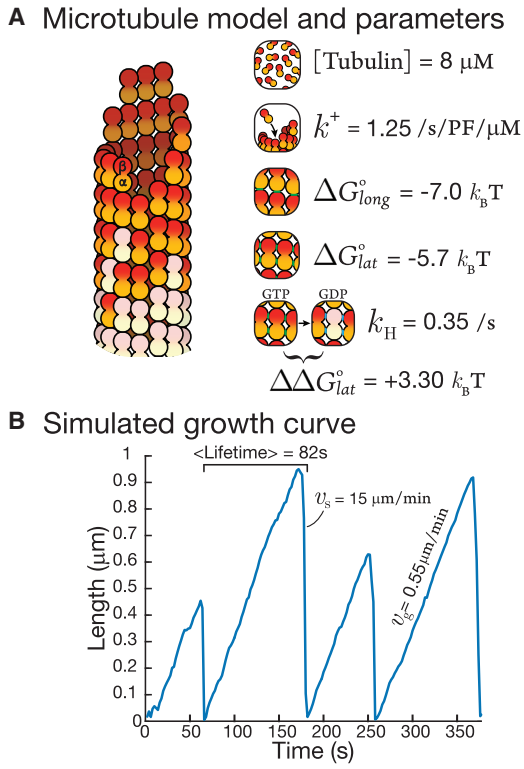

FIGURE 5 Microtubule dynamics. (*A*) Shown is the base model of our microtubule simulation after VanBuren et al. (2). (*B*) Shown is a plot of length versus time from our microtubule simulation. To see this figure in color, go online.

tubulin (31). If this GTP cap is "lost," the polymer switches to rapid shrinkage in an event known as a "catastrophe," the hallmark of dynamic instability (20).

There are many subtleties and caveats to models of dynamic instability. For example, which bonds are weakened by GTP hydrolysis is not well established (32,33), and the transition from GTP-tubulin to GDP-tubulin may have substeps (33). These subtleties are discussed in Appendix 5 in the Supporting Material. We used the direct method of the Gillespie algorithm (19), which is a different implementation than the one in VanBuren et al. (2) and Ayaz et al. (25). To validate our Gillespie algorithm, we used the parameters found in Ayaz et al. (25) and confirmed that our simulation produces identical results. The details of our simulation method and the benchmarking of our algorithm against published data can be found in Appendix 5 in the Supporting Material and Fig. S2.

To compress our model, we varied all five parameters (ΔG^o_{long}, ΔG^o_{lat}, k_H, k^+, and $\Delta\Delta G^o_{lat}$) around their initial values. We measured four independent observables of the simulations that correspond to the experimental data used in our parameter sweep (30): 1) the length of the microtubule (Fig. 6 *A*); 2) the decay constant that describes the conversion of GTP-tubulin into GDP-tubulin ("GTP cap size," Fig. 6 *B*) (34); 3) the microtubule lifetime (Fig. 6 *C*); and 4) the postcatastrophe shrinkage rate (Fig. 6 *D*). Two of these observables can be tracked continuously over the time course of the simulation—namely, the length of microtubule and the decay constant. The second column of Fig. 6, *A* and *B* shows the eigenvalues over time for these observables. The other two observables—namely, microtubule lifetimes and the postcatastrophe shrinkage rate—are not tracked continuously because they require postsimulation analysis to determine when catastrophes occurred (see Appendix 6 in the Supporting Material). The second column of Fig. 6, *C* and *D* shows the eigenvalues for these observables at the conclusion of the simulation, when the distributions have reached stationarity. This framework allowed us to apply our numerical PSC method to our model of microtubule dynamic instability.

For three out of four observables, one eigenvalue dominates the others by at least one order of magnitude (note the log-scale for eigenvalues). This dominance implies that the distribution of these observables is determined by a single effective parameter. This result is not obvious: one expects that the mean and the variance of any given distribution are described by independent parameters, as was the case for the random walk (14). Rather, three of our microtubule observables are similar to the number of proteins in the protein production/degradation model, in which both the mean and variance of the distributions are determined by a single effective parameter. The only exception we observe to this rule is the microtubule lifetime distribution; even though one eigenvalue strongly dominates this distribution, a second eigenvalue is a bit less than one order of magnitude below the first one. This smaller difference suggests that although the lifetime distribution is mostly determined by one effective parameter, another parameter mildly modulates it.

As for the protein production/degradation case, the single effective parameter determining the distribution of each observable is a priori a complex function of the initial parameters. As before, the relative influence of each initial parameter is given by the eigenvector components of the dominant eigenvalue (see column three of Fig. 6, *A–D*; for the lifetime distribution, we also show the eigenvector components for the second eigenvalue). Importantly, we can also see which parameters are not important for a given observable because these parameters will be insignificant components of the dominating eigenvalue.

The important components for microtubule length are the lateral bond, ΔG^o_{lat}, followed closely at later times by the longitudinal bond, ΔG^o_{long}. These components are not surprising considering that the bond energies are what drive polymerization. The important components for the decay constant are more interesting. In addition to the obvious parameter of the GTP hydrolysis rate constant, k_H, the decay constant is also determined by k^+ and ΔG^o_{long}. A simple interpretation of this result is that a microtubule that forms stronger bonds (and hence grows faster) will have a larger

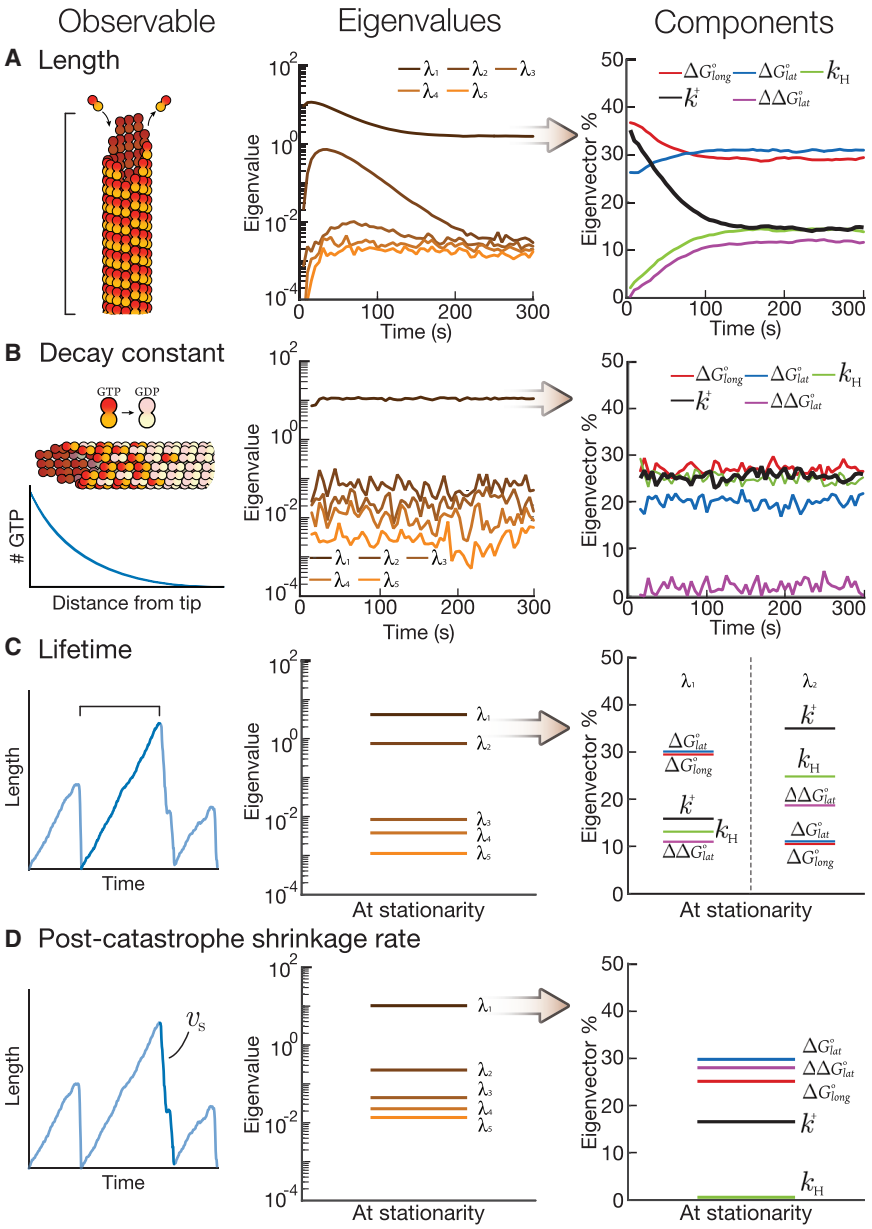

FIGURE 6 Numerical PSC for the five-parameter bovine microtubule model. (*A–D*) Eigenvalues and eigenvector components (percent) for four observables: (*A*) length (the exact number of dimers in each microtubule), (*B*) decay constant of GTP from the tip, (*C*) average lifetime of the microtubule, and (*D*) postcatastrophe shrinkage rate. Note: the eigenvector percent is the absolute value of the parameter component. To see this figure in color, go online.

GTP cap. Consistently, microtubules that grow faster have larger GTP caps when end-binding proteins are used as reporters of GTP cap size (34). The lifetime distribution depends on the two bond energies and is further modulated by k^+ and k_H via the second eigenvalue. However, for the postcatastrophe shrinkage rate, the most important parameter is the lateral bond ΔG^o_{lat}, followed closely by $\Delta\Delta G^o_{lat}$ and ΔG^o_{long}. Table 1 summarizes the parameters that are important for each observable.

As explained in Scaling and Algorithm, we calculated the FIM by varying our parameters by 0.05 k_BT. When we used smaller variations (0.01 k_BT), the results were similar in magnitude but clearly noisier (see Fig. S6) because the distributions of observables were shifted to a lesser extent.

The results above used a parameter set that reproduces data from mammalian microtubules using tubulin purified from brain tissue (e.g., (30)). We wondered whether a different parameter set would give the same results in terms

TABLE 1 This Table Demonstrates the Importance of Parameters on Different Observables

	Ranking		
Observables	Important	Intermediate	Less Important
Length	$\Delta G^o_{lat}\Delta G^o_{long}$	NA	$k^+k_H\Delta\Delta G^o_{lat}$
Decay constant	$\Delta G^o_{long}k_Hk^+$	ΔG^o_{lat}	$\Delta\Delta G^o_{lat}$
Lifetime	$\Delta G^o_{lat}\Delta G^o_{long}$	NA	$k^+k_H\Delta\Delta G^o_{lat}$
Postcatastrophe shrinkage rate	$\Delta G^o_{lat}\Delta\Delta G^o_{lat}\Delta G^o_{long}$	k^+	k_H

of the number of dominating eigenvalues and their components. In other words, do our results apply only to a local region of parameter space, or do they apply globally? To answer this question, we used a parameter set that reproduces different data—namely, the dynamic instability of *Caenorhabditis elegans* microtubules (30). *C. elegans* microtubules are among the most divergent measured to date in that they grow faster and have shorter lifetimes than microtubules from several other species (35–37). Our PSC predicts that faster growth requires more negative values for ΔG^o_{long} and ΔG^o_{lat}, and indeed, the *C. elegans* parameter set is shifted accordingly (see Fig. S4).

With this *C. elegans* parameter set, we performed PSC on our model. The results were quite similar; e.g., as with the bovine parameter set, the length distribution, decay constant, and postcatastrophe shrinkage rate had one dominating eigenvalue (see Fig. S4), indicating that these observables are controlled by a single effective parameter. Similarly, the lifetime distribution had two dominating eigenvalues (see Fig. S4). Indeed, the eigenvector components of these eigenvalues were nearly identical in every case, indicating that the low-dimensional structure of our model is conserved between the brain and *C. elegans* parameter sets. The one difference we observed was in the eigenvector components that describe the effective parameter for the postcatastrophe shrinkage rate (see Fig. S4). In the *C. elegans* case, the association rate constant k^+ is a significant component. We interpret this result in light of our previous observation that *C. elegans* tubulin is more "active" in solution (30): a more active dimer may influence the rate of shrinkage through its binding to microtubule ends. Despite this difference, the PSC results for brain microtubules and *C. elegans* microtubules are broadly similar. We conclude that our PSC results are weakly dependent on specific parameter values and/or the local position in parameter space.

Parameter dependencies of distributions

As shown above, the distribution of most observables can be described by a single effective parameter, which is specified by the eigenvector components of the dominating eigenvalue. We can illustrate this phenomenon by plotting the distribution of microtubule lengths (Fig. 7 *A*); the distribution is a nearly perfect exponential (as predicted from simple analytical models (38,39)). Exponential probability distributions are described by one parameter, which in the microtubule case is the average length, $\langle L \rangle$, thus defining the effective parameter.

In contrast to the length distribution, the microtubule lifetime distribution should be controlled by two parameters according to our analysis. Consistent with this idea, it has been suggested that this lifetime follows a Γ-distribution (40). We thus plotted our computed lifetime distribution (Fig. S3). We indeed observe an increase followed by an exponential tail (consistent with a multistep process as proposed in (40)), but our distribution rises slightly faster than the best Γ-fit. To visualize the effects of the two effective parameters, we follow the two eigendirections, as illustrated in Fig. S3 B. To compare properly the shapes of the distribution, 1) we further rescale the distributions by their maximal probability, and 2) we adjust the magnitude of the changes so that the most probable lifetime is the same in both directions. It is then very clear that although the leftmost part of the distribution is similar, the exponential tail differs. Contrary to the length case above, we thus cannot define the distribution with only one parameter (such as the most probable lifetime) because the tail of the distribution clearly requires another parameter.

Estimating the dimensionality of the system

Our eigenvector components tell us which directions matter in parameter space, similar to the "hyper-ribbon" notion described in (17). But is the important direction for microtubule length, e.g., the same direction that is important for the other observables? Or do we need five orthogonal directions to describe the full model? For three of our observables, the eigenvector components are very similar, suggesting a common effective parameter. More specifically, ΔG^o_{lat}, ΔG^o_{long}, and k^+ are dominating eigenvector

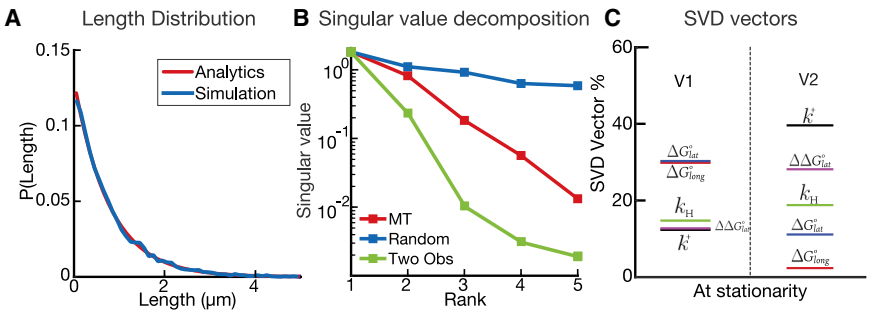

FIGURE 7 Identifying the dimensionality of the models. (*A*) Shown is the probability of length plot for our simulation versus analytic expression from (38,39). (*B*) Shown is the plot of the singular values for three different cases: in red, our five parameter microtubule model; in blue, five random five-dimensional vectors; in green, length and decay constant observables. (*C*) The first column is the leading vector for the highest singular value for the microtubule system. The lateral bond and longitudinal bond are the controlling parameters. The second column is the leading vector for the second highest singular value for the microtubule system. The on-rate constant and the energy penalty after hydrolysis are the controlling parameters. Note: the SVD vector percent plotted is the absolute value of the parameter component. To see this figure in color, go online.

components for microtubule lengths, lifetimes, and the post-catastrophe shrinkage rates. In contrast, the decay constant has a significant eigenvector component from k_H, indicating that its effective parameter may differ from the others. So what is the true dimensionality of our model?

To perform a rigorous estimation of the dimensionality, we computed the singular value decomposition (SVD) on all the eigenvectors of all the observables. The number of large singular values in an SVD analysis indicates how many dimensions are sufficient to describe the original matrix with good precision. To perform our SVD analysis, we rescaled the eigenvectors proportionally to their respective eigenvalues so that the largest eigenvalue for each observable has a weight of 1 (41). This rescaling allows us to capture the global sensitivity with respect to parameters of all observables simultaneously. We compared the SVD computation of our model with two cases: 1) five random vectors in five directions and 2) combinations of eigenvectors for the length and lifetime observables in our microtubule system. The singular values for these computations are shown in Fig. 7 *B*. The random case gives a baseline for what might be expected from a five-dimensional system: random vectors are almost orthogonal, so we obtained five large singular values. The two-observable case gives us a baseline for a low-dimensional system: the singular values decay rapidly from the first large singular value, indicating that only one effective parameter defines the distributions of these two observables.

SVD for the full model shows an intermediate result with a sloppy distribution of singular values: there is one large singular value, a second singular value roughly three times smaller, and three even smaller singular values at least one order of magnitude below. It is also visually clear that the relative positions of the first two singular values are roughly comparable to the first two singular values of the random case, indicating that the system is close to two-dimensional.

Thus, the presence of two large singular values means that two effective parameters are essentially enough to fit the data. A standard estimation of the precision of this dimensional reduction can be done by computing ratios of Frobenius norms of the singular matrix $r_k = (\sum_{i<k}\lambda_i^2 / \sum_i \lambda_i^2)$, where λ_i are singular values ordered from top to bottom rank (42). The closer to 1, the better the reduction, and a good rule of thumb is $r_k > 0.9$. For $k = 1$, one finds $r_1 = 0.82$, whereas for $k = 2$, one finds $r_2 = 0.99$, indicating an excellent dimensional reduction if we keep the first two modes.

Our SVD analysis demonstrates rigorously that the full dimensionality of our model is essentially equal to two. The vectors that correspond to the two dominating singular values are shown in Fig. 7 *C*. The first vector is primarily composed of the lattice bond energies. The second vector combines parameters associated with the binding and unbinding of tubulin dimers, with major contributions from k^+ and $\Delta\Delta G^o_{lat}$. We observed similar results for an SVD analysis of our model when we used our *C. elegans* parameter values (Fig. S5). The *C. elegans* model was also two-dimensional, but the second singular value included a significant contribution from k^+, which is consistent with our PSC results for the worm model described above. Thus, our results generalize to different parameter sets while revealing how new parameter dependencies can appear.

DISCUSSION

As biophysicists, we want to capture the complexity of biology in the simplest possible terms, even if those terms are themselves quite complex. Our work has demonstrated the power of numerical PSC as a method for identifying the essential parameters and low-dimensional structure of complex models. We first validated our method against two analytically solvable models and then applied it to a well-known computational model of microtubule dynamic instability. Thus, our method opens the door to the simplification of many computational models in biophysics. Computation of effective parameters is made rigorous and possible by the use of our approach. PSC of microtubule dynamics is very reminiscent of classical examples, such as random walks (14), in which a sloppy distribution of eigenvalues is also observed and two effective parameters naturally appear (drift and diffusion, Fig. 1 *C*). It is remarkable that a biological phenomenon as complex as dynamic instability can be compressed into a two-parameter system as well.

Our analysis of microtubule dynamic instability revealed that almost all data simulated here can be described with only two effective parameters: 1) a "polymerization" parameter, which includes ΔG^o_{long} and ΔG^o_{lat}, and 2) a GTP cap parameter, which includes k_H, k^+, and $\Delta\Delta G^o_{lat}$. These two parameters form a two-dimensional "ribbon" within the five-dimensional space of parameters (16). The polymerization parameter makes physical sense if most of the binding sites at the end of a microtubule are shaped like "corners," where the incoming tubulin subunit will form one longitudinal bond and one lateral bond. Because these bonds are formed at the same time in the model, they are strongly coupled, as our analysis shows. The GTP cap parameter makes physical sense because it couples the size of the GTP cap (determined by k_H and k^+) and the strength of the GTP cap (determined by $\Delta\Delta G^0_{lat}$, which encodes the extra bond strength found in the cap versus the lattice).

This ribbon structure also provides direct insight into catastrophe dynamics. For example, we were surprised to find that k_H had a lesser influence on the lifetime distribution than ΔG^o_{lat} and ΔG^o_{long}. This suggests that catastrophe might be more efficiently prevented by making stronger bonds rather than by slowing down hydrolysis. Our interpretation is that stronger bonds help prevent pfs from losing their terminal GTP-tubulin dimers, which would cause the pf to become "fully uncapped." Bowne-Anderson et al. argued

that uncapping of pfs is the irreversible event that leads to catastrophe (43). Similarly, poisoning of pf ends with the drug eribulin has a very strong effect on catastrophe frequency (44). Therefore, our results showing the importance of ΔG^o_{long} and ΔG^o_{lat} are consistent with the emerging concept that "pf destabilization" is a root cause of catastrophe. Our results may also explain why depolymerases and catastrophe factors work by disrupting lattice bonds rather than acting as GTPase activating proteins, which are common in the regulation of other GTPases.

From the modeling standpoint, our results imply that a basic quantitative understanding of dynamic instability might not require many parameters but, rather, only the effective ones. This simplified structure is illustrated by the one-parameter fit of the length distribution in Fig. 7 *A*. But one limitation of PSC is the lack of a universal method for converting these discoveries into a coarse-grained model (e.g., a two-parameter simulation or analytical model). Machine-learning approaches and principal component analysis face a similar problem: how to interpret the lower-dimensional model or principal components. In the microtubule PSC case, we have made our best effort to interpret our results in physically meaningful ways that will facilitate the development of simpler models.

It is important to point out that the addition of new parameters might add new dependencies on the corresponding eigenvectors/eigenvalues, meaning that those parameters would matter (in the sense that they influence the effective parameters). However, they might change neither the nature nor the number of the effective parameters controlling the dynamics. A more interesting situation would be when adding a new biochemical parameter also adds a new effective parameter to the system, increasing the net dimensionality from two and three. Additionally, we can add new observables to our analysis (e.g., the "taper length" that describes the difference in length between the shortest and longest profotilaments). These new observables may demand new effective parameters. Our approach could help experimentalists identify the types of data that are necessary and sufficient to define such effective parameters. Which parameters of a model are stiff and which are sloppy depends critically on the observables that the model attempts to reproduce.

Our ability to distinguish between models in science is always limited by the availability of hard data. In biophysics, the rigor of physical modeling collides against the complexity of biological interactions. A coupling of theory and experiment is necessary to disentangle this complexity. PSC tightens this coupling by improving the interpretability of models, which in turn identifies the key experiments that drive theory forward.

SUPPORTING MATERIAL

Supporting Material can be found online at https://doi.org/10.1016/j.bpj.2020.01.023.

AUTHOR CONTRIBUTIONS

G.J.B. and P.F. conceptualized the project. C.-T.H. did all simulations, calculations, and analysis. C.-T.H., G.J.B., and P.F. wrote the manuscript. G.J.B. and P.F. administered and supervised the project.

ACKNOWLEDGMENTS

The authors thank Sami Chaaban, Claire Edrington, Dr. Hadrien Mary, Félix Proulx-Giraldeau, Thomas Rademaker, Laurent Jutras-Dubé, and Dr. Adrien Henry for feedback on this project and comments on the manuscript.

C.-T.H. acknowledges support from the Milton Leung fund in McGill Physics and from Fonds de recherche du Quebec - Nature et technologies (FRQNT) Bourse. G.J.B. is supported by the Canadian Institutes of Health Research (137055 and PJT-148702) and the Natural Sciences and Engineering Research Council of Canada (372593). P.F. is supported by a Simons Foundation fellowship in Mathematical Modelling of Biological Systems and NSERC (2016-06501). Lastly, this work was supported by a FRQNT Projet de Recherche en Equipe (FRQ-NT191128).

REFERENCES

1. Mitra, A., and D. Sept. 2008. Taxol allosterically alters the dynamics of the tubulin dimer and increases the flexibility of microtubules. *Biophys. J.* 95:3252–3258.
2. VanBuren, V., D. J. Odde, and L. Cassimeris. 2002. Estimates of lateral and longitudinal bond energies within the microtubule lattice. *Proc. Natl. Acad. Sci. USA.* 99:6035–6040.
3. VanBuren, V., L. Cassimeris, and D. J. Odde. 2005. Mechanochemical model of microtubule structure and self-assembly kinetics. *Biophys. J.* 89:2911–2926.
4. Castle, B. T., S. McCubbin, …, D. J. Odde. 2017. Mechanisms of kinetic stabilization by the drugs paclitaxel and vinblastine. *Mol. Biol. Cell.* 28:1238–1257.
5. Brugués, J., and D. Needleman. 2014. Physical basis of spindle self-organization. *Proc. Natl. Acad. Sci. USA.* 111:18496–18500.
6. Dyson, F. 2004. A meeting with Enrico Fermi. *Nature.* 427:297.
7. Mayer, J., K. Khairy, and J. Howard. 2010. Drawing an elephant with four complex parameters. *Am. J. Phys.* 78:648–649.
8. Gunawardena, J. 2014. Models in biology: 'accurate descriptions of our pathetic thinking'. *BMC Biol.* 12:29.
9. Gutenkunst, R. N., J. J. Waterfall, …, J. P. Sethna. 2007. Universally sloppy parameter sensitivities in systems biology models. *PLoS Comput. Biol.* 3:1871–1878.
10. Vajda, S., P. Valko, and T. Turányi. 1985. Principal component analysis of kinetic models. *Int. J. Chem. Kinet.* 17:55–81.
11. Transtrum, M. K., and P. Qiu. 2014. Model reduction by manifold boundaries. *Phys. Rev. Lett.* 113:098701.
12. Proulx-Giraldeau, F., T. J. Rademaker, and P. François. 2017. Untangling the hairball: fitness-based asymptotic reduction of biological networks. *Biophys. J.* 113:1893–1906.
13. Daniels, B. C., and I. Nemenman. 2015. Automated adaptive inference of phenomenological dynamical models. *Nat. Commun.* 6:8133.
14. Machta, B. B., R. Chachra, …, J. P. Sethna. 2013. Parameter space compression underlies emergent theories and predictive models. *Science.* 342:604–607.
15. Waterfall, J. J., F. P. Casey, …, J. P. Sethna. 2006. Sloppy-model universality class and the Vandermonde matrix. *Phys. Rev. Lett.* 97:150601.
16. Transtrum, M. K., B. B. Machta, and J. P. Sethna. 2011. Geometry of nonlinear least squares with applications to sloppy models and

optimization. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 83:036701.

17. Transtrum, M. K., B. B. Machta, …, J. P. Sethna. 2015. Perspective: sloppiness and emergent theories in physics, biology, and beyond. *J. Chem. Phys.* 143:010901.
18. Ross, S. M. 2019. The exponential distribution and the poisson process. Introduction to Probability Models. Academic Press, pp. 293–374.
19. Gillespie, D. T. 1977. Exact stochastic simulation of coupled chemical reactions. *J. Chem. Phys.* 81:2340–2361.
20. Mitchison, T., and M. Kirschner. 1984. Dynamic instability of microtubule growth. *Nature.* 312:237–242.
21. Chen, Y. D., and T. L. Hill. 1985. Monte Carlo study of the GTP cap in a five-start helix model of a microtubule. *Proc. Natl. Acad. Sci. USA.* 82:1131–1135.
22. Bayley, P., M. Schilstra, and S. Martin. 1989. A lateral cap model of microtubule dynamic instability. *FEBS Lett.* 259:181–184.
23. Zakharov, P., N. Gudimchuk, …, E. L. Grishchuk. 2015. Molecular and mechanical causes of microtubule catastrophe and aging. *Biophys. J.* 109:2574–2591.
24. McIntosh, J. R., E. O'Toole, …, N. Gudimchuk. 2018. Microtubules grow by the addition of bent guanosine triphosphate tubulin to the tips of curved protofilaments. *J. Cell Biol.* 217:2691–2708.
25. Ayaz, P., S. Munyoki, …, L. M. Rice. 2014. A tethered delivery mechanism explains the catalytic action of a microtubule polymerase. *eLife.* 3:e03069.
26. Ledbetter, M. C., and K. R. Porter. 1964. Morphology of microtubules of plant cell. *Science.* 144:872–874.
27. Löwe, J., H. Li, …, E. Nogales. 2001. Refined structure of alpha beta-tubulin at 3.5 A resolution. *J. Mol. Biol.* 313:1045–1057.
28. Mandelkow, E. M., E. Mandelkow, and R. A. Milligan. 1991. Microtubule dynamics and microtubule caps: a time-resolved cryo-electron microscopy study. *J. Cell Biol.* 114:977–991.
29. Maurer, S. P., F. J. Fourniol, …, T. Surrey. 2012. EBs recognize a nucleotide-dependent structural cap at growing microtubule ends. *Cell.* 149:371–382.
30. Chaaban, S., S. Jariwala, …, G. J. Brouhard. 2018. The structure and dynamics of C. elegans tubulin reveals the mechanistic basis of microtubule growth. *Dev. Cell.* 47:191–204.e8.
31. Carlier, M. F., and D. Pantaloni. 1981. Kinetic analysis of guanosine 5′-triphosphate hydrolysis associated with tubulin polymerization. *Biochemistry.* 20:1918–1924.
32. Zhang, R., and E. Nogales. 2015. A new protocol to accurately determine microtubule lattice seam location. *J. Struct. Biol.* 192:245–254.
33. Manka, S. W., and C. A. Moores. 2018. The role of tubulin-tubulin lattice contacts in the mechanism of microtubule dynamic instability. *Nat. Struct. Mol. Biol.* 25:607–615.
34. Bieling, P., L. Laan, …, T. Surrey. 2007. Reconstitution of a microtubule plus-end tracking system in vitro. *Nature.* 450:1100–1105.
35. Katsuki, M., D. R. Drummond, …, R. A. Cross. 2009. Mal3 masks catastrophe events in *Schizosaccharomyces pombe* microtubules by inhibiting shrinkage and promoting rescue. *J. Biol. Chem.* 284:29246–29250.
36. Vemu, A., J. Atherton, …, A. Roll-Mecak. 2016. Structure and dynamics of single-isoform recombinant neuronal human tubulin. *J. Biol. Chem.* 291:12907–12915.
37. Walker, R. A., E. T. O'Brien, …, E. D. Salmon. 1988. Dynamic instability of individual microtubules analyzed by video light microscopy: rate constants and transition frequencies. *J. Cell Biol.* 107:1437–1448.
38. Verde, F., M. Dogterom, …, S. Leibler. 1992. Control of microtubule dynamics and length by cyclin A- and cyclin B-dependent kinases in Xenopus egg extracts. *J. Cell Biol.* 118:1097–1108.
39. Dogterom, M., and S. Leibler. 1993. Physical aspects of the growth and regulation of microtubule structures. *Phys. Rev. Lett.* 70:1347–1350.
40. Gardner, M. K., M. Zanic, …, J. Howard. 2011. Depolymerizing kinesins Kip3 and MCAK shape cellular microtubule architecture by differential control of catastrophe. *Cell.* 147:1092–1103.
41. Wall, M. E., A. Rechtsteiner, and L. M. Rocha. 2003. Singular value decomposition and principal component analysis. *In* A Practical Approach to Microarray Data Analysis. D. P. Berrar, W. Dubitzky, and M. Granzow, eds. Springer US, pp. 91–109.
42. Golub, G. H., and C. F. Van Loan. 1966. Matrix Analysis. Johns Hopkins University Press, Baltimore, MD.
43. Bowne-Anderson, H., M. Zanic, …, J. Howard. 2013. Microtubule dynamic instability: a new model with coupled GTP hydrolysis and multistep catastrophe. *BioEssays.* 35:452–461.
44. Doodhi, H., A. E. Prota, …, M. O. Steinmetz. 2016. Termination of protofilament elongation by eribulin induces lattice defects that promote microtubule catastrophes. *Curr. Biol.* 26:1713–1721.

Biophysical Journal, Volume *118*

Supplemental Information

Numerical Parameter Space Compression and Its Application to Biophysical Models

Chieh-Ting (Jimmy) Hsu, Gary J. Brouhard, and Paul François

Manuscript submitted to **Biophysical***Journal*
Article

Supplement: Numerical parameter space compression and its application to biophysical models

Chieh-Ting (Jimmy) Hsu[1], Gary J. Brouhard[2,1*], and Paul François[1,2*]

[1]McGill University, Department of Physics, 3600 Rue University, Montréal, QC H3A 2T8
[2]McGill University, Department of Biology, 1205 Ave Docteur Penfield, Montréal, QC H3A 1B1
[*]Correspondence: gary.brouhard@mcgill.ca, paul.francois2@mcgill.ca

APPENDIX S1: DERIVATION OF THE FISHER INFORMATION MATRIX (FIM) EXPRESSION FOLLOWING TRANSTRUM *ET AL.* 2011 (1)

The FIM expression presented in the main text is a way to estimate how model parameters can be fitted to data, and by extension how models with different parameters are distinguishable. We will use the FIM in the latter sense, and in the following recall its connection to data fitting.

Assume we have a mathematical model of a biological system, with parameters $\vec{\theta}$, giving the probability $y\left(\vec{\theta}, x_i\right)$ of observable x to take the value $x = x_i$. We call d_i the experimentally measured probability of value x_i and assume that:

$$d_i = y\left(\vec{\theta}, x_i\right) + \sigma_i r_i \tag{1}$$

where we assume r_i to be a random Gaussian noise of mean 0 and variance 1, and σ_i a local variance so that:

$$r_i\left(\vec{\theta}\right) = \frac{d_i - y\left(\vec{\theta}, x_i\right)}{\sigma_i} \tag{2}$$

Assuming all r_i's are independent, the total probability $P(\vec{r}, \vec{\theta})$ of experimentally observing the values $\vec{d} = \{d_i\}$ given the residuals $\vec{r} = \{r_i\}$,is thus :

$$P\left(\vec{r}, \vec{\theta}\right) = \frac{1}{(2\pi)^{M/2}} \exp\left(-\frac{1}{2}\sum_{i=1}^{M} r_i\left(\vec{\theta}\right)^2\right) \tag{3}$$

where M is the number of points where we try to fit data.

The Fisher Information Matrix (FIM) defines the amount of information that the residuals r_i contain on parameters. Intuitively, the FIM tells us about the distinguishability of two parameter sets given the data. It is given by:

$$\begin{aligned} I_{\mu,\nu} &= \left\langle -\frac{\partial^2 \log P\left(\vec{r}, \vec{\theta}\right)}{\partial\theta_\mu \partial\theta_\nu} \right\rangle \\ &= -\int d\vec{r} P\left(\vec{r}, \vec{\theta}\right) \frac{\partial^2 \log P\left(\vec{\theta}, r_i\right)}{\partial\theta_\mu \partial\theta_\nu} \end{aligned} \tag{4}$$

Since r_i are Gaussian distributed, one can explicitly perform the computation of the FIM, which can then be interpreted as a metric $g_{\mu,\nu}$ quantifying the ability to distinguish between different parameter sets. Following Machta *et al.* we have:

$$\begin{aligned} g_{\mu,\nu} &= \left\langle -\frac{\partial^2 \log P\left(\vec{\theta}, \xi\right)}{\partial\theta_\mu \partial\theta_\nu} \right\rangle = \left\langle \frac{\partial^2 \sum_i \frac{1}{2} r_i^2}{\partial\theta_\mu \partial\theta_\nu} \right\rangle \\ &= \sum_i \left\langle r_i \frac{\partial^2 r_i}{\partial\theta_\mu \partial\theta_\nu} + \frac{\partial r_i}{\partial\theta_\mu}\frac{\partial r_i}{\partial\theta_\nu} \right\rangle \\ &= \sum_i \left\langle r_i \frac{\partial^2 r_i}{\partial\theta_\mu \partial\theta_\nu} \right\rangle + \sum_i \left\langle \frac{\partial r_i}{\partial\theta_\mu}\frac{\partial r_i}{\partial\theta_\nu} \right\rangle \end{aligned} \tag{5}$$

We substitute equation 2 into equation 5. The first sum cancels out because r_i is independent of $y\left(\vec{\theta}, x_i\right)$, d_i is independent of $\vec{\theta}$ and the expectation value of the residue r_i itself is zero. The second term becomes fully deterministic since with the same assumptions $\frac{\partial r_i}{\partial \theta_\nu} = \frac{\partial y\left(\vec{\theta}, x_i\right)}{\partial \theta_\nu}$. We arrive at the final expression for the Fisher Information Matrix of the model:

$$g_{\mu,\nu} = \sum_x \frac{\partial y\left(\vec{\theta}, x\right)}{\partial \theta_\mu} \frac{\partial y\left(\vec{\theta}, x\right)}{\partial \theta_\nu} \tag{6}$$

where we assume all σ_i to be equal (and rescaled to 1). The remarkable result is that this metric, while initially computed by averaging over experiments, is a pure function of the model y, and as a consequence, can be used independently of actual experiments to estimate in a deterministic way the distinguishability of models.

Biological models mix rates and energies, so the parameters are potentially of very different nature. Mixing units might yield purely dimensional effects in the analysis of important directions in parameter space. Energy is the most fundamental quantity and kinetic rates are exponentially related to the rescaled energy (in unit of k_BT). We take the derivatives with respect to the log of the parameter (*i.e.* $\frac{\partial y\left(\vec{\theta}, x_i\right)}{\partial \log \theta_{mu}} = \frac{\partial y\left(\vec{\theta}, x_i\right)}{\partial \theta_{mu}} \theta_{mu}$) to express every parameters in the Fisher Information Matrix with the same effective unit. Therefore, the final expression becomes:

$$g_{\mu,\nu} = \sum_x \frac{\partial y\left(\vec{\theta}, x\right)}{\partial \theta_\mu} \frac{\partial y\left(\vec{\theta}, x\right)}{\partial \theta_\nu} \theta_\mu^{\alpha_\mu} \theta_\nu^{\alpha_\nu} \tag{7}$$

where $\alpha_\mu = 0$ if θ_μ is an energy and $\alpha_\mu = 1$ if θ_μ is a kinetic rate.

APPENDIX S2: CONVERSION FOR ENERGY k_BT TO RATES s^{-1} IN THE MICROTUBULE MODEL

As shown in Appendix S1, biological models mix energies and rate constants as parameters. While energies are the most fundamental quantities for our Fisher Information Matrix computation, computing derivatives by incremental changes of energies might be challenging since small changes of energies could potentially give big changes of rates and thus of probability distributions. For numerical computations, we vary rate constants by a small increment α, and shift energies such as ΔG^o_{long}, ΔG^o_{lat} and $\Delta\Delta G^o_{lat}$ logarithmically. In order to ensure a smoother change of probability distribution, we need to come up with a conversion between changing bond energy and the change in rate.
As an example, the off rate of a dimer for the microtubule model is given as:

$$k_{off} = \frac{k^+}{e^{-\Delta G^o_{tot}}} \tag{8}$$

where k^+ is the apparent on rate and ΔG^o_{tot} is the total bond energy associated with the particular dimer.
Let us define the original off rate k^-_{ori} and the new off rate after changing the bond strength k^-_{new}. Assume the difference factor is defined as the following:

$$\alpha = \frac{k^-_{ori} - k^-_{new}}{k^-_{ori}} \tag{9}$$

which implies that

$$\begin{aligned} k^-_{new} &= k^-_{ori}(1-\alpha) \\ \Rightarrow \frac{k^+}{e^{-\Delta G^o_{new}}} &= \frac{k^+}{e^{-\Delta G^o_{ori}}}(1-\alpha) \\ \Rightarrow e^{\Delta G^o_{new}} &= e^{\Delta G^o_{ori}}(1-\alpha) \\ \Rightarrow \ln\left(e^{\Delta G^o_{new}}\right) &= \ln\left(e^{\Delta G^o_{ori}}(1-\alpha)\right) \\ \Rightarrow \Delta G^o_{new} &= \Delta G^o_{ori} + \ln(1-\alpha) \\ \Rightarrow \Delta\left(\Delta G^o\right) &= \ln(1-\alpha) \end{aligned} \tag{10}$$

Equation 10 gives us the change in energy corresponding to a relative change α. Practically, when computing derivatives such as the ones in equation 7 for the Fisher Information Matrix, we thus use a relative rate of α for kinetic rates, and of $\ln(1-\alpha)$ for energies. Notice that as α goes to 0 the relative change of energy also becomes zero. However, we can not pick an α that is too small due to numerical imprecision. Fig. S1 shows the relationship between the change in energy and the ratio of change in rates. For example, a 5% change in rate constant is around 0.05 k_BT change in energy.

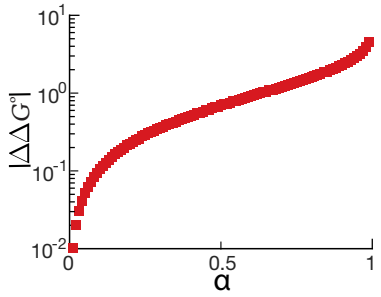

Figure S1. The conversion plot for various percentage change for parameters that have a unit of k_BT. The conversion rate is universal and it does not depend on the value of the parameter itself. Note: a 5% change in parameter value is about $0.05\ k_BT$.

APPENDIX S3: ANALYTIC CALCULATION OF EIGENVALUES FOR A ONE DIMENSIONAL DIFFUSION SYSTEM AT THE FIRST TIME STEP

For the one dimensional random walk where the particles diffuse uniformly, the probability density after one time step is proportional to the number of particles at each possible lattice site. Therefore, $P(N_\mu) = \frac{N_\mu}{N_{total}} = \kappa$ where N_μ is the number of particles that are in each possible lattice sites and total number of particles $N_{total} = \sum_\mu N_\mu$.

Using equation 7 from Appendix S1 and taking into account that after the first time step only the diagonal elements of the Fisher Information Matrix are nonzero, equation 7 simplifies to:

$$g_{\mu,\nu} = \frac{\partial y}{\partial \theta_\mu}\frac{\partial y}{\partial \theta_\nu}\theta_\mu\theta_\nu\delta_{\mu\nu} \tag{11}$$

where $\delta_{\mu\nu}$ is the Kronecker delta.
Therefore, the corresponding eigenvalues are:

$$\begin{aligned}
\lambda_\mu = g_{\mu,\mu} &= \left[\frac{y_{\theta+\Delta\theta} - y_{\theta-\Delta\theta}}{2\Delta\theta}\right]^2\theta^2 \\
&= \left[\frac{\frac{N_\mu+\Delta N_\mu}{N_{tot}} - \frac{N_\mu-\Delta N_\mu}{N_{tot}}}{\frac{2\Delta N_\mu}{N_{tot}}}\right]^2 \\
&= \left[(N_\mu + \Delta N_\mu) - (N_\mu - \Delta N_\mu)\right]^2 \frac{N_\mu^2}{4\Delta N_\mu^2 N_{tot}^2} \\
&= (2\Delta N_\mu)^2\left(\frac{N_\mu^2}{4\Delta N_\mu^2 N_{tot}^2}\right) \\
&= \frac{N_\mu^2}{N_{tot}^2} \\
&= P(N_\mu)^2
\end{aligned} \tag{12}$$

This shows that the eigenvalue of a one dimensional random walk at the first time step is equal to the probability at a given lattice site squared. The result for both the drift right diffusion and uniform diffusion with different number of sites are shown in Figure 2C and D.

APPENDIX S4: ANALYTIC CALCULATION OF THE DOMINANT EIGENVALUE FOR THE PROTEIN PRODUCTION AND DEGRADATION SYSTEM

The stationary distribution of a simple production and degradation system is given by the Poisson distribution (2):

$$P(\rho, \delta, n) = \frac{1}{n!}e^{-\frac{\rho}{\delta}}\left(\frac{\rho}{\delta}\right)^n \tag{13}$$

Therefore, the Fisher Information Matrix for this system from equation 7 becomes:

$$\begin{aligned}&\begin{bmatrix} \sum_n \left(\frac{\partial P_n}{\partial \rho}\rho\right)^2 & \sum_n \frac{\partial P_n}{\partial \rho}\rho\frac{\partial P_n}{\partial \delta}\delta \\ \sum_n \frac{\partial P_n}{\partial \rho}\rho\frac{\partial P_n}{\partial \delta}\delta & \sum_n \left(\frac{\partial P_n}{\partial \delta}\delta\right)^2 \end{bmatrix} \\ &= \begin{bmatrix} \sum_n \alpha_n^2 & \sum_n \alpha_n\beta_n \\ \sum_n \alpha_n\beta_n & \sum_n \beta_n^2 \end{bmatrix}\end{aligned} \tag{14}$$

where $\alpha_n = \frac{\partial P_n}{\partial \rho}\rho$ and $\beta_n = \frac{\partial P_n}{\partial \delta}\delta$.

To find the eigenvalues λ of this matrix, we subtract the identity matrix with diagonal value λ and set the determinant equals to zero:

$$\det \begin{bmatrix} \sum_n \alpha_n^2 - \lambda & \sum_n \alpha_n\beta_n \\ \sum_n \alpha_n\beta_n & \sum_n \beta_n^2 - \lambda \end{bmatrix} = 0 \tag{15}$$

We can calculate α_n and β_n analytically and show that the magnitudes of the two terms are equal to each other:

$$\alpha_n = |\beta_n| = P(n)\frac{(\rho - n\delta)}{\delta} \tag{16}$$

Thus the determinant matrix becomes the following form, where $A = -\alpha_n\beta_n$:

$$\det \begin{bmatrix} A - \lambda & -A \\ -A & A - \lambda \end{bmatrix} = 0 \tag{17}$$

The eigenvalues for this matrix are easy to compute with only one nonzero eigenvalue:

$$\lambda_1 = 2A = 2\sum_n P(n)^2\frac{(\rho - n\delta)^2}{\delta^2} \tag{18}$$

To calculate analytically the nonzero eigenvalue: λ_1, we approximate the Poisson distribution with a Gaussian distribution and change the integral into a summation with $\gamma = \frac{\rho}{\delta}$:

$$\begin{aligned}\lambda_1 &= 2\sum_n P^2\frac{(\rho - n\delta)^2}{\delta^2} \\ &\simeq 2\int_0^\infty \left(\frac{1}{\sqrt{2\pi\gamma}}e^{\frac{(x-\gamma)^2}{2\gamma}}\right)^2\left(\frac{(\rho - x\delta)^2}{\delta^2}\right)dx \\ &= \frac{1}{\pi}\left[\frac{1}{4}\sqrt{\pi\gamma}erf\left(\frac{x}{\sqrt{\gamma}} - \sqrt{\gamma}\right) + \frac{1}{2\delta}e^{-\frac{(x-\gamma)^2}{\gamma}}(\rho - \delta x)\right]_0^\infty \\ &= \frac{1}{2}\sqrt{\frac{\gamma}{\pi}} - \frac{\gamma}{2\pi}e^{-\gamma}\end{aligned} \tag{19}$$

If the value of $\gamma \gg 0$, the second term becomes negligible due to the exponential. Thus, the final expression for the eigenvalue is:

$$\lambda_1 \simeq \frac{1}{2}\sqrt{\frac{\gamma}{\pi}} = \frac{1}{2}\sqrt{\frac{\rho}{\delta\pi}} \tag{20}$$

APPENDIX S5: MICROTUBULE SIMULATION

We use VanBuren *et al.* 2002 as our inspiration for our simulation of microtubule dynamic instability (3).

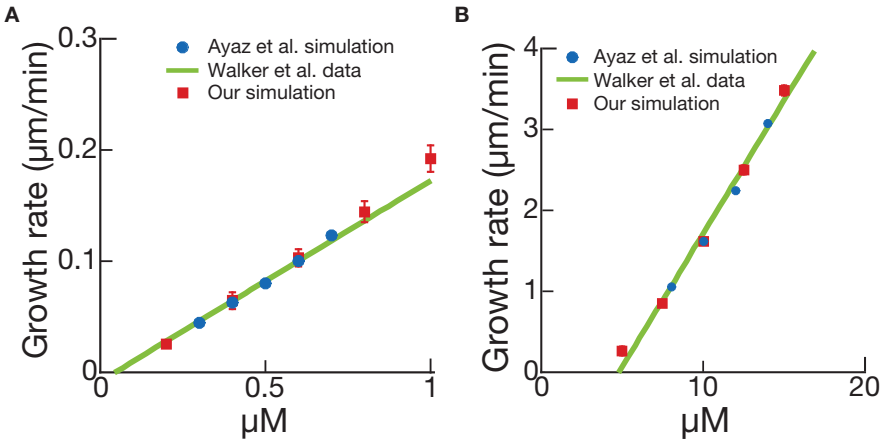

Figure S2. Microtubule simulation bench mark against computer simulation of Ayaz *et al.* (4) and experimental data of Walker *et al.* (5) (A) GMPCPP tubulin (B) GTP tubulin.

Parameters

The microtubule has 13 protofilaments with a three monomer offset at the seam. The base parameters in the model are: (1) k^+, the association rate constant for tubulin subunts to associate with the end of a protofilament; (2) ΔG^o_{long}, the longitudinal bond energy between dimers; (3) ΔG^o_{lat}, the lateral bond energy between tubulin subunits in a B-lattice configuration ($\alpha - \alpha$ and $\beta - \beta$); (4) k_H, the hydrolysis rate constant for the conversion of GTP-tubulin to GDP-tubulin; (5) $\Delta\Delta G^o_{lat}$, the change in free energy associated with GTP hydrolysis, which is assigned to each lateral bond; and (6) $[Tubulin]$, the concentration of tubulin.

Coupled Random Hydrolysis

An α-tubulin contributes catalytic residues to the GTP pocket of the β-tubulin that sits below it. Therefore, a tubulin subunit cannot hydrolyze its GTP unless another tubulin subunit is above it; in other words, only non-terminal subunits can hydrolyze GTP. The GTP hydrolysis reaction for all non-terminal subunits occurs at random time intervals (see Gillespie algorithm below). This implementation of GTP hydrolysis is known as coupled random hydrolysis (6). We do note, however, that earlier models of dynamic instability used other implementations of GTP hydrolysis, e.g., where the hydrolysis reaction was obligate after association of a new terminal subunit, known as vectorial hydrolysis. But coupled random hydrolysis is the standard among contemporary models.

Our model assumes the transition from GTP-tubulin to GDP-tubulin occurs in a single step, with no intermediates in the hydrolysis pathway. A single step is surely a simplification, as GTPases often have important GDP-Pi intermediate states. Recently, Manka et al. solved a cryo-EM structure of the putative GDP-Pi state (7), and very recent models have incorporated a GDP-Pi state explicitly (8). Testing the relevance of a GDP-Pi state will be the subject of future studies.

Lateral Weakening

The effect of GTP hydrolysis in the VanBuren model is a weakening of lateral bonds. This choice is based on the observation that protofilaments peel outward after a catastrophe (9), which indicates that the lateral bonds rupture first. More recent observations suggest that longitudinal bonds may also be affected by GTP hydrolysis; more specifically, the N-domain of α-tubulin appears to compact down into the β-tubulin below it.

Gillespie algorithm

We use the direct method of the Gillespie algorithm to simulate all the possible events for the microtubule at a given time (10). The possible events for the microtubule simulation are: association of dimers, dissociation of dimers, and hydrolysis of GTP-tubulin to GDP-tubulin. The association and dissociation events only happen at the ends of protofilaments. The association rate is equal on the top of each protofilament and is defined as the association constant k^+ multiplies by the concentration of tubulin:

$$k_{on,PF} = k^+[Tubulin] \tag{21}$$

The dissociation rate depends on ΔG^o_{total}: the total bond energies the tubulin dimer have with its neighboring dimers:

$$k_{off} = \frac{k^+}{e^{-\frac{\Delta G^o_{total}}{k_B T}}} \tag{22}$$

The total number of hydrolysis events depend on the number of GTP tubulin that are present in the lattice and a nonterminal GTP can be hydrolyzed into a GDP at a fixed first order rate k_H which leads to a total rate of hydrolysis at any given time as:

$$k_{hyd} = k_H \times N(GTP) \tag{23}$$

The effect of hydrolysis is a decrease in the bond strength laterally by $\Delta\Delta G^o$. This is commonly known as the coupled random hydrolysis model.

We sum up all the possible rates r_i for all the possible events αand generate two random number R_1, R_2 between zero and one. We choose the events that satisfy the following condition:

$$\frac{\sum_0^{i-1} r_i}{\alpha} \leq R_1 < \frac{\sum_0^{i} r_i}{\alpha} \tag{24}$$

and we increment the simulation time t by τ, where τ is defined as

$$\tau = \frac{1}{\alpha} \log\left(\frac{1}{R_2}\right) \tag{25}$$

Benchmarking

Because we use the direct method of the Gillespie algorithm, we benchmarked our Gillespie simulation against published results of Ayaz *et al.* (4) and VanBuren *et al.* (3) to ensure that our simulation was properly implemented. We used their precise parameter values and simulated microtubule growth in the absence of GTP hydrolysis (which is handled differently in the two models). Our simulation recovered their results exactly (see Fig. S2A and S2B). Therefore our simulation is well-executed.

Reproduction of Experimental Data

Simple models like those described here can microtubule growth rates and post-catastrophe shrinkage rates across a range of tubulin concentrations. The models can reproduce the mean lifetime at a single tubulin concentration, but then the trouble begins. These models cannot reproduce the mean lifetime across a range of tubulin concentrations. The models are much too sensitive to $[Tubulin]$–at high $[Tubulin]$, catastrophes become exceedingly rare, in contrast with experimental observations. More complex models do better (11–13).

APPENDIX S6: MICROTUBULE SIMULATION ANALYSIS

We use an in house Matlab code to analyze the microtubule trajectories over time. The algorithm smoothies the trajectories and identifies local maxima and minima. The maxima correspond to potential catastrophe positions and minima correspond to potential rescue positions. Lifetime and post-catastrophe shrinkage rate is then determined. Note: to avoid fluctuation of the stochastic simulation, a threshold of 250 *nm* is used (since this is the point spread function of our microscope) such that any lifetime that is counted towards the probability distribution starts from a minima to a maxima while passing 250 *nm*. Similarly, for the post-catastrophe shrinkage rate but instead it goes from a maxima to a minima.

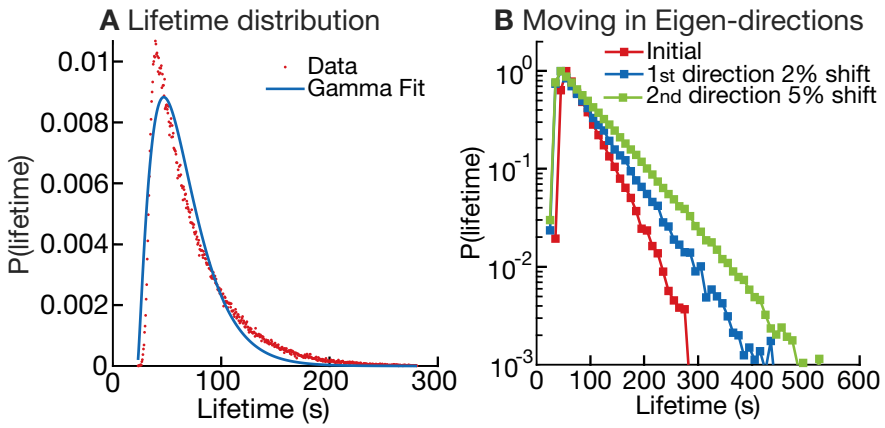

Figure S3. (A)The lifetime distribution is close to a gamma distribution which indicates it is indeed two parameters (B) Moving along the two dominating Eigenvector directions. The distribution changes subtly but clearly different which indicate it is a two direction distribution.

REFERENCES

1. Transtrum, M. K., B. B. Machta, and J. P. Sethna, 2011. Geometry of nonlinear least squares with applications to sloppy models and optimization. *Physical Review E* 83:036701.

2. Klipp, E., W. Liebermeister, C. Wierling, and A. Kowald, 2016. Systems biology : a textbook. Wiley-VCH, Weinheim.

3. VanBuren, V., D. J. Odde, and L. Cassimeris, 2002. Estimates of lateral and longitudinal bond energies within the microtubule lattice. *Proc Natl Acad Sci U S A* 99:6035–40.

4. Ayaz, P., S. Munyoki, E. A. Geyer, F.-A. Piedra, E. S. Vu, R. Bromberg, Z. Otwinowski, N. V. Grishin, C. A. Brautigam, and L. M. Rice, 2014. A tethered delivery mechanism explains the catalytic action of a microtubule polymerase. *eLife* 3:e03069.

5. Walker, R. A., E. T. O'Brien, N. K. Pryer, M. F. Soboeiro, W. A. Voter, H. P. Erickson, and E. D. Salmon, 1988. Dynamic instability of individual microtubules analyzed by video light microscopy: rate constants and transition frequencies. *J Cell Biol* 107:1437–48.

6. Bowne-Anderson, H., M. Zanic, M. Kauer, and J. Howard, 2013. Microtubule dynamic instability: a new model with coupled GTP hydrolysis and multistep catastrophe. *Bioessays* 35:452–61.

7. Manka, S. W., and C. A. Moores, 2018. The role of tubulin-tubulin lattice contacts in the mechanism of microtubule dynamic instability. *Nat Struct Mol Biol* 25:607–615.

8. Kim, T., and L. M. Rice, 2018. A role for long-range, through-lattice coupling in microtubule catastrophe. *bioRxiv* .

9. Mandelkow, E. M., E. Mandelkow, and R. A. Milligan, 1991. Microtubule dynamics and microtubule caps: a time-resolved cryo-electron microscopy study. *J Cell Biol* 114:977–91.

10. Gillespie, D. T., 1977. Exact stochastic simulation of coupled chemical reactions. *The Journal of Physical Chemistry* 81:2340–2361.

11. Gardner, M., B. Charlebois, I. Jánosi, J. Howard, A. Hunt, and D. Odde, 2011. Rapid Microtubule Self-Assembly Kinetics. *Cell* 146:582–592.

12. Castle, B. T., S. McCubbin, L. S. Prahl, J. N. Bernens, D. Sept, and D. J. Odde, 2017. Mechanisms of kinetic stabilization by the drugs paclitaxel and vinblastine. *Mol Biol Cell* 28:1238–1257.

13. Zakharov, P., N. Gudimchuk, V. Voevodin, A. Tikhonravov, F. I. Ataullakhanov, and E. L. Grishchuk, 2015. Molecular and Mechanical Causes of Microtubule Catastrophe and Aging. *Biophys J* 109:2574–91.

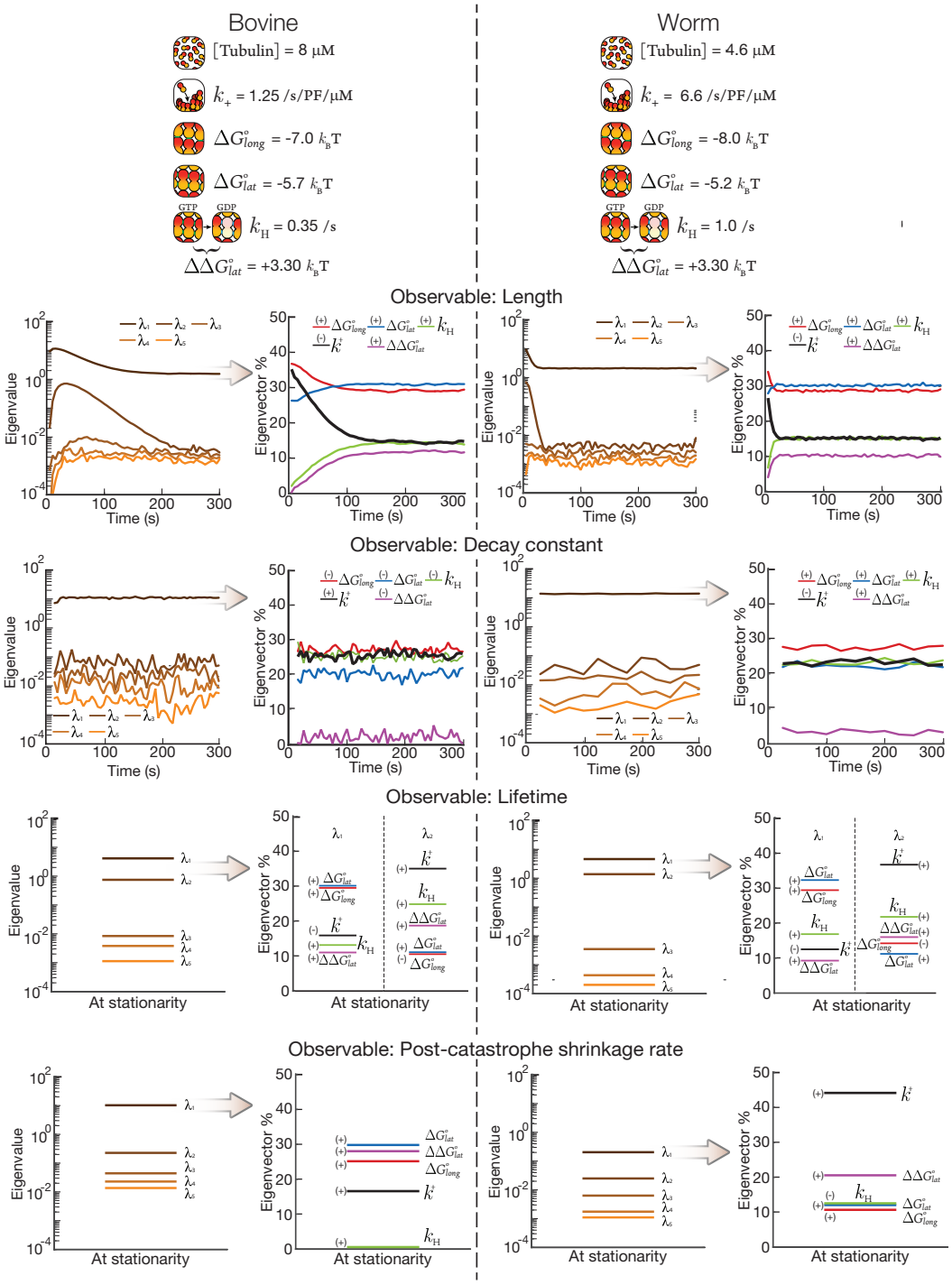

Figure S4. The comparison between the parameters and the FIM analysis for bovine and worm. The result between the two parameter sets are consistent with one another even given the completely different parameter value.

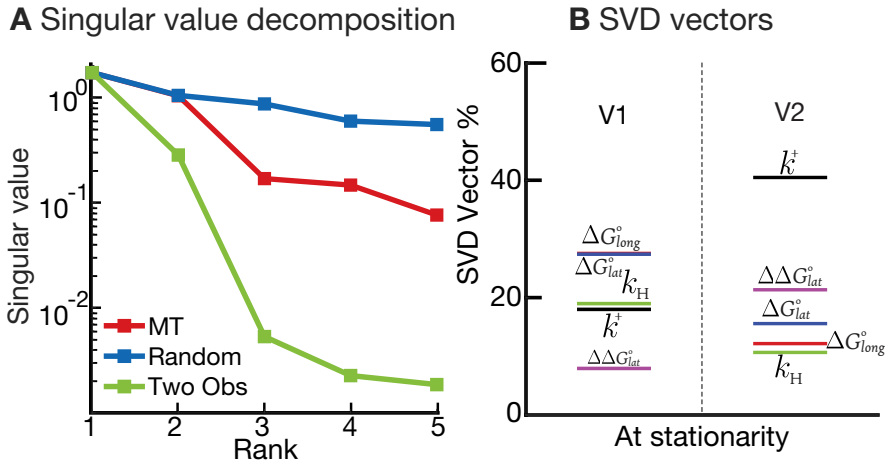

Figure S5. SVD analysis for the worm tubulin dataset. The system is two dimensional like the bovine data set, however, the components are not exactly the same with k^+ much more dominating in this system.

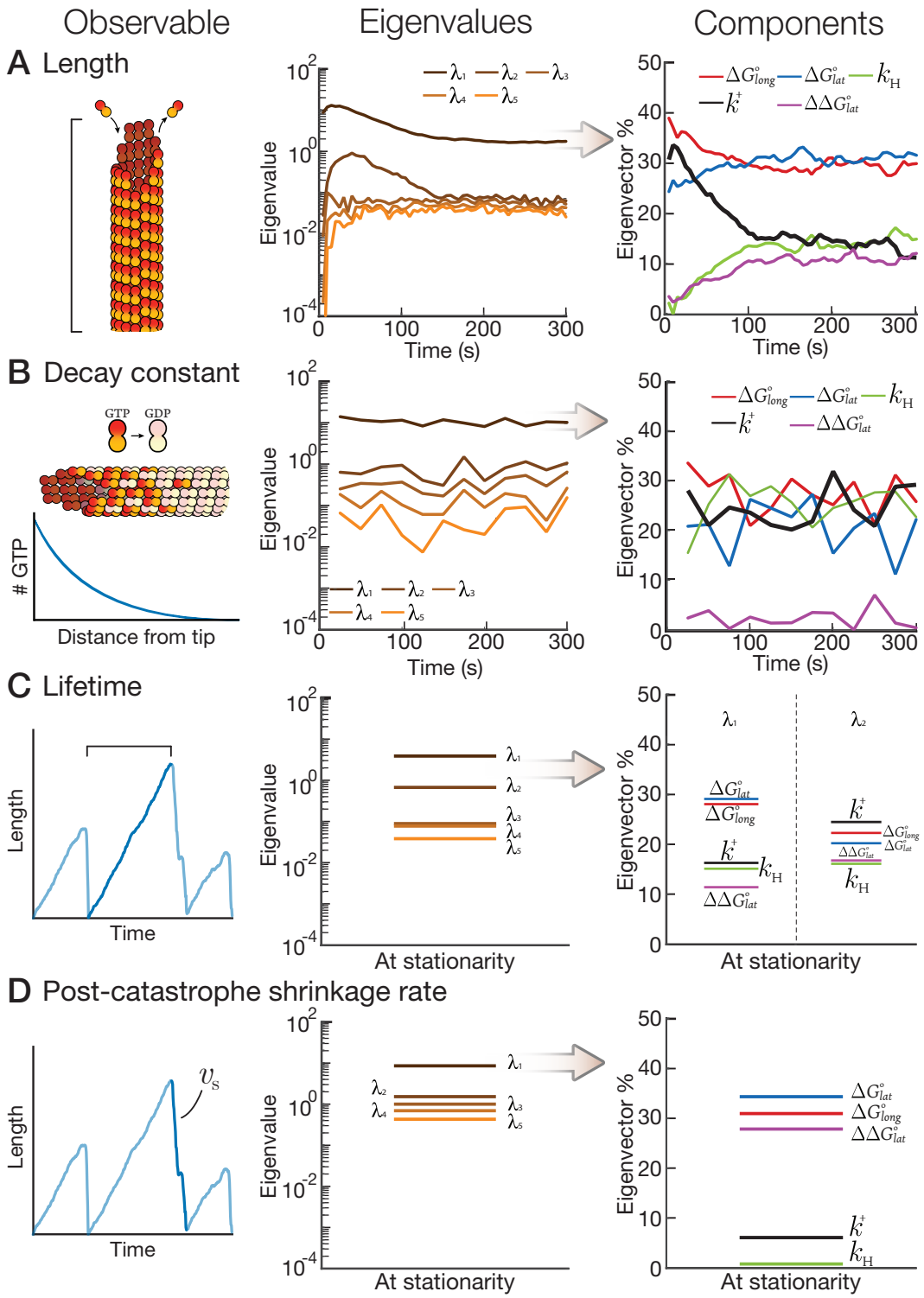

Figure S6. shows the FIM analysis of the microtubule system when shifting the parameter for $0.01\ k_BT$. The result is a lot nosier compare with our choice of $0.05\ k_BT$.