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.

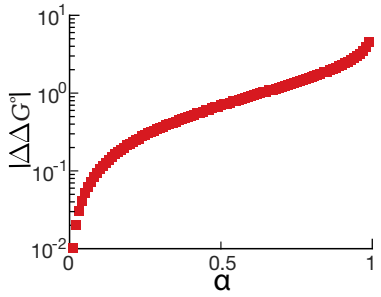

Figure S1. The conversion plot for various percentage change for parameters that have a unit of k_BT. The conversion rate is universal and it does not depend on the value of the parameter itself. Note: a 5% change in parameter value is about $0.05\ k_BT$.

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).

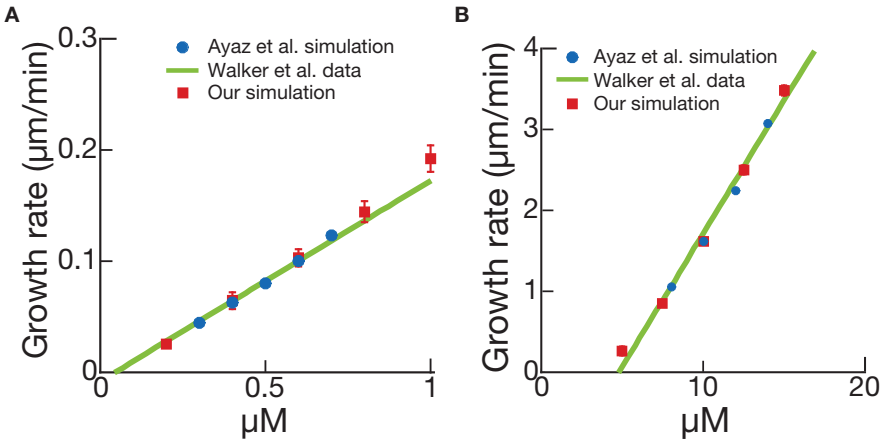

Figure S2. Microtubule simulation bench mark against computer simulation of Ayaz *et al.* (4) and experimental data of Walker *et al.* (5) (A) GMPCPP tubulin (B) GTP tubulin.

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.

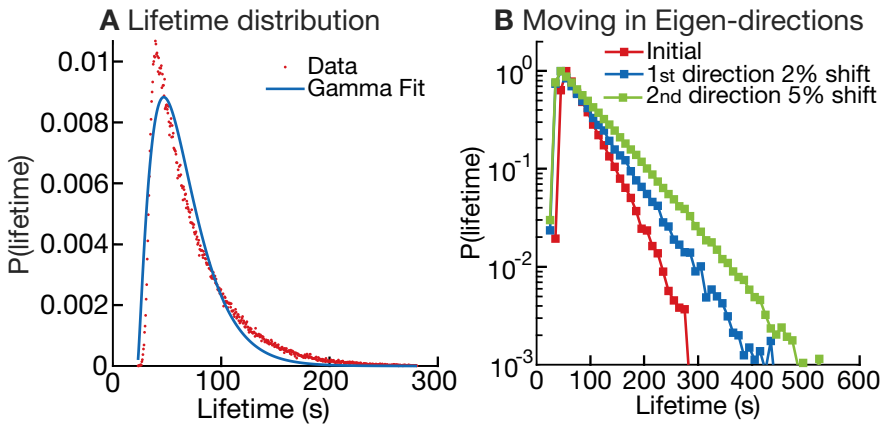

Figure S3. (A)The lifetime distribution is close to a gamma distribution which indicates it is indeed two parameters (B) Moving along the two dominating Eigenvector directions. The distribution changes subtly but clearly different which indicate it is a two direction distribution.

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.

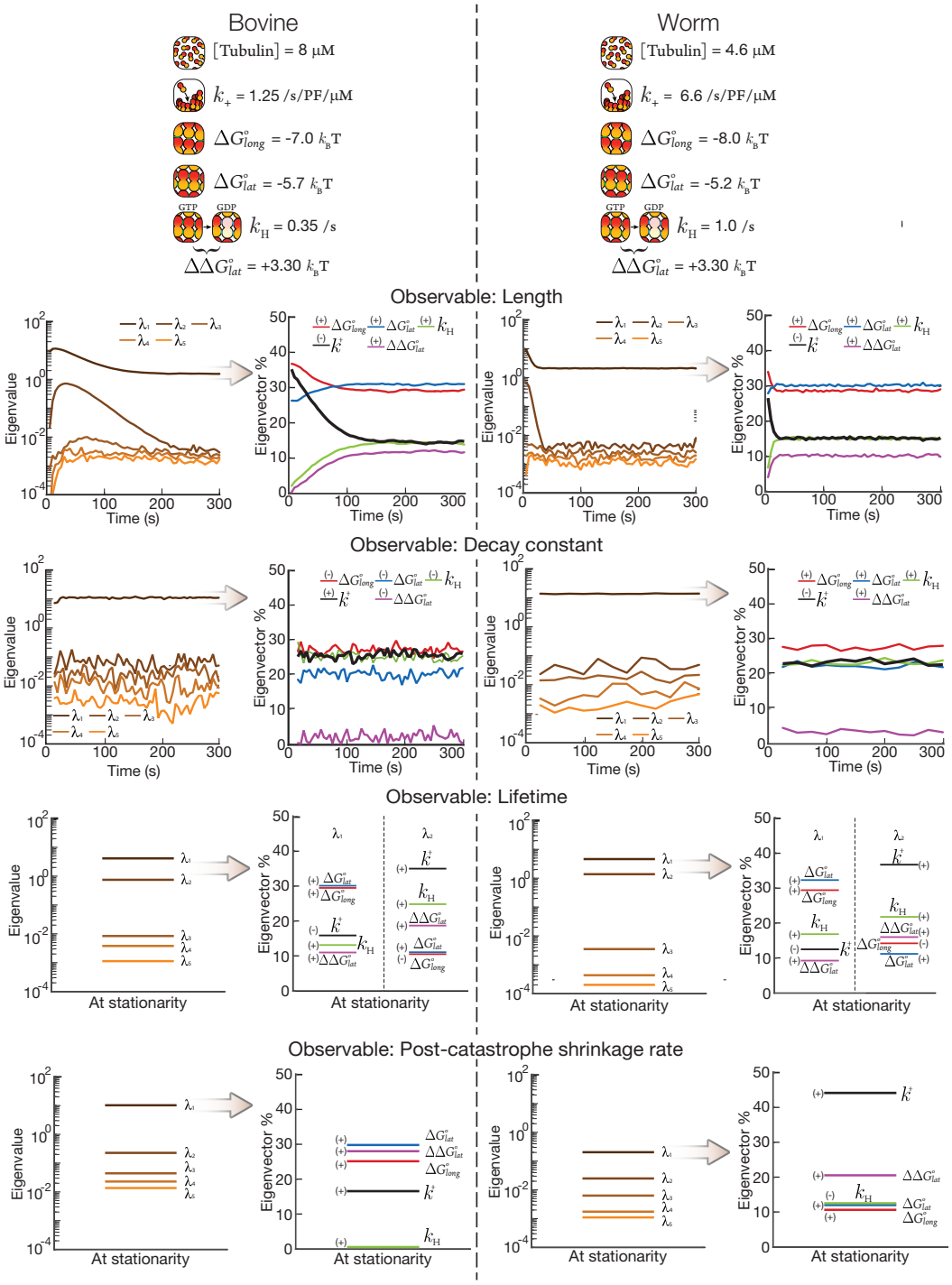

Figure S4. The comparison between the parameters and the FIM analysis for bovine and worm. The result between the two parameter sets are consistent with one another even given the completely different parameter value.

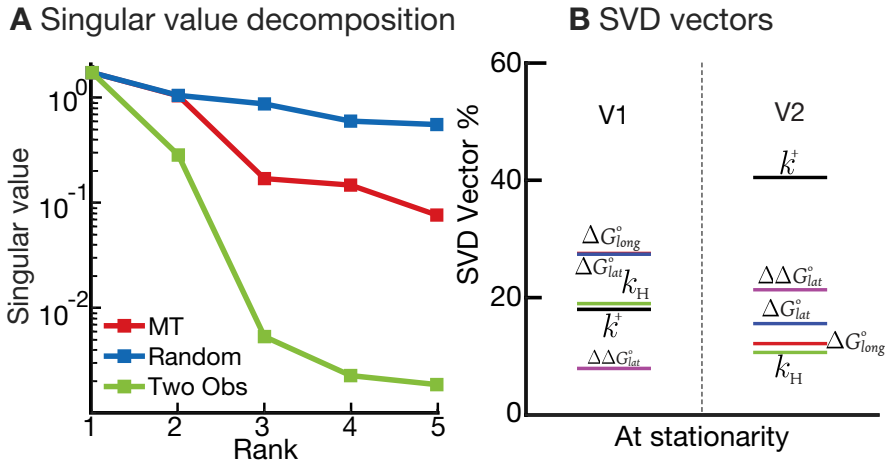

Figure S5. SVD analysis for the worm tubulin dataset. The system is two dimensional like the bovine data set, however, the components are not exactly the same with k^+ much more dominating in this system.

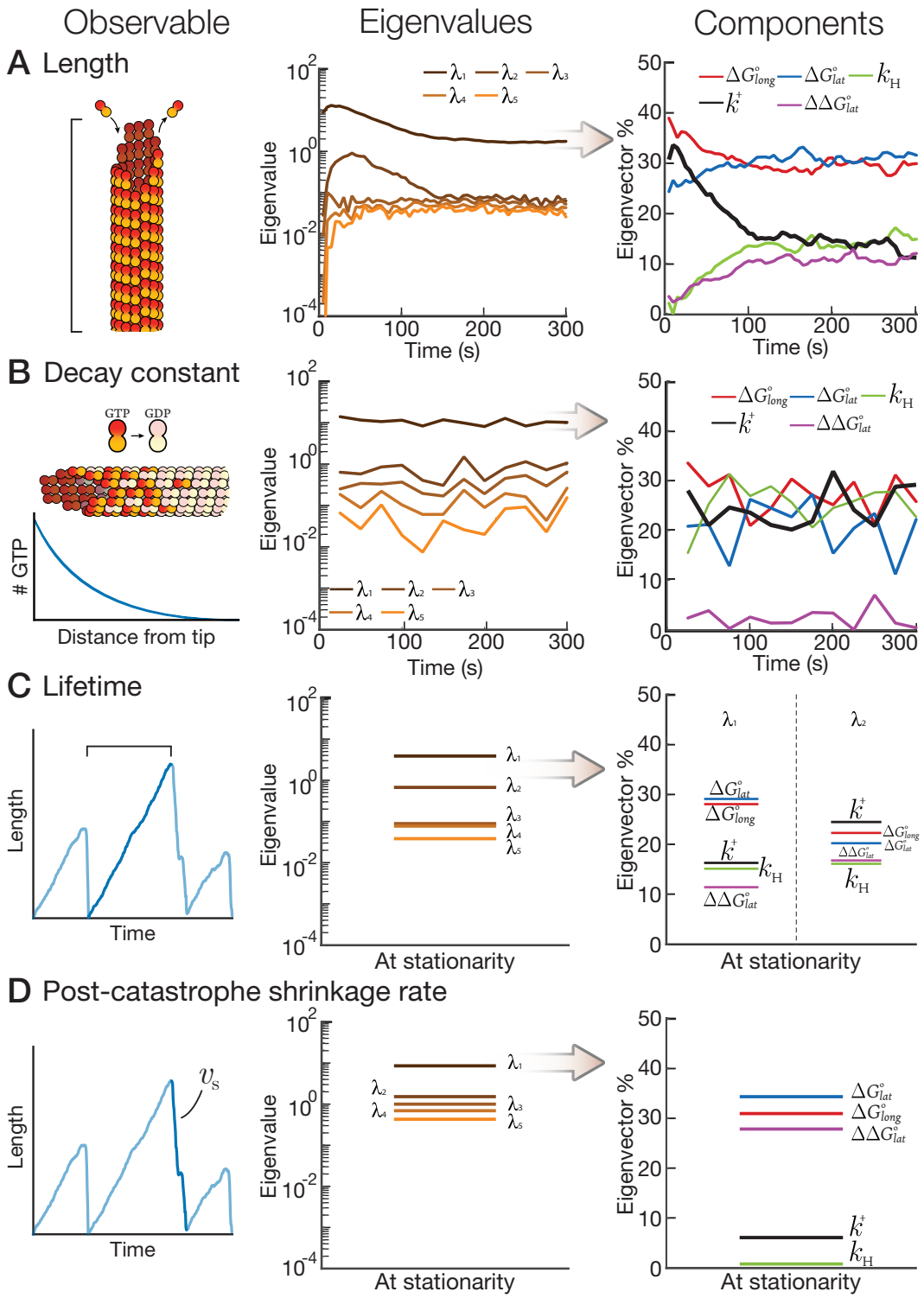

Figure S6. shows the FIM analysis of the microtubule system when shifting the parameter for $0.01\ k_BT$. The result is a lot nosier compare with our choice of $0.05\ k_BT$.