Spectral Analysis Methods for the Robust Measurement of the Flexural Rigidity of Biopolymers

David Valdman,[†] Paul J. Atzberger,[†¶*] Dezhi Yu,[‡] Steve Kuei,[§] and Megan T. Valentine[¶*]
[†]Department of Mathematics, [‡]Department of Materials, [§]Research Internships in Science and Engineering, Materials Research Laboratory, and [¶]Department of Mechanical Engineering, University of California, Santa Barbara, California

ABSTRACT The mechanical properties of biopolymers can be determined from a statistical analysis of the ensemble of shapes they exhibit when subjected to thermal forces. In practice, extracting information from fluorescence microscopy images can be challenging due to low signal/noise ratios and other artifacts. To address these issues, we develop a suite of tools for image processing and spectral data analysis that is based on a biopolymer contour representation expressed in a spectral basis of orthogonal polynomials. We determine biopolymer shape and stiffness using global fitting routines that optimize a utility function measuring the amount of fluorescence intensity overlapped by such contours. This approach allows for filtering of high-frequency noise and interpolation over sporadic gaps in fluorescence. We use benchmarking to demonstrate the validity of our methods, by analyzing an ensemble of simulated images generated using a simulated biopolymer with known stiffness and subjected to various types of image noise. We then use these methods to determine the persistence lengths of taxol-stabilized microtubules. We find that single microtubules are well described by the wormlike chain polymer model, and that ensembles of chemically identical microtubules show significant heterogeneity in bending stiffness, which cannot be attributed to sampling or fitting errors. We expect these approaches to be useful in the study of biopolymer mechanics and the effects of associated regulatory molecules.

INTRODUCTION AND BACKGROUND

Cytoskeletal polymers, including actin and microtubules, are stiff, multistranded filaments that are essential to cell organization, motility, and division; to the transport of intracellular cargos by motor proteins; and to the generation and transmission of forces within and across cells. Because of their important role in establishing and regulating cellular mechanics, the elasticity of filaments and entangled cytoskeletal networks has been studied extensively. However, many important questions remain about the relationships between structure and mechanics (1). In particular, although in vitro measurements of single filament elasticity have consistently shown significant variations in stiffness over roughly an order of magnitude, the molecular origins of these variations are incompletely understood (2–7). This is largely a result of an inability to distinguish real heterogeneity in elasticity from variations that arise from sources of experimental and statistical uncertainty.

For microtubules (MTs), the stiffest cytoskeletal filaments, distinguishing between signals and noise is particularly challenging because the exhibited bending amplitudes are small and often of a comparable magnitude to experimental noise. MTs are formed from the head-to-tail polymerization of tubulin dimers in long protofilaments that interact laterally to form a closed tubular structure, with outer diameter of roughly 25 nm (8–11). Structural studies have demonstrated that the number of MT protofilaments varies within in vitro and in vivo systems, and can even change along the length of a single MT (12–14). Because the bending stiffness of a biopolymer scales as the fourth power of its radius, even small changes in the effective radius of the MT could have a large mechanical effect. Under some in vitro conditions, MT stiffnesses appear to depend on both the length and the polymerization velocity, suggesting that lattice shear and structural defects may also play an important role (5,15–17). Unfortunately, the large variation in experimental estimates of MT stiffness values has severely compromised our ability to correlate changes in mechanical response and filament composition, and has led to an incomplete understanding of the origins and regulation of MT mechanics.

To help address these issues, we have developed what to our knowledge are new methods to determine and analyze the motions of stiff, isolated, thermally fluctuating biopolymers visualized using fluorescence microscopy. Using approaches from statistical mechanics, we then infer mechanical properties from a spectral analysis of the ensemble of biopolymer configurations at thermal equilibrium. A central challenge in using this approach is the sensitivity of the spectral analysis methods to experimental noise (4,18). Previous methods to characterize the fluctuation spectra of biopolymer filaments made use primarily of local information in the fluorescent image, and typically involved three distinct operations: Step 1. Filament tracing to determine biopolymer configurations. Step 2. Calculation of spectral components from estimated contours. Step 3. Analysis of the spectral components using statistical mechanics to ascertain mechanical properties (4,5,18). Step 1 typically

Submitted June 28, 2011, and accepted for publication January 24, 2012.

*Correspondence: atzberg@math.ucsb.edu or valentine@engineering.ucsb.edu

Editor: Gijsje Hendrika Koenderink.

0006-3495/12/03/1144/10 $2.00

doi: 10.1016/j.bpj.2012.01.045

involved interpolating individual configuration points that were determined by manual selection or by local fitting of the cross-sectional intensity profiles of fluorescently labeled filaments (4,5,18). Because each control node is fit independently, any local aberration (such as irregularity in fluorophore labeling or the interference of a nearby physical object) in the image near a control point could have a large influence on the entire fitted contour. Though these approaches have been successfully employed to provide important mechanical information for many types of biopolymers, for the stiffest biopolymers, results vary by roughly an order of magnitude, motivating the development of improved fitting and analysis methods (7,20).

In this work, we develop a unified approach that combines Steps 1 and 2 into a single procedure, thus minimizing the introduction of errors. This is achieved by representing the biopolymer shape in terms of a contour expanded in an orthogonal polynomial basis. We fit the coefficients of the expansion directly to the fluorescence image by optimizing a utility function that measures the overlap of the entire contour with fluorescence intensity. This global approach is more robust to local disturbances in the fluorescence signal, and allows for the determination of biopolymer shape even in the presence of significant noise and artifacts in the images. To estimate the flexural rigidity of the biopolymer, we develop a statistical mechanics theory based on an energy formulated directly in terms of the coefficients of our spectral biopolymer representation.

Our approach has a number of distinct advantages. The description of the biopolymer by a smooth contour that is fit to the entire image at once naturally handles sporadic gaps in fluorescence intensity along the biopolymer through interpolation, and shows improvement in both filament tracing and determination of persistence length as compared to pointwise tracing. Because these methods are more robust against common sources of experimental noise, it is possible to analyze a larger number of images, giving rise to smaller sampling errors and improving measurement precision. Additionally, we achieve a more accurate description of the biopolymer near the end-points through the use of a basis of orthogonal polynomials. This is in contrast to the use of a trigonometric Fourier basis that implicitly requires a periodic function, and thus a no-curvature condition at the contour end-points to reduce spurious oscillations arising from Gibb's Phenomena (4,18,21). Our method can, in principle, allow for curvature near the end-points, and could potentially facilitate studies of a wider class of biopolymers.

To assess the sensitivity of the proposed methods to experimental noise and sampling error, we introduce a benchmarking approach in which an ensemble of simulated images is generated from simulations of a fluctuating biopolymer with known mechanical properties. We introduce noise and artifacts into the images that are similar to those observed in experimental data. This approach allows for the systematic study of the roles played by different types of experimental noise, and the resultant uncertainty of estimated mechanical properties. The benchmarking approach we propose provides a potentially powerful metric for rating different spectral analysis methods and for understanding the statistical significance of differences reported in experimental results. We then apply the techniques to an experimental data set of fluorescence images of fluctuating MTs. For even these stiff biopolymers, we find our methods reliably estimate the flexural rigidity and produce modal covariances in agreement with a wormlike-chain (WLC) model for MT mechanics. Interestingly, we find that the level of heterogeneity in persistence length in an ensemble of chemically identical MTs vastly exceeds that predicted from algorithmic error analysis, indicating that significant structural heterogeneity may be present in these samples. We anticipate that these approaches will facilitate the development of more sensitive assays based on the thermal fluctuations of biopolymers, and will enable central questions concerning the molecular origins of cytoskeletal mechanics to be answered.

METHODOLOGY

Variational contour fitting method

In developing this method, we have designed our algorithm for use with images of stiff, isolated filaments, such as MTs, in which a direct determination of the tangent-angle decorrelation function is not possible. Instead, we use information obtained from a description of the biopolymer shape in terms of a curve $\mathbf{x}(s)$ of length L, where $s \in [0,L]$ is the arc-length parameter of the contour. To measure how well the contour $\mathbf{x}$ overlaps with the fluorescence signal of the biopolymer, we use the utility function

$$U[\mathbf{x}, I] = -\int_0^L \int_\Omega k(|\mathbf{y} - \mathbf{x}(s)|) I(\mathbf{y}) d\mathbf{y} ds. \qquad (1)$$

The fluorescence image intensity is given by $I = I(\mathbf{y})$ parameterized over the spatial domain Ω and $k(r)$ is a smoothing kernel vanishing for $r > r_0$. The inner-integral of Eq. 1 gives the average intensity in a region near the location $\mathbf{x}(s)$ by using the weighting specified by $k(r)$. The outer-integral collects these values to provide a measure of the total amount of overlap of a contour with the biopolymer fluorescence signal. The convolution by $k(r)$ with I filters the high spatial-frequency noise inherent in the image intensity.

We use the following kernel function (see Fig. 1), which has a number of desirable properties when used for discrete pixel maps and lattice models (see (22,23)):

$$k(r) = \begin{cases} \alpha[1 + \cos(\pi r / r_0)] & r \le r_0 \\ 0 & r > r_0. \end{cases} \qquad (2)$$

Here, r_0 is chosen to be approximately equal to the width of the imaged polymer. In the case of MTs, with a diameter of ~25 nm, r_0 is taken to be approximately equal to the width of the point-spread function of the microscope. The α is a normalization constant ensuring the kernel function weighting integrates to one.

The contour configuration that minimizes the utility function U provides a fit that maximizes the overlap within the image between the contour and

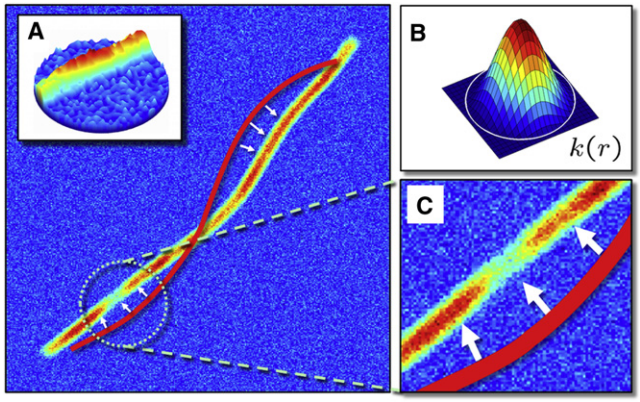

FIGURE 1 Schematic of the methodology. (*A*) The intensity of the fluorescence signal of the biopolymer (*inset*) and the fluorescence microscopy image being fit by a trial contour (*solid line*). (*Arrows*) Direction of evolution of the trial contour when using the method of steepest descent for the utility function given in Eq. 1. (*B*) The radially symmetric kernel function $k(r)$ having cut-off radius r_0 (indicated by *dotted circle*). (*C*) Close-up of fluorescence image data within the range of influence given by r_0. In this case, the fluorescence image exhibits gaps in intensity along the biopolymer that are handled naturally by the utility function and the inherent interpolation of the trial contour.

the biopolymer fluorescence signal. This minimization requires a representation for the contour amenable to calculations, so we represent the contour by its tangent angle $\theta(s)$ along the length and by a reference point at the left end-point $\mathbf{x}_0 = \mathbf{x}(0)$. These physically meaningful degrees of freedom uniquely specify the contour curve $\mathbf{x}(s)$,

$$\mathbf{x}(s) = \int_0^s \boldsymbol{\tau}(\theta(s'))ds' + \mathbf{x}_0. \tag{3}$$

The tangent vector τ for a given angle θ is given by

$$\boldsymbol{\tau}(\theta) = (\cos(\theta), \sin(\theta)).$$

This representation is used to minimize the utility function by evolving the degrees of freedom $(\theta(s), \mathbf{x}_0)$ using the steepest-descent dynamical equations

$$\begin{aligned} \frac{\partial\theta(s)}{\partial t} &= -\frac{\delta U}{\delta\theta}(s) \\ \frac{\partial\mathbf{x}_0}{\partial t} &= -\nabla_{\mathbf{x}_0} U. \end{aligned} \tag{4}$$

The term $\delta U/\delta\theta$ denotes the variational derivative of the utility function, which generalizes the usual vector derivative and captures how values change when the entire collection of tangent vectors along the contour are varied (24). The $\nabla_{\mathbf{x}_0} U$ denotes the usual vector derivative in $\mathbf{x}_0$ of the utility function.

It can be shown that for all possible variations of $(\theta(s), \mathbf{x}_0)$, the direction in configuration space giving the most rapid decrease in the utility function U is the negative of the gradient

$$-\nabla_{\theta(s),\mathbf{x}_0} U = -\left(\frac{\delta U}{\delta\theta}, \nabla_{\mathbf{x}_0} U\right).$$

This motivates the choice for the dynamics, which ensures the contour configuration moves in a manner that steadily decreases the value of U over time. The limiting contour configuration that is stationary under these dynamics has a zero gradient and is a critical point of the utility function U. Such a contour is a candidate for minimizing U (24).

To work with this description in practice, we expand the tangent angles in an orthogonal polynomial basis

$$\theta(s) = \sum_n a_n T_n(s). \tag{5}$$

Each $T_n(s)$ is a polynomial of degree n satisfying the orthonormal inner-product condition $\langle T_i, T_j\rangle = \delta_{ij}$, where δ_{ij} is the Krönecker delta-function. The T_n values are shown in Fig. 2. In practice, this expansion is truncated after a finite number of terms N, allowing for representation up to degree N polynomials. Dynamical equations are obtained readily for the coefficients $a_n(t)$ by plugging this expansion into Eq. 4 and projecting the direction of evolution on the polynomials up to degree N (see the Supporting Material).

A particularly useful feature of this coefficient representation is that even when only a finite number of coefficients are used, the contour recovered by Eqs. 3 and 5 has total arc-length L throughout the minimization procedure. Experimentally, this is a very reasonable constraint, as MT length is typically held constant by the use of small molecule inhibitors (such as taxol) during the course of the measurement. Although the total length of the polymer is held constant, the fitted contour can slide freely within the fluorescence signal of the illuminated filament during fitting. In practice, we determine the filament length from the average of the length of the contours obtained from our initial fits, then truncate by a fixed amount (typically by 10–20% of the initial length) to obtain a working length for final fitting. This truncation is implemented to avoid possible high-frequency oscillations that can sometimes arise due to compression of the fitted contour within the polymer after photobleaching has dimmed the fluorescence signal at the polymer ends. Even in the absence of spurious oscillations, the fluorescence signal at the ends of the microtubule tends to be of poor quality. By using a slightly shortened fitting contour, we avoid corrupting the spectral results with MT blurring and other imaging artifacts. Using benchmarking, we have determined the effect of contour truncation on the calculated persistence length, and for the simulated and experimental images analyzed here, the errors are typically <2–3%.

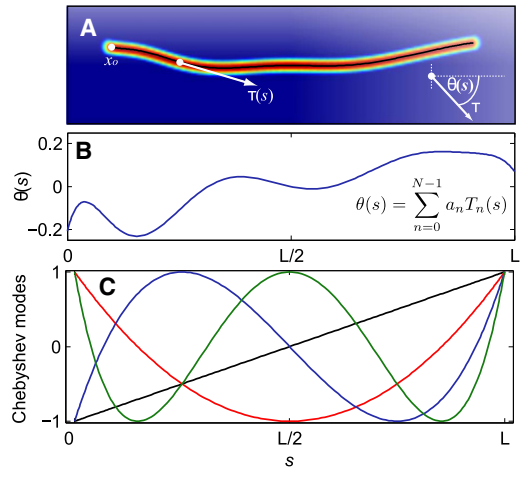

FIGURE 2 Representation of the biopolymer shape using orthogonal polynomials. (*A*) The contour tracing out the shape of the biopolymer is described by parameterizing the shape using arc-length s and the local angle $\theta(s)$ of the tangent vector τ with respect to the x axis. The point $\mathbf{x}_0$ is used to uniquely determine the position of the contour. (*B*) To work with this description in practice, we expand $\theta(s)$ in an orthogonal polynomial basis to obtain coefficients a_n. (*C*) We use the Chebyshev orthogonal polynomial basis for this purpose. The first four nonconstant orthogonal polynomial modes are shown.

In presenting our approach, we use throughout the Chebyshev orthogonal polynomials defined by

$$T_n(s) = \cos\left(n \arccos\left(\left(\frac{2s}{L}\right) - 1\right)\right)$$

(see Trefethen (21)). This choice was motivated by the ability to take advantage of fast Fourier transform methods; however, other orthogonal polynomial bases could also be used. For more details concerning the particular forms of the variational derivatives used and truncation of the dynamics for a finite number of coefficients, see the Supporting Material.

Determining persistence length from the spectrum of biopolymer fluctuations

Using results from equilibrium statistical mechanics, we can estimate the elastic properties of isolated, thermally fluctuating biopolymers (4). We focus here on determining the persistence length; however, these methods can be applied more generally to other mechanical moduli. To describe the elastic responses of biopolymers, we use the WLC model (25), which associates to a given biopolymer configuration $\mathbf{x}(s)$ a bending energy

$$E_{bend}[\mathbf{x}] = \frac{EI}{2}\int_0^L \left(\dot{\theta}(s)\right)^2 ds. \tag{6}$$

The EI denotes the flexural rigidity and $\dot{\theta}(s) = d\theta/ds$ denotes the derivative of the tangent angle in s. For an isotropic elastic structure, EI is the product of the Young's modulus E and the geometric moment of inertia I. At thermodynamic equilibrium, the biopolymer thermal fluctuations have a Gibbs-Boltzmann distribution with the probability density

$$\rho_{bend}[\mathbf{x}] = \frac{1}{Z}\exp\left[\frac{-E_{bend}[\mathbf{x}]}{k_B T}\right], \tag{7}$$

where T is the temperature, k_B is the Boltzmann constant, and Z denotes the partition function (see Reichl (26)).

Using our representation of the biopolymer configuration $\mathbf{x}(s)$ in terms of coefficients of the orthogonal polynomial expansion (see Eq. 5), the energy can be expressed as

$$\begin{aligned} E_{bend}[\mathbf{a}] &= \frac{EI}{2}\mathbf{a}^T S\mathbf{a} \\ S_{ij} &= \int_0^L \dot{T}_i(s)\dot{T}_j(s)ds. \end{aligned} \tag{8}$$

The term $\mathbf{a}$ denotes the composite vector of coefficients with $[\mathbf{a}]_n = a_n$ and S denotes the stiffness matrix of the biopolymer modes. The entries S_{ij} are given by the L^2-inner product of the orthogonal polynomials with index i and j, so the matrix is not necessarily diagonal. For example, for our choice of Chebyshev polynomials the off-diagonal entries of S are nonzero.

The Gibbs-Boltzmann distribution can be expressed using this coefficient representation as

$$\rho_{bend}[\mathbf{a}] = \frac{1}{\tilde{Z}}\exp\left[-\frac{1}{2}L_p\mathbf{a}^T S\mathbf{a}\right], \tag{9}$$

where $L_p = EI/k_B T$ gives the persistence length of the correlations of fluctuations along the contour and $\tilde{Z}$ is the partition function of this representation. Throughout, we treat the metric of the generalized coordinates as constant. In this form, we see that ρ_{bend} has the convenient form of a multivariate Gaussian with mean zero and covariance

$$\left\langle \mathbf{a}\mathbf{a}^T \right\rangle = \frac{1}{L_p}S^{-1}. \tag{10}$$

In our analysis, we find it convenient for finite contours of length L to define a nondimensional persistence length $\ell_p = L_p/L$. The covariance structure for biopolymer fluctuations derived from the WLC model and Eq. 10 is given by

$$C_{wlc}(\ell_p) = \frac{1}{\ell_p}\tilde{S}^{-1}. \tag{11}$$

We have used the nondimensional WLC stiffness matrix defined by $\tilde{S} = LS$. This provides a covariance structure predicted by the WLC model when the nondimensional persistence length is ℓ_p.

In experiments, the covariance is estimated by fitting contours to the fluorescence images and estimating modal coefficients. For M samples, the covariance is estimated by

$$C_{exp} = \frac{1}{M}\sum_m \mathbf{a}^{(m)}\mathbf{a}^{(m)T}. \tag{12}$$

The $\mathbf{a}^{(m)}$ denotes the m^{th} sampled modal coefficient.

A central relation we use to interpret experimental fluctuations of a biopolymer and to infer its mechanical properties is

$$C_{exp} = C_{wlc}(\ell_p). \tag{13}$$

This expression provides the key link between observed biopolymer fluctuations (left-hand side) and the biopolymer mechanical properties (right-hand side). To infer mechanical properties in experiments, we seek to find a value of ℓ_p so that C_{wlc} matches, to a good approximation, the covariance of the experimentally observed biopolymer fluctuations C_{exp}.

This requires minimizing the least-squares error given by

$$V(\ell_p) = \sum_n \left(c_n - \frac{1}{\ell_p}d_n\right)^2. \tag{14}$$

In principle, any one entry in the covariance matrix could be used to determine persistence length, although it is possible to make use of redundant terms to increase accuracy (because there is sampling error in the coefficients), at the expense of additional computation time. We find in practice that it is sufficient to consider just the diagonal entries, thus the $c_n = [C_{exp}]_{n,n}$ are the diagonal entries of the covariance matrix for the experimentally observed biopolymer fluctuations. The $d_n = [\tilde{S}^{-1}]_{n,n}$ are the coefficients used for representing the covariances obtained from the WLC model (see Eq. 11). The fit for ℓ_p is obtained by minimizing $V(\ell_p)$ and is given by

$$\ell_p = \frac{\|\mathbf{d}\|^2}{(\mathbf{d}\cdot\mathbf{c})}. \tag{15}$$

We use composite vector notation for the experimental covariance data $[\mathbf{c}]_n = c_n$ and for the coefficients of the WLC model $[\mathbf{d}]_n = d_n$,

$$\|\mathbf{d}\|^2 = \sum_n d_n^2 \text{ and } \mathbf{d}\cdot\mathbf{c} = \sum_n c_n d_n.$$

We remark that other approaches making use of additional features beyond the covariance matrix of the WLC model could also be used to estimate the persistence length, such as the maximum-likelihood method (27).

RESULTS AND DISCUSSION

Role of sampling error on estimated persistence length

There are two fundamentally independent sources of error to consider for this analysis. The first is the inherent problem of sampling error, where the derivation of mechanical quantities leads to errors due to an averaging over a finite number of experimental observations. The second is the effect of different types of noise in the microscopy images on the estimated values of physical quantities.

To determine the error in the estimated biopolymer persistence length ℓ_p in terms of the number of sampled images M, we assume a biopolymer exhibits fluctuations given by the WLC model, and estimate the covariance using Eq. 12. The sampling error can be expressed as

$$\tilde{C}_M = C_{wlc}(\ell_p) + \boldsymbol{\xi}. \tag{16}$$

The $C_{wlc}(\ell_p)$ denotes the covariance structure obtained from the WLC model in Eq. 7. For a sufficient number of samples, ξ is approximately a Gaussian with mean zero and covariance $\langle \boldsymbol{\xi}\boldsymbol{\xi}^T \rangle = \mathrm{Cov}(\mathbf{a}\mathbf{a}^T)/M$. Expressions for this covariance can be obtained by computing the fourth moments of the Gaussian distribution given in Eq. 7. The tilde notation will be used throughout to distinguish variables that model quantities that would be estimated experimentally.

To simplify the presentation, we assume that the components of $[\mathbf{a}]_n$ can be treated as statistically independent. Additionally, we describe our theory only for estimates of the diagonal entries of the covariance matrix, which are the only entries used in the least-squares fitting. We denote the diagonal entries of the covariance by the vectors $\tilde{\mathbf{c}}_M = \mathrm{diag}(\tilde{C}_M)$ and $\mathbf{c}_{wlc} = \mathrm{diag}(C_{wlc})$. We model the covariance estimates for M samples by

$$\tilde{\mathbf{c}}_M = \mathbf{c}_{wlc}(\ell_p) + \boldsymbol{\xi}. \tag{17}$$

The ξ denotes a Gaussian with independent components each having mean zero and covariance $\langle \boldsymbol{\xi}\boldsymbol{\xi}^T \rangle = D_M$. The covariance has diagonal entries

$$[D_M]_{n,n} = E[\xi_n^2] = 2[\mathbf{c}_{wlc}(\ell_p)]_n^2/M \tag{18}$$

with the off-diagonal entries zero.

This model can be used to study how sampling errors from the estimate of the modal covariances $\tilde{\mathbf{c}}_M$ propagate into the estimation of the persistence length $\tilde{\ell}_p$. In particular, from the least-squares fit of Eq. 15, we have

$$\tilde{\ell}_p^{-1} = \frac{(\mathbf{d}\cdot\tilde{\mathbf{c}}_M)}{\|\mathbf{d}\|^2} = \ell_p^{-1} + \sum_n \frac{\xi_n d_n}{\|\mathbf{d}\|^2}. \tag{19}$$

This shows the estimated inverse persistence length $\tilde{\ell}_p^{-1}$ is a Gaussian-distributed quantity with mean $\tilde{\mu} = \ell_p^{-1}$ and variance $\tilde{\sigma}_M^2 = \mathbf{d}^T D_M \mathbf{d}/\|\mathbf{d}\|^4$.

The estimated persistence length $\tilde{\ell}_p$ has the probability distribution

$$\rho_M(\ell_p) = \frac{\ell_p^{-2}}{\sqrt{2\pi\tilde{\sigma}_M^2}} \exp\left[-\frac{(\ell_p^{-1} - \tilde{\mu})^2}{2\tilde{\sigma}_M^2}\right]. \tag{20}$$

Note that this distribution is not Gaussian; instead, it has long-tails as a consequence of the ${\ell_p}^{-2}$ term and yields an infinite variance. For different values of M, this distribution is shown in Fig. S1 in the Supporting Material.

The non-Gaussian form of the distribution requires that some care is taken when characterizing how the sampling errors influence the estimated value of ℓ_p. We can no longer make use of the standard deviation to give the magnitude of errors because the second moment is infinite. Instead, we use a confidence interval based on the above probability distribution ρ_M. Interestingly, although the second moment is infinite, the distribution ρ_M can be well approximated by a Gaussian distribution for M sufficiently large ($M > 100$; see Fig. S1). Through an asymptotic analysis of Eq. 20, as M becomes large, we find ρ_M is approximated by a Gaussian with mean $\mu = \ell_p$ and variance $\sigma_M^2 = \ell_p^4\tilde{\sigma}_M^2 = \ell_p^4\mathbf{d}^T D_M \mathbf{d}/\|\mathbf{d}\|^4$. This can be used in practice to obtain confidence intervals for errors in estimates of $\tilde{\ell}_p$.

Generating simulated fluorescence images with controlled levels of noise and artifacts

To investigate effects of image noise on persistence length determination, we develop a systematic benchmarking approach, in which we generate an ensemble of simulated images from the fluctuations of a simulated biopolymer with known mechanical properties. We introduce in these images simulated background noise and other artifacts. We then use our contour tracing and spectral analysis methods to determine the persistence length. A comparison of the measured and actual persistence lengths provides a well-controlled test for the reliability of the proposed methods, and enables experimental conditions to be optimized to improve measurement precision and accuracy.

To generate noise and artifacts similar to those encountered in experiments, we consider primarily two types of noise: 1), background noise in which randomly varying levels of intensity are seen throughout the image, and 2), sporadic gaps in which intensity varies in the fluorescence signal along the biopolymer. The background noise is motivated by contributions from ambient light sources and out-of-focus fluorophores, whereas gap artifacts are motivated by the uneven binding of fluorescent labels along the biopolymer, inherent fluorescence excitation inefficiencies, and stochastic photobleaching.

To simulate the configurations of a biopolymer having a specified persistence length, we generate modal coefficients for our orthogonal polynomial representation of the

contour using the multivariate Gaussian distribution with mean zero and covariance given by Eq. 10. Throughout our presentation, we use $\ell_p = 10$ and calculate values for the first eight modes ($N = 8$). To obtain an image with a fluorescence signal for the biopolymer of thickness r_0, we use the kernel function of Eq. 2 to trace along the biopolymer contour. For convenience, we normalize all fluorescence intensity values to lie between zero and one. We have also simulated filament contours using a random walk generator, and we find no significant differences in algorithm performance, thereby verifying that there is no bias introduced when generating the contours by using the Chebyshev basis.

To introduce background noise, we perturb each pixel value using two different Gaussian distributions. For pixels that are located on the contour, we perturb by a random value having mean $\mu_c = 0$ and variance ${\sigma^2}_c$. For pixels not located on the contour, we perturb using mean $\mu_b = 0$ and variance ${\sigma^2}_b = {\sigma^2}_c$. To characterize the noise, we define I_c to be the characteristic intensity difference between the contour and background $I_c = \langle I(\mathbf{y})\rangle_c - \langle I(\mathbf{y})\rangle_b$, where $\langle\cdot\rangle_\ell$, $\ell \in \{b,c\}$ gives, respectively, the average of intensity values over the contour or background pixels. To characterize the level of background noise in our images, we use the ratio σ_b/I_c.

To introduce gap noise along the contour, we modulate the fluorescence signal by a weight function obtained from a cosine series

$$c(s) = \sum_{k=0}^{K-1} w_k \cos\left(\frac{2\pi k s}{L}\right).$$

The random weight coefficients w_k are chosen so that the integral of $c(s)$ over $[0,L]$ has mean μ_g and variance ${\sigma^2}_g$. The form of $c(s)$ necessitates that $w_0 = \mu_g$. To achieve a variance proportional to ${\sigma^2}_g$, we use $(w_1,\ldots,w_{K-1})$ uniformly distributed over the surface of a $K-1$ dimensional sphere of radius $\sqrt{2}\sigma_g$. To control how oscillatory the gap artifacts appear in the image, we vary the number of modes K. The random coefficients ensure random amplitude and phases for each of the modes, creating an irregular, realistic gap pattern. We find that $K = 25$ provides a modulation that agrees well with what is seen in experimental fluorescence images. To characterize the level of gap artifacts in our images we use the ratio σ_g/μ_g.

Benchmarking studies for different levels of noise and artifacts

To investigate the robustness of our methodology, we numerically generate ensembles of fluorescence images with prescribed noise conditions, similar to those observed in experiments (see Fig. 4). In contrast to prior works that have modeled the role of noise on fitting using theoretical assumptions (4), this benchmarking approach provides a direct and realistic comparison with actual microscopy data. The ensembles contain $M = 1000$ images, simulating the thermal fluctuations of a biopolymer with persistence length $L_p = 9.45\ L$, where L is the polymer length (taken to be "1" in arbitrary units). To minimize the effects of sampling errors in these studies, we use the same underlying configurations for the simulated biopolymer to generate each ensemble of images. A sample image from each of these ensembles is shown in Fig. 3.

To initialize the contour tracing routine, an initial contour is chosen as the diagonal of a bounding box that encloses the Z brightest pixels (Z is typically chosen to be 2500). Because our algorithm is optimized to identify stiff, isolated filaments, the bounding box is automatically drawn around the filament of interest. We calculate the average intensity at each corner, and choose the diagonal that connects the

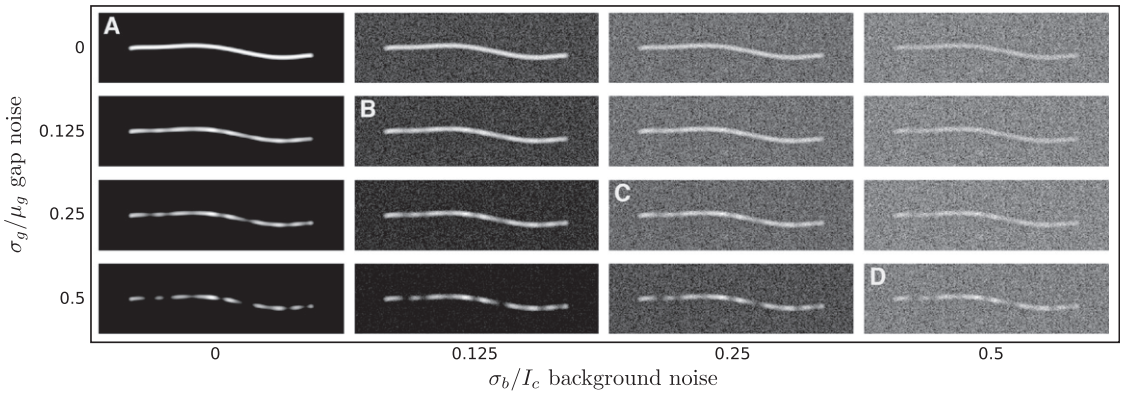

FIGURE 3 Ensembles of simulated fluorescence images. To investigate our spectral analysis methods when subjected to the types of noise and artifacts found in experimental images, we simulate ensembles of images from a simulated biopolymer with known mechanical properties. Shown here are images generated from biopolymer configurations using our contour representation with $\ell_p = 10, N = 8$. The ensembles correspond to different levels of background noise throughout the image and gap artifacts along the biopolymer contour. The background noise increases left to right while the gap artifacts increase from top to bottom. The level of background noise is characterized by the ratio σ_b/I_c for the perturbed background pixel standard deviation σ_b and the average difference of intensity between the contour and background I_c. The level of gap artifacts is characterized by the ratio σ_g/μ_g obtained by integrating the random cosine modulations used to generate the artifacts to obtain an effective mean μ_g and standard deviation σ_g.

two brightest corners as our initial contour for fitting. This diagonal typically overlays the actual filament, with numerous points of intersection and the length of the diagonal is roughly equal to the actual filament length.

To determine the effect of initial conditions on filament tracing, we tested 300 perturbations to the best initial fit of a single noisy numerically generated fluorescence image, then recorded the maximal difference between the final fit and the actual filament position to determine the robustness of our tracking algorithm. For the perturbations, the angle of the initial fit was chosen to lie within a range of $[-\pi/8, \pi/8]$ around the actual filament position, whereas the midpoint was independently perturbed to lie within a circle of radius 20 pixels from the actual center. For noise levels, we chose values that were similar to those observed experimentally. Of these, 90% of the initial fits converged to within two pixels (maximal deviation) of the correct contour, whereas 5% fit a portion of the filament well but failed to track the filament ends, and another 5% were drawn in regions dominated by noise and thus failed to converge to anything reasonable. This reproducibility demonstrates that our steepest-descent algorithm requires only a single point of intersection between the initial contour and fluorescence signal of the polymer to reliably trace the polymer shape.

Using an unperturbed initial fit, we find that in the absence of any explicitly introduced background noise or gap artifacts, our methods yield an estimated nondimensional persistence length of $\tilde{\ell}_p = 9.38$. We use this value (which is well within the expected range given our sampling error of 5.2% for $M = 1000$ images) as our reference when reporting relative errors to remove the baseline sampling error from the reported results and better reflect the differences in the levels of noise and artifacts in the images. For each simulated ensemble of images, we performed the spectral analysis using the initial five Chebyshev modes, which, for the stiff filaments we consider here, represent physically meaningful fluctuations. We report the results of the spectral analysis using our methods for each of the image ensembles in Table 1. When performing fits, we find that introduction of noise and gap artifacts can, in a small number of cases, result in images of insufficient quality. For such images (typically <5% of the ensemble), the contour fitting does not converge in a manner that is independent of the initial trial contour. These images are ignored when estimating the modal covariances and persistence length. In the case of experimental data, there is an initial manual review of images to eliminate those which show significant blurring of the filament contour or the presence of a very large interfering object. Additionally, we use only those mode variances that lie within ≈ 3 standard deviations of the mean, to eliminate the possibility that a rare fitting artifact could disproportionately change our stiffness results. In practice, very few images (<10%) are rejected based on these criteria.

As shown in Table 1, our global fitting approach enables robust determination of persistence length, even in the presence of substantial background noise and gap artifacts, which are frequently encountered in experiments. We find relative errors for the estimated persistence length to be <1% for a majority of low-to-moderate noise cases, and <10% even for the largest background and gap noise levels probed. We find similarly good agreement for simulated filaments that are an order-of-magnitude stiffer and tested using an even larger range of gap and background noise conditions and sample sizes (see Table S3 in the Supporting Material).

To further demonstrate the level of improvement in filament tracing and persistence length determination in comparison to more conventional approaches, we generated and analyzed ensembles of 500 images of filaments with known stiffness, and subjected them to known noise-background and gap-noise artifacts. To assess the quality of contour fitting, we used two methods: the global contour tracing routines described above, and a pointwise fitting routine in which the filament position was determined by fitting a Gaussian distribution to the cross-sectional intensity profile of the filament at 200 points along the contour. We compared the coordinates outputted from each fit routine

TABLE 1 Results for the simulated ensembles of fluorescence images

Gap noise	L_P	% Err	% Conv	L_P	% Err	% Conv	L_P	% Err	% Conv	L_P	% Err	% Conv
0	9.38	0.00	100	9.34	0.48	100	9.37	0.10	100	10.1	6.99	98
0.125	9.39	0.07	100	9.36	0.22	100	9.42	0.43	99	9.84	4.66	98
0.25	9.39	0.15	99	9.38	0.03	99	9.45	0.77	98	10.2	8.22	96
0.5	9.43	0.55	94	9.42	0.45	94	9.71	3.35	93	10.6	10.7	89
						Background noise						
		0			0.125			0.25			0.5	

To test the robustness of our spectral analysis methods, we simulated ensembles of fluorescence images with varying levels of background noise and gap artifacts generated from the fluctuations of a simulated biopolymer with known persistence length, $L_p = 9.45\ L$, where L is the polymer length (taken to be "1" in arbitrary units). At each condition we use $M = 1000$ simulated images. To simulate the configurations of a biopolymer having a specified persistence length, for each image we generate modal coefficients for our orthogonal polynomial representation of the contour using the multivariate Gaussian distribution with mean zero and covariance given by Eq. 10. We calculate values for the first eight Chebyshev modes ($N = 8$). We report the relative errors (% Err) in these estimates and the percentage of images found to be of sufficient quality to allow for convergent fits independent of the initial trial contour (% Conv). Relative errors are reported with respect to the baseline case of no background noise or gap artifacts. The table indices correspond to the same ordering and ratios used for characterizing the ensembles of images shown in Fig. 3.

(x_{fit}, y_{fit}) to the coordinates used to generate the filaments (x_o, y_o) and calculated the root mean square (rms) error at each pixel position to be

$$E_{rms} = \sqrt{(x_{fit} - x_o)^2 + (y_{fit} - y_o)^2}.$$

For both fitting methods, we found the mean value of E_{rms} to be <1 pixel; however, we find the fit errors to be substantially lower when using our global fitting routine (as shown in Table S1). We see a factor-of-2 improvement in contour fitting under low noise conditions, and a 10–20% improvement in even the noisiest cases. To complete the comparison, we then input the filament coordinates determined via pointwise fitting into a Fourier-based spectral analysis routine. We found reasonable agreement with the known persistence length (<10% deviation) for all noise levels when using only the first Fourier mode. However, we found the higher modes to be unreliable, particularly under levels of high background and gap noise (see Table S2). Noise corruption in the higher modes limits the use of averaging to improve measurement accuracy and prevents the identification of deviations from WLC behavior. The ability to average the response of the higher-order modes, each of which fluctuates independently, is particularly important, as sampling error dominates measurement uncertainty in many cases.

Experimental results for microtubules

To demonstrate our approach in practice, we apply our filament tracing and spectral analysis methods to characterize the persistence length of an isolated microtubule (MT) imaged using total internal reflection fluorescence microscopy. The MT was labeled with rhodamine dyes to enable visualization, and confined to move within a thin, well-sealed sample chamber to ensure that the fluctuating filament remained in focus throughout the experiment and only thermal forces acted upon the MT. The MTs were polymerized in the presence of the stabilizing compound taxol, rendering their length constant over the time course of the measurement, which is typically limited by fluorophore bleaching to be <100 s. The experimental details are further described in the Supporting Material.

MTs were visualized using an electron-multiplying charge-coupled device camera in full frame transfer mode at a frame rate of 10 Hz and using an exposure time of 0.1 s, which is comparable to the relaxation time of the slowest hydrodynamic mode. To verify that filament motion during this exposure time was not influencing our measurement, we reduced the exposure time to 0.066 s or 0.033 s (at a fixed frame rate of 10 Hz), and found no obvious trend in our persistence length data that would indicate a systematic bias. We did, however, find that image noise increased substantially for shorter exposure times, with an ~40% increase in background noise for the 0.033-s exposure time as compared to the 0.1-s exposure time. A smaller fraction of images have convergent fits under these high noise conditions, consistent with the benchmarking results shown in Table 1.

As a test of our approach, an ensemble of $M = 147$ fluorescence microscopy images was analyzed using our contour fitting method and least-squares estimator to analyze the first four modes. From these, we determine a nondimensional persistence length of $\tilde{\ell}_p = 176.6$ using the spectral characteristics of the fluctuating filament (see Fig. 4). In this test case, the MT contour length was estimated to be $\tilde{L} = 19.6\ \mu$m. This gives an actual dimensional persistence length of $\tilde{L}_p = 3.45$ mm. To determine the uncertainty in this measurement, we consider errors arising from both sampling and image artifacts. We estimate the sampling error when using only $M = 147$ images to be ± 0.48 mm in $\tilde{L}_p$. From analysis of the experimental fluorescence images, we estimate an average effective background noise of $\sigma_b/I_c = 0.143$ and average effective gap noise of $\sigma_g/\mu_g = 0.184$. From the results of a benchmarking study performed under similar noise conditions, we estimate the contribution to $\tilde{L}_p$-uncertainty to be ± 0.2 mm. We anticipate that the uncertainties due to sampling errors and noise artifacts behave in a fairly independent manner. Under this assumption, the total uncertainty can be estimated by adding these two contributions in quadrature. Therefore, we estimate $\tilde{L}_p = 3.45 \pm 0.52$ mm. These results represent a measurement uncertainty of $\approx 15\%$. Our measured persistence length is well within the range of previously reported values (7,20). The dashed black line in Fig. 4 shows the variances for the Chebyshev modes predicted using a WLC model with the same mean bending stiffness. Although deviations from WLC behavior have been reported for short MTs, for the

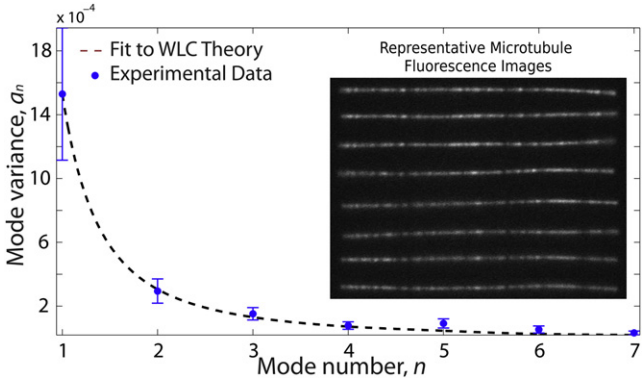

FIGURE 4 Experimental results for a microtubule. We analyze, using our spectral analysis methods, the thermal fluctuations of a microtubule for $M = 147$ images. (*Inset*) Exhibited thermal fluctuations in the microtubule shape. The exhibited modal covariances of the microtubule fluctuations agree well with a WLC model (*dashed line*). The error bars correspond to the 95.5% confidence intervals using our sampling error analysis in Role of Sampling Error on Estimated Persistence Length (see main text). Our methods yield an actual dimensional persistence length of $\tilde{L}_p = 3.45 \pm 0.52$ mm.

range of MT lengths ($\approx$8–20 μm) reported here, we find the predicted good agreement between the WLC model and our spectral data, for at least six independent modes (28). This validates our fitting algorithm under experimental conditions and relatively small sample sizes.

We then applied this method to the study of an ensemble of 54 different MTs, generated under chemically identical conditions, and subjected to identical temperature, storage, and handling conditions. For each MT, ~300 images were analyzed to determine the single-filament persistence length. Given the image noise and sampling errors, we would predict a measurement uncertainty of $\approx$10%. As shown in Fig. 5, we measure a much larger range of persistence lengths, indicating that real sample heterogeneity, not image processing uncertainty, is driving this broad distribution. Within this large variation, we find the persistence length to be independent of the filament length, for the range of lengths (10–20 μm) studied here (see Fig. S2).

In the limit of homogeneous isotropic materials at fixed ambient temperature, the persistence length depends only on the Young's modulus E, contour length L, and filament radius, r. L_p varies linearly with E, which has been found to be in the range of 2–3 GPa for a wide range of proteins, including tubulin and actin, likely due to van der Waals interactions along the protein backbone (29). By contrast, L_p varies as r^4, indicating that even small polydispersity in MT radius could give rise to substantial variations in stiffness. Interestingly, MTs have been shown to form with variable numbers of protofilaments both in vivo and in the presence of taxol in reconstituted systems (30). Thus, we favor a model in which an ensemble of MTs with differing radii give rise to a broad distribution of MT persistence lengths. A less likely possibility, in our view, is that the assumption of homogeneity and isotropy fails under some conditions, in which case it is more difficult to relate changes in L_p to a molecular mechanism. We would expect that in this limit, the fluctuation spectra would not follow a WLC model, so we would argue that the assumption of homogeneity and isotropy is appropriate in our analysis.

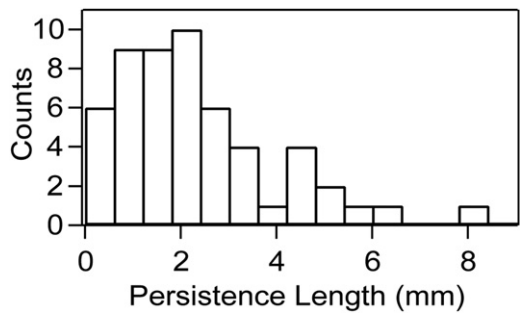

FIGURE 5 Experimental results for an ensemble of microtubules. Histogram of the persistence lengths of an ensemble of 54 different MTs, generated under chemically identical conditions, and subjected to identical temperature, storage, and handling conditions. For each MT, ~300 images were analyzed to determine the single-filament persistence length. Given the image noise and sampling errors, we would predict a measurement uncertainty of $\approx$10%, yet we measure a much larger range of persistence lengths, indicating that real sample heterogeneity, not image processing uncertainty, is driving this broad distribution.

SUMMARY

We present what to our knowledge are new spectral analysis methods for the measurement of biopolymer flexural rigidity from observations of the biopolymer thermal fluctuations. Our approach was based on global fitting of an entire trial contour at once to the fluorescence image. We used a contour representation expressed in terms of the curve tangent angles parameterized by arc-length and expanded in a basis of orthogonal polynomials. Using this representation, we performed statistical analysis of the modal coefficients to infer a flexural rigidity for the biopolymer. Our spectral analysis methods were found to work very well even in the case of images exhibiting significant background noise and gap artifacts. The benchmarking approach we propose provides a potentially powerful metric for rating different spectral analysis methods and for understanding the statistical significance of differences reported in experimental results. Our data indicate that these spectral analysis methods provide a substantial improvement in precision for measurements of stiffness based on observed fluctuations of a biopolymer. We expect these approaches will enable future studies of the differential effects of polymerization conditions and the binding of various regulatory molecules on microtubule mechanics.

SUPPORTING MATERIAL

Additional sections with two figures and three tables are available at http://www.biophysj.org/biophysj/supplemental/S0006-3495(12)00162-2.

Authors thank Professor Leslie Wilson (University of California, Santa Barbara) for providing tubulin proteins and Anna Simon for useful discussions.

We gratefully acknowledge support from the National Science Foundation (CAREER Award No. 0956210 to P.J.A., IGERT DGE-0221715 for D.V., and University of California, Santa Barbara Materials Research Laboratory grant No. DMR-0520415 for S.K. and D.Y.), a Burroughs Wellcome Fund Career Award at the Scientific Interface (to M.T.V.), and a research award from CurePSP (to M.T.V.).

REFERENCES

1. Fletcher, D. A., and R. D. Mullins. 2010. Cell mechanics and the cytoskeleton. *Nature.* 463:485–492.
2. Felgner, H., R. Frank, …, M. Schliwa. 1997. Domains of neuronal microtubule-associated proteins and flexural rigidity of microtubules. *J. Cell Biol.* 138:1067–1075.
3. Felgner, H., R. Frank, and M. Schliwa. 1996. Flexural rigidity of microtubules measured with the use of optical tweezers. *J. Cell Sci.* 109:509–516.
4. Gittes, F., B. Mickey, …, J. Howard. 1993. Flexural rigidity of microtubules and actin filaments measured from thermal fluctuations in shape. *J. Cell Biol.* 120:923–934.

5. Janson, M. E., and M. Dogterom. 2004. A bending mode analysis for growing microtubules: evidence for a velocity-dependent rigidity. *Biophys. J.* 87:2723–2736.

6. Mickey, B., and J. Howard. 1995. Rigidity of microtubules is increased by stabilizing agents. *J. Cell Biol.* 130:909–917.

7. van Mameren, J., K. C. Vermeulen, …, C. F. Schmidt. 2009. Leveraging single protein polymers to measure flexural rigidity. *J. Phys. Chem. B.* 113:3837–3844.

8. Nogales, E., S. G. Wolf, …, K. H. Downing. 1995. Structure of tubulin at 6.5 Å and location of the taxol-binding site. *Nature.* 375:424–427.

9. Nogales, E., M. Whittaker, …, K. H. Downing. 1999. High-resolution model of the microtubule. *Cell.* 96:79–88.

10. Nogales, E., S. G. Wolf, and K. H. Downing. 1998. Structure of the $\alpha\beta$ tubulin dimer by electron crystallography. *Nature.* 391:199–203.

11. Wang, H. W., and E. Nogales. 2005. Nucleotide-dependent bending flexibility of tubulin regulates microtubule assembly. *Nature.* 435: 911–915.

12. Bohm, K., W. Vater, …, E. Unger. 1984. Effect of microtubule-associated proteins on the protofilament number of microtubules assembled in vitro. *Biochim. Biophys. Acta.* 800:119–126.

13. Chrétien, D., F. Metoz, …, R. H. Wade. 1992. Lattice defects in microtubules: protofilament numbers vary within individual microtubules. *J. Cell Biol.* 117:1031–1040.

14. Pierson, G. B., P. R. Burton, and R. H. Himes. 1978. Alterations in number of protofilaments in microtubules assembled in vitro. *J. Cell Biol.* 76:223–228.

15. Pampaloni, F., G. Lattanzi, …, E. L. Florin. 2006. Thermal fluctuations of grafted microtubules provide evidence of a length-dependent persistence length. *Proc. Natl. Acad. Sci. USA.* 103:10248–10253.

16. van den Heuvel, M. G., S. Bolhuis, and C. Dekker. 2007. Persistence length measurements from stochastic single-microtubule trajectories. *Nano Lett.* 7:3138–3144.

17. Van den Heuvel, M. G., M. P. de Graaff, and C. Dekker. 2008. Microtubule curvatures under perpendicular electric forces reveal a low persistence length. *Proc. Natl. Acad. Sci. USA.* 105:7941–7946.

18. Brangwynne, C. P., G. H. Koenderink, …, D. A. Weitz. 2007. Bending dynamics of fluctuating biopolymers probed by automated high-resolution filament tracking. *Biophys. J.* 93:346–359.

19. Reference deleted in proof.

20. Kikumoto, M., M. Kurachi, …, H. Tashiro. 2006. Flexural rigidity of individual microtubules measured by a buckling force with optical traps. *Biophys. J.* 90:1687–1696.

21. Trefethen, L. 2000. Spectral Methods in MATLAB. Society for Industrial and Applied Mathematics, Philadelphia, PA.

22. Peskin, C. 2002. The immersed boundary method. *Acta Numerica.* 11:479–517.

23. Atzberger, P. J., P. R. Kramer, and C. S. Peskin. 2007. A stochastic immersed boundary method for fluid-structure dynamics at microscopic length scales. *J. Comput. Phys.* 224:1255–1292.

24. Gelfand, I. M., and S. V. Fomin. 2000. Calculus of Variations. Dover, Mineola, NY.

25. Saitô, N., K. Takahashi, and Y. Yunoki. 1967. The statistical mechanical theory of stiff chains. *J. Phys. Soc. Jpn.* 22:219–225.

26. Reichl, L. E. 1997. A Modern Course in Statistical Physics. John Wiley and Sons, New York.

27. Cam, L. L. 1990. Maximum Likelihood: An Introduction. *Intl. Stat. Rev.* 58:153–171.

28. Taute, K. M., F. Pampaloni, …, E. L. Florin. 2008. Microtubule dynamics depart from the wormlike chain model. *Phys. Rev. Lett.* 100:028102.

29. Howard, J. 2001. Mechanics of Motor Proteins and the Cytoskeleton. Sinauer, Sunderland, MA.

30. Díaz, J. F., J. M. Valpuesta, …, J. M. Andreu. 1998. Changes in microtubule protofilament number induced by Taxol binding to an easily accessible site. Internal microtubule dynamics. *J. Biol. Chem.* 273: 33803–33810.