End-to-end distance vector distribution with fixed end orientations for the wormlike chain model

Andrew J. Spakowitz and Zhen-Gang Wang

Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California 91125, USA

We find exact expressions for the end-to-end distance vector distribution function with fixed end orientations for the wormlike chain model. This function in Fourier-Laplace space adopts the form of infinite continued fractions, which emerges upon exploiting the hierarchical structure of the moment-based expansion. Our results are used to calculate the root-mean-square end displacement in a given direction for a chain with both end orientations fixed. We find that the crossover from rigid to flexible chains is marked by the root-mean-square end displacement slowly losing its angular dependence as the coupling between chain conformation and end orientation wanes. However, the coupling remains strong even for relatively flexible chains, suggesting that the end orientation strongly influences chain conformation for chains that are several persistence lengths long. We then show the behavior of the distribution function by a density plot of the probability as a function of the end-to-end distance vector for a wormlike chain in two dimensions with one end pointed in a fixed direction and the other end free (in its orientation). As we progress from high to low rigidity, the distribution function shifts from being peaked at a location near the full contour length of the chain in the forward direction, corresponding to a straight configuration, to being peaked near zero end separation, as in the Gaussian limit. The function exhibits double peaks in the crossover between these limiting behaviors.

I. INTRODUCTION

The wormlike chain model describes a polymer molecule that resists bending deformation embedded in a thermal environment [1], which is the simplest model for a semiflexible polymer. This model acts as a fundamental tool for addressing the effects of semiflexibility in a variety of instances including polymer liquid crystals [2–4], polyelectrolytes [5,6], protein networks [7,8], and the response of individual DNA molecules to tension [9].

A key descriptor of the statistical behavior of a polymer chain is the end-to-end distance distribution function. While the end-to-end distance distribution function for a flexible polymer is well known to be Gaussian, an exact solution for the corresponding distribution function for the wormlike chain is not as trivially found. Since the introduction of the wormlike chain model [1], many studies have addressed the statistical behavior of a semiflexible polymer as defined by this model and its several variants; these studies provide deep insight into the behavior of individual chains and a range of collective effects.

In the flexible limit, the chain statistics tend to a Gaussian distribution function; corrections to this limiting behavior for finite rigidity capture the first moments up to order l_p/L, where l_p is the persistence length and L is the total chain length [10,11]. Chain statistics in the rigid-rod limit are obtained using the path integral formalism in the WKB approximation including fixed end orientations [12] and by analytically evaluating the partition function by summing over transverse fluctuations about a nearly straight chain [13]. Furthermore, the limiting behavior of the chain statistics for a rigid chain are found by performing the mathematical manipulations in Laplace space (Laplace transform of the chain length) [14]. The chain statistics over the entire spectrum of chain rigidities is approximated by a fitting function that is found using variational methods that match the first

several moments [15]. The statistical behavior of the wormlike chain model including twist, bending discrepancies, and natural helicity is numerically found through a harmonic analysis on the rotation group [16].

The wormlike chain model when defined as a continuous space curve has the mathematically challenging constraint that the tangent vector is of unit magnitude at all points along the chain [17]. Relaxing this constraint such that the chain length is globally constrained [18], the tangent is constrained at one end [19,20], or the tangent length fluctuates about a unit average [21–24] results in analytically tractable expressions for the chain statistics that are accurate up to several moments. Alternative models to the wormlike chain model permit the evaluation of chain statistics. For example, a Dirac chain exhibits a similar crossover in behavior from the rigid to flexible limits as the wormlike chain model and captures several average quantities [25,26].

An exact solution for the end-to-end distance distribution function for the wormlike chain was not available until very recently [27–31]. In particular, we have developed a compact continued-fraction representation of the end-to-end distance distribution function in Fourier-Laplace space as well as single-chain partition functions in Laplace space for a wormlike chain in dipole and quadrupole fields [30]. These advances allow a number of important chain properties, such as its structure factor, to be studied, and reveal some unexpected behaviors. For example, the end-to-end distance distribution function exhibits two peaks for intermediate values of the chain rigidity when the chain length is the same order of magnitude as the persistence length [27,28].

A more complete characterization of the semiflexible chain conformation requires specifying both the end-to-end distance vector and the orientations of the chain ends. For a polymer chain with intermediate to large rigidity, the end-to-end distance vector will be strongly coupled to the orientations of the chain ends. Although the end-to-end distance

distribution provides information on the average spatial dimension of the chain, the shape of a polymer is better described by the probability distribution of the end-to-end distance vector relative to the orientation of the starting end. Furthermore, a number of important phenomena cannot be addressed unless the end-point orientations are included in the end-to-end distance vector distribution. For example, the closure of a DNA plasmid [32–35] requires the chain ends to be oriented with parallel tangents in order to form a smooth closed circle. Theoretical treatments of the thermodynamics of ring closure generally rely on quadratic-order fluctuations of the conformation about a minimum-energy reference conformation [36,37], an approximation that is valid for short chain lengths (or large rigidity) but fails for intermediate to long chain lengths (or small rigidity). Also, gene regulation in the cell is orchestrated by DNA-binding proteins; many of which loop DNA with fixed end orientations that are dictated by the structure of the proteins [38–41]. Fixing the chain ends at specified spatial locations and orientations leads to more complicated conformations than those prevalent in a smoothly closed ring. In the case of DNA bound to *lac* repressor, for example, energy minimization of elastic threads predicts the dominant conformation of the bound DNA [42,43]. However, the free energy of the bound DNA including conformation fluctuations away from the dominant structure requires methods to address the statistical behavior of a semiflexible polymer with both specified end-to-end distance and fixed end orientations that is accurate over the entire chain-rigidity spectrum.

In this paper, we present the exact analytical solution for the end-to-end distance vector distribution function with fixed end orientations (the Green function) for the wormlike chain model. In Sec. II, we describe the methods and derive results for the Green function in two and three dimensions in Fourier-Laplace space. In Sec. III, we address the statistical behavior of a polymer chain in two dimensions with fixed end orientations. Specifically, we consider the root-mean-square end displacement of a chain with both ends clamped as well as the behavior of the end-to-end distance vector distribution relative to the orientation of one end. These results are used to provide insight into the waning influence of fixed end orientations on the chain statistics as the rigidity decreases and the double-peaked structure of the end-to-end distance distribution function found in previous studies [27,28]. We conclude this paper with a summary of our results in Sec. IV.

II. FULL GREEN FUNCTION FOR THE WORMLIKE CHAIN MODEL

The formal definition of the wormlike chain model requires a definition of both the spatial geometry of the chain and the resulting deformation energy. The chain geometry is described by an inextensible space curve $\vec{r}(s)$ in d dimensions that is parametrized by the arc length parameter s, which runs from zero at one end of the chain to the contour length L at the opposite end. The chain is assumed inextensible, so that the tangent $\vec{u}(s) = \partial_s \vec{r}(s)$ is a unit vector at all positions along the chain $[|\vec{u}(s)| = 1$ for all $s]$. The bending

deformation energy is taken to be quadratic in the curvature of the chain conformation, as in linear elasticity theory [44,45], and is given by [17]

$$\beta \mathcal{H}_0 = \frac{l_p}{2} \int_0^L ds \left(\frac{\partial \vec{u}}{\partial s} \right)^2, \tag{1}$$

where l_p is the persistence length and $\beta = 1/(k_B T)$. Additional terms can be added to the Hamiltonian to account for other effects including multibody interactions, external fields, and constraints on the chain dimensions.

We now turn to the statistical mechanics of this model by subjecting the elastic chain to thermal fluctuations, which cause the chain to explore a range of conformations. A key characterization of the statistical behavior of the fluctuating chain is given by the Green function $G(\vec{R}, \vec{u} | \vec{u}_0; L)$ that gives the probability that a chain of length L with one end fixed at the origin with orientation $\vec{u}_0$ will have the other end located at position $\vec{R}$ with orientation $\vec{u}$. The Green function is found by performing a sum over all possible paths consistent with these specifications, awarding each conformation a Boltzmann weight. Mathematically,

$$G(\vec{R}, \vec{u} | \vec{u}_0; L) = \int_{\vec{u}_0}^{\vec{u}} \mathcal{D}[\vec{u}(s)] \delta \left(\vec{R} - \int_0^L ds \vec{u}(s) \right) e^{-\beta \mathcal{H}_0 [\vec{u}(s)]}, \tag{2}$$

where $\mathcal{D}[\vec{u}(s)]$ indicates path integration over the tangent-orientation paths of the chain with initial and final orientations $\vec{u}_0$ and $\vec{u}$, respectively [46]. The path integration over the tangent vector field is strictly restricted to values of the tangent vector with unit magnitude, thus enforcing the chain inextensibility constraint. The delta function in Eq. (2) counts only those conformations with tangent vectors adding up to $\vec{R}$.

To evaluate Eq. (2), we first perform a Fourier transform from the position variable $\vec{R}$ to the Fourier variable $\vec{k}$; the Green function is now written as

$$G(\vec{k}, \vec{u} | \vec{u}_0; L) = \int_{\vec{u}_0}^{\vec{u}} \mathcal{D}[\vec{u}(s)] \exp \left(-\beta \mathcal{H}_0 [\vec{u}(s)] + i\vec{k} \cdot \int_0^L ds \vec{u}(s) \right)$$

$$= \left\langle \exp \left(i\vec{k} \cdot \int_0^L ds \vec{u}(s) \right) \right\rangle_{\vec{u}_0}^{\vec{u}}, \tag{3}$$

where the angle brackets $\langle \cdots \rangle_{\vec{u}_0}^{\vec{u}}$ indicate an average with respect to the Hamiltonian $\beta \mathcal{H}_0$ [given by Eq. (1)] with initial and final tangent orientations $\vec{u}_0$ and $\vec{u}$, respectively. The form of the Green function given in Eq. (3) is equivalent to finding the statistical behavior of a wormlike chain under the influence of an imaginary dipole field with fixed initial and final chain orientations.

Our previous work on the behavior of a wormlike chain in an external field serves as the foundation for this current study [30]. In Ref. [30], we provide exact results for the partition function, Laplace transformed from the dimensionless length $N = L/(2l_p)$ to the Laplace variable p, of a wormlike chain imbedded in a dipole and quadrupole field, in both

two and three dimensions. We adapt the techniques developed in Ref. [30] to address the statistical behavior of a semiflexible chain with fixed end orientations.

Before delving into the details of our solutions, we outline the main procedures of our derivation. As previously mentioned, Eq. (3) reduces our problem to a wormlike chain in an external imaginary dipole field. We note that the Hamiltonian $\beta\mathcal{H}_0$ is invariant to rotation and perform a rotation of the tangent vector such that the Fourier vector $\vec{k}$ is aligned with the x axis in two dimensions and the z axis in three dimensions, which greatly simplifies the calculation of the average with respect to the Hamiltonian $\beta\mathcal{H}_0$. We perform a Taylor expansion of the exponential, which results in the infinite moment-based expansion of the Green function. The moments are calculated using the Green function for the Hamiltonian $\beta\mathcal{H}_0$, which is a function of the tangent vector only. Noting the convolution structure of the moment-based expansion, we take the Laplace transform from the chain length variable $N=L/(2l_p)$ to the Laplace variable p, which converts the product of integrals within each moment to a simple algebraic product of terms. We then utilize a diagrammatic representation of the algebra first developed by Yamakawa [47,48] and establish the diagrammatic rules that are relevant to our current study. We show that the diagrams from the moment-based expansion can be sorted in a meaningful manner that results in a hierarchical structure for the solution for the Green function in Fourier-Laplace space.

We now give the detailed derivation for the Green function of a wormlike chain in two and three dimensions, following the steps outlined above; these derivations run parallel to the progression in Ref. [30].

A. Two-dimensional solution

In two dimensions, the orientation of the tangent vector is defined by the angle θ between the tangent vector and the x axis. We require the orientation-only Green function $G_0(\theta|\theta_0;L)$ that gives the conditional probability that a chain of length L with initial tangent orientation θ_0 will have the final orientation θ. The Green function can be written as an eigenfunction expansion

$$G_0(\theta|\theta_0;L) = \sum_{m=-\infty}^{\infty} \frac{1}{2\pi} e^{im(\theta-\theta_0)} \exp\left(-\frac{m^2 L}{2l_p}\right). \quad (4)$$

Equation (4) permits the evaluation of the averages with respect to $\beta\mathcal{H}_0$ present in the full Green function $G(\vec{R},\vec{u}|\vec{u}_0;L)$ [Eq. (3)].

The Fourier variable $\vec{k}$ in two dimensions is written as $\vec{k}=k\cos\chi\hat{x}+k\sin\chi\hat{y}$, where k is the magnitude and χ gives the orientation relative to the x axis. We now perform a rotation of the tangent orientation vector of the form $\vec{u}=\mathbf{\Gamma}\cdot\vec{u}'$ such that $\vec{k}\cdot\mathbf{\Gamma}\cdot\vec{u}'=k\hat{x}\cdot\vec{u}'$, thus $\vec{k}\cdot\mathbf{\Gamma}=k\hat{x}$. This is accomplished by a rotation of the coordinate axis by an angle χ. Since the Hamiltonian is invariant to such a transformation, the effect of the rotation merely resets the initial and final orientation of the tangent vector to $\theta_0-\chi$ and $\theta-\chi$, respectively. Equation (3) is now written as

$$G(\vec{k},\vec{u}|\vec{u}_0;L) = \left\langle \exp\left(ik\hat{x}\cdot\int_0^L ds\,\vec{u}(s)\right)\right\rangle_{\theta_0-\chi}^{\theta-\chi}$$

$$= \sum_{n=0}^{\infty} \frac{(ik)^n}{n!}\left\langle\left(\int_0^L ds\,\cos\theta(s)\right)^n\right\rangle_{\theta_0-\chi}^{\theta-\chi}, \quad (5)$$

the second form in Eq. (5) is the moment-based expansion. To simplify the subsequent steps, we nondimensionalize the chain length and the Fourier variable to $N=L/(2l_p)$ and $K=2l_p k$, respectively.

The average in the nth term of the moment-based expansion is rewritten as

$$n!\left\langle\prod_{i=1}^{n}\int_0^{s_{i+1}} ds_i\,\cos\theta(s_i)\right\rangle_{\theta_0-\chi}^{\theta-\chi}, \quad (6)$$

where $s_{n+1}=N$, and the factor of $n!$ arises from the "time" ordering of the integrations within the average. The Markovian nature of the statistics permits the evaluation of this average using the Green function $G_0(\theta|\theta_0;N)$ given by Eq. (4); $G_0[\theta(s_{i+1})|\theta(s_i);s_{i+1}-s_i]$ is sandwiched between each successive factor in the product of Eq. (6), and the result is integrated over the intermediate angular variables, fixing the initial and final orientations. The nth term requires $n+1$ propagators and thus $n+1$ intermediate m indices.

Since the eigenfunctions in the propagator [Eq. (4)] are complete and orthonormal, only certain m-index values will survive the integration over the intermediate orientations in Eq. (6). Noting that $\cos\theta e^{im\theta}=e^{i(m+1)\theta}/2+e^{i(m-1)\theta}/2$, the m index values that contribute a nonzero value to the average in Eq. (6) are selected by the criteria that each intermediate m index value is offset from its neighbor by either $+1$ or -1. Since the initial and final orientations are fixed, the initial m-index m_1 and the final m-index m_{n+1} are free to take any value. As an example to illustrate these steps, we consider the average in the $n=2$ term in Eq. (6). After performing the integrals over the intermediate tangent orientations, we have

$$2!\left\langle\prod_{i=1}^{2}\int_0^{s_{i+1}} ds_i\,\cos\theta(s_i)\right\rangle_{\theta_0-\chi}^{\theta-\chi}$$

$$= \frac{1}{2(2\pi)^2}\int_0^N ds_2\int_0^{s_2} ds_1 \sum_{m_1,m_2,m_3} e^{im_3(\theta-\chi)} C_{m_3}(N-s_2)$$

$$\times(\delta_{m_3,m_2+1}+\delta_{m_3,m_2-1})C_{m_2}(s_2-s_1)$$

$$\times(\delta_{m_2,m_1+1}+\delta_{m_2,m_1-1})C_{m_1}(s_1)e^{-im_1(\theta_0-\chi)}, \quad (7)$$

where the sums over the m-indices run from $-\infty$ to ∞, and $C_m(N)=\exp(-m^2 N)$. The Kronecker delta functions $\delta_{m,n}$ in Eq. (7) restrict the allowed m-index values such that each m-index is offset from its neighbor by $+1$ or -1. The sum over the intermediate m-index m_2 results in initial and final m-index values that satisfy either $|m_3-m_1|=2$ or $m_3-m_1=0$.

Similar manipulations as described in the previous paragraph yield the nth term in the moment-based expansion [Eq. (6)], which now reads

$$n! \left\langle \prod_{i=1}^{n} \int_0^{s_{i+1}} ds_i \cos\theta(s_i) \right\rangle_{\theta_0-\chi}^{\theta-\chi}$$

$$= \frac{n!}{2^n(2\pi)^2} \prod_{i=1}^{n} \sum_{\{m_i\}} e^{im_{n+1}(\theta-\chi)} \int_0^{s_{i+1}} ds_i C_{m_{i+1}}(s_{i+1}-s_i)$$

$$\times (\delta_{m_{i+1},m_i+1} + \delta_{m_{i+1},m_i-1}) e^{-im_1(\theta_0-\chi)} C_{m_1}(s_1), \qquad (8)$$

where $s_{n+1}=N=L/(2l_p)$, and the sum over $\{m_i\}$ indicates summation over all m_i ($i=1,2,\ldots,n+1$) from $-\infty$ to ∞. Evaluating Eq. (8) requires finding the m-index values that survive the Kronecker delta functions upon summing over the m-index values; this essentially involves finding the m-index values that correspond to paths between m_1 and m_{n+1} with n steps of unit magnitude separating them. The convolution structure of Eq. (8) suggests a simple solution when we perform a Laplace transform from N to p. Equation (8) is now written as

$$\mathcal{L}\left(n! \left\langle \prod_{i=1}^{n} \int_0^{s_{i+1}} ds_i \cos\theta(s_i) \right\rangle_{\theta_0-\chi}^{\theta-\chi} \right)$$

$$= \frac{n!}{2^n(2\pi)^2} \prod_{i=1}^{n} \sum_{\{m_i\}} \frac{1}{P_{m_i}} \frac{1}{P_{m_{n+1}}} (\delta_{m_{i+1},m_i+1} + \delta_{m_{i+1},m_i-1})$$

$$\times e^{im_{n+1}(\theta-\chi)} e^{-im_1(\theta_0-\chi)}, \qquad (9)$$

where $\mathcal{L}$ indicates a Laplace transform from N to p, and $P_m=p+m^2$. Solving Eq. (9) involves finding the m-index paths that connect m_1 to m_{n+1}, which are subsequently summed from $-\infty$ to ∞.

With Eq. (9), we write the Laplace transform of Eq. (5) as

$$G(\vec{k},\vec{u}|\vec{u}_0;p) = \frac{1}{(2\pi)^2} \sum_{n=0}^{\infty} \left(\frac{iK}{2}\right)^n \sum_{m_1,m_{n+1}} e^{im_{n+1}(\theta-\chi)} e^{-im_1(\theta_0-\chi)}$$

$$\times \sum_{\text{paths } \mu}^{m_1 \to m_{n+1}} \prod_{i=1}^{n+1} \frac{1}{P_{m_i^{(\mu)}}}, \qquad (10)$$

where $K=2l_p k$. The sum over the paths μ involves finding the m-index paths that connect m_1 to m_{n+1} with n intermediate steps of unit magnitude. In its current form, the Green function is found by summing terms from the moment-based expansion according to the power of the term; however, it is easier to collect all the terms that begin and end at the same m-index values, which we identify as m_0 for the initial m index and m_f for the final m index. The Green function is now written as

$$G(\vec{k},\vec{u}|\vec{u}_0;p) = \frac{1}{(2\pi)^2} \sum_{m_0,m_f} e^{im_f(\theta-\chi)} e^{-im_0(\theta_0-\chi)}$$

$$\times \sum_{\text{paths } \mu}^{m_0 \to m_f} \prod_{\{m_i^{(\mu)}\}} \frac{iK/2}{P_{m_i^{(\mu)}}}, \qquad (11)$$

where the product over $\{m_i^{(\mu)}\}$ indicates a product of terms for each path that connects m_0 to m_f. We now turn to finding the sum of all m-index paths that connect m_0 to m_f.
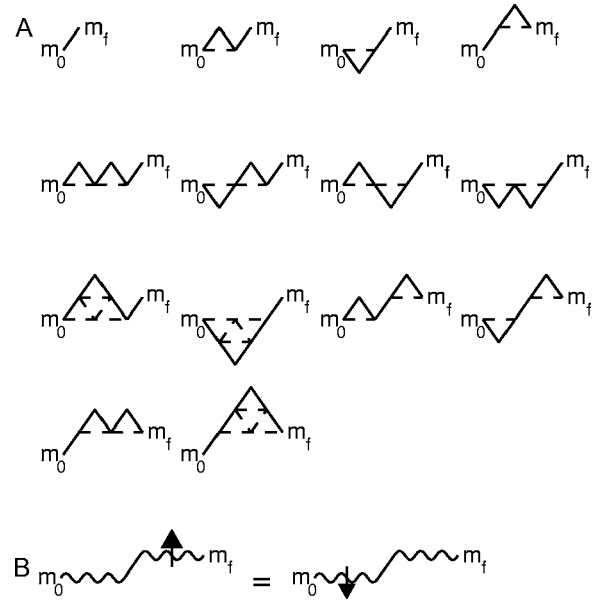

FIG. 1. The first 14 diagrams for the case $m_f=m_0+1$ are shown in A. The infinite sum of diagrams for $m_f=m_0+1$ collapse into the two equivalent diagrams shown in B. In B, the wavy line represents the sum of diagrams that begin and end at that level, the wavy line with the upward pointing arrow represents the sum of diagrams that begin and end at that level without going beneath that level, and the wavy line with the downward pointing arrow represents the sum of diagrams that begin and end at that level without going above that level (see text for details).

Following Yamakawa [47,48], we represent the m-index paths diagrammatically using a stone-fence representation; this is done by plotting the values of the m indices and connecting each successive value by a line segment. The benefit of casting our current problem into the diagrammatic representation is in reducing the complicated algebra to a simpler representation and in identifying and collecting similar terms in the summation. We define $G_{m_0}^{m_f}(K;p)$ as the sum over all possible m-index paths that connect m_0 to m_f [equal to the second line in Eq. (11)]. The selection rule for the m indices requires that each m index differs from its neighbor by either $+1$ or -1, so a valid diagram that connects m_0 to m_f zigzags from m_0 to m_f without violating this selection rule. The first contribution for $|m_0-m_f|=n$ is a straight line that requires n steps, and subsequent contributions involve adding an equal number of up and down steps so as to not upset the end values. As an illustration, we consider the first 14 diagrams that contribute to the term $m_f=m_0+1$ in Fig. 1(A); these diagrams represent all possible paths that connect m_0 to m_f using either one, three, or five steps of unit magnitude in between. These diagrams, along with the infinite number of diagrams with greater than five steps, are summed together to find $G_{m_0}^{m_0+1}$. The diagrams in Fig. 1(A) are easily converted to their corresponding mathematical expressions by noting that each m-index value contributes a factor of $P_{m_i}^{-1}$ and each step contributes $iK/2$ to the product. For example, the first diagram in Fig. 1(A) represents the expression $P_{m_0}^{-1}(iK/2)P_{m_0+1}^{-1}$, and the second diagram represents $P_{m_0}^{-1}(iK/2)P_{m_0+1}^{-1}(iK/2)P_{m_0}^{-1}(iK/2)P_{m_0+1}^{-1}$. With these diagram-

matic rules in place, any diagram is easily converted back to its mathematical expression in Laplace space and vice versa.

We now recall several results found in Ref. [30] in order to further simplify Eq. (11). Our previous study on the statistical behavior of a wormlike chain without fixed end orientations involves a similar m-index path summation as found in Eq. (10). We found that the infinite sum of terms arising from m-index paths with equal initial and final m-index values $m_0 = m_f = n$ is given by

$$W_n = \frac{1}{\dfrac{K^2 w_{n-1}^{(-)}}{4} + P_n + \dfrac{K^2 w_{n+1}^{(+)}}{4}}, \qquad (12)$$

where $P_n = p + n^2$, and $w_n^{(-)}$ and $w_n^{(+)}$ obey the recursive relationships $w_n^{(-)} = 1/(P_n + K^2 w_{n-1}^{(-)}/4)$ and $w_n^{(+)} = 1/(P_n + K^2 w_{n+1}^{(+)}/4)$. This result adopts the form of an infinite continued fraction, which emerges naturally by a collection of similar terms within the infinite summation that is best visualized using the diagrammatic representation [30]. The partial continued fractions $w_n^{(-)}$ and $w_n^{(+)}$ are collections of particular m-index paths. The sum of all m-index paths between n and n with no intermediate m-index values greater than n is given by $w_n^{(-)}$, and the sum of all m-index paths between n and n with no intermediate m-index values less than n is given by $w_n^{(+)}$.

The results found in Ref. [30] provide the necessary elements to find the solution for $G_{m_0}^{m_f}$. For the case $m_0 = m_f$, the sum of diagrams $G_{m_0}^{m_0}$ is exactly equal to W_{m_0}. For the case $m_0 \neq m_f$, the sum of diagrams $G_{m_0}^{m_f}$ requires special care to avoid double counting the diagrams; however, the result is easily constructed with our results from Ref. [30]. In Fig. 1, we consider the case $m_f = m_0 + 1$. The summation from m_0 to m_f is found by performing a sum of all diagrams from m_0 to m_0, stepping up to $m_0 + 1$, and then performing a sum of diagrams from $m_f (= m_0 + 1)$ to m_f with the requirement that no intermediate values are less than m_f. This requirement ensures that no diagrams are double counted in $G_{m_0}^{m_0+1}$. For example, the third diagram in the second line of Fig. 1(A) is decomposed into one part that goes from m_0 to m_0 (four steps, up→down→down→up) and a step up to m_f; this diagram is counted only once in $G_{m_0}^{m_0+1}$ if we sum all diagrams between m_0 and m_0, step up to m_f, and sum all diagrams between m_f and m_f without going below m_f. We do not arrive at $G_{m_0}^{m_0+1}$ if we sum all diagrams between m_0 and m_0, step up to m_f, and sum all diagrams between m_f and m_f since many diagrams are multiply counted, including the third diagram in the second line of Fig. 1(A). By this technique, the left four steps in this diagram are contributed by the sum of all diagrams between m_0 and m_0, and stepping up to m_f completes this diagram. However, the third diagram in the second line of Fig. 1(A) is also constructed by first stepping up to m_f from m_0, and the right four steps in this diagram are contributed by the sum of all diagrams between m_f and m_f. This demonstrates the need to restrict the sum of diagrams between m_f and m_f to those diagrams that do not extend below m_f to avoid this double counting of diagrams. Alternatively, the summation from m_0 to m_f can be found by summing

from m_0 to m_0 without going above m_0, then stepping up to $m_f (= m_0 + 1)$, and then summing all diagrams that begin and end at m_f. In Fig. 1(B), we show modified diagrams that demonstrate how we collect these terms to find $G_{m_0}^{m_0+1}$. The wavy line in Fig. 1(B) represents the sum of all diagrams that begin and end at that level (W_n), the wavy line with the upward pointing arrow represents the sum of diagrams that begin and end at that level without ever going beneath that level ($w_n^{(+)}$), and the wavy line with the downward pointing arrow represents the sum of diagrams that begin and end at that level without ever going above that level ($w_n^{(-)}$). Therefore, we find $G_{m_0}^{m_0+1} = W_{m_0}(iK/2)w_{m_0+1}^{(+)} = w_{m_0}^{(-)}(iK/2)W_{m_0+1}$; the equivalence of these two expressions can be shown using the results from Ref. [30] provided in the preceding paragraph.

This procedure can be extended to arbitrary m_0 and m_f using similar arguments as the preceding paragraph. As previously mentioned, the case where $m_0 = m_f$ gives

$$G_{m_0}^{m_0} = W_{m_0}. \qquad (13)$$

For the case $m_0 < m_f$, the sum of diagrams from m_0 to m_f is given by

$$G_{m_0}^{m_f} = W_{m_0}\left(\frac{iK}{2}\right)^{|m_f - m_0|} \prod_{n=1}^{|m_f - m_0|} w_{m_0+n}^{(+)}, \qquad (14)$$

and for the case $m_0 > m_f$, we have

$$G_{m_0}^{m_f} = W_{m_0}\left(\frac{iK}{2}\right)^{|m_f - m_0|} \prod_{n=1}^{|m_f - m_0|} w_{m_0-n}^{(-)}. \qquad (15)$$

Equation (11) is now written as

$$G(\vec{k}, \vec{u} | \vec{u}_0; p) = \frac{1}{(2\pi)^2} \sum_{m_0, m_f} e^{im_f(\theta - \chi)} e^{-im_0(\theta_0 - \chi)} G_{m_0}^{m_f}. \qquad (16)$$

This completes our formal solution of the full Green function of a wormlike chain in two dimensions.

B. Three-dimensional solution

The solution for the Green function for the wormlike chain model in three dimensions is found using a virtually identical progression of steps as in the two-dimensional derivation. However, there are several nontrivial differences that we will draw attention to along the way. Perhaps the most substantial difference in the three-dimensional derivation lies in the added complexity of having more degrees of freedom associated with both the end orientations and the end position; nonetheless, the added complexity does not render the problem intractable.

In three dimensions, the orientation of the tangent vector is defined in spherical coordinates by the polar angle θ and the azimuthal angle ϕ, which are defined according to the usual convention [49]. To perform averages with respect to $\beta\mathcal{H}_0$ [Eq. (1)], we require the orientation-only Green function $G_0(\vec{u} | \vec{u}_0; L)$ that gives the joint probability that a chain of length L will have initial tangent orientation $\vec{u}_0$ and final tangent orientation $\vec{u}$ regardless of the end position in space. The Green function is given by

$$G_0(\vec{u}|\vec{u}_0;L) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} Y_l^m(\vec{u}) Y_l^{m*}(\vec{u}_0) C_l(N), \qquad (17)$$

where $C_l(N) = \exp[-l(l+1)N]$, $N = L/(2l_p)$, and Y_l^m are the spherical harmonics [49].

The Fourier variable in three dimensions is written in spherical coordinates as $\vec{k} = k \sin\chi \cos\omega \hat{x} + k \sin\chi \sin\omega \hat{y} + k \cos\chi \hat{z}$, where k is the magnitude of $\vec{k}$, and the orientation of $\vec{k}$ is defined by the polar angle χ and the azimuthal angle ω. To aid in the calculation of the average in Eq. (3), we perform a rotation of the coordinate system such that $\vec{k}$ is aligned along the z axis. This is done by setting the tangent vector to $\vec{u} = \Gamma \cdot \vec{u}'$ such that $\vec{k} \cdot \vec{u} = \vec{k} \cdot \Gamma \cdot \vec{u}'$ is equal to $k\hat{z} \cdot \vec{u}'$, thus we have $\vec{k} \cdot \Gamma = k\hat{z}$. Since the Hamiltonian $\beta\mathcal{H}_0$ is invariant under arbitrary coordinate rotation, this operation merely resets the fixed initial and final tangent orientations from $\vec{u}_0, \vec{u}$ to $\vec{u}_0', \vec{u}'$ as determined by the inverse of the rotation matrix Γ, i.e., $\vec{u}' = \Gamma^{-1} \cdot \vec{u}$ and $\vec{u}_0' = \Gamma^{-1} \cdot \vec{u}_0$. The rotation that draws the Fourier variable $\vec{k}$ into coincidence with the z axis involves the two-step process of rotation of $\vec{u}$ about the y axis by the angle χ and then rotation of $\vec{u}$ about the z axis by the angle ω. The inversion of our rotation is expressible in terms of the Euler angles α, β, γ [49,50] according to $\alpha = -\omega$, $\beta = -\chi$, and $\gamma = 0$, which gives the rotated initial and final tangent orientations $\vec{u}_0'$ and $\vec{u}'$. Upon performing such a rotation, the Green function in three dimensions is given by

$$\begin{aligned}
G(\vec{k},\vec{u}|\vec{u}_0;L) &= \left\langle \exp\left(ik\hat{z} \cdot \int_0^L ds \vec{u}(s) \right) \right\rangle_{\vec{u}_0'}^{\vec{u}'} \\
&= \sum_{n=0}^{\infty} \frac{(ik)^n}{n!} \left\langle \left(\int_0^L ds \cos\theta(s) \right)^n \right\rangle_{\vec{u}_0'}^{\vec{u}'}. \quad (18)
\end{aligned}$$

As in two dimensions, we simplify the derivation by nondimensionalizing the chain length and Fourier variable to $N = L/(2l_p)$ and $K = 2l_p k$, respectively.

The average in Eq. (18) is performed by a virtually identical procedure as in two dimensions. The three-dimensional version of Eq. (6) is identical except that the constraints on the average are altered to the initial and final orientations $\vec{u}_0'$ and $\vec{u}'$. The average is found using the Markovian nature of the statistics by inserting the Green function given by Eq. (17) between each successive term and integrating over the intermediate tangent orientations. In two dimensions, we found that certain values of the m indices of the Green functions contributed to the average. However, in three dimensions, the Green functions have both l indices and m indices that must be accounted for when determining the contributing terms. We note that the spherical harmonics obey the relationship $\cos\theta Y_l^m = a_{l+1}^m Y_{l+1}^m + a_l^m Y_{l-1}^m$ where $a_l^m = \sqrt{(l-m)(l+m)/(4l^2-1)}$ [49]. Therefore, the l indices that contribute to the averages in question must be offset from their neighbors by either $+1$ or -1 (same as m indices in two dimensions), and the m indices are constant for any contributing m-index path. The task of finding l, m-index paths that

contribute to the average is no more complicated than the two-dimensional case since the initial, intermediate, and final m indices are all equal.

The diagrammatic representation of the Laplace-space calculation of the average in three dimensions is virtually identical to the two dimensional case described in the previous section and shown in Fig. 1 with only minor modifications. First, we switch all labels of m in the two-dimensional case to l indices. Second, due to the properties of spherical harmonics, the l indices cannot be less than the magnitude of the m index ($|m|$), which is a constant for a given diagram. This restriction has some consequences on the diagrammatic summation. For $l_0 \geq |m| + 2$, the three-dimensional analogue of Fig. 1(A) is identical to Fig. 1(A). For $l_0 = |m| + 1$, the second diagram in the third line of Fig. 1(A) does not exist since the third l index is less than $|m|$. Similarly, for $l_0 = |m|$, the third diagram in the first line, the second, third, and fourth diagrams of the second line, and the second and fourth diagrams of the third line of Fig. 1(A) do not contribute since at least one intermediate l index is less than $|m|$.

The diagrammatic representation is easily converted into the Laplace-space expressions that they represent; a diagram corresponding to the particular l-index path $\{l_1, l_2, \ldots, l_{n+1}\}$ (n steps separate the initial and final l indices) with m-index value m is equal to $P_{l_1}^{-1} \prod_{i=2}^{n+1} iKa_{l_i+t_i}^m P_{l_i}^{-1}$ where $P_l = p + l(l+1)$, and $t_i = 0$ if $l_i = l_{i-1} + 1$ (step up) and $t_i = 1$ if $l_i = l_{i-1} - 1$ (step down). For example, the three-dimensional expression for the first diagram of the third line of Fig. 1(A) is given by

$$\frac{1}{P_{l_0}} \frac{iKa_{l_0+1}^m}{P_{l_0+1}} \frac{iKa_{l_0+2}^m}{P_{l_0+2}} \frac{iKa_{l_0+2}^m}{P_{l_0+1}} \frac{iKa_{l_0+1}^m}{P_{l_0}} \frac{iKa_{l_0+1}^m}{P_{l_0+1}}, \qquad (19)$$

where we express all l indices in terms of l_0 and $l_f = l_0 + 1$.

Following similar steps as in the two-dimensional solution, the Fourier-Laplace transformed Green function is written as

$$G(\vec{k},\vec{u}|\vec{u}_0;p) = \frac{1}{4\pi} \sum_{l_0,l_f,m} Y_{l_f}^m(\vec{u}') Y_{l_0}^{m*}(\vec{u}_0') G_{l_0}^{l_f,m}, \qquad (20)$$

where we define the sum of all diagrams that connect l_0 to l_f with a fixed value of m as $G_{l_0}^{l_f,m}(K;p)$. The indices l_0 and l_f in the sum run from zero to infinity; however, the index m runs from $-\min(l_0,l_f)$ to $\min(l_0,l_f)$ where $\min(l_0,l_f)$ is the minimum value of l_0 and l_f.

We now recall and adapt several results from Ref. [30] to complete our derivation. The infinite sum of diagrams with equal initial and final l-index values of l- and m-index values m for all diagrams is given by

$$W_l^m = \frac{1}{(a_l^m K)^2 w_{l-1}^{m(-)} + P_l + (a_{l+1}^m K)^2 w_{l+1}^{m(+)}}, \qquad (21)$$

where $P_l = p + l(l+1)$, $a_l^m = \sqrt{(l-m)(l+m)/(4l^2-1)}$, and $K = 2l_p k$. The terms $w_l^{m(+)}$ and $w_l^{m(-)}$ represent the sums of all diagrams that start and end at l with no intermediate l indices below l ($w_l^{m(+)}$) or above l ($w_l^{m(-)}$) with m-index value m for all diagrams. These terms are continued fractions that are governed by the recursive relationships $w_l^{m(+)}$

$=1/[P_l+(a_{l+1}^m K)^2 w_{l+1}^{m(+)}]$ $(l \geqslant m)$, $w_l^{m(-)}=1/[P_l+(a_l^m K)^2 w_{l-1}^{m(-)}]$ $(l \geqslant m+1)$, and $w_l^{m(-)}=1/P_l$ $(l=m)$. Using similar arguments as the two-dimensional case, the sum of diagrams with initial l-index value l_0 and final l-index value l_f is constructed using the partial sums W_l^m, $w_l^{m(+)}$, and $w_l^{m(-)}$. For the case $l_0=l_f$, the sum of diagrams is

$$G_{l_0}^{l_0,m} = W_{l_0}^m. \tag{22}$$

For the case $l_0 < l_f$, the sum of diagrams is given by

$$G_{l_0}^{l_f,m} = W_{l_0}^m \prod_{n=1}^{|l_f-l_0|} iK a_{l_0+n}^m w_{l_0+n}^{m(+)}, \tag{23}$$

and for the case $l_0 > l_f$, the sum of diagrams is given by

$$G_{l_0}^{l_f,m} = W_{l_0}^m \prod_{n=1}^{|l_f-l_0|} iK a_{l_0+1-n}^m w_{l_0-n}^{m(-)}. \tag{24}$$

These expressions and Eq. (20) complete our solution for the full Green function in Fourier-Laplace space for a wormlike chain in three dimensions.

III. CHAIN STATISTICS WITH FIXED END ORIENTATIONS

In the preceding section, we derive the full Green function for the wormlike chain model in two and three dimensions in Fourier-Laplace space, which describes the statistical behavior of a semiflexible polymer chain with fixed orientations of the end tangents. In this section, we demonstrate the utility of our results by examining the statistical behavior of a chain with fixed end orientations (either both ends clamped or one end clamped).

Many studies of the statistical behavior of semiflexible polymers focus on the end-to-end distance distribution function [10,11,13–16,18,19,22–28], which gives the probability that the ends of a semiflexible chain are separated by a fixed distance regardless of the end tangent orientations. These studies demonstrate the role that semiflexibility plays in the spatial separation of the chain ends. For rigid chains, the ends are extended to nearly the contour length of the chain, indicative of the chain conformation being essentially straight with small wiggles in the transverse direction. For flexible chains, the end-to-end distance distribution function tends to a Gaussian distribution, as a result of the polymer trajectory tracing a random walk through space once the tangent orientations become uncorrelated by thermal fluctuations. As we progress from high rigidity to low rigidity, the end-to-end distance distribution function shifts from being peaked at an end separation of the contour length L to the peak being located at zero end separation. The crossover between these two limiting behaviors exhibits a double-peaked distribution function for intermediate chain rigidities in both two and three dimensions [27,28], suggesting that such a chain could exist in an open and a closed state. The implications of this double-peak feature on the dynamic behavior of semiflexible polymers are yet to be explored.

The end-to-end distance distribution function provides a clear picture of the fluctuating separation between the chain ends, but it only gives indirect information about the fluctuating conformation. For semiflexible polymers, the chain conformation is strongly coupled to the end orientations, thus the influence of thermal fluctuations on the chain conformation is best examined by observing the statistics of the end-to-end vector while the ends have fixed orientations.

We begin by considering the orientation-dependent end displacement for a semiflexible polymer with fixed end orientations. The full Green function in two dimensions [Eq. (16)] permits the evaluation of average quantities with fixed end orientations. For example, the nth moment of the full Green function gives the following average quantity:

$$M_n = \langle (\vec{R} \cdot \hat{e}(\chi))^n \rangle_{\theta_0}^\theta = \frac{(-i)^n}{q(N)} \mathcal{L}^{-1}\left(\left. \frac{\partial^n G(\vec{k},\theta|\theta_0;p)}{\partial K^n}\right|_{K=0}\right), \tag{25}$$

where $\mathcal{L}^{-1}$ indicates an inverse Laplace transform from p to N. Equation (25) gives the average end displacement in the $\hat{e}(\chi)$ direction raised to the nth power for a chain with fixed initial orientation θ_0 and final orientation θ, where length is nondimensionalized by $2l_p$. The partition function q in Eq. (25) normalizes the average and is given by

$$q(N) = \frac{1}{(2\pi)^2} \sum_{m=-\infty}^{\infty} \exp[im(\theta-\theta_0)]\exp(-m^2 N). \tag{26}$$

Application of Eq. (25) for $n=2$ gives the mean-square displacement along $\hat{e}(\chi)$ of a chain with initial orientation θ_0 and final orientation θ. In Fig. 2, we show the scaled root-mean-square displacement $\sqrt{M_2/N^2}$ along the direction $\vec{e}(\chi)$ of a chain in two dimensions with the initial tangent pointing up along the x axis and the other end with orientation θ for several values of $N=L/(2l_p)$. The upper left-hand figure shows the minimum-energy conformations of a chain with fixed end orientation $\theta=0$ (straight chain marked A), $\theta=\pi/2$ (curved chain marked B), and $\theta=\pi$ (curved chain marked C). For a given value of N, we show the root-mean-square displacement as the radial distance in a polar plot over the angle χ for $\theta=0$ (solid curve), $\pi/2$ (dashed curve), and π (dashed-dotted curve).

In the limit of $N \to 0$, the chain conformation tends to the minimum-energy conformations presented in the upper-left part of Fig. 2; therefore, the root-mean-square displacement behaves as

$$\sqrt{M_2/N^2} \to \left| \frac{\sin(\theta-\chi)-\sin(\theta_0-\chi)}{\theta-\theta_0} \right|. \tag{27}$$

For $N=0.1$ in Fig. 2, the root-mean-square displacement only deviates slightly from Eq. (27). As N increases, thermal undulations in the chain cause the root-mean-square displacement to lose its anisotropy, eventually leading to a rotationally symmetric root-mean-square displacement given by the random walk behavior $\sqrt{M_2/N^2}=1/\sqrt{N}$.

The crossover from rigid to flexible chain in Fig. 2 demonstrates that a signature of semiflexibility is the strong coupling between chain conformation and end orientation. Notably, the root-mean-square end displacement for $N=10$ still deviates appreciably from the rotationally invariant behavior
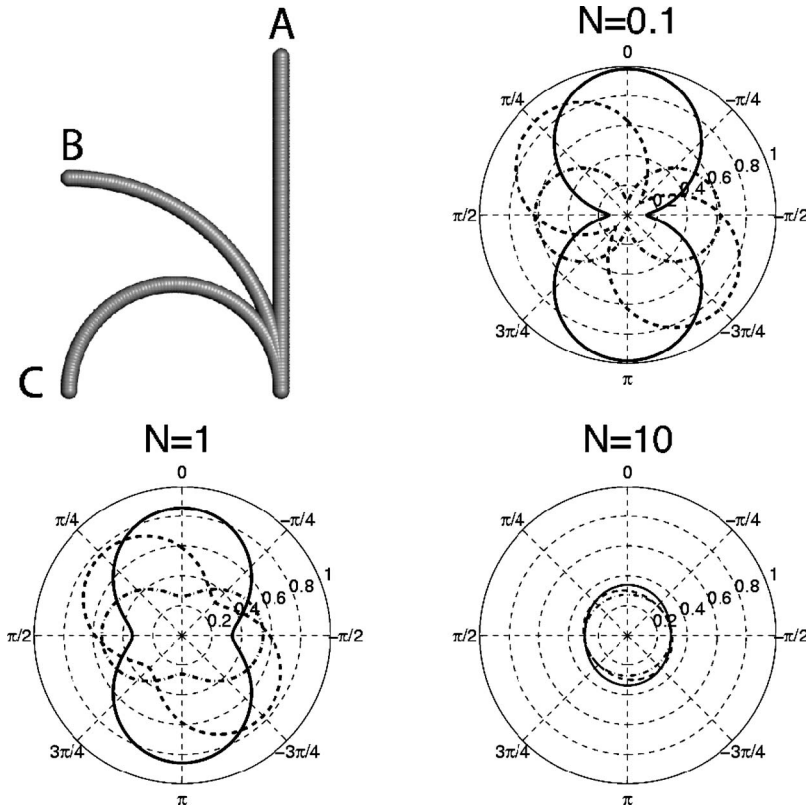
FIG. 2. The root-mean-square displacement $\sqrt{M_2/N^2}$ (length nondimensionalized by the contour length L) projected along the direction $\vec{e}(\chi)$ of a chain in two dimensions with the initial tangent pointing up along the x axis and the other end with orientation θ. The upper left-hand figure shows minimum energy conformations for $\theta=0$ (A), $\theta=\pi/2$ (B), and $\theta=\pi$ (C). We plot the root-mean-square displacement as the radial distance (radial distance ranges from zero to one) against the polar angle χ for $\theta=0$ (solid curve), $\theta=\pi/2$ (dashed curve), and $\theta=\pi$ (dashed-dotted curve) for the dimensionless chain length $N=L/(2l_p)$ equal to 0.1, 1, and 10.

prevalent for perfectly flexible chains despite the fact that length of the polymer is 20 times the persistence length. To quantify the waning coupling between chain conformation and end orientation, we define the following metric:

$$\Delta_{\parallel} = \frac{1}{\pi^2/4 - 1}\left(\frac{M_2(\chi=\pi/2, \theta=\pi, \theta_0=0)}{M_2(\chi=0, \theta=0, \theta_0=0)} - 1\right). \quad (28)$$

The metric $\Delta_{\parallel}$ gives the relative discrepancy between the maximum displacement for chain ends that are oriented parallel and antiparallel to each other and tends to one as $N\rightarrow 0$ and zero as $N\rightarrow\infty$. In Fig. 3, we plot $\Delta_{\parallel}$ over several decades of N. We see in Fig. 3 that the shoulder in $\Delta_{\parallel}$ occurs at $N\sim 3$. For $N>3$, $\Delta_{\parallel}$ tends to zero as N^{-1}, in agreement with the first-order correction to Gaussian chain statistics that couples end position to end orientations [10,11]. Figure 3 suggests the end orientations substantially influence the end displacement for semiflexible polymers, even for relatively large values of $N=L/(2l_p)$. For example, the mean-square end displacement for a chain that is 20 persistence lengths long ($N=10$) exhibits a 24% difference between the case of ends oriented parallel and the case of ends oriented antiparallel.

We now consider the end distribution function of a worm-like chain in two dimensions with one end fixed at the origin, pointed along the x axis, and the other end free. We note that this problem was previously addressed using Monte Carlo simulation [51] and analytical theory in the stiff-chain limit [52]; our results are consistent with these results, but reveal several features that were not addressed in these studies. In addressing this problem, we demonstrate the role that correlation between the orientation of the tangent vector of the

fixed chain end and the vector distance of the free end plays in the chain statistics and provide further insight into the cause of the anomalous statistical behavior at intermediate chain rigidity [27,28]. Since this anomalous behavior exists
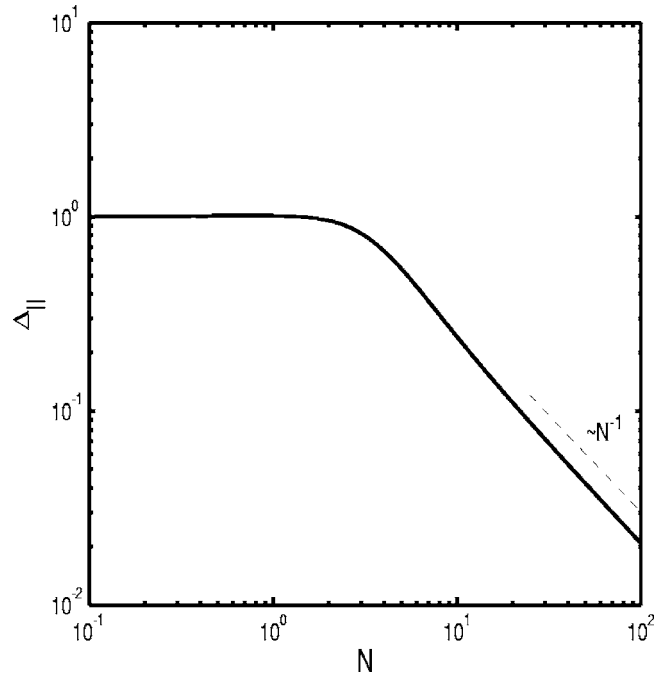

FIG. 3. The metric $\Delta_{\parallel}$ [defined by Eq. (28)] versus the chain length $N=L/(2l_p)$, showing the waning discrepancy between the maximum end displacement for chain ends that are oriented parallel and antiparallel to each other.

in both two and three dimensions, we restrict our analysis to the two-dimensional case for simplicity.

The statistical behavior of a semiflexible polymer with one end clamped at the origin and its tangent oriented in the direction $\hat{x}$ is governed by the Green function averaged over the orientation of the other end. Using results from the preceding section, the Green function for our current problem is

$$G(\vec{k}|\vec{u}_0;p) = \frac{1}{2\pi}G_0^0(K;p) + \frac{1}{\pi}\sum_{m=1}^{\infty} \cos\chi G_0^m(K;p), \quad (29)$$

where $\vec{u}_0 = \hat{x}$, χ gives the orientation of $\vec{k}$, $K = 2l_p k$, and the functions G_0^m are the infinite continued fractions derived in the preceding section.

The Green function in real space is found by Fourier and Laplace inversions of Eq. (29), which requires approximate evaluation of the infinite continued fractions. In our calculations, we truncate the infinite continued fractions at the tenth level, thus the summation over m in Eq. (29) stops at $m=10$. The Laplace inversion is performed on the resulting function by numerically finding the residues of the truncated continued fraction, which are all simple poles. The Fourier inversion involves an integral over the angle χ and an integral over the $\vec{k}$-vector magnitude K. The former is performed analytically, and the latter is performed numerically from $K=0$ to $K=K_{\max}$, where $K_{\max}$ is a large-K cutoff value that is sufficiently large to achieve convergence.

We use several metrics to determine the numerical accuracy; for example, the moments $\langle R^{2n}\rangle$ and $\langle(\vec{R}\cdot\vec{u}_0)^n\rangle$ are numerically calculated using our results and compared to analytical solutions. The utility of our solutions as moment generating functions is demonstrated for the moment $\langle R^{2n}\rangle$ in Ref. [30] and for the moment $\langle(\vec{R}\cdot\vec{u}_0)^n\rangle$ using the following expression:

$$\langle(\vec{R}\cdot\vec{u}_0)^n\rangle = (-i)^n 2\pi \left.\frac{\partial^n G(\vec{k}|\vec{u}_0;p)}{\partial K^n}\right|_{K=0,\chi=0}, \quad (30)$$

where $\vec{R}$ is nondimensionalized by $2l_p$. We compare our results for $n=1, 2, 3, 4$ with the analytical expressions and find good agreement for both $\langle R^{2n}\rangle$ (within 1.5% difference for $0.5 < N < 10.0$) and $\langle(\vec{R}\cdot\vec{u}_0)^n\rangle$ (within 0.08% difference for $0.5 < N < 10.0$). In Fig. 4, we show our results for the moments $\langle(\vec{R}\cdot\vec{u}_0)^n\rangle$ for $n=1, 2, 3, 4$ compared with the analytical values found using Eq. (30), over the range $0.5 < N < 10.0$. In principle, the order of the calculation determines the number of moments that are accurately calculated, i.e., the order is equal to the highest moment that is exactly captured in the statistics. All higher moments are approximate, though with better accuracy for larger chain length. Therefore, our results capture the first 10 moments exactly and all higher moments approximately, with the extremely slight inaccuracy displayed in Fig. 4 attributed to the numerical Fourier inversion.

In Fig. 5, we present our results for the end distance vector distribution function for a semiflexible polymer with one end fixed at the origin with tangent orientation $\hat{x}$ (pointing up) and the other end located at position (x, y) in space over
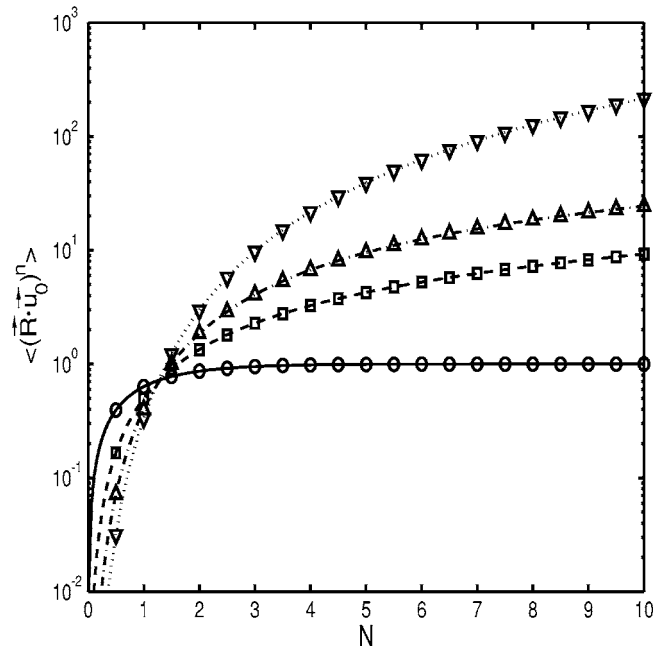

FIG. 4. The first four moments of the end distance vector distribution function for a semiflexible polymer with one end fixed at the origin pointed in the x direction ($\vec{u}_0 = \hat{x}$). The figure includes both analytical results (lines) and values calculated using the numerical Fourier-Laplace inversion of Eq. (29) (data points) for $\langle(\vec{R}\cdot\vec{u}_0)^n\rangle$ with $n=1$ ($\bigcirc$), $n=2$ ($\square$), $n=3$ ($\triangle$), and $n=4$ (∇).

a range of chain rigidities ($N=1.0$ to $N=6.5$). The end distance vector distribution functions in Fig. 5 have finite values for $R < 1$ (nondimensionalized by the chain length L) and are zero for $R > 1$; therefore, these density plots, which show high probability in white (red online) and low probability in black (blue online), only display values for $R \leq 1$. As $N \to 0$, the location of the chain end is fixed at a distance $R=1$ away from the origin along the direction of the initial orientation $\vec{u}_0 = \hat{x}$, indicating a perfectly straight chain. As the rigidity decreases (or as N increases), the end distance vector distribution function fans out around the edge of the circle of influence ($R \leq 1$) due to transverse fluctuations of the chain. Simultaneously, the peak position along the x-axis retracts towards the origin since the transverse fluctuations cause the ends to pull towards each other due to the fixed contour length of the chain. From $N=1.0$ to $N=2.5$, the end distance vector distribution function fans out until it essentially wraps around towards the origin. The origin is the entropically favored position for the free end since there are many more paths that the chain can take.

Near $N=2.6$, a second peak in the distribution emerges when the tips of the fan touch each other near the origin. We identify this peak as the entropic peak and the other peak, located further from the origin along the direction of the initial tangent, as the energetic peak. For $2.6 < N < 5.6$, both the entropic peak and the energetic peak are present in the end distance vector distribution function; as the rigidity decreases, the dominant peak shifts from the energetic peak to the entropic peak. As the entropic peak grows, the fanning of the distribution function begins to meld into a circularly symmetric feature. This is indicative of the bending energy
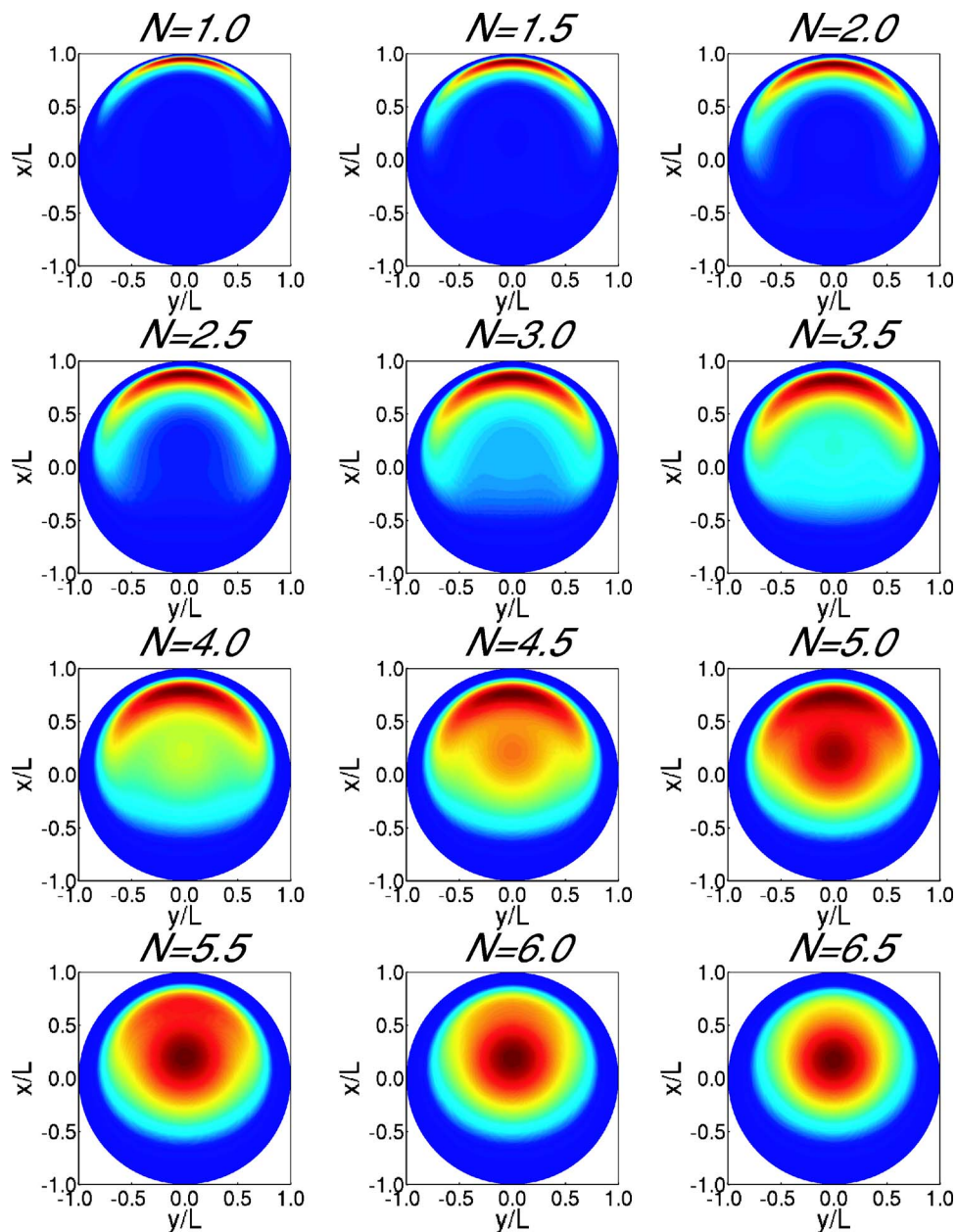
FIG. 5. (Color online) The probability that a chain in two dimensions with one end fixed at the origin with tangent orientation $\hat{x}$ (pointing up) will have the other end located at position (x, y) in space (unrestricted end orientation) for various values of the dimensionless length $N = L/(2l_p)$. These density plots show high probability in white (red) and low probability in black (blue). We provide an animated version of this figure at http://nature.berkeley.edu/~ajspakow

having less influence over the chain conformation and the entropy dominating the end distance vector distribution function. The energetic peak shrinks with N as the entropic peak emerges and grows, until the energetic peak is quelled by the growing entropic peak at $N=5.6$.

For $N>5.6$, the end distance vector distribution function has a single peak associated with the entropic peak described in the preceding paragraph. The location of the peak is slightly away from the origin in the direction of the initial tangent vector; as $N \rightarrow \infty$, the peak location slowly creeps toward the origin. This is due to the waning influence of the chain rigidity over the conformation statistics. We note, however, that the peak remains eccentric for relatively flexible

chains; for example, the peak location for $N=10.0$ is at $x/L=0.1111$.

In Fig. 6, we show the location and the value of the probability for the maxima in the end distance vector distribution versus the dimensionless chain length $N=L/(2l_p)$. Figure 6(A) shows the location of the entropic peak (solid curve) and the energetic peak (dashed curve) along the x axis, and Fig. 6(B) gives the value of the end distance vector distribution function at each peak. As $N \rightarrow 0$, the energetic peak proceeds to $x/L=1$ in a linear fashion as the chain straightens out. This coincides with the end distance vector distribution function becoming a delta function at $x/L=1$ and $y/L=0$, and the peak height in the distribution function increases
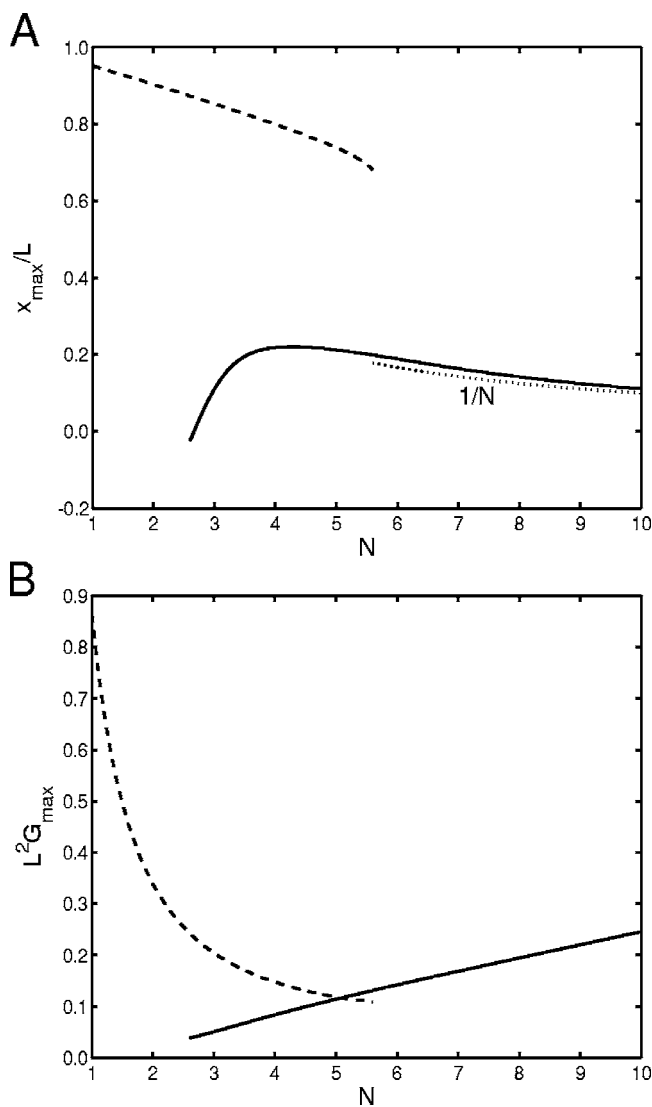
FIG. 6. Location (found on the x axis) and peak value of the maxima in the end distance vector distribution function for a chain in two dimensions with one end fixed at the origin with tangent orientation $\hat{x}$ versus the dimensionless length $N=L/(2l_p)$. For $2.6<N<5.6$, there are two peaks in the probability distribution, which we refer to as the energetic peak (dashed curve) and the entropic peak (solid curve). (A) shows the position of the peaks along the x axis along with the long-chain behavior of the peak location $1/N$ (dotted curve). (B) gives the value of the probability evaluated at the peaks.

exponentially as $N \rightarrow 0$. For $2.6<N<5.6$, both the energetic peak and the entropic peak are present in the end distance vector distribution function. At $N=2.6$, the energetic peak dominates the distribution function, and as N increases, the dominant peak shifts to the entropic peak (equal probability occurs at $N=5.095$). As the rigidity decreases, the energetic peak moves monotonically toward the origin with N; however, the entropic peak first shifts away from the origin from $N=2.6$ to $N=4.3$ and then back toward the origin for $N>4.3$. For $N>5.6$, the energetic peak disappears, and the end distance vector distribution function has only the entropic peak. The location of the peak slowly tends to the

origin as the contribution of the bending rigidity is overwhelmed by the conformational entropy that prefers the distribution to be a Gaussian function that is centered at the origin. Since the initial orientation is specified, the chain end position tends to be off-centered by a Kuhn segment length $2l_p$ in the direction of the initial orientation for chains that are much longer than the persistence length l_p. Therefore, we expect the location of the peak x_{max}/L to tend to zero as $1/N$; we include this long-chain behavior in Fig. 6(A) (dotted curve).

We observe a double-peaked end distance vector distribution function between $N=2.6$ and $N=5.6$, similar to the behavior observed in the end-to-end distance distribution function with both ends free to rotate [27,28]. However, the end-to-end *distance* distribution does not unambiguously clarify whether the effect is real because the double-peak feature in the end-to-end distance distribution could potentially arise from the rotational average. Our results verify that the double-peaked end-to-end distance distribution functions in Refs. [27,28] indeed reflect the underlying conformation of the polymer and were not artificial effects from the rotational averaging.

The double-peaked end-to-end distribution affects the equilibrium behavior of semiflexible polymers with potentially important consequences on the biological function of DNA. The process of looping DNA plays an important role in gene regulation in the cell, as many regulatory proteins function by looping short segments of DNA with fixed end positions and orientations [38–41]. At intermediate chain lengths, the double-peaked end-to-end distribution implies a nonmonotonic force history as the chain ends are drawn into coincidence [27,28,53,54]. The behavior of DNA at short chain lengths may require more elaborate models than the wormlike chain model [54–57]; however, such models tend to enhance the looping probability and the nonmonotonicity of the force history during loop formation [54]. Since the probability modulates only slightly between the two dominant peaks, the force remains relatively small as the chain proceeds from open to closed. As such, the chain ends are expected to experience large fluctuations for chain lengths that exhibit the double-peaked distribution.

The pathway that the chain end traverses during the looping process is governed by the free energy as determined by the end-to-end distribution function. Although the end-to-end *distance* distribution function gives the thermodynamics of end separation, the pathway of the chain end during looping cannot be deduced unless the initial orientation is specified, as in Fig. 5. The formation of a loop for a rigid chain occurs by the chain looping around while maintaining a smoothly bent conformation; this sweeping trajectory is clearly exhibited in the fanlike probability function in $N=1.0$ in Fig. 5. For chains of moderate rigidity where the double-peaked distribution occurs, the trajectory of the chain end follows more closely along the direction of the initial orientation, indicating that the looping is more of an end retraction rather than a sweeping loop (for example, see $N=5.0$ in Fig. 5). The process of looping a flexible polymer occurs by a diffusive process of the chain end exploring its surroundings, as indicated by the limiting behavior of the end-to-end distribution tending to a Gaussian as $N \rightarrow \infty$.

The force-extension behavior of the wormlike chain model is monotonic for sufficiently long chain lengths [9]. However, short chain lengths exhibit a variety of possible behaviors depending on the specific conditions imposed at the ends; for example, the force-extension behavior of a chain under constant tension is different than that of a chain under fixed end separation [53]. In this paper, we demonstrate the effect of fixing the end orientations on the chain end statistics. Understanding the behavior of semiflexible polymers at small length scales requires special consideration for the particular conditions at the chain ends, as demonstrated in Ref. [53] and in our present work.

IV. SUMMARY

In this paper, we find the exact solution for the end-to-end distance vector distribution function for a wormlike chain in two and three dimensions that incorporate both the end position and the tangent orientation at the end-points. This study extends our previous work on a wormlike chain in an aligning field [30], which made use of diagrammatic techniques [47,48] to find exact expressions for the partition function of a wormlike chain in an aligning field. Our present results, as well as those found in Ref. [30], adopt the form of infinite continued fractions, which naturally emerge upon noting the hierarchical structure of the moment-based expansions employed in the derivation. The diagrammatic representation results in a formalism that simplifies mathematical manipulations and provides a visually convenient means of identifying similar terms in the expansion. These methods can be useful for addressing a wide range of physical problems involving semiflexible polymers.

The chain statistics derived in this paper permit the evaluation of spatial averages for a chain with one or both ends fixed. We present the root-mean-square displacement in a given direction for a chain with both end orientations fixed. For rigid chains $[N=L/(2l_p)<1]$, the root-mean-square displacement tends to that of the energy-minimum shapes, which are anisotropic and show a strong coupling between end orientation and chain conformation. As $N \rightarrow \infty$, the influence of the end orientations over the chain conformation diminishes, marked by the crossover to a root-mean-square displacement that is rotationally invariant. However, this crossover occurs at intermediate values of N, and the coupling between end orientations and conformation remains strong for relatively flexible chains.

We further demonstrate our results by considering the end distribution function of a wormlike chain in two dimensions with one end clamped at the origin and pointed in a fixed direction. Due to the rapid convergence of the continued fractions, we are able to achieve good numerical accuracy by truncation after just a few layers, as demonstrated by good agreement with the analytical solutions for $\langle (\vec{R} \cdot \vec{u}_0)^n \rangle$ over a wide range of chain rigidities $(0.5 < N < 10.0)$. The simplicity and conciseness of our solutions should thus facilitate their use in studying various aspects of the statistical mechanics of semiflexible polymers.

The end distance vector distribution function of a wormlike chain with one clamped end in two dimensions exhibits a double-peaked structure for intermediate chain rigidities $(2.6 < N < 5.6)$. These results are consistent with previous results for the end-to-end distance distribution function with two free ends [27,28] and one free end [51,52]. The progression of the end distance vector distribution function from a rigid chain to a flexible chain exhibits an interesting crossover behavior where the distribution fans out from the straight chain peak until the ends of the fan touch, a second peak emerges at the point where the fan tips meet, and the second peak grows until it drowns out the original peak. As the chain rigidity decreases further, the end distance vector distribution function becomes circularly symmetric and the remaining peak slowly creeps toward the origin. We find that the end distance vector distribution function remains eccentric even for relatively flexible chains; for example, the peak in the distribution function remains a distance of 10% the contour length away from the origin for $N=10$, suggesting the distribution function is still substantially non-Gaussian for $N=10$, which agrees with our results for the root-mean-square end displacement with both ends fixed. These results suggest that semiflexibility plays an important role over a wide range of chain rigidities.

[1] O. Kratky and G. Porod, Recl. Trav. Chim. Pays-Bas **68**, 1106 (1949).

[2] X. J. Wang and M. Warner, J. Phys. A **19**, 2215 (1986).

[3] A. J. Liu and G. H. Fredrickson, Macromolecules **26**, 2817 (1993).

[4] A. J. Spakowitz and Z.-G. Wang, J. Chem. Phys. **119**, 13113 (2003).

[5] K. Ghosh, G. A. Carri, and M. Muthukumar, J. Chem. Phys.

115, 4367 (2001).

[6] K. Ghosh and M. Muthukumar, J. Polym. Sci., Part B: Polym. Phys. **39**, 2644 (2001).

[7] F. C. MacKintosh, J. Käs, and P. A. Janmey, Phys. Rev. Lett. **75**, 4425 (1995).

[8] D. A. Head, A. J. Levine, and F. C. MacKintosh, Phys. Rev. E **68**, 061907 (2003).

[9] J. F. Marko and E. D. Siggia, Macromolecules **28**, 8759

(1995).

[10] H. E. Daniels, Proc. R. Soc. Edinburgh, Sect. A: Math. Phys. Sci. **63**, 290 (1952).

[11] H. Kleinert, *Path Integrals in Quantum Mechanics, Statistics, Polymer Physics, and Financial Markets* (World Scientific, Singapore, 2004).

[12] H. Yamakawa and M. Fujii, J. Chem. Phys. **59**, 6641 (1973).

[13] J. Wilhelm and E. Frey, Phys. Rev. Lett. **77**, 2581 (1996).

[14] J. J. Hermans and R. Ullman, Physica (Amsterdam) **18**, 951 (1952).

[15] B. Hamprecht, W. Janke, and H. Kleinert, Phys. Lett. A **330**, 254 (2004).

[16] G. S. Chirikjian and Y. F. Wang, Phys. Rev. E **62**, 880 (2000).

[17] N. Saito, K. Takahashi, and Y. Yunoki, J. Phys. Soc. Jpn. **22**, 219 (1967).

[18] R. A. Harris and J. E. Hearst, J. Chem. Phys. **44**, 2595 (1966).

[19] Y. Tagami, Macromolecules **2**, 8 (1969).

[20] K. F. Freed, J. Chem. Phys. **54**, 1453 (1971).

[21] M. G. Bawendi and K. F. Freed, J. Chem. Phys. **83**, 2491 (1985).

[22] J. B. Lagowski, J. Noolandi, and B. Nickel, J. Chem. Phys. **95**, 1266 (1991).

[23] R. G. Winkler, P. Reineker, and L. Harnau, J. Chem. Phys. **101**, 8119 (1994).

[24] R. G. Winkler, L. Harnau, and P. Reineker, Macromol. Theory Simul. **6**, 1007 (1997).

[25] A. L. Kholodenko, Ann. Phys. (N.Y.) **202**, 186 (1990).

[26] A. L. Kholodenko, J. Stat. Phys. **65**, 291 (1991).

[27] J. Samuel and S. Sinha, Phys. Rev. E **66**, 050801(R) (2002).

[28] A. Dhar and D. Chaudhuri, Phys. Rev. Lett. **89**, 065502 (2002).

[29] S. Stepanow and G. M. Schütz, Europhys. Lett. **60**, 546 (2002).

[30] A. J. Spakowitz and Z.-G. Wang, Macromolecules **37**, 5814 (2004).

[31] S. Stepanow, Eur. Phys. J. B **39**, 499 (2004).

[32] D. M. Crothers, J. Drak, J. D. Kahn, and S. D. Levene, Methods Enzymol. **212**, 3 (1992).

[33] R. Schleif, Annu. Rev. Biochem. **61**, 199 (1992).

[34] D. Shore, J. Langowski, and R. L. Baldwin, Proc. Natl. Acad. Sci. U.S.A. **78**, 4833 (1981).

[35] T. E. Cloutier and J. Widom, Mol. Cell **14**, 355 (2004).

[36] J. Shimada and H. Yamakawa, Macromolecules **17**, 689 (1984).

[37] Y. L. Zhang and D. M. Crothers, Biophys. J. **84**, 136 (2003).

[38] M. Geanacopoulos, G. Vasmatzis, V. B. Zhurkin, and S. Adhya, Nat. Struct. Biol. **8**, 432 (2001).

[39] J. Plumbridge and A. Kolb, Nucleic Acids Res. **26**, 1254 (1998).

[40] K. Rippe, P. H. Vonhippel, and J. Langowski, Trends Biochem. Sci. **20**, 500 (1995).

[41] M. Brenowitz, A. Pickar, and E. Jamison, Biochemistry **30**, 5986 (1991).

[42] A. Balaeff, L. Mahadevan, and K. Schulten, Phys. Rev. Lett. **83**, 4900 (1999).

[43] E. Villa, A. Balaeff, and K. Schulten, Biophys. J. **84**, 462A (2003).

[44] A. E. H. Love, *A Treatise on the Mathematical Theory of Elasticity* (Dover, New York, 1944).

[45] L. D. Landau and E. M. Lifshitz, *Theory of Elasticity* (Pergamon, New York, 1986).

[46] R. P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integrals* (McGraw-Hill, New York, 1965).

[47] H. Yamakawa, J. Chem. Phys. **59**, 3811 (1973).

[48] H. Yamakawa, *Helical Wormlike Chains in Polymer Solutions* (Springer-Verlag, New York, 1997).

[49] G. B. Arfken and H. J. Weber, *Mathematical Methods for Physicists* (Academic, San Diego, 1995).

[50] V. Devanathan, *Angular Momentum Techniques in Quantum Mechanics* (Kluwer Academic, Boston, 1999).

[51] G. Lattanzi, T. Munk, and E. Frey, Phys. Rev. E **69**, 021801 (2004).

[52] P. Benetatos and E. Frey, Phys. Rev. E **70**, 051806 (2004).

[53] S. Sinha and J. Samuel, Phys. Rev. E **71**, 021104 (2005).

[54] P. Ranjith, P. B. S. Kumar, and G. I. Menon, Phys. Rev. Lett. **94**, 138102 (2005).

[55] J. Yan and J. F. Marko, Phys. Rev. Lett. **93**, 108108 (2004).

[56] P. A. Wiggins, R. Phillips, and P. C. Nelson, Phys. Rev. E **71**, 021909 (2005).

[57] P. A. Wiggins and P. C. Nelson, cond-mat/0508155v1 (2005).