

**Crucial role of the intrinsic twist rate for the size of an intrinsically curved semiflexible biopolymer**Zicong Zhou \**Department of Physics, Tamkang University, New Taipei City, 25137 Taiwan, Republic of China*

(Received 26 June 2020; accepted 24 August 2020; published 8 September 2020)

We study the effects of the intrinsic curvature (IC), intrinsic twist rate (ITR), anisotropic bending rigidities, sequence disorder, and temperature ( $T$ ) on the persistence length ( $l_p$ ) of a two- or three-dimensional semiflexible biopolymer. We develop some general expressions to evaluate exactly these effects. We find that a moderate IC alone reduces  $l_p$  considerably. Our results indicate that the centerline of the filament keeps as a helix in a rather large range of  $T$  when ITR is small. However, a large ITR can counterbalance the effect of IC and the result is insensitive to the twist rigidity. Moreover, a weak randomness in IC and ITR can result in an “overexpanded” state. Meanwhile, when ITR is small,  $l_p$  is not a monotonic function of  $T$  but can have either minimum or maximum at some  $T$ , and in the two-dimensional case the maximum is more obvious than that in the three-dimensional case. These results reveal that to obtain a proper size at a finite  $T$  for an intrinsically curved semiflexible biopolymer, proper values of bending rigidities and ITR are necessary but a large twist rigidity may be only a by-product. Our findings are instructive in controlling the size of a semiflexible biopolymer in organic synthesis since the mean end-to-end distance and radius of gyration of a long semiflexible biopolymer are proportional to  $l_p$ .

DOI: [10.1103/PhysRevE.102.032405](https://doi.org/10.1103/PhysRevE.102.032405)**I. INTRODUCTION**

Different semiflexible biopolymers exhibit different thermal and mechanical properties. One of the most significant properties for a semiflexible biopolymer is its statistical measure of size, characterized by mean end-to-end distance or mean radius of gyration or persistence length ( $l_p$ ).  $l_p$  represents an effective statistical segment length of a coarse-grained model, in which one replaces the semiflexible biopolymer by a random walk with the same contour length and mean end-to-end distance. In other words,  $l_p$  is approximately the nondeformable length of a semiflexible biopolymer, and the mean end-to-end distance and radius of gyration of a long semiflexible biopolymer are proportional to  $l_p$ . A proper magnitude of  $l_p$  is therefore crucial to the function of semiflexible biopolymers. For instance, in a highly compacted semiflexible biopolymer it may be difficult to identify the start site in transcription, replication, recombination, and repair processes. In contrast, a loose semiflexible biopolymer with a large  $l_p$  may have a too-large volume and may encounter too many attacks from environment so is fragile.

Intuitively, the bending rigidity should dominate  $l_p$ , since the larger the bending rigidity, the larger the force or energy required to bend the semiflexible biopolymer. The result obtained from the simplest model for a semiflexible biopolymer, the wormlike chain (WLC) model, supports this conjecture [1–4]. However, many other parameters can also affect the value of  $l_p$ , such as temperature ( $T$ ), a finite intrinsic curvature (IC) or twist rate (ITR), anisotropic bending rigidities, and randomness in sequences. Having a finite IC means that free

of external force or torque, the natural shape of the semiflexible biopolymer is a curve so the semiflexible biopolymer has a well-defined curvature in its ground-state configuration (GSC, or the configuration with the lowest energy). Meanwhile, having a finite ITR means that free of external force or torque, the semiflexible biopolymer is naturally twisted so has a unique twist rate in its GSC. Many semiflexible biopolymers, such as some double-stranded DNA (dsDNA), have a finite IC or ITR. Consequently, what parameter can be as important as the bending rigidity is a significant topic.

$T$  is certainly the first factor we should consider and it is indeed very important in many systems. For instance, it has been found that a flexible polymer has three phases below, at or above the  $\theta$  temperature [5], which is the temperature at which the phase transition takes place. At  $\theta$  temperature, the conformation of a polymer can be modeled as a random walk of monomer subunits, and such a polymer is referred as an ideal chain or a free-jointed chain. Correspondingly, a solvent at the  $\theta$  temperature is called  $\theta$  solvents [5]. On the other hand, in good solvent or above the  $\theta$  temperature, the chain is in a swollen state and can be described by a self-avoiding random walk [5–10]. Moreover, in poor solvent or below  $\theta$  temperature, a polymer chain collapses into a compact globule state. In these three phases, the end-to-end distance of polymers obey different power law [5–10].  $T$  also affects greatly the mechanical property of a polymer. For swollen or ideal polymer chains, the polymer extends progressive with increasing stretching force so that no sharp transition in extension occurs [11–14]. However, in poor solvent, the extension of a polymer can subject to a first-order transition under a stretching force [15].

There is no way to define a finite IC for a flexible polymer since it can be arbitrarily bent. In contrast, a semiflexible

\*zzhou@mail.tku.edu.tw

biopolymer, such as dsDNA, actin, and microfilament, always has a unique GSC or IC. Due to its slender shape, a semiflexible biopolymer is usually modeled as a filament. The simplest model for a semiflexible biopolymer is the WLC model which views a filament as an inextensible chain with a finite bending rigidity but a zero cross section, and it has been used to account for the entropic elasticity of some semiflexible biopolymers [1–4]. Another useful model is the WLRC model [1–3,16–18] which regards a filament as a chain with a finite ITR and a circular cross section. The importance of semiflexible biopolymers stimulates lots of researches in these two models and some relevant models [1–4,16–46].

Both WLC and WLRC models are intrinsically straight or have a zero IC, i.e., the natural shape of their centerlines are straight lines. However, intrinsically curved or twisted semiflexible biopolymers are also ubiquitous. For instance, special sequence orders favor a finite IC for some short dsDNA chains so they look like circles [47–51]. It has also been reported that a long-range correlated dsDNA has a macroscopic (intrinsic) curvature so that the WLC model fails to account for its property [19]. Moreover, it has been found that an intrinsically curved filament has complete different conformational and mechanical properties from that of WLC and WLRC [20–27]. In the three-dimensional (3D) case, both exact theoretical analysis and experimental observation reveal that an intrinsically curved semiflexible biopolymer can form a stable helix and under an applied force the extension of the helix can subject to a discontinuous transition [20–26]. Meanwhile, in the 2D case it has been shown that a finite IC alone can induce a discontinuous transition in extension [27].

On the other hand, at a first approximation a filament is assumed to be homogeneous. However, the structure of a semiflexible biopolymer is usually sequence dependent. Sequence disorder results in disordered IC and ITR and affects considerably properties of semiflexible biopolymers [19,31–36]. Without correlation or with short-range correlation (SRC) in sequences and free of external force, it has been shown rigorously that the effects of sequence disorder on IC and ITR can be incorporated into a model with well-defined mean IC and ITR as well as renormalized bending and twist rigidities [19,28,29,31–33]. In contrast, with long-range correlation (LRC) in sequences, a simple correction to bending and twist rigidities is invalid [19] but the model with a finite IC can remedy the problem [46].

Relationships among  $l_p$ , IC, and ITR have been studied [28–30,44,45]. These works present some general expresses and offer some intrigue results, but the knowledge on the dominative parameters for  $l_p$  is yet incomplete. In particular, the role of sequence disorder to  $l_p$  is not transparent. In this work we clarify this problem. We find that a moderate IC alone reduces  $l_p$  considerably. In contrast, a large ITR can depress the effect of IC so yield a  $l_p$  that is comparable to the  $l_p$  of the WLC model, and the result is insensitive to the twist rigidity. Moreover, the couple between IC or ITR and randomness can result in an “overexpanded” state. Our results also suggest that the elasticity of a filament with a large ITR should be the same as that of a WLC. Therefore, a finite ITR plays a crucial role for the size of an intrinsically curved semiflexible biopolymer.

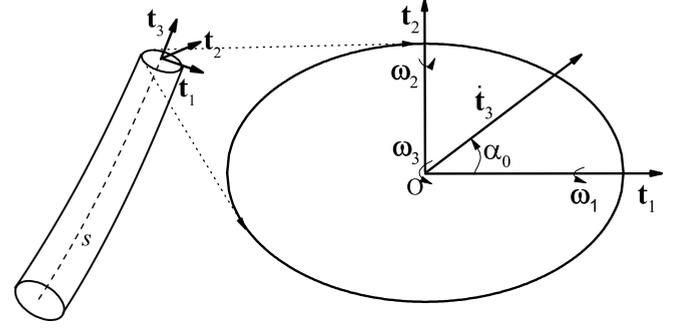


FIG. 1. Schematic diagram of a 3D filament (left) and a cross section (right) of it showing the notation used. The dashed line in the left figure is the centerline of the filament, and the origin of the coordinate system for the cross section is at the centerline of the filament,  $\mathbf{t}_3$  is the tangent of the centerline and is perpendicular to the cross section.  $\omega_i = \omega_i \mathbf{t}_i$  ( $i = 1, 2, 3$ ) rotates around  $\mathbf{t}_i$ .

The paper is organized as follows. In the next section we set up models and present some basic expressions. The Sec. III presents the results for a 3D filament. The Sec. IV discusses the 2D system. Finally, the conclusions and discussions conclude the paper.

## II. MODELS AND BASIC EXPRESSIONS

### A. Three-dimensional model

The configuration of a 3D filament can be described by a triad of unit vectors  $\{\mathbf{t}_i(s)\}_{i=1,2,3}$  [21,30,38], where  $\mathbf{t}_3 = \dot{\mathbf{r}}$  is the tangent to the locus of centerline  $\mathbf{r}(s)$  of the filament,  $\mathbf{t}_1$  and  $\mathbf{t}_2$  orient along the principal axes of the cross section,  $s$  is the arclength of the centerline, and the symbol “ $\dot{\quad}$ ” represents the derivative with respect to  $s$ . The triad obeys the generalized Frenet equations  $\dot{\mathbf{t}}_i = \boldsymbol{\omega} \times \mathbf{t}_i$  [21,25,26,30,38], where  $\boldsymbol{\omega} = \omega_1 \mathbf{t}_1 + \omega_2 \mathbf{t}_2 + \omega_3 \mathbf{t}_3$  is a vector in which  $\omega_1$  and  $\omega_2$  are components of curvature  $c = \sqrt{\omega_1^2 + \omega_2^2}$ , and  $\omega_3$  is the twist rate. The direction of the curvature is given by  $\mathbf{t}_3$  [25]. Figure 1 shows the relations in these quantities.

The elastic energy of a filament with a finite IC and ITR can be written as [21,30,38]

$$E_{3D} = \frac{1}{2} \int_0^L \sum_{i=1}^3 k_i (\omega_i - \zeta_i)^2 ds, \quad (1)$$

where  $k_1$  and  $k_2$  are bending rigidities along two principle axes on the cross section and  $k_3$  is the twist rigidity.  $\zeta_1$  and  $\zeta_2$  are components of IC ( $= \zeta_0 = \sqrt{\zeta_1^2 + \zeta_2^2}$ ) and  $\zeta_3$  is ITR.  $\zeta_0$  is the natural curvature of the centerline, i.e., the curvature of the centerline when the filament is free of external force and at its GSC. Similarly,  $\zeta_3$  characters the natural twist rate of the cross section. The intrinsic radius is  $R_0 = 1/\zeta_0$ . In the model  $L$  is the contour length of the centerline and is a constant so that the filament is inextensible. This is a reasonable approximation since the contour length of a semiflexible biopolymer is decided by the covalent bonds between monomers and it is uneasy to change the bond-length. When  $k_1 = k_2 = k$ , the

filament is isotropic or has a circular cross section if it is uniform, i.e., has a constant density. In contrast,  $k_1 \neq k_2$  results from an elliptical cross section for a uniform filament. When  $\zeta_1 = 0$  and  $k_1 = k_2$ , the model becomes the helical wormlike chain model [44,45]. Moreover, when  $\zeta_0 = 0$  and  $k_1 = k_2$ , the model is reduced into the WLRC model. Meanwhile, when  $\zeta_3 = 0$ , the WLRC model becomes the WLC model. In both WLC and WLRC models,  $l_p = l_p^0 \equiv k/k_B T$ . We also let  $\zeta_1 = \zeta_0 \cos \alpha_0$  and  $\zeta_2 = \zeta_0 \sin \alpha_0$  since  $l_p$  is independent of  $\alpha_0$  when  $k_1 = k_2 = k$ .  $\alpha_0$  is the angle between IC and  $\mathbf{t}_1$ , as shown in Fig. 1.

Experimentally, bending rigidity of a semiflexible biopolymer can be measured by embedding one end of a short segment of the semiflexible biopolymer into a hard substrate so to fix the initial angle, applying a force to another end to bend it, and then measuring the changes of the bending angle. Meanwhile, twist rigidity can be measured by sticking the cross section of one end tightly on a nondeformable surface, applying a torque to another end, and then measuring the changes of the twisting angle. These measurements can be realized by some single molecule techniques, such as laser or magnetic tweezers, atomic force microscopy, and so on. Larger bending rigidity are usually resulted from stronger covalent bonds between neighbor monomers. Similarly, larger twist rigidity can come from more hydrogen bonds inside a semiflexible biopolymer or larger energy barrier between gauche and trans conformations in molecules.

If  $\zeta_i$  and  $k_i$  are well-defined functions of  $s$  (i.e., free of randomness in sequences or monomers), then a macroscopic quantity is defined as the average over all possible conformations or a path integral for the continuous system in the form [21,29,30,38,52]

$$B = \langle B[\{\omega_i(s)\}] \rangle = \frac{1}{\mathcal{Z}_{3D}} \int \mathcal{D}[\omega_i] B[\{\omega_i(s)\}] e^{-\beta E_{3D}}, \quad (2)$$

$$\mathcal{Z}_{3D} = \int \mathcal{D}[\omega_i] e^{-\beta E_{3D}}, \quad (3)$$

where  $\beta = 1/k_B T$  and  $k_B$  is the Boltzmann constant.  $\omega_i(0)$  and  $\omega_i(L)$  are free in Eq. (2) [52]. The detailed form of  $B[\{\omega_i(s)\}]$  depends on what quantity we want to study and may be a very complex function of  $\omega_i$ . For instance, if  $B[\{\omega_i\}] = \mathbf{t}_j(s_1) \cdot \mathbf{t}_k(s_2)$ , then we find the orientational correlation function between  $\mathbf{t}_j$  and  $\mathbf{t}_k$ ; if  $B[\{\omega_i\}] = |\mathbf{r}_L - \mathbf{r}_0|^2$ , then we obtain the mean square end-to-end distance, where  $\mathbf{r}_L = \mathbf{r}(L)$  and  $\mathbf{r}_0 = 0$ . The partition function  $\mathcal{Z}_{3D}$  is essentially a Gaussian integral so is independent of  $\zeta_i$  but dependent on  $k_i$ .  $B$  is in general a function of  $k_i$ ,  $\zeta_i$ ,  $T$ , and  $L$ .

For a dsDNA chain at room temperature  $T = T_r = 298$  K,  $l_p^0 \approx 50$  nm,  $k_3/k_B T_r \approx 75$  nm, and  $\zeta_3 \approx 1.76$  nm<sup>-1</sup> [1,18]. When  $\zeta_0 = 0$ , the GSC of a dsDNA is a twisted straight cylinder, i.e., its centerline is straight but its cross section is twisted. In this work, our numerical calculations are based on these parameters since we focus on a semiflexible biopolymer and dsDNA is a typical one.

### B. Two-dimensional model

In 2D case, the configuration of a filament is determined by its tangent vector,  $\mathbf{t} = \dot{\mathbf{r}} = \{\cos \phi, \sin \phi\}$ , to its contour line, where  $\mathbf{r} = (x, y)$  is the locus of the filament and  $\phi(s)$  is the

angle between the  $x$  axis and  $\mathbf{t}$ . One end of the chain (at  $s = 0$ ) is fixed at  $\mathbf{r} = 0$ , so the elastic energy of the filament is [19,28,37,53]

$$E_{2D} = \frac{1}{2} \int_0^L ds k (\dot{\phi} - \zeta_0)^2, \quad (4)$$

where  $\dot{\phi} \equiv \dot{x}\ddot{y} - \ddot{x}\dot{y}$  is the signed curvature and  $\zeta_0$  is the intrinsic signed curvature.

If  $k$  and  $\zeta_0$  are well-defined functions of  $s$ , a macroscopic quantity is given by [28,37,52]

$$B(k, \zeta_0, T, L) = \frac{1}{\mathcal{Z}_{2D}} \int \mathcal{D}[\phi] B[\{\phi(s)\}] e^{-\beta E_{2D}}, \quad (5)$$

$$\mathcal{Z}_{2D} = \int \mathcal{D}[\phi] e^{-\beta E_{2D}}. \quad (6)$$

$\phi(0)$  and  $\phi(L)$  are free in Eq. (5) [52]. The partition function  $\mathcal{Z}_{2D}$  is also a Gaussian integral so is independent of  $\zeta_0$  but dependent on  $k$ .

### C. Randomness average

Equations (2) and (5) are defined for a random-free filament. However, semiflexible biopolymers are often sequence dependent or heterogeneous. The disorder in sequences can be characterized by  $s$ -dependent IC or ITR. Suppose that the filament has a well-defined mean IC or ITR; in a coarse-grained model the distribution of IC and ITR can be approximated as a Gaussian distribution [29,46]

$$W(\{\zeta_i\}) = \exp \left\{ - \sum_{i=1}^3 \frac{1}{2} \int k'_i [\zeta_i(s) - \omega_{i0}]^2 ds \right\}, \quad (7)$$

where  $\omega_{i0}$  is the mean of  $\zeta_i$ . In this case, we need to average over  $\zeta_i$  again so an observable is calculated by

$$B(k_i, k'_i, \omega_{i0}) = \frac{\int \mathcal{D}[\zeta_i] W(\{\zeta_i\}) (B[\{\omega_i(s)\}])}{\int \mathcal{D}[\zeta_i] W(\{\zeta_i\})}. \quad (8)$$

The larger the  $k'_i$ , the smaller the standard deviations of the  $\zeta_i$ s so the less the randomness. When  $k'_i \rightarrow \infty$ , it recovers a randomness-free system as given by Eq. (2). Note that  $k_i$  and  $k'_i$  have different units. To study the effect of  $T$  we assume that  $k_i$  and  $k'_i$  are independent of  $T$  since they characterize the intrinsic property of a filament.  $c_0 = \sqrt{\omega_{10}^2 + \omega_{20}^2}$  is the mean IC and the mean intrinsic radius is  $R_0 = 1/c_0$ . We also let  $\omega_{10} = c_0 \cos \alpha_0$  and  $\omega_{20} = c_0 \sin \alpha_0$ .

The physical quantity  $B[\{\omega_i(s)\}]$  is independent of  $\zeta_i$ s so that the integration over  $\zeta_i$ s in Eq. (8) is in fact Gaussian. Therefore, applying a standard path integral technique [54,55] to integrate  $\zeta_i$ s in Eq. (8), we can find exactly [29]

$$B(k_i, k'_i, \omega_{i0}) = \frac{\int \mathcal{D}[\omega_i] B[\{\omega_i(s)\}] e^{-\mathcal{H}_{3D}}}{\int \mathcal{D}[\omega_i] e^{-\mathcal{H}_{3D}}}, \quad (9)$$

$$\mathcal{H}_{3D} = \frac{1}{2} \int_0^L \sum_{i=1}^3 \kappa_i [\omega_i(s) - \omega_{i0}]^2 ds, \quad (10)$$

$$\kappa_i = \beta k_i k'_i / (\beta k_i + k'_i) = k_i k'_i / (k_i + k'_i k_B T). \quad (11)$$

Equations (9)–(11) are valid even if  $k_i$ ,  $k'_i$ , and  $\omega_{i0}$  are  $s$  dependent.

For a 2D system, the distribution of  $\zeta_0$  is [28]

$$W(\{\zeta\}) = \frac{1}{2} \int_0^L ds k'_1 (\zeta_0 - c_0)^2. \quad (12)$$

Similarly to the 3D system, replacing  $k$  by  $\kappa_1$  and  $\zeta_0$  by  $c_0$  in Eqs. (4) and (5), we obtain the results for the 2D sequence-dependent system [28].

#### D. Mean-square end-to-end distance and persistence length

The end-to-end vector of the filament is given by  $\mathbf{r}(L) = \int_0^L \mathbf{t}_3 ds$  so that the mean-square end-to-end distance can be found by

$$\langle \mathbf{r}^2(L) \rangle = 2 \int_0^L \int_0^{s_1} ds_2 ds_1 \langle \mathbf{t}_3(s_1) \cdot \mathbf{t}_3(s_2) \rangle. \quad (13)$$

The persistence length  $l_p$  can be interpreted as an effective segment length of a coarse-grained model, in which one replaces the filament by a random walk with the same  $L$  and mean end-to-end distance, i.e., [30]

$$\begin{aligned} l_p &= \lim_{L \rightarrow \infty} \frac{1}{2L} \langle \mathbf{r}^2(L) \rangle \\ &= \lim_{L \rightarrow \infty} \frac{1}{L} \int_0^L \int_0^{s_1} ds_2 ds_1 \langle \mathbf{t}_3(s_1) \cdot \mathbf{t}_3(s_2) \rangle. \end{aligned} \quad (14)$$

Therefore,  $\langle \mathbf{r}^2(L) \rangle \approx 2l_p L$  for a long filament. Moreover, since the radius of gyration has the same scaling behavior as  $\langle \mathbf{r}^2(L) \rangle$ , it should be proportional to  $l_p L$ .

#### E. Some general expressions

When  $k_i$  and  $\zeta_i$  are  $s$  independent, in 3D case it has been shown rigorously that the orientational correlation functions can be calculated by ( $s_1 > s_2$ ) [30]

$$\langle \mathbf{t}_i(s_1) \cdot \mathbf{t}_j(s_2) \rangle = [e^{-\Lambda(s_1-s_2)}]_{ij}, \quad (15)$$

$$\Lambda_{ij} = \gamma_i \delta_{ij} + \sum_{l=1}^3 \epsilon_{ijl} \zeta_l, \quad (16)$$

$$\gamma_i = k_B T \left( \sum_{j=1}^3 \frac{1}{2k_j} - \frac{1}{2k_i} \right), \quad (17)$$

where  $\epsilon_{ijk}$  is the antisymmetric tensor.

We can diagonalize  $\Lambda$  to obtain

$$\Lambda = P \underline{\Lambda} P^{-1}, \quad \underline{\Lambda}_{ij} = \lambda_i \delta_{ij}, \quad (18)$$

where  $P^{-1}$  is the inverse matrix of  $P$  and  $\lambda_i s$  are eigenvalues of  $\Lambda$ .

$k_i$  and  $\zeta_i$  can be  $s$  dependent but this work does not consider this case because it requires to replace Eq. (15) by a  $s$ -dependent expression so to perform exact calculations becomes difficult for a long filament [30]. Meanwhile,  $s$ -dependent  $k_i$  and  $\zeta_i$  may come from the sequence-disorder so at a first approximation they may be absorbed to the randomness average. Consequently,  $P$  and  $\lambda_i$  are also  $s$  independent

so we have  $\Lambda^n = (P \underline{\Lambda} P^{-1})^n = P \underline{\Lambda}^n P^{-1}$  and

$$\begin{aligned} e^{-\Lambda(s_1-s_2)} &= \sum_{n=0}^{\infty} \frac{1}{n!} (-\Lambda)^n (s_1 - s_2)^n \\ &= P \left[ \sum_{n=0}^{\infty} \frac{1}{n!} (-\underline{\Lambda})^n (s_1 - s_2)^n \right] P^{-1} = P \Gamma P^{-1}, \end{aligned} \quad (19)$$

with  $\Gamma_{ij} = e^{-\lambda_i(s_1-s_2)} \delta_{ij}$ . Equation (15) is then simplified to

$$\langle \mathbf{t}_i(s_1) \cdot \mathbf{t}_j(s_2) \rangle = [P \Gamma P^{-1}]_{ij}. \quad (20)$$

Consequently, when  $L \rightarrow \infty$ ,

$$\frac{1}{L} \int_0^L \int_0^{s_1} ds_2 ds_1 \Gamma_{ii} \rightarrow \frac{1}{\lambda_i}, \quad (21)$$

$$l_p = [P \underline{\Lambda}^{-1} P^{-1}]_{33}. \quad (22)$$

If  $\zeta_i s$  are random, then we only need to replace  $k_i$  by  $\kappa_i$  and  $\zeta_i$  by  $\omega_{i0}$  in Eqs. (15)–(22) [29]. It is easy to evaluate Eqs. (15)–(22) exactly and in some cases, we can even find a simple closed-form expression, such as Eq. (26).

For a 2D filament, we can find some short closed-form expressions, presented in Eqs. (32) and (33), for  $l_p$  [37].

#### F. Elasticity

Applying a uniaxial force (along the  $x$  axis)  $F$  to another end at  $s = L$ , we need to add another term,  $-F x_L \equiv -F x_L$ , to either  $E_{3D}$  or  $E_{2D}$  [19,25,26,37,53]. From Eqs. (2) and (5), it is straightforward to find

$$\langle x_L \rangle = \frac{\partial \ln \mathcal{Z}}{\partial f} \quad \text{and} \quad \frac{\partial \langle x_L \rangle}{\partial f} = \langle x_L^2 \rangle - \langle x_L \rangle^2, \quad (23)$$

where  $f \equiv \beta F$  and  $\langle x_L \rangle$  is the extension. Adopting hinged-hinged boundary conditions (BCs) so the orientations of the filament at both  $s = 0$  and  $s = L$  are free, we have  $\langle x_L \rangle = 0$  so  $\langle x_L^2 \rangle = \langle \mathbf{r}^2(L) \rangle / D$  when  $F = 0$ , where  $D = 2$  for a 2D system or  $D = 3$  for a 3D system. Consequently, at a low force, the relationship between applied force and extension can be obtained from

$$\langle x_L \rangle = f \langle x_L^2 \rangle_{f=0} = f \langle \mathbf{r}^2(L) \rangle_{f=0} / D. \quad (24)$$

The stretching strength ( $\mu$ ) is thus given by

$$\mu = \frac{1}{L} \frac{\partial \langle x_L \rangle}{\partial f} = \frac{\langle x_L^2 \rangle - \langle x_L \rangle^2}{L}. \quad (25)$$

Therefore, at a low force and for a long filament,  $\mu = 2l_p / D$  so that it has the same behavior as  $l_p$ .

Note that Eqs. (15)–(18) do not include applied force so that they only can be applied to evaluate the elasticity of a filament under a low force.

### III. EFFECTS OF IC AND ITR ON A 3D FILAMENT

#### A. Mean energy and specific heat

From equipartition theorem, it is straightforward to show that in both 2D and 3D systems, the mean energy is  $\langle E \rangle \propto k_B T L$ , independent of  $k_i$ ,  $k'_i$ , and  $\omega_{i0}$ . It follows that the specific heat ( $= \partial \langle E \rangle / \partial T / L$ ) is independent of  $k_i$ ,  $k'_i$ ,  $\omega_{i0}$ , and  $T$ .

In numerical calculations we also find that  $l_p$  is always a smooth function. Therefore, there is no phase transition for the system. However, we can find that magnitudes of IC and ITR have significant impact on the value of  $l_p$ .

### B. $l_p$ for the disorder-free system

When  $k_1 = k_2 = k$  or for an isotropic filament and  $\zeta_3 = 0$ , from Eqs. (15)–(22) we can obtain a short closed-form expression (see Appendix),

$$l_p = \frac{k(k + k_3)k_B T}{(k + k_3)k_B^2 T^2 + 2k^2 k_3 \zeta_0^2}. \quad (26)$$

When  $\zeta_0 = 0$ , Eq. (26) recovers  $l_p = l_p^0 = k/k_B T$ .

Equation (26) indicates that  $l_p$  is dependent on  $\zeta_0$  but independent of  $\zeta_1/\zeta_2$  or  $\alpha_0$ . Numerically the same conclusion applies to the case with  $k_1 = k_2$  but  $\zeta_3 \neq 0$  though it still lacks an exact proof. Moreover,  $l_p$  is a smooth function of all parameters so that no phase transition occurs for the system.

From Eq. (26) it is straightforward to show that  $l_p$  decreases monotonically with increasing  $\zeta_0$  or  $k_3$  so that the larger the  $k_3$  or  $\zeta_0$ , the smaller the  $l_p$ . Meanwhile,  $l_p$  is not a monotonic function of  $T$  but has a maximum at

$$T = T^{\max} = \frac{k\zeta_0\sqrt{2k_3}}{k_B\sqrt{k+k_3}}, \quad l_p = l_p^{\max} = \frac{\sqrt{k+k_3}}{2\sqrt{2k_3}\zeta_0}. \quad (27)$$

Equation (26) also reveals that  $l_p = 0$  at  $T = 0$  when  $\zeta_0 \neq 0$  so is different from that in the WLC model. This is due to the GSC of the filament is a circle of radius  $R_0$  so  $l_p \propto R_0/L \rightarrow 0$  when  $L \rightarrow \infty$ . The circle expands with increasing  $T$  so  $l_p$  also increases to  $T^{\max}$ . But at high  $T$ , the thermal fluctuation dominates the behavior so that the filament contracts and  $l_p$  decreases with increasing  $T$ , similarly to the behavior of  $l_p^0$ .  $T^{\max}$  gives a boundary between these two tendencies. These are also common characters even if  $k_1 \neq k_2$  and with randomness when  $\zeta_3 = 0$ , as we can see in Figs. 2–7.

Figure 2 displays some typical results obtained from Eq. (26) and it reveals that a moderate IC results in a quite different behavior between  $l_p$  and  $l_p^0$ . For instance, when  $k = k^0 \equiv 50 \text{ nm} \times k_B T_r = 50k_B T_r \text{ nm}$ ,  $k_3 = k_3^0 \equiv 75k_B T_r \text{ nm}$ , and  $R_0 = 50 \text{ nm}$ , we obtain  $l_p = 22.7 \text{ nm} \ll l_p^0$  at  $T_r$  and  $l_p^{\max} = 22.8 \text{ nm}$  at  $T = 326 \text{ K}$ , as we can see from the solid black line in Fig. 2. To recover  $l_p = l_p^0 = 50 \text{ nm}$  at  $T_r$ , it requires  $k = 180k_B T_r \text{ nm}$  and  $k_3 = 20k_B T_r \text{ nm}$ , as we can see from the purple dotted line in Fig. 2. Meanwhile, when  $k = k^0$ ,  $k_3 = k_3^0$ , and  $R_0 = 100 \text{ nm}$ , we obtain  $l_p = 38.5 \text{ nm} < l_p^0$  at  $T = T_r$ , as we can find from the blue dash-dotted line in Fig. 2, and  $l_p^{\max} = 45.6 \text{ nm}$  at  $T = 163 \text{ K}$ . In other words, to obtain a proper  $l_p$ , it requires a much larger  $k$  but a much smaller  $k_3$ . Note that  $R_0 = 50 \text{ nm}$  gives a circumference of 314 nm or 900 base pairs for a DNA chain, or the filament is only moderately curved in this case.

However, a finite  $\zeta_3$  leads to very different results, as shown in Fig. 3 for some typical results. In this case, there is not a short closed-form expression like to Eq. (26) but we still can calculate  $l_p$  exactly from Eqs. (15)–(22). When  $\zeta_0 = 0$ ,  $l_p = l_p^0$  is independent of  $\zeta_3$ , but a finite  $\zeta_0$  results in a couple between  $\zeta_0$  and  $\zeta_3$  so a much more complicate behavior. At first,  $l_p \propto 1/T$  at low  $T$  and  $l_p \rightarrow \infty$  when  $T \rightarrow 0$ , similarly to  $l_p^0$  since the shape of the centerline of the

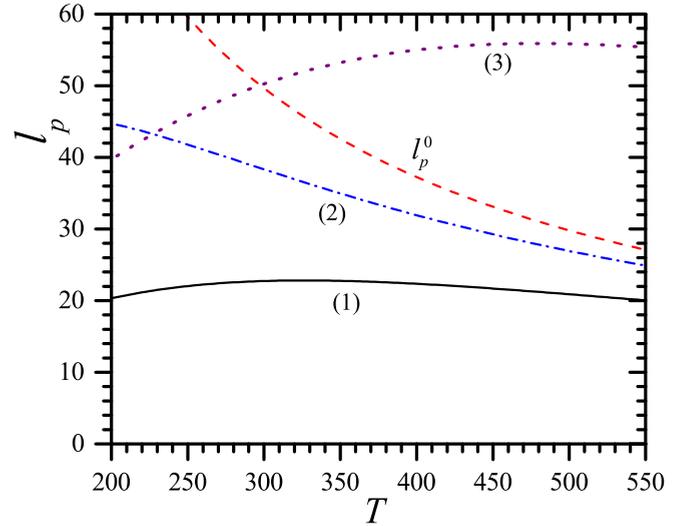


FIG. 2.  $l_p^0$  (red dashed and  $k = k^0 = 50$ ) and  $l_p$  vs.  $T$  for a 3D filament. For  $l_p$ ,  $k_1 = k_2 = k$  and  $\zeta_3 = 0$  and (1)  $R_0 = 1/\zeta_0 = 50$ ,  $k = k^0$ , and  $k_3 = k_3^0 = 75$  (solid black); (2)  $R_0 = 100$ ,  $k = k^0$ , and  $k_3 = k_3^0 = 75$  (blue dash-dotted); and (3)  $R_0 = 50$ ,  $k = 180$ , and  $k_3 = 20$  (purple dotted). Black and purple lines display maxima in  $l_p$ . The unit of length is nm, the unit of  $T$  is K, and the unit of  $k_i$  is  $k_B T_r \text{ nm}$  with  $T_r = 298 \text{ K}$ .

filament in its GSC is a helix of pitch  $= 2\pi/(\zeta_0^2 + \zeta_3^2)$  and height  $= h_{\text{helix}} = \zeta_3/\sqrt{\zeta_0^2 + \zeta_3^2}L$  [22,23], instead of a circle at  $\zeta_3 = 0$ . Furthermore, from Fig. 3, we can see that  $l_p$  has both a minimum and a maximum when  $\zeta_3$  is small, shown as the purple dotted line in the figure, so has quite different tendency and magnitude from  $l_p^0$ . The purple dotted line in Fig. 3 also

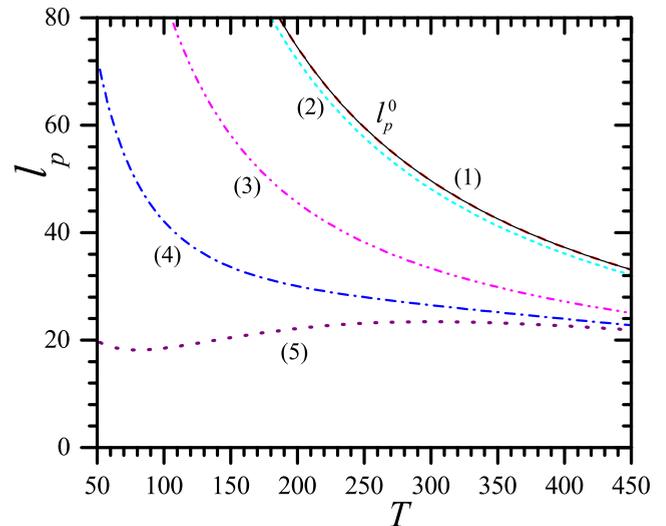


FIG. 3.  $l_p^0$  (red dashed and  $k = k^0 = 50$ ) and  $l_p$  vs.  $T$  for a 3D filament. For  $l_p$ ,  $k_1 = k_2 = k^0$ ,  $k_3 = k_3^0$ ,  $R_0 = 50$  and (1)  $\zeta_3 = 1.76$  (solid black and is overlapped with  $l_p^0$ ); (2)  $\zeta_3 = 0.1$  (cyan short-dashed); (3)  $\zeta_3 = 0.02$  (magenta dash-dot-dotted); (4)  $\zeta_3 = 0.01$  (blue dash-dotted); and (5)  $\zeta_3 = 0.004$  (purple dotted). The purple dotted line has both a minimum and a maximum. The unit of length is nm, the unit of  $T$  is K, and the unit of  $k_i$  is  $k_B T_r \text{ nm}$ .

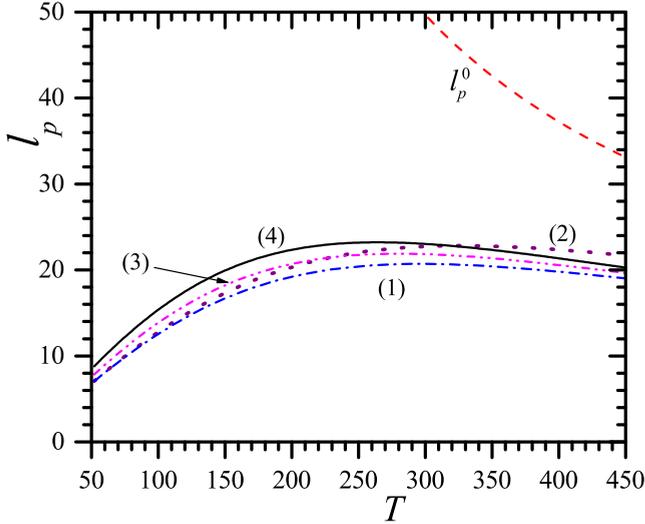


FIG. 4.  $l_p^0$  (red dashed and  $k = k^0 = 50$ ) and  $l_p$  vs.  $T$  for a 3D filament. For  $l_p$ ,  $k_1 = k^0$ ,  $k_3 = k_3^0$ ,  $R_0 = 50$ ,  $\zeta_3 = 0$  and (1)  $k_2 = 35$  and  $\alpha_0 = 0$  (blue dash-dotted); (2)  $k_2 = k_1$  (purple dotted); (3)  $k_2 = 35$  and  $\alpha_0 = 0.8$  (magenta dash-dot-dotted); and (4)  $k_2 = 35$  and  $\alpha_0 = \pi/2$  (solid black). The unit of length is nm, the unit of  $T$  is K, and the unit of  $k_i$  is  $k_B T_r$  nm.

indicates that  $l_p$  varies little in a rather large range of  $T$  so the filament should keep as a helix in this range, especially note that  $h_{\text{helix}}/L = 0.1961 \approx l_p$  in this case. It also suggests that increasing  $T$  enlarges the radius of the helix but changes little on its pitch, since for a long helix the pitch dominates the end-to-end distance. However, increasing  $\zeta_3$  makes  $l_p$  close to  $l_p^0$  but it has always  $l_p < l_p^0$ . When  $\zeta_3 = 0.1 \text{ nm}^{-1}$ ,  $l_p$  is already very close to  $l_p^0$  even at  $T = 5 \text{ K}$ , shown as the cyan short-dashed line in Fig. 3; when  $\zeta_3 = 1.76 \text{ nm}^{-1}$ ,  $l_p$  becomes indistinguishable from  $l_p^0$  in a rather large range of  $T$ , shown as the solid black line in Fig. 3. This is also not a surprise since the GSC of the filament is a twisted straight cylinder. In fact, when  $\zeta_3 = 0.2 \text{ nm}^{-1}$ ,  $l_p$  is already indistinguishable from  $l_p^0$ . We find further that the value of  $k_3$  has little effect at a large  $\zeta_3$ . In other words,  $\zeta_3$  favors a large  $l_p$  and plays an opposite role to  $\zeta_0$ , and a large  $\zeta_3$  can offset the effect of  $\zeta_0$  completely so makes  $l_p \approx l_p^0$  in a rather large range of  $T$ .

When  $k_1 \neq k_2$  but  $\zeta_3 = 0$ , we find that the effect of a finite  $c_0$  is still serious so  $l_p < l_p^0$  but the anisotropic effect is complicate and  $\alpha_0$  dependent. Figure 4 displays some typical results. When  $k_1 > k_2$ ,  $l_p$  reaches minimum at  $\alpha_0 = 0$  or  $\pi$  ( $\zeta_2 = 0$ ), but becomes maximum at  $\alpha_0 = \pi/2$  or  $3\pi/2$  ( $\zeta_1 = 0$ ). Clearly, when  $k_1 < k_2$ , the  $\alpha_0$ s for maximum and minimum  $l_p$  exchange the roles. In other words, a larger  $\zeta_2$  or nearly coaxial between IC and the hardest principal axis favors a larger  $l_p$  since the thermal fluctuation is more serious in the softer direction. At the minimum case,  $l_p$  is always smaller than that of the isotropic case, i.e., the case with  $k = k_m = \text{maximum}(k_1, k_2)$ , as we can see by a comparison between purple dotted ( $k = k_m$ ) and blue dash-dotted ( $\alpha_0 = 0$ ) lines in Fig. 4; but beginning from a moderate  $\alpha_0$ ,  $l_p$  can be larger than that of the isotropic case with  $k = k_m$  in a range of  $T$ , shown as the magenta dash-dot-dotted and solid black lines in Fig. 4.

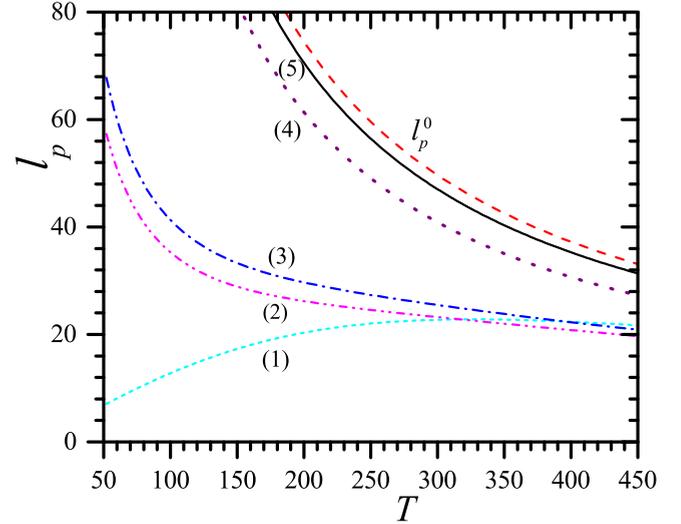


FIG. 5.  $l_p^0$  (red dashed and  $k = k^0 = 50$ ) and  $l_p$  vs.  $T$  for a 3D filament. For  $l_p$ ,  $k_1 = k^0$ ,  $k_3 = k_3^0$ ,  $R_0 = 50$  and (1)  $k_2 = k^0$  and  $\zeta_3 = 0$  (cyan short-dashed and the same as the solid black line in Fig. 2); (2)  $k_2 = 35$ ,  $\zeta_3 = 0.01$ , and  $\alpha = 0$  (magenta dash-dot-dotted); (3)  $k_2 = 35$ ,  $\zeta_3 = 0.01$ , and  $\alpha = \pi/2$  (blue dash-dotted); (4)  $k_2 = 35$ ,  $\zeta_3 = 1.76$ , and  $\alpha = \pi/2$  (purple dotted); and (5)  $k_2 = 45$ ,  $\zeta_3 = 1.76$ , and  $\alpha = \pi/2$  (solid black). The unit of length is nm, the unit of  $T$  is K, and the unit of  $k_i$  is  $k_B T_r$  nm.

Now we consider the case with  $k_1 \neq k_2$  and  $\zeta_3 \neq 0$  and Fig. 5 displays some typical results. Similarly to the isotropic case, a finite  $\zeta_3$  favors a larger  $l_p$  and a large  $\zeta_3$  can remove the effect of  $\zeta_0$  completely but  $k_3$  still plays little role. Five different lines in Fig. 5 demonstrate clearly that the smaller the  $\zeta_3$ , the smaller the  $l_p$ . Moreover, at a large  $\zeta_3$  the value of  $l_p$  is dominated by  $k_m$ . But the smaller the  $k'_m = \text{minimum}(k_1, k_2)$ , the smaller the  $l_p$ , as we can see from the solid black and purple dotted lines in Fig. 5. Meanwhile,  $l_p$  is dependent on  $\alpha_0$  in the same way as that in the isotropic case, as we can see from the blue dash-dotted and magenta dash-dot-dotted lines in Fig. 5. When  $k_2 = 35k_B T_r$  nm,  $k_3 = k_3^0$ ,  $\zeta_3 = 1.76 \text{ nm}^{-1}$  and  $R_0 = 50$  nm, to recover  $l_p = l_p^0$  requires  $k_1 = 85k_B T_r$  nm; when  $k_2 = 45k_B T_r$  nm,  $k_3 = k_3^0$ ,  $\zeta_3 = 1.76 \text{ nm}^{-1}$  and  $R_0 = 50$  nm, to recover  $l_p = l_p^0$  requires  $k_1 = 56k_B T_r$  nm.

In all cases, the sign of  $\zeta_3$  does not affect the result.

### C. $l_p$ for the disordered system

Equation (11) indicates that  $\kappa_i$  decreases monotonically with increasing  $T$ ,  $\kappa_i \leq \beta k_i$  and  $\kappa_i$  reaches maximum when  $k'_i = \infty$ . In other words, sequence disorder weakens rigidity. When  $\omega_{30} = 0$ ,  $k_1 = k_2$ , and  $k'_1 = k'_2$ , we can find exact express for  $l_p$ , i.e.,

$$l_p = k_1 k'_1 [k_1 k_3 (k'_1 + k'_3) + k'_1 (k_1 + k_3) k'_3 k_B T] / B, \quad (28)$$

$$B = k_1^2 k_3 (k'_1 + k'_3 + 2c_0^2 k_1^2 k'_3) + k_1^2 (k_1 + k_3) k'_3 k_B^2 T^2 + k_1 k'_1 [k'_1 k_3 + (k_1 + 2k_3) k'_3] k_B T. \quad (29)$$

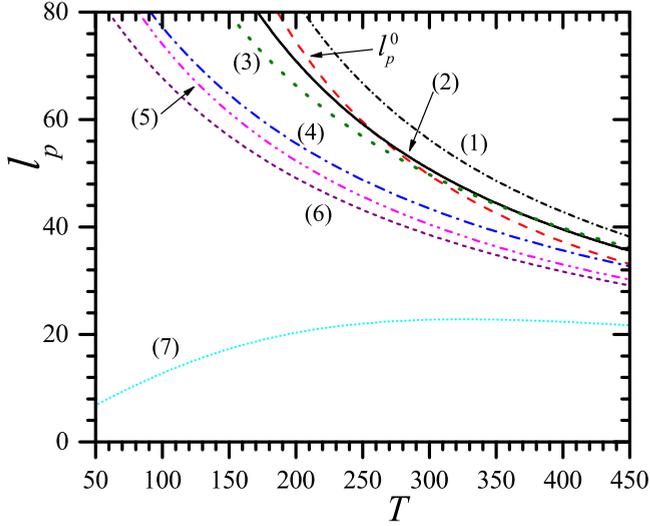


FIG. 6.  $l_p^0$  (red dashed and  $k = k^0 = 50$ ) and  $l_p$  vs.  $T$  for a 3D filament. For  $l_p$ ,  $k_3 = k_3^0$ ,  $R_0 = 50$ ,  $\omega_{30} = 1.76$  and (1)  $k_1 = k_2 = k^0$ ,  $k'_1 = k'_2 = 800$ , and  $k'_3 = 800$  (black short-dash-dotted); (2)  $k_1 = k_2 = k^0$ ,  $k'_1 = k'_2 = 400$ , and  $k'_3 = 300$  (solid black); (3)  $k_1 = k_2 = 60$ ,  $k'_1 = k'_2 = 200$ , and  $k'_3 = 200$  (olive dotted); (4)  $k_1 = k_2 = 60$ ,  $k'_1 = k'_2 = 400$ , and  $k'_3 = 75$  (blue dash-dotted); (5)  $k_1 = k_2 = k^0$ ,  $k'_1 = k'_2 = 400$ , and  $k'_3 = 75$  (magenta dash-dot-dotted); and (6)  $k_1 = k_2 = k^0$ ,  $k'_1 = k'_2 = 200$ , and  $k'_3 = 75$  (purple short-dashed). For comparison, we also show a sample (7) with  $k_1 = k_2 = k^0$ ,  $k_3 = k_3^0$ ,  $R_0 = 50$ ,  $k'_1 = k'_2 = k'_3 = \infty$ , and  $\omega_{30} = 0$  (cyan short-dotted and the same as the solid black line in Fig. 2). The unit of length and  $k'_i$  are nm, the unit of  $T$  is K, and the unit of  $k_i$  is  $k_B T_r$  nm.

$l_p$  has a maximum at

$$T^{\max} = \frac{k_1[\sqrt{2(k_1 + k_3)}k_3k'_1k'_3c_0 - k_3(k'_1 + k'_3)]}{(k_1 + k_3)k'_1k'_3k_B}, \quad (30)$$

$$l_p^{\max} = \frac{(k_1 + k_3)k'_1k'_3}{2\sqrt{2(k_1 + k_3)}k_3k'_1k'_3c_0 + k_1k'_3 - k'_1k_3}. \quad (31)$$

When  $k'_i = \infty$ , Eqs. (30) and (31) are the same as Eq. (27). Again, there is not a short closed-form expression for  $l_p$  when  $\omega_{30} \neq 0$ .

Figure 6 displays some typical results. Our results support the conclusion that a finite IC and anisotropic rigidities reduce  $l_p$ , but a finite ITR enhances  $l_p$ . Moreover, the effect of randomness is complicate. We find that a strong randomness or small  $k'_i$ 's reduce  $l_p$  even with a large  $\omega_{30}$ , shown as blue dash-dotted, magenta dash-dot-dotted, and purple short-dashed lines in Fig. 6. This is a natural result since the smaller the  $k'_i$ 's, the smaller the  $\kappa_i$ 's and the more flexible the filament. Meanwhile, though  $\kappa_i$  increases monotonically with increasing  $k'_i$ ,  $l_p$  does not follow such a behavior, as we can evaluate exactly from Eqs. (28) and (29) when  $\omega_{30} = 0$ ,  $k_1 = k_2$ , and  $k'_1 = k'_2$ .  $l_p$  also can become larger (smaller) than its randomness-free counterpart at low (high)  $T$ . 2D case has the similar results, displayed as cyan short-dashed, purple short-dash-dotted, and solid black lines in Fig. 7.

A special and maybe strange state is observed at a small randomness or at large  $k'_i$ 's. In this case, a large  $\omega_{30}$  can even make  $l_p > l_p^0$  at a rather large range of  $T$ , shown as black

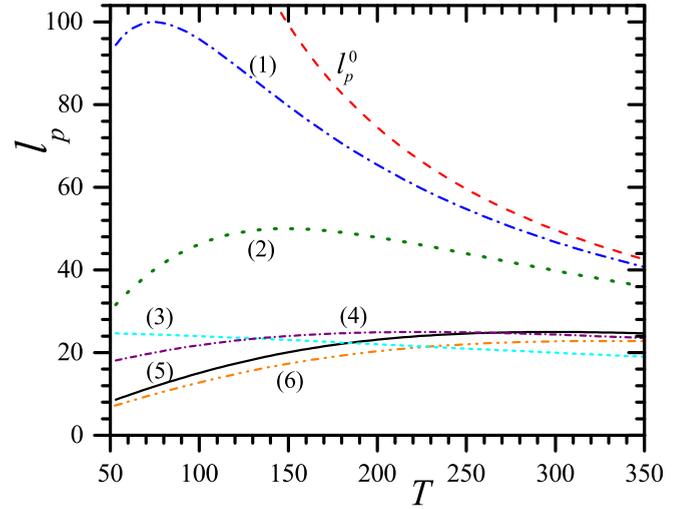


FIG. 7.  $l_p^0$  (red dashed and  $k = k^0 = 50$ ) and  $l_p$  vs.  $T$  for a 2D filament. For  $l_p$ , (1)  $k = k^0$  and  $R_0 = 200$  (blue dash-dotted); (2)  $k = k^0$  and  $R_0 = 100$  (olive dotted); (3)  $k = k^0$ ,  $R_0 = 50$ , and  $k'_1 = 25$  (cyan short-dashed); (4)  $k = k^0$ ,  $R_0 = 50$ , and  $k'_1 = 200$  (purple short-dash-dotted); (5)  $k = k^0$  and  $R_0 = 50$  (solid black). For comparison, we also show a 3D result (6) (orange dash-dot-dotted) with  $k_1 = k_2 = k^0$ ,  $k_3 = k_3^0$ , and  $R_0 = 50$ . The unit of length and  $k'_i$  are nm, the unit of  $T$  is K, and the unit of  $k_i$  is  $k_B T_r$  nm.

short-dash-dotted, solid black, and olive dotted lines in Fig. 6. In particular, the black short-dash-dotted line indicates that around  $T = T_r$ , a proper randomness can increase  $l_p$  considerably. We also find that  $l_p$  has a up bound at large  $k'_i$ . For instance, doubling  $k'_i$  in black short-dash-dotted line in Fig. 6 makes almost no difference in  $l_p$ . We refer to the state with  $l_p > l_p^0$  as an “overexpanded” state since it has always  $l_p < l_p^0$  in disorder-free system and  $k'_i \rightarrow \infty$  recovers a disorder-free system, i.e.,  $l_p \rightarrow l_p^0$  at very large  $k'_i$ 's and this is also conformed in our numerical calculations. The formation of such an overexpanded state should be due to a couple between IC or ITR and randomness, as we can see from Eqs. (28) and (29). Our calculations show further that  $k'_3$  is rather important for the “overexpanded” state though  $k_3$  still has less effect. For instance, in the magenta dash-dot-dotted line,  $k'_3$  is smaller than that in the solid black line, so  $l_p$  is also smaller. Finally,  $\alpha_0$  plays the same role as in the disorder-free system.

The  $l_p$  of several short (from 79 to 789 base pairs) and intrinsically straight DNA chains in different solutions are evaluated by transient electric birefringence measurements at temperatures between 4° and 43°C and follows a fitting to five different models [43]. The magnitude of  $l_p$  is solution and model dependent but the experimental data show an intrigue common behavior that  $l_p$  exhibits a shallow maximum at 20°C and decreases slowly on either increasing or decreasing  $T$  [43]. The existence of such a maximum and its  $T$ -dependent behavior are consistent with our results, as we can see from solid-black line in Fig. 2 and purple dotted lines in Figs. 3 and 4. Our results occur at a finite IC but a vanishing or small ITR so that such a consistent suggests that the solutions may induce a finite IC but weaken the effect or magnitude of ITR and it may deserve a further observation.

#### D. Elasticity

From the Sec. II F, we know that  $\mu \propto l_p$  at a low force. Meanwhile, it has been reported that extension of an intrinsically curved filament can subject to a discontinuous transition under a stretching force if it is very rigid or at low  $T$  [20–26]. Can such a transition occur at a finite  $T$  for a filament with a nonvanishing ITR? Our results suggest that at a finite  $T$  such a transition disappears for a filament with a large  $\omega_{30}$  since such a filament is similar to a WLC. The lack of such a transition in the WLRC model is a convincing evidence for the conclusion. It also reveals that the IC dominates the transition and agrees well with the results obtained from 2D systems [22,23,25–27].

### IV. IMPACT OF IC ON A 2D SYSTEM

#### A. Free of disorder

In the 2D case, from Eq. (46) in Ref. [37], we can find exactly

$$l_p = \frac{l_0}{1 + l_0^2 \zeta_0^2} = \frac{2kk_B T}{k_B^2 T^2 + 4k^2 \zeta_0^2}, \quad (32)$$

where  $l_0 \equiv 2k\beta$  is clearly the persistence length at  $\zeta_0 = 0$ . High  $T$  makes  $l_p \approx l_p^0$  or the effect of  $\zeta_0$  is negligible. At  $T = T_r$ ,  $l_p^0 = 50$  nm so  $k = k^0 \equiv 25k_B T_r$  nm.  $l_p \leq l_p^0$  and has a maximum at  $T = T^{\max} = 2\zeta_0 k/k_b$  with a  $k$ -independent  $l_p^{\max} = 1/2\zeta_0$ . Meanwhile,  $l_p$  has a maximum at  $k = k_B T/2\zeta_0$  with a  $T$ -independent  $l_p^{\max} = 1/5\zeta_0$ .

From Eq. (32), we obtain similar conclusion as that for a 3D filament, i.e., beginning from a moderate IC the behavior and magnitude of  $l_p$  is different from that of  $l_p^0$ , as shown in Fig. 7. Due to a smaller fluctuation,  $l_p$  is slightly larger than that in a 3D filament with the same parameters, as we can see from a comparison between solid black and orange dash-dot-dotted lines in Fig. 7. Meanwhile, around  $l_p^{\max}$ , the curve of  $l_p$  vs.  $T$  is sharper than that in 3D cases, as we can see from Figs. 2–7.

#### B. With disorder

When IC is sequence dependent, we need only to replace  $\beta k$  and  $\zeta_0$  in Eq. (32) by  $\kappa_1$  and  $c_0$  [28], so

$$l_p = \frac{2kk'_1(k + k'_1 k_B T)}{k^2(1 + 4k_1^2 c_0^2) + 2kk'_1 k_B T + k_1^2 k_B^2 T^2}. \quad (33)$$

$l_p$  has a maximum at  $T = T^{\max} = k_1(2c_0 k'_1 - 1)/k_B k'_1$  with  $l_p^{\max} = 1/2c_0$ .  $T^{\max}$  increases with increasing  $k_1$  or  $c_0$  or  $k'_1$ , but  $l_p^{\max}$  is independent of  $k$  and  $k'_1$ . Meanwhile,  $l_p$  has a maximum at  $k = k'_1 k_B T/(2c_0 k'_1 - 1)$  or  $k'_1 = k/(2c_0 k - k_B T)$  and  $l_p^{\max} = 1/2c_0$ . In other words, though  $\kappa_1 < k$  and  $\kappa_1$  increases monotonically with increasing  $k'_1$ ,  $l_p$  does not follow such a behavior, and can become larger or smaller than its randomness-free counterpart at low or high  $T$ , shown as cyan short-dashed, purple short-dash-dotted, and solid black lines in Fig. 7. From Eq. (33) we know that this is because  $c_0$  is coupled with randomness so that the phenomenon does not occur when  $c_0 = 0$ , similar to that we have reported in Sec. III C for a 3D filament. Moreover, at a small  $c_0$  and at some special  $T$ , randomness can result in  $l_p^{\max} > l_p^0$  or give an overexpanded state, similar to that occurs in the 3D system.

### V. CONCLUSION AND DISCUSSION

In summary, we investigate the effects of IC, ITR, anisotropic  $k_i$ , randomness in sequences and  $T$  on the  $l_p$  of a semiflexible biopolymer. We develop some general expressions to evaluate these effects exactly and easily. We show rigorously that the specific heat of the semiflexible biopolymer is independent of all of these parameters and our exact numerical calculations indicate that  $l_p$  is always a smooth function of them, so that there is no phase transition. We reveal that a moderate IC solely can reduce  $l_p$  remarkably and result in a very different relationship between  $l_p$  and  $T$ , in comparison with the WLC model. It follows that for a filament with a moderate or large IC but a small ITR, to obtain a large  $l_p$  it requires a rather large bending rigidity but a relative small twist rigidity. When ITR is small, our results suggest that the centerline of the filament keeps a helical shape in a rather large range of  $T$  so it may play the role of a Hookean spring. However, our calculations suggest that a large ITR can cancel the effect of IC completely. Another intriguing fact is that  $k_3$  has little role when ITR is sufficient large. This is because the main effect of a large  $k_3$  is to depress the fluctuation of the cross-section distortion so is decoupled from the variation of the end-to-end distance, similar to that occurs in the WLRC model. Meanwhile, we find that  $l_p$  is not a monotonic function of  $T$ , can have either minimum and maximum at some special  $T$ , and in the 2D case the maximum is more clear than that in the 3D case. This is also significant since the environment of a semiflexible biopolymer is usually crowded so may subject to some geometrical constraints. Our result also suggest that with a large ITR, an intrinsically curved filament should have the same mechanical property as that of a WLC.

Besides preserving all above results, anisotropy in bending rigidities and randomness in sequences add something new and intrigue. At a moderate  $T$  the anisotropy can result in a larger  $l_p$  than its isotropic counterpart even if one of bending rigidities is smaller than its isotropic counterpart. On the other hand, a strong randomness shrinks  $l_p$  remarkably since it reduces the effective bending rigidity considerably. However, a weak randomness can enlarge  $l_p$  considerably and lead to an overexpanded state. The overexpanded state is unexpected or “abnormal,” and reveals that there exists a complicate cooperation between IC or ITR and randomness, and a large ITR can stress the cooperation. The properties and applications, such as its influences to elasticity, electric, and magnetic responses, of such an “abnormal” state should be an intriguing topic and deserve a further study.

We should point out that our conclusions are valid in the temperature range  $340 \text{ K} > T > 240 \text{ K}$  except for the minima in  $l_p$  usually occur at a lower  $T$  ( $\sim 100 \text{ K}$ ). This range of  $T$  is of special meaningful since the dsDNA solution freezes below  $240 \text{ K}$ , while its basepairs internally melt beyond  $340 \text{ K}$ .

Maybe the most significant finding in this work is that to have a large  $l_p$  at a finite  $T$  for an intrinsically curved semiflexible biopolymer, a large ITR and a weak randomness in IC and ITR are as important as a large bending rigidity, but the value of the twist rigidity is unimportant.

In this work we ignore the excluded volume effect. The excluded volume tends to increase the distant between

monomers so plays a crucial role in thermal and mechanical property of a long filament. Especially, the excluded volume effect should be serious if the bending rigidity is small or the temperature is high. Therefore, it deserves a further study.

Our numerical calculations focus on dsDNA since it is a typical natural twisted semiflexible biopolymer so that the main conclusions are also instructive to other semiflexible biopolymers. We should also stress that it is rather easy to perform calculations for other semiflexible biopolymers by using our basic expressions. Since the mean end-to-end distance and radius of gyration of a long semiflexible biopolymer are proportional to  $l_p$ , our findings provide a full picture to the influences of different parameters on the size of a semiflexible biopolymers and are instructive in controlling the size of some semiflexible biopolymers in organic synthesis. Experimentally, required bending and twist rigidities can be realized by controlling the strength of covalent bonds or the numbers of hydrogen bonds or the energy barrier between gauche and trans conformations so it is feasible.

#### ACKNOWLEDGMENT

This work has been supported by the MOST of the Republic of China.

#### APPENDIX: DERIVATION OF EQ. (26)

When  $k_1 = k_2 = k$  and  $\zeta_3 = 0$ , Eqs. (16) and (17) become

$$\gamma_1 = \gamma_2 = k_B T (1/k + 1/k_3)/2, \quad \gamma_3 = k_B T/k, \quad (\text{A1})$$

$$\Lambda = \begin{pmatrix} \gamma_1 & 0 & -\zeta_0 \sin \alpha_0 \\ 0 & \gamma_1 & \zeta_0 \cos \alpha_0 \\ \zeta_0 \sin \alpha_0 & -\zeta_0 \cos \alpha_0 & \gamma_3 \end{pmatrix}. \quad (\text{A2})$$

The eigenvalues of  $\Lambda$  are  $\lambda_1 = \gamma_1$  and  $\lambda_{2,3} = (\gamma_1 + \gamma_3 \mp \gamma_0)/2$  with  $\gamma_0 \equiv \sqrt{(\gamma_1 - \gamma_3)^2 - 4\zeta_0^2}$ . Applying a similarity transformations  $\Lambda = P \underline{\Lambda} P^{-1}$  to diagonalize  $\Lambda$ , we have  $\underline{\Lambda}_{ij} = \lambda_i \delta_{ij}$  and so  $\underline{\Lambda}_{ij}^{-1} = \lambda_i^{-1} \delta_{ij}$ , as well as

$$P = \begin{pmatrix} \cot \alpha_0 & \frac{2\zeta_0 \sin \alpha_0}{\gamma_1 - \gamma_3 + \gamma_0} & \frac{2\zeta_0 \sin \alpha_0}{\gamma_1 - \gamma_3 - \gamma_0} \\ 1 & \frac{-2\zeta_0 \cos \alpha_0}{\gamma_1 - \gamma_3 + \gamma_0} & \frac{-2\zeta_0 \cos \alpha_0}{\gamma_1 - \gamma_3 - \gamma_0} \\ 0 & 1 & 1 \end{pmatrix}, \quad (\text{A3})$$

$$P^{-1} = \begin{bmatrix} \sin(2\alpha_0)/2 & \sin^2 \alpha_0 & 0 \\ -\zeta_0 \sin \alpha_0 / \gamma_0 & \zeta_0 \cos \alpha_0 / \gamma_0 & (\gamma_1 - \gamma_3 + \gamma_0)/(2\gamma_0) \\ \zeta_0 \sin \alpha_0 / \gamma_0 & -\zeta_0 \cos \alpha_0 / \gamma_0 & (\gamma_3 - \gamma_1 + \gamma_0)/(2\gamma_0) \end{bmatrix}. \quad (\text{A4})$$

Finally, combining Eqs. (22) and (A1)–(A4), we obtain Eq. (26)

$$l_p = [P \underline{\Lambda}^{-1} P^{-1}]_{33} = \frac{k(k + k_3)k_B T}{(k + k_3)k_B^2 T^2 + 2k^2 k_3 \zeta_0^2}. \quad (\text{A5})$$

- 
- [1] J. F. Marko and E. D. Siggia, *Science* **265**, 506 (1994).  
[2] C. Bustamante, J. F. Marko, E. D. Siggia, and S. Smith, *Science* **265**, 1599 (1994).  
[3] J. F. Marko and E. D. Siggia, *Macromolecules* **28**, 8759 (1995).  
[4] C. Bouchiat, M. D. Wang, J. F. Allemand, T. Strick, S. M. Block, and V. Croquette, *Biophys. J.* **76**, 409 (1999).  
[5] M. Doi and S. F. Edwards, *The Theory of Polymer Dynamics* (Clarendon Press, Oxford, 1986).  
[6] C. Domb, *J. Chem. Phys.* **38**, 2957 (1963).  
[7] D. S. McKenzie, *Phys. Rep.* **27**, 35 (1976).  
[8] P. G. de Gennes, *Scaling Concept in Polymer Physics* (Cornell University Press, Ithaca, NY, 1979).  
[9] J. W. Essam, *Rep. Prog. Phys.* **43**, 833 (1980).  
[10] S. M. Bhattacharjee, A. Giacometti, and A. Maritan, *J. Phys.: Condens. Matter* **25**, 503101 (2013).  
[11] P. Pincus, *Macromolecules* **9**, 386 (1976).  
[12] M. Daoud and P. G. de Gennes, *J. Phys. (Paris)* **38**, 85 (1976).  
[13] F. Brochard-Wyart, *Europhys. Lett.* **23**, 105 (1993).  
[14] F. Brochard-Wyart, H. Hervet, and P. Pincus, *Europhys. Lett.* **26**, 511 (1994).  
[15] P.-Y. Lai, *Phys. Rev. E* **58**, 6222 (1998).  
[16] B. Fain, J. Rudnick, and S. Östlund, *Phys. Rev. E* **55**, 7364 (1997).  
[17] J. F. Marko, *Phys. Rev. E* **55**, 1758 (1997).  
[18] C. Bouchiat and M. Mézard, *Phys. Rev. Lett.* **80**, 1556 (1998).  
[19] J. Moukhtar, E. Fontaine, C. Faivre-Moskalenko, and A. Arnéodo, *Phys. Rev. Lett.* **98**, 178101 (2007).  
[20] B. Smith, Y. V. Zastavker, and G. B. Benedek, *Phys. Rev. Lett.* **87**, 278101 (2001).  
[21] D. A. Kessler and Y. Rabin, *Phys. Rev. Lett.* **90**, 024301 (2003).  
[22] Z. Zhou, P.-Y. Lai, and B. Joós, *Phys. Rev. E* **71**, 052801 (2005).  
[23] Z. Zhou, B. Joós, P.-Y. Lai, Y.-S. Young, and J.-H. Jan, *Mod. Phys. Lett. B* **21**, 1895 (2007).  
[24] E. L. Starostin and G. H. M. van der Heijden, *Phys. Rev. Lett.* **101**, 084301 (2008).  
[25] Z. Zhou, *J. Phys. Commun.* **2**, 035008 (2018).  
[26] Z. Zhou, *Eur. Biophys. J.* **48**, 329 (2019).  
[27] Z. Zhou, F.-T. Lin, C.-Y. Hung, H.-Y. Wu, and B.-H. Chen, *J. Phys. Soc. Jpn.* **83**, 044802 (2014).  
[28] Z. Zhou and B. Joós, *Phys. Rev. E* **80**, 061911 (2009).  
[29] Z. Zhou and B. Joós, *Phys. Rev. E* **77**, 061906 (2008).  
[30] S. V. Panyukov and Y. Rabin, *Phys. Rev. E* **62**, 7135 (2000).  
[31] P. C. Nelson, *Phys. Rev. Lett.* **80**, 5810 (1998).  
[32] D. Bensimon, D. Dohmi, and M. Mezard, *Europhys. Lett.* **42**, 97 (1998).  
[33] Y. O. Popov and A. V. Tkachenko, *Phys. Rev. E* **76**, 021901 (2007).  
[34] G. Zuccheri, A. Scipioni, V. Cavaliere, G. Gargiulo, P. De Santis, and B. Samor, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 3074 (2001).  
[35] J. Van Noort, T. van der Heijden, M. de Jager, C. Wyman, R. Kanaar, and C. Dekker, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 7581 (2003).  
[36] S. Rappaport and Y. Rabin, *Macromolecules* **37**, 7847 (2004).  
[37] Z. Zhou, *Phys. Rev. E* **76**, 061913 (2007).

- [38] S. V. Panyukov and Y. Rabin, *Phys. Rev. E* **64**, 011909 (2001).
- [39] A. P. Fields, E. A. Meyer, and A. E. Cohen, *Nucl. Acids Res.* **41**, 9881 (2013).
- [40] P. Cong, L. Dai, H. Chen, J. R. C. van der Maarel, P. S. Doyle, and J. Yan, *Biophys J.* **109**, 2338 (2015).
- [41] S. K. Nomidis, F. Kriegel, W. Vanderlinden, J. Lipfert, and E. Carlon, *Phys. Rev. Lett.* **118**, 217801 (2017).
- [42] S. K. Nomidis, E. Skoruppa, E. Carlon, and John F. Marko, *Phys. Rev. E* **99**, 032414 (2019).
- [43] Y. Lu, B. Weers, and N. C. Stellwagen, *Biopolymers* **61**, 261 (2002).
- [44] H. Yamakawa and M. Fujii, *J. Chem. Phys.* **64**, 5222 (1976).
- [45] H. Yamakawa and T. Yoshizaki, *Helical Wormlike Chains in Polymer Solutions* (Springer-Verlag, Berlin, 1996).
- [46] C.-Y. Hung, Z. Zhou, F.-T. Lin, and Y.-S. Young, *Int. J. Mod. Phys. B* **26**, 1250131 (2012).
- [47] M. Dlakic, K. Park, J. D. Griffith, S. C. Harvey, and R. E. Harrington, *J. Biol. Chem.* **271**, 17911 (1996).
- [48] W. Han, S. M. Lindsay, M. Dlakic, and R. E. Harrington, *Nature* **386**, 563 (1997).
- [49] E. Segal, Y. Fondufe-Mittendorf, L.-Y. Chen, A. Thåström, Y. Field, I. K. Moore, J.-P. Z. Wang, and J. Widom, *Nature* **442**, 772 (2006).
- [50] W. Han, M. Dlakic, Y. J. Zhu, S. M. Lindsay, and R. E. Harrington, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 10565 (1997).
- [51] H. R. Drew and A. A. Travers, *J. Mol. Biol.* **186**, 773 (1985).
- [52] Z. Zhou and B. Joós, *Chin. Phys. B* **25**, 088701 (2016).
- [53] C. Vaillant, B. Audit, and A. Arnéodo, *Phys. Rev. Lett.* **95**, 068101 (2005).
- [54] H. Kleinert, *Path Integrals in Quantum Mechanics, Statistics, and Polymer Physics* (World Scientific, Singapore, 1990).
- [55] D. C. Khandekar, S. V. Lawande, and K. V. Bhagwat, *Path Integrals Methods and Their Applications* (World Scientific, Singapore, 1993).