



HAL
open science

Free energy measurements by the generalized fluctuation theorems: Theory and numerical study of a model filament

Min-Kyung Chae, Yunha Kim, Nam-Kyung Lee, A. Johner, Jeong-Man Park

► **To cite this version:**

Min-Kyung Chae, Yunha Kim, Nam-Kyung Lee, A. Johner, Jeong-Man Park. Free energy measurements by the generalized fluctuation theorems: Theory and numerical study of a model filament. European Physical Journal E: Soft matter and biological physics, 2020, 43 (9), 10.1140/epje/i2020-11984-5 . hal-02992581

HAL Id: hal-02992581

<https://cnrs.hal.science/hal-02992581v1>

Submitted on 10 Nov 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Free energy measurements by the generalized fluctuation theorems: theory and numerical study of a model filament

Min-Kyung Chae¹, Yunha Kim¹, Nam-Kyung Lee^{1,2a}, A. Johner² and Jeong-Man Park^{3b}

¹ Department of Physics and Astronomy, Sejong University, Seoul 05006, South Korea,

² Institute Charles Sadron, CNRS 23 Rue du Loess, 67034 Strasbourg cedex 2, France,

³ Department of Physics, The Catholic University of Korea, Bucheon 14662, South Korea

June 30, 2020

Abstract. We measure the free energy of a model filament, which undergoes structural transitions, as a function of its extension, *in silico*. We perform Brownian Dynamics (BD) simulations of pulling experiments at various speeds, following a protocol close to experimental ones. The results from the fluctuation theorems are compared with the estimates from Monte Carlo (MC) simulation, where the rugged free energy landscape is produced by the density of states method. The fluctuation theorems (FT) give accurate estimates of the free energy up to moderate pulling speeds. At higher pulling speeds, the work distributions do not efficiently sample the domain of small work and FT slightly overestimates free energy. In order to comprehend the differences, we analyze the work distributions from the BD simulations in the framework of trajectory thermodynamics. The measured work - free energy relation is consistent with the results obtained from the generalized fluctuation theorems that take into account informations (relative entropies). We discuss operational methods to improve the estimates at high pulling speed.

PACS. 0 5.40.-a Fluctuation phenomena, random processes, noise, and Brownian motion, 05.07.-a Thermodynamics 82.37.-j Single molecule kinetics

1 Introduction

For last decades, the notion of the second law of thermodynamics has been extended to small scale systems where thermal fluctuations are relevant for work, heat exchange, and internal energy. For meso-scopic systems embedded in a heat bath with well-defined temperature, stochastic energetics has been developed and entropy can be defined along the fluctuating trajectory [1, 2]. Theoretical relations like the fluctuation theorems (FT) such as the Jarzynski equality (JE) [3–6] and the Crooks fluctuation theorem (CFT) [7, 8] involve the probability distribution of external work and allow us to obtain the free energy from non-equilibrium measurements. These theorems have been put into experimental test, using biomolecules under mechanical control [9, 10], as well as colloidal particles in time-dependent optical traps [11]. The case of an underdamped particle was examined in Ref. [12].

In RNA folding/unfolding experiments, the free energy is successfully obtained from the JE and the CFT [9, 10]. In these experiments, the domain undergoing folding/unfolding transition can be treated essentially as a microscopic two-level system. The work distributions are

basically (single) Gaussian and the standard deviations remain comparable to the thermal energy $\sim k_B T$.

Despite of the remarkable success of the JE, it is yet a difficult task to measure thermodynamic functions for many processes that involve structural transformations, especially when there are energy barriers associated. One controls a few macroscopic parameters such as extension or total twist angle. By changing the macroscopic parameters, conformational transitions recognized by macroscopic observables are followed. For example, under a tension of $\sim 80 - 100$ pN, B-DNA undergoes a transition to overstretched S-form, so-called B-S transition [13]. For this pulling experiment, the control parameter is the end-to-end distance. The sequence specific B-Z transition [14, 15], where right handed B-DNA is converted to left handed Z-DNA under torsional stress, is another example. For this transition, the control parameter would be the torsion angle [16–18]. These transitions often involve large energy barriers corresponding to the domain wall energy of $\sim 8k_B T$ [19]. The JE can be applied to systems of any size driven arbitrarily away from equilibrium provided that the initial state is in equilibrium. However, a difficulty on applying the JE to processes involving energy barriers lies in that only a limited number of experimental measurements are available in practice, which may not allow to sample the most relevant paths with small work efficiently [20].

^a E-mail: namklee@sejong.edu

^b E-mail:jmanpark@catholic.ac.kr

Some efforts to handle barrier crossing problem in the fluctuation relations have been set forward by several theoretical studies [21–23] and tested experimentally [22, 24, 23, 25] and numerically [26] for biomolecular systems. In these studies, initial and final states are well-separated by energy barriers and the modified fluctuation relations allow to estimate the free energy between these well-separated states. Here, we consider the system with multiple macroscopic variables with the expanded phase space, over which initial and final states cannot be disjointly separated. This situation is more common in the polymeric systems. In this contribution, we propose the generalized expression suitable for estimating the free energy between the states not well-separated with multiple degrees of freedom.

For the purpose of applying the FTs, we consider a stiff helical filament, nearly circular when confined into a plane without other constraints (Fig. 1). Much is known on this model filament, which has been studied previously [27–29]. **The system retains twist and bending degrees of freedoms.** When elongated, it undergoes structural changes accompanied by curvature flips, so called twist-kinks. Among many degrees of freedom, we retain two macroscopic observables: the end-to-end distance (d_s) and the number of curvature flips (n_k). Following the experimental protocols, d_s is controlled during the computer experiments by changing the distance between the two handle beads. We perform Brownian Dynamics (BD) simulations to obtain forward/backward trajectories of stretching/releasing cycles and measure the work done on the system for each trajectory. We analyze the work distributions from the BD computer experiments to find the work-free energy relation by using the FTs. Estimates from the FTs are accurate to $k_B T$ or off by several $k_B T$ s depending on the pulling speeds. In order to comprehend this speed dependence, we introduce the generalized Jarzynski equality (GJE) and the generalized Crooks fluctuation theorem (GCFT) by taking into account informations (relative entropies) in the scheme of trajectory thermodynamics [30]. By quantitatively evaluating the information **in microscopic variables**, we put the GJE and the GCFT to a test *in silico*.

2 Theory : Generalization of the FTs

We consider a classical Hamiltonian system with a heat bath at temperature T . Below, we will set $k_B = 1$ and the temperature $T = 1/\beta$ is assumed to be a constant. The JE allows us to compute the free energy difference between two equilibrium states, ΔF . The exponential average of the work done is related to the equilibrium free energy difference by

$$\langle e^{-\beta W} \rangle_{\mathcal{N}} = e^{-\beta \Delta F}. \quad (1)$$

Specializing to the pulling experiments, $\langle \dots \rangle_{\mathcal{N}}$ denotes averaging over \mathcal{N} trajectories. For each i -th trajectory, work W_i ($i = 1, 2, \dots, \mathcal{N}$) to bring the system from X_0 to X_f is the integral of the time-dependent force f_i , $W_i = \int_{X_0}^{X_f} f_i(X') dX'$.

The Crooks fluctuation theorem (CFT) also allows us to estimate the equilibrium free energy from the hysteresis

obtained from stretching and releasing cycle. The CFT relates the work distributions, $\rho_F(W)$ and $\rho_R(-W)$, of the work applied to the system during stretching (forward, F) and releasing (backward reverse, R) to the free energy difference between the initial and final states, ΔF , by

$$\frac{\rho_F(W)}{\rho_R(-W)} = \exp(\beta(W - \Delta F)). \quad (2)$$

The original JE and CFT can be derived from the Crooks' general form of the FT (cite Crooks PRE 2000)

$$\langle \mathcal{F} e^{-\beta(W - \Delta F)} \rangle_{F=} \langle \hat{\mathcal{F}} \rangle_{R}, \quad (3)$$

where \mathcal{F} stands for an arbitrary functional of the forward trajectories and $\hat{\mathcal{F}}$ is the time-reversal of \mathcal{F} . The averages $\langle \dots \rangle_{F(R)}$ are taken over the ensemble of all possible forward (backward reverse) trajectories. The FTs require that the initial states in both the forward and the backward reverse processes must be sampled from global equilibrium distributions [31].

This is a serious limitation when it is difficult to observe states in full equilibrium conditions. There have been many efforts to modify the FTs in order to overcome this limitation. Most of these modifications start with the Crooks' general form of the FT (Eq.(.)).

Among many, one of the prominent modifications is the differential fluctuation theorem (DFT) derived by Maragakis et al. (cite J. Phys. Chem. B 2008). Maragakis et al. used the Crooks' general form of the FT to derive the DFT identity

$$P_F(W, a \rightarrow b) e^{-\beta W} = P_R(-W, b \rightarrow a) e^{-\beta \Delta F}, \quad (4)$$

where $P_F(W, a \rightarrow b)$ is the joint probability of performing work W during the forward process starting from value a of the function $a(x)$ of microstate $x(0)$ at the start and taking value b of the function $b(x)$ of microstate $x(\tau)$ at the end of the forward process, and $P_R(-W, b \rightarrow a)$ is the corresponding reverse trajectory quantity.

By expanding the joint probabilities in terms of the conditional probabilities, the above DFT can be written as

$$P_F(W|a \rightarrow b) p_F(b|a) p_A(a) e^{-\beta W} = P_R(-W|b \rightarrow a) p_R(a|b) p_B(b) e^{-\beta \Delta F}, \quad (5)$$

where $P_F(W|a \rightarrow b)$ is the probability to perform work W given that the forward trajectory starts at value a of $a(x)$ and ends in value b of $b(x)$, $p_F(b|a)$ is the probability to end in b given that the forward trajectory started from a , and $p_A(a)$ is the probability to measure $a = a(x)$ in the equilibrium distribution in state A . The right-hand side contains the corresponding reverse trajectory quantities that start from equilibrium distribution in state B and end in state A .

The DFT allows us to estimate the free energy difference ($\Delta F = F_B - F_A$) between states A and B by using the work data restricted to measurements between two subsets a and b of the full states A and B . All the conditional

probabilities in the DFT equation are trajectory-averaged quantities that depend on the processes between a small domain a of A and b of B .

However, the probabilities $p_A(a)$ and $p_B(b)$ are trajectory-independent equilibrium quantities. When there are high barrier in between states A and B , the sampling of the entire equilibrium ensembles of A and B through molecular dynamics would be very time-consuming. Then, the equilibrium probabilities $p_A(a)$ and $p_B(b)$ should be obtained by separated different experiments or simulations.

Another significant modification is the extended fluctuation theorem (EFT) derived by Junier et al. (cite PRL 2009). Both the EFT and the DFT have considered the subsets of the full phase space. Whereas Maragakis et al. still used global equilibrium probabilities $p_A(a)$ and $p_B(b)$ for the initial states, which should be obtained by some other experiments than the forward and the backward reverse processes, Junier et al. used partial equilibrium for the sampling of the initial states, which means that the initial states are Boltzmann-Gibbsian but restricted over a subset of the full phase space.

Junier et al. defined a 'kinetic state' as a partially equilibrated region of the phase space and states inside each region are sampled according to the Boltzmann-Gibbs equilibrium distribution restricted to such region. The EFT can be written as

$$P^{S_0 \rightarrow S_1}(W)p^{S_0 \rightarrow S_1}e^{-\beta W} = P^{S_1 \rightarrow S_0}p^{S_1 \rightarrow S_0}e^{-\beta \Delta G_{S_0, \lambda_0}^{S_1, \lambda_1}},$$

where $P^{S_0 \rightarrow S_1}$ is the work distribution for the trajectory starting in partially equilibrated state S_0 and ending in S_1 , $p^{S_0 \rightarrow S_1}$ stands for the probability to be in S_1 at the end of the forward process started in S_0 , and $\Delta G_{S_0, \lambda_0}^{S_1, \lambda_1} = G_{S_1}(\lambda_1) - G_{S_0}(\lambda_0)$ is the free energy difference between partially equilibrated states S_0 and S_1 . The right-hand side contains the corresponding reverse trajectory quantities that start in S_1 and end in S_0 .

The EFT equation has the same form as that of the DFT. The only difference is that the DFT is derived for the case of global equilibrium initial conditions, but the EFT is derived under the partial equilibrium over a subset of the phase space.

In this paper, we generalize the FTs to non-equilibrium initial states, that is, to the case where the initial states are not Boltzmann-Gibbsian over the full phase space. First we generalize the JE to non-equilibrium initial states [30, 32–34]. Having prepared the initial state of equilibrium or non-equilibrium, we allow the system to evolve over time from the state A at $t = 0$ to the state B at $t = K\Delta t$, according to the predetermined protocol. The microscopic history of a system during this process is described by a trajectory $\Pi = \{x_0, x_1, \dots, x_{K-1}, x_K\}$ evolving under Hamilton's equation in the phase space. x_k denotes a point in the phase space of the system at time $t = k\Delta t$. Since the initial condition x_0 uniquely determines a trajectory, the exponential average of the work done during this process

can be evaluated as

$$\begin{aligned} \langle e^{-\beta W} \rangle &= \int d\Pi \mathcal{P}(\Pi) e^{-\beta W[\Pi]} = \int dx_0 p(x_0) e^{-\beta W} \\ &= \int dx_0 p^{eq}(x_0) e^{-\beta W + D(p_i | p_0^{eq})} = \langle e^{-\beta(W - TD_i)} \rangle_{eq}, \end{aligned} \quad (6)$$

where $\mathcal{P}(\Pi)$ is the probability for a trajectory Π and $D_i = D(p_i | p_0^{eq}) = \ln[p(x_0)/p^{eq}(x_0)]$ is the relative entropy (or Kullback-Leibler divergence) of an arbitrary initial probability distribution $p_i = p(x_0)$ to the equilibrium probability distribution $p_0^{eq} = p^{eq}(x_0)$ [35, 36]. The relative entropy can be identified with the information as a measure of the distance between the actual distribution and the equilibrium distribution in the information-theoretic interpretation [37–39]. Then the average we evaluate becomes

$$\begin{aligned} \langle e^{-\beta(W + TD_i)} \rangle &= \int dx_0 p(x_0) e^{-\beta W - \ln[p(x_0)/p^{eq}(x_0)]} \\ &= \int dx_0 p^{eq}(x_0) e^{-\beta W} = \langle e^{-\beta W} \rangle_{eq}, \end{aligned} \quad (7)$$

and by combining Eqs. (1) and (6), we obtain

$$\langle e^{-\beta(W + TD_i)} \rangle = \langle e^{-\beta W} \rangle_{eq} = e^{-\beta \Delta \tilde{F}}, \quad (8)$$

where we introduced $\Delta \tilde{F}$ to indicate the free energy difference evaluated from the modified work distributions, $W' = W + TD_i$. This equation is referred as the generalized Jarzynski equality (GJE) [30, 34]. The equality Eq. (8) suggests that if the system under consideration is initially not at equilibrium, the work - free energy relation should be extended to include initial state information.

Similarly, we can generalize the CFT for non-equilibrium initial states (of forward/ backward processes). The forward process is the one described above, in which the work parameter is varied from A at $t = 0$ to B at $t = K\delta t$ using a protocol $\mathcal{D}_F(t)$. During the backward reverse process, \mathcal{D}_R is varied from B at $t = 0$ to A at $t = K\delta t$ by using the time-reversed protocol. At the start of each process, the system is prepared in an arbitrary state, corresponding to $\mathcal{D}_{F(R)} = A(B)$, at temperature T . The notation $\Pi_F = \{x_0, x_1, \dots, x_{K-1}, x_K\}$ denotes a trajectory that might be observed during a realization of the forward process and $\Pi_R = \{x_K, x_{K-1}, \dots, x_1, x_0\}$ is its conjugate reverse process. The ratio of probabilities to observe the Hamiltonian trajectories Π_F during the forward process and Π_R during the backward reverse process can be written as

$$\begin{aligned} \frac{\mathcal{P}_F[\Pi_F]}{\mathcal{P}_R[\Pi_R]} &= \frac{p_F(x_0)P_F(x_0 \rightarrow x_1) \cdots P_F(x_{K-1} \rightarrow x_K)}{p_R(x_K)P_R(x_K \rightarrow x_{K-1}) \cdots P_R(x_1 \rightarrow x_0)} \\ &= \frac{e^{\ln[p_F(x_0)/p^{eq}(x_0)]} p^{eq}(x_0) P_F(x_0 \rightarrow x_1) \cdots P_F(x_{K-1} \rightarrow x_K)}{e^{\ln[p_R(x_K)/p^{eq}(x_K)]} p^{eq}(x_K) P_R(x_K \rightarrow x_{K-1}) \cdots P_R(x_1 \rightarrow x_0)} \\ &= e^{D(p_i^F | p_0^{eq}) - D(p_i^R | p_K^{eq})} \cdot e^{\beta(W_F - \Delta F)} \\ &= e^{\beta(W_F + TD_i^F - TD_i^R) - \beta \Delta F}, \end{aligned}$$

where $D_i^F = D(p_i^F | p_0^{eq}) = \ln[p_F(x_0)/p^{eq}(x_0)]$, $D_i^R = D(p_i^R | p_K^{eq}) = \ln[p_R(x_K)/p^{eq}(x_K)]$ and we used the CFT. Then the distribution $\rho_F(W')$ for the modified work $W' = W_F +$

$TD_i^F - TD_i^R$ can be evaluated as

$$\begin{aligned} \rho_F(W') &= \int d\Pi_F \mathcal{P}_F[\Pi_F] \delta(W' - (W_F + TD_i^F - TD_i^R)) \\ &= \int d\Pi_F e^{\beta(W_F + TD_i^F - TD_i^R) - \beta\Delta F} \mathcal{P}_R[\Pi_R] \delta(W' - (W_F \\ &= e^{\beta(W' - \Delta F)} \int d\Pi_F \mathcal{P}_R[\Pi_R] \delta(W' - (W_F + TD_i^F - T. \\ &= e^{\beta(W' - \Delta F)} \int d\Pi_R \mathcal{P}_R[\Pi_R] \delta(W' + W_R - TD_i^F + TD_i^R)) \\ &= e^{\beta(W' - \Delta F)} \rho_R(-W'), \end{aligned}$$

where we used $d\Pi_F = d\Pi_R$ because the Jacobian is 1 and $W_F[\Pi_F] = -W_R[\Pi_R]$ because the work done to the system during the forward process and the work done by the system during the conjugate reverse process are the same. Thus we obtain the generalized Crooks fluctuation theorem (GCFT) for the work $W' = W + TD(p_F(x_0)|p^{eq}(x_0)) - TD(p_R(x_K)|p^{eq}(x_K))$ modified by the information of non-equilibrium initial states

$$\frac{\rho_F(W')}{\rho_R(-W')} = e^{\beta(W' - \Delta F)}. \quad (11)$$

We can derive our generalized FTs, the GJE and the GCFT, from the Crooks' general form of the FT (cite Crooks PRE 2000) We use the forward path functional

$$\mathcal{F}_F = \delta(W' - (W[x(0 \rightarrow \tau)] + TD_i^F - TD_i^R)) e^{-D_i^F}, \quad (12)$$

where $x(0 \rightarrow \tau)$ denotes a forward trajectory, and its time-reversed path function

$$\mathcal{F}_R = \delta(W' - (W[x(\tau \rightarrow 0)] + TD_i^R - TD_i^F)) e^{-D_i^R}, \quad (13)$$

where $x(\tau \rightarrow 0)$ denotes a time-reversed trajectory. Inserting these path functions into Eq. (3) gives

$$\begin{aligned} & \langle \delta(W' - (W[x(0 \rightarrow \tau)] + TD_i^F - TD_i^R)) e^{-D_i^F} e^{-\beta(W - \Delta F)} \rangle \\ &= \langle \delta(W' - (W[x(\tau \rightarrow 0)] + TD_i^R - TD_i^F)) e^{-D_i^R} \rangle_R, \end{aligned} \quad (14)$$

and we obtain the generalization of the CFT

$$\rho_F(W') e^{-\beta(W' - \Delta F)} = \rho_R(-W'). \quad (15)$$

Finally, by integrating the above equation over W' , we obtain the generalization of the JE,

$$\langle e^{-\beta W'} \rangle = e^{-\beta \Delta \bar{F}}, \quad (16)$$

where $W' = W + TD_i$.

3 Pulling a Filament *in Silico*

3.1 Model

We consider numerical stretching experiments of a filament of length $S = Nb$, consisting of N monomers of size

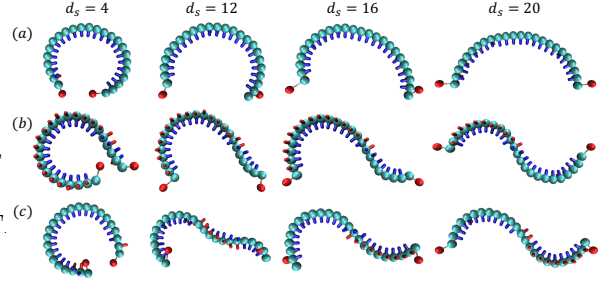


Fig. 1. Conformations of confined model filaments consisting of 24 monomers of size b sampled from equilibrium states (MC). Beads are attached to both ends via harmonic springs and the distance between the two beads (d_s) is controlled during pulling/releasing cycles. The three representative paths from $d_s = 4b$ to $d_s = 20b$ are shown: (a) continuous deformations ($n_k = 0$ state) and (b-c) structural transitions to the final state with a curvature flip ($n_k = 1$ state). The transitions are via (b) a single twist kink insertion at one of the filament ends or (c) a pairwise kink/anti-kink insertion ($n_k = 2$) in the middle of the filament. Normal vectors are shown as dark blue lines along the filament and their directions flip where twist kinks are located.

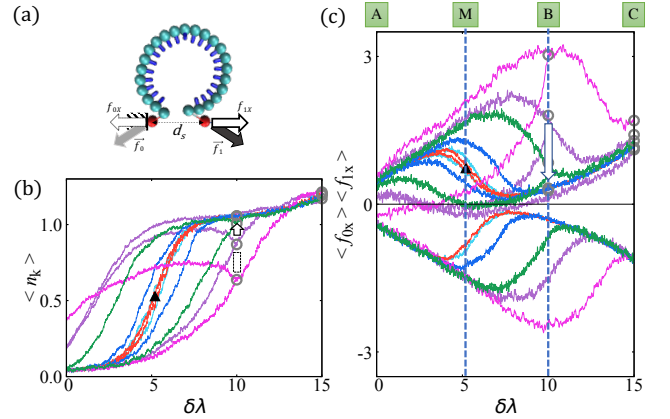


Fig. 2. (a) Schematics of pulling experiment. (b) n_k - extension ($\delta\lambda$) relation and (c) force (f_x)-extension ($\delta\lambda$) relation. The projected forces in both moving bead ($f_{1X} > 0$) and fixed bead ($f_{0X} < 0$) are averages of 1000 trajectories for different pulling speeds $v = 0.0002$ (red), 0.0004 (cyan), 0.002 (blue), 0.01 (green), 0.02 (purple) and 0.04 (magenta). The initial state ($\lambda_A = 4.0$), two final states ($\lambda_B = 14$ or $\lambda_C = 19$), and the midpoint ($\lambda_M = 9.2$) are designated as A, B, C and M, respectively. The symbol \blacktriangle represents the midpoint values of $\langle n_k \rangle$ and $\langle f_{1X} \rangle$. Relaxation process (for $v = 0.02 b/t_0$) after arrival at the final position is indicated as thick arrows.

b , that has both bending and twist degrees of freedoms [27, 28]. The elastic energy is the sum of its bending and twist energy contributions,

$$\frac{\mathcal{H}_{el}}{k_B T} = \frac{1}{2} \int^S ds [l_p \{(\Omega_1(s) - \omega)^2 + \Omega_2(s)^2\} + l_t (\Omega_3(s) - \tau)^2] \quad (17)$$

where ω and τ are the intrinsic curvature and torsion. Along the filament, the strain vector $\boldsymbol{\Omega}(s) = \{\Omega_1(s), \Omega_2(s), \Omega_3(s)\}$ is related to the local curvature $\kappa(s)$ and torsion $\tau(s)$ by $\kappa(s) = \sqrt{\Omega_1^2(s) + \Omega_2^2(s)}$ and $\tau(s) = \Omega_3(s)$. The components of $\boldsymbol{\Omega}(s)$ are optimized to meet the preferred values for the 3D ground state which is helical. The bending and twist fluctuations around the preferred values are governed by bending and twist moduli, $l_p k_B T$ and $l_t k_B T$ and the fluctuations in bending angles and twist angles are expected to be of $\sim l_p^{-1}$ and $\sim l_t^{-1}$, respectively.

Details of the model can be found in previous publications [27–29]. In order to apply constraints to the filament, we introduce two beads attached to the filament by springs of strength $k_s = 5k_B T/b^2$. These springs mimic molecular handles whose separation will be imposed/monitored and obey Gaussian fluctuations ($V = k_s(r - r_0)^2$, $r_0 = 1.5b$). Unlike molecular handles, which are usually much longer than the size of the target molecule, the handles of the current simulation has a monomeric size (no additional entropy correction is needed) and it is introduced to allow monomeric length fluctuation for the end positions of the filament.

We choose parameters $l_p = 32b$, $l_t = 4b$ so that the large bending energy cost penalizes deformation involving curvature change from optimal value within its contour length $S = 24b$. The 3D ground state of the filament is a helix. We reduce degrees of freedom by confining the filament in effective 2D space using a harmonic potential $V(z) = \frac{1}{2}kz^2$ with $k = 50k_B T/b^2$, where the potential has minimum at the surface $z = 0$ [28]. When molecules are squeezed, the ground state conformation of the filament (up to boundary effects) is the circular shape of a constant curvature, all twist being expelled. The twist angle may change by $\pm\pi$ if curvature flips ($\Delta n_k = \pm 1$). The corresponding change of twist number is $\Delta T w = \pm 0.5$. The conformation of 2D filament is then characterized by its extension d_s and the number of curvature flips (twist kink) n_k . The energy for inserting a single twist kink in a *continuous* infinite unconstrained filament can be estimated as $E_k = \pi(\sqrt{\gamma} - 1)l_t k_B T \tau$, where $\gamma = \frac{4}{\pi^2} \frac{l_p w^2}{l_t \tau^2}$ measures the ratio between the bending energy cost and twist energy cost [27]. With prescribed values of $\omega = 0.195b^{-1}$ and $\tau = 0.174b^{-1}$, twist kinks of energy $E_k = 2.22k_B T$ can be added by thermal activation. If a small end-to-end distance is imposed, the energy for inserting a twist-kink is larger. In contrast a twist-kink may form spontaneously at larger extensions. As the localized curvature flip is usually induced by the extension of the filament, increasing end-to-end distance (or applying tension) triggers the transition between conformations of different number of twist kinks.

In Brownian Dynamics simulations, the pulling speeds can be controlled by changing the distance between the two handle beads d_s after a given duration time t_D . Whilst it is possible to control the macroscopic variable d_s (and possible to measure n_k), other internal degrees of freedom related to the elastic energy of the filament are difficult to control in a simple way.

3.2 Simulation Scheme

We carry out the non-equilibrium pulling experiments on a model filament, *in silico*. The filament undergoes structural transitions between two states separated by the free energy barrier. By utilizing the BD simulations, we first establish both force-extension and n_k -extension relations during the stretching-releasing cycles. The projected forces in the moving bead ($f_{1X} > 0$) and fixed bead ($f_{0X} < 0$) are recorded at given extension parameters $\lambda = d_s/b$ (Fig. 2(a)). The work distributions are evaluated by integrating $f_{1X}(\lambda)$ for each of the 1000 trajectories. Fig. 2 shows force(f_X)-extension($\delta\lambda$) and n_k - extension($\delta\lambda$) relations averaged for all trajectories, where the final displacements (marked as open circles) are chosen to be $\delta\lambda = 10$ (B) or 15 (C). As expected the force-extension relation shows marked hysteresis. For slow pulling speeds, the mean number of twist-kink $\langle n_k \rangle$ is ~ 0.5 at extension $\delta\lambda_M = 5.2$. For fast pulling, most of filaments are found at $n_k = 0$ state, when the displacement reaches at position M (λ_M) and follow the path illustrated in Fig. 1(a). Some transition paths nonetheless follow the two-step process (Fig. 1(c)) which involves pair-wise twist-kink insertion (i.e. via $n_k = 2$ state) rather than a single kink-insertion process via the filament ends (Fig. 1(b)). The former process involves usually larger energy barrier. The mid-points, where two types of conformations are equally populated, are obviously shifted to larger displacements in dynamical processes. The average projected force f_{1X} goes through a maximum at the point M[†], which we refer to as the transition point. For slow pulling speeds $v < 0.002b/t_0$, the projected pulling force f_{1X} reaches to 0 close to B (at the local free energy minimum) but remains at non-vanishing value for fast pulling speeds. Force f_{1X} is slightly larger than force f_{0X} when the system is not mechanically equilibrated. For each trajectory, the work done by the moving bead to the system is evaluated as $W = \int \mathbf{f}_1 \cdot d\mathbf{r}$ and the work distribution at the given pulling speed is obtained. From this work distribution, we evaluate $\langle e^{-\beta W} \rangle$ and $\langle e^{-\beta(W+TD_i)} \rangle$ to apply the JE and GJE where D_i stands for the information to be discussed next.

In parallel, we measure the free energy variation upon elongation in a Monte Carlo simulation by using the Density of States (DoS) method [40,41] suited for systems involving energy barriers. For this purpose, the phase space covered by the DoS needs to be enlarged compared with the previous work [27]. The detailed simulation methods are described in [Appendix](#).

Evaluation of information - We present how we access to the informations $D_i^{F/R}$ in the generalized FT. The phase space of the filament (with controlled extension) is its configuration-space of high dimension. In Fig. 3, we show distributions of some specific internal variables together with relaxed (equilibrium) distributions. The distributions of middle monomer ($n = 12$) positions and the bond angle of the corresponding link (Fig. 3(b)) show the reduction in the volume of their phase spaces when the filaments are brought at M ($\delta\lambda = 4.2$) by finite speed pulling procedures. The distributions of total twist, ac-

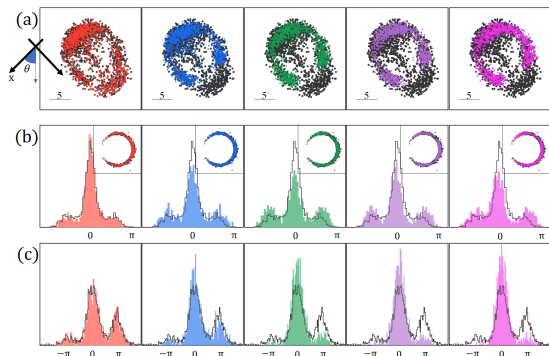


Fig. 3. Distributions obtained from 1000 trajectories at the midpoint extension M . (a) Distributions of the middle monomer ($n = 12$) positions. The positions of each monomer are represented in x - y plane. The pulling direction is indicated as $+\hat{x}$. (b) Distributions of orientational link angles $\theta_{n=12}$ and (c) accumulated twist angles. Colors indicate different pulling speeds: red, blue, green, purple and magenta represent $v = 0.0002, 0.002, 0.01, 0.02$ and $0.04 b/t_0$, respectively (from left to right) and the relaxed distributions at the given extension are shown in black.

cumulated twist along the chain, manifest similar effects (Fig. 3(c)). The equilibrium distribution has several peaks, one around 0 corresponding to $\langle Tw \rangle = 0$ - state (no twist-kink or kink/anti-kink pair) and the others around $\pm\pi$ corresponding to $\langle Tw \rangle = \pm 0.5$ - state (single twist-kink). At high pulling speeds, one of the twist-kink states is mainly populated due to the energy barrier associated with the transition. Because fast relaxing variables, at the speeds considered, do not contribute significantly to the information, we use an operational definition based on a set of independent slow variables.

We evaluate information in the middle monomer positions ($n = 12$), by counting the number of trajectories where the position of the middle monomer $\mathbf{r}_{n=12}$ falls in a cell $\mathbf{R}(l, m)$ of area of $(dx \times dy)$ in 2-dimensional space (one could also find probability distributions of x - and y -coordinates), and by computing the logarithms of relative probabilities $i_{p,n=12} = \sum_{l,m} \ln(p(\mathbf{r}_n \in \mathbf{R}(l, m))/p_{eq}(\mathbf{r}_n \in \mathbf{R}(l, m)))$ with respect to the equilibrium distributions. We also evaluate informations in the cumulated twist by $i_{Acc.tw} = \ln(p(\sum_n \psi_n)/p_{eq}(\sum_n \psi_n))$. Below we compute information at each trajectory at given time by the sum¹ of $i_{p,n=12}$ and $i_{Acc.tw}$,

$$D_i = i_{p,n=12} + i_{Acc.tw}. \quad (18)$$

¹ Alternatively, either the central monomer orientation or the center of mass position can be considered together with the accumulated twist for the evaluation of the information. The former choice slightly underestimates the information correction while the later choice slightly overestimates it.

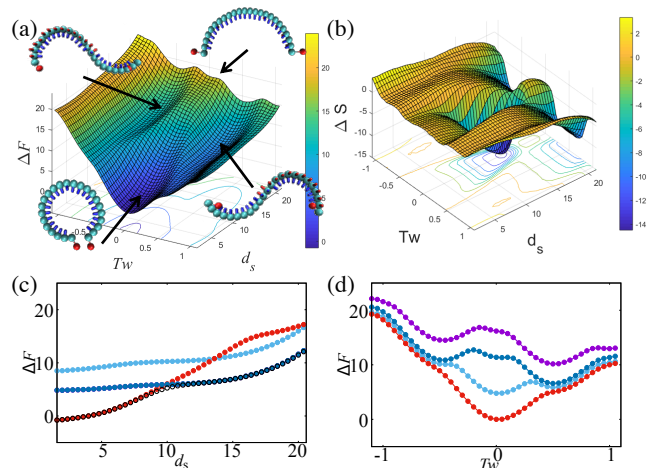


Fig. 4. (a) Free energy landscape $F(d_s, Tw)$ obtained by the DoS method of MC simulation. (Representative conformations are shown.) (b) Mean entropy $\Delta S(d_s, Tw)$. (c) Free energy difference $\Delta F(d_s)$ as a function of d_s (averaged over Tw) (represented by black \circ) and $\Delta F_{Tw}(d_s)$ for given values of $Tw = 0, 0.5$ and -0.5 (red, blue, right blue). The free energies of $Tw = 0$ and $Tw = 0.5$ cross at $d_s \approx 10b$. (d) $\Delta F(Tw)$ measured at given d_s as a function of Tw . Different colors represent different values of $d_s = 4b, 9b, 14b$ and $19b$ from bottom to top.

4 Results

4.1 Equilibrium free energy landscape, MC simulation

In Fig. 4(a), we show equilibrium free energy landscape obtained by the DoS method of MC simulations for the model filament. Several local free energy minima are found at Tw values that are multiples of 0.5, the twist increment per curvature flip. The lowest free energy state is found at $n_k = 0$ around the extension of $d_s = 4b$ and a region of local minimum is found at $n_k = 1$ ($Tw = 0.5$) along the extension of 0 - 20b. The corresponding entropy map is shown in Fig. 4(b). In Fig. 4(c) the free energy difference $\Delta F(d_s)$ as a function of d_s (reference to $d_s = 4b$) and $\Delta F_{Tw}(d_s)$ for fixed values of $Tw = 0$ and ± 0.5 are plotted. Fig. 4(d) shows the free energy $\Delta F(Tw)$ for fixed values of d_s . At small extension $d_s \sim 4b$, the free energy has a well defined minimum at $Tw = 0$ ($n_k = 0$). At larger extensions, the minimum shifts to larger value of n_k . The $Tw = 0.5$ ($n_k = 1$) conformation is the absolute minimum for extension $d_s > 10b$. Typically, these local minima are separated by energy barriers $\sim 5 - 10 k_B T$.

4.2 BD simulation of non-equilibrium measurements

In Fig. 5(a), we show the free energy differences $\Delta F^{(1)}(\delta\lambda)$ obtained using the JE (Eq. 8). The results from three slow pulling speeds are in agreement with each other, and collapse into a single line, referred as MD line below. It turns out that the MD line is slightly below the MC estimates (see SI) but remains within $\sim k_B T$ over the whole range of

tested displacement. The estimation of ΔF for all pulling speeds collapse into a single line for $\lambda < \sim \lambda_M$ but split at larger displacements. For pulling speeds (0.0002 - 0.002 b/t_0), we obtain $\Delta F_{AB}^{(1)} \sim 5.8\text{-}5.9 k_B T$, but for the fastest pulling cases (0.02 - 0.04 b/t_0) the $\Delta F_{AB}^{(1)}$ deviates by up to $\sim 3k_B T$ at larger displacement $\delta\lambda \sim 15$.

To get the thermodynamic free energy from the fast processes, we apply the JE in two steps. In step (1), we bring the molecule to M starting from equilibrated state A. And in step (2) we bring the molecule to B starting from M after equilibration. We measure $\Delta F_1 = \Delta F_{AM}$ and $\Delta F_2 = \Delta F_{MB}$ using Eq. (8). The obtained total free energies $\Delta F_{AB}^{(2)} = \Delta F_1 + \Delta F_2$ are shown in Fig. 5(b), where data from all pulling speeds collapse to a single line.

To obtain ΔF_{AB} from the CFT (Eq. (2)), we evaluate work distributions from forward/backward work distributions. The reverse processes begin after some relaxation time $\sim 20,000 t_0$ at B following the same protocol as the corresponding forward processes. The overlapping forward (right) and backward (left) work distributions are shown in Fig. 5(c). Applying the CFT, the cross points of two distributions (indicated as dashed lines) fall between 5.6 - 6.0 $k_B T$, in fair agreement with the JE estimate ΔF_{AB} of slow pullings (Fig. 5(a)) or two step JE (represented as MD line, Fig. 5(b)). For comparison, we show reverse work distributions $\rho_R(-W)$ which takes the final state of the forward process as the initial state of the backward process without equilibration (shown in grey). Naively looking for the cross point with $\rho_F(W)$ produces large overestimate of ΔF_{AB} for fast pulling/releasing processes.

Single step JE estimates of fast pulling processes also deviate from the JE of slow measurements for $\delta\lambda > \delta\lambda_M$ (MD line). We attribute the deviation of the JE and CFT estimate to the extra work needed in practice for crossing large energy barriers that can not be easily overcome by thermal fluctuations. The corresponding information is accumulated and is to be dissipated after crossing the barrier. We did also test CFT for ΔF_{AC} and reached good agreement without any correction. This is because the internal degrees of freedom can be easily relaxed over shallow barrier along $d_s = 19$ line (see Fig. 4(d)).

In Fig. 6(a), we show the distributions of corresponding informations (Eq. 18) at the midpoint M and the final points B and C, respectively. At small pulling rates, the informations have narrow distributions with a vanishing mean. At larger pulling speeds, informations accumulated at B (peak at +4) disappear at C (Fig. 5(a)).

We check the influence of information taking advantage of the generalized form of the CFT. We obtain modified work distributions $W'_{F/R} = W_{F/R} - TD_{i,B}$ for fast pulling cases taking into account informations at the final state (B) while $D_{i,A} = 0$ (Fig. 6(c)). The most evident shift is found in the case of $v = 0.04b/t_0$. The large overestimated free energy ($\sim 9.6 k_B T$) found from the cross point of the non-equilibrium work distribution is shifted to smaller value $6.1 \pm 0.5k_B T$, when the final point information is considered. The estimate of free energy by the GCFT at smaller pulling speeds also converges to 5.5 \pm

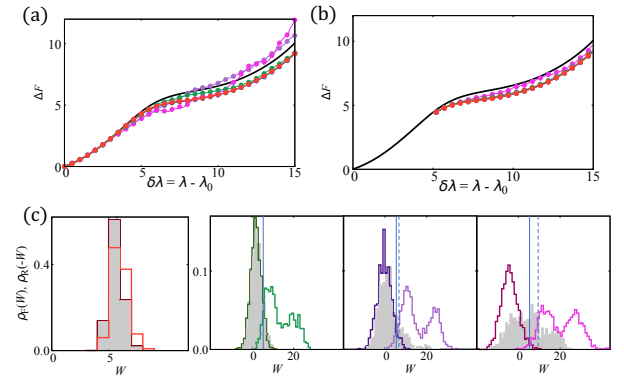


Fig. 5. (a) The estimated free energy difference $\Delta F^{(1)}(\delta\lambda)$ as a function of $\delta\lambda$ evaluated from the JE and (b) $\Delta F^{(2)}(\delta\lambda)$ by using the JE in two steps. Different pulling speeds $v = 0.0002, 0.0004, 0.002, 0.01, 0.02$ and $0.04 b/t_0$ are represented by red, cyan, blue, green, purple and magenta, respectively. The data from three slow pullings are on top of each other. The estimates from MC simulations are shown as a black solid line for comparison. (c) The forward/backward work distributions from pulling/releasing experiments with various pulling speeds. The final state is B with extension $\delta\lambda = 10$. The slowest case $v = 0.0002 b/t_0$ and three faster cases $v = 0.01, 0.02$ and $0.04 b/t_0$ are shown in increasing order from left to right. The lines are work distributions with equilibrated initial states, A (forward) and B (backward). The cross points ($\rho_F(W) = \rho_B(-W)$) providing ΔF_{AB} via CFT are represented as solid lines. The distributions shown in grey are work distributions $-W$ of the reverse processes which take the final state of the forward process as the initial state of the backward process without an equilibration. The corresponding cross points $\rho_F(W) = \rho_B(-W)$ are indicated as dashed lines.

$0.5k_B T$. For slow pulling speeds $v < 0.01b/t_0$, as expected, the influence of information is almost negligible.

To evaluate the free energy for the barrier crossing problem using the GJE, we split the process into two steps, and apply the GJE in the 2nd step. We measure $\Delta F_1 = \Delta F_{AM^\dagger}$ with respect to A. In the second step, we measure $\Delta F_2 = \Delta \tilde{F}_{M^\dagger B}$ at B with respect to M^\dagger using Eq. (8). We include the information of M^\dagger as the initial state information in the work distributions $W'_{M^\dagger B} = W_{M^\dagger B} + TD_{i,M^\dagger}$ and evaluate the free energy $\Delta \tilde{F}_{M^\dagger B}$ from $W'_{M^\dagger B}$. If we choose the splitting point as the thermodynamic midpoint M ($\delta\lambda = 5.2$), the work distribution of the second step W_2 should be modified to $W'_2 = W_{MB} + TD_{i,M}$. By taking into account the information, we obtain consistent values of the free energy $\Delta \tilde{F}_{AB}$ for all pulling speeds $v \leq 0.02b/t_0$ but not for $v = 0.04b/t_0$ (Fig. 6(b)). The exchange of work and information occurs until the system reaches to the transition state M^\dagger where the average force $\langle f_{1X} \rangle$ is largest. We may evaluate the free energy of the second step ΔF_2 starting from M^\dagger . In the second panel of Fig. 6(b), we plot the total free energy $\Delta \tilde{F}_{AB}$ with information at the transition point for each pulling speed. (For $v = 0.01, 0.02$ and $0.04 b/t_0$, we locate the transition points as $\delta\lambda_{M^\dagger} = 7, 9, 10$, respectively.) With these informations, we find good

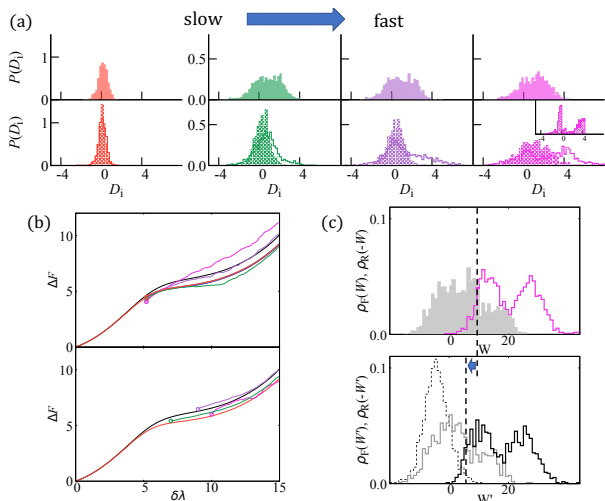


Fig. 6. (a) Distributions of informations $P(D_i)$ (Eq. (18)). The relevant positions are midpoint M (top), B(bottom,empty) C (bottom, filled). The pulling rates are $v = 0.0002, 0.01, 0.02$ and $0.04 b/t_0$ left to right. For $v = 0.04 b/t_0$, at M, the information distribution of the accumulated twist alone is shown in the inset. (b) The estimates of free energy difference by using the GJE: $\Delta \tilde{F}(\delta\lambda_M)$ (top), and $\Delta \tilde{F}(\delta\lambda_{M^+})$ (bottom). Different pulling speeds $v = 0.0002, 0.0004, 0.002, 0.01, 0.02$ and $0.04 b/t_0$ are represented by red, cyan, blue, green, purple and magenta, respectively. (c) The estimates of free energy difference by using the GCFT. The work distributions for the pulling rates: $v = 0.04 b/t_0$. The distributions involving the non-equilibrium state B (top) and the work distributions ($\rho_F(W'), \rho_R(-W')$) corrected by the contribution of information by $-TD_{i,B}$ (bottom). The cross point of two distributions are shifted to smaller value when the information is taken into account.

agreement between the evaluated free energy difference from the fast pullings and the quasi-static MD line.

5 Conclusions

The formal beauty of the fluctuation theorems like the JE and CFT resides in their universality. It is enough to measure work distributions and no a priori knowledge of the system is required, in order to estimate equilibrium quantities such as the free energy. In practice it may be difficult to measure those distributions with the required precision for some systems [20]. To remedy this difficulty, extrapolation schemes for the low work part of the distribution have been devised [42]. We evaluate the free energy differences of a model filament possessing many degrees of freedom from non-equilibrium *in silico* experiments. As a reference we produce the rugged free energy landscape represented in the parameter space reduced to two macroscopic quantities (elongation and number of kinks) by MC simulations. Following a constant speed protocol, we perform numerical extension experiments with BD. A set of thousand experiments is done for different speeds and work distributions are evaluated. Upon applying the JE to the obtained

work distributions, we found that there are distinctive two regimes of pulling speeds. For slow pulling processes, the JE provides a good estimate and recovers the thermodynamic free energy. However, for fast pulling processes the JE results deviate from the estimates expected in the thermodynamic limit. (According to the Kramers's rate equation, the time required for the thermal equilibration increases exponentially and the required pulling speed for efficient sampling for the JE is expected to decrease exponentially with the height of the barrier.)

To resolve this discrepancy induced by fast pulling speeds upon crossing energy barriers, we extend the JE to the GJE where the condition of the initial states being in equilibrium is no longer required. We extend the FT to take into account the stored information in internal degrees of freedom at the transition state and evaluate the free energy in piece-wise manner.

It would be very difficult to (experimentally) extract the information in all internal variables for a poorly known system. Instead, we consider a few slow (uncontrolled) variables which contribute most to the information. On the other hand, instead of applying the GJE taking into account informations at the transition state, one may consider equilibrating at the transition point before proceeding to the 2nd step of pulling. Similarly, we did extend the CFT to the GCFT. Polymeric systems provide a good platform to further test the generalized form of fluctuation theorems. We plan to report similar analysis for coil globule transitions [43–46] and also to give different (equivalent) ways for evaluating informations.

Our work is supported by National Research Foundation grants provided by Korean government, NRF-2017R1A2B4010632 (NKL), NRF-2020R1A2B5B01002041 (NKL) and NRF-2018R1A2B6007007 (JMP). JMP is also supported by the Catholic University of Korea research fund 2019.

References

1. J.M. Parrondo, J.M. Horowitz, T. Sagawa, *Nature Physics* **11**(2), 131 (2015)
2. U. Seifert, *Phys. Rev. Lett.* **95**, 040602 (2005)
3. C. Jarzynski, *Phys. Rev. E* **56**(5), 5018 (1997)
4. C. Jarzynski, *J. Stat. Mech.: Theor. Exp.* **2004**(09), P09005 (2004)
5. C. Jarzynski, *Annual Review of Condensed Matter Physics* **2**, 329 (2011)
6. E. Boksenbojm, B. Wynants, C. Jarzynski, *Physica A* **389**, 4416 (2010)
7. G. Crooks, *J. Stat. Phys.* **90**, 1481 (1998)
8. G. Crooks, *Phys. Rev. E* **60**, 2721 (1999)
9. J. Liphardt, S. Dumont, S.B. Smith, I. Tinoco, C. Bustamante, *Science* **296**(5574), 1832 (2002), ISSN 0036-8075
10. D. Collin, F. Ritort, C. Jarzynski, S.B. Smith, I. Tinoco Jr, C. Bustamante, *Nature* **437**, 231 (2005)
11. U. Seifert, *Eur. Phys. J. B* **64**, 423 (2008)
12. J. Fargo, *J. Stat. Phys.* **107**, 781 (2002)
13. S. Cocco, J. Yan, D. Leger, J.-F. Chatenay, J.F. Marko, *Phys. Rev. E* **70**, 011910 (2004)

14. A. Herbert, A. Rich, J. Biol. Chem. **271**, 11595 (1996)
15. L. Peck, J. Wang, Proc. Natl. Acad. Sci. USA **80**, 6206 (1983)
16. F.C. Oberstrass, L.E. Fernandes, P. Lebel, Z. Bryant, Phys. Rev. Lett **110**, 178103 (2013)
17. A. Kwon, N.K. Lee, S. Hong, J. Fierling, A. Johner, Biophys. J. **108**, 2562 (2015)
18. M. Lee, S.H. Kim, S.C. Hong, Proc. Natl. Acad. Sci. USA **107**, 4985 (2010)
19. R.R. Sinden, *DNA Structure and Function* (Elsevier, San Diego, 1994)
20. C. Jarzynski, Phys. Rev. E **73**(4), 046105 (2006)
21. G.E. Crooks, Phys. Rev. E **61**(3), 2361 (2000)
22. I. Junier, A. Mossa, M. Manosas, F. Ritort, Phys. Rev. Lett. **102**(7), 070602 (2009)
23. A. Alemany, M. Ribezzi-Crivellari, F. Ritort, New Journal of Physics **17**(7), 075009 (2015)
24. A. Alemany, A. Mossa, I. Junier, F. Ritort, Nature Physics **8**(9), 688 (2012)
25. J. Camunas-Soler, A. Alemany, F. Ritort, Science **355**(6323), 412 (2017)
26. P. Maragakis, M. Spichty, M. Karplus, J. of Phys. Chem. B **112**(19), 6168 (2008)
27. G. Nam, N.K. Lee, H. Mohrbach, A. Johner, I.M. Kulić, EPL **100**, 28001 (2012)
28. M.K. Chae, Y. Kim, A. Johner, N.K. Lee, Soft Matter **48**(14), 2346 (2018)
29. M.K. Chae, Y. Kim, A. Johner, N.K. Lee, Polymers **12**(1), 192 (2020)
30. C. Van den Broeck, Physics of Complex Colloids pp. 155–193 (2013)
31. H. Park, JKPS **72**, 1413 (2018)
32. U. Seifert, Rep. Prog. Phys. **75**, 1 (2012)
33. J. Horowitz, C. Jarzynski, J. Stat. Mech.: Theor. Exp. **2007**(11), P11002 (2007)
34. Z. Gong, H.T. Quan, Phys. Rev. E **92**(1), 012131 (2015)
35. S. Kullback, R.A. Leibler, The annals of mathematical statistics **22**(1), 79 (1951)
36. T.M. Cover, J.A. Thomas, *Elements of information theory* (John Wiley & Sons, 2012)
37. H.H. Hasegawa, J. Ishikawa, K. Takara, D.J. Driebe, Phys. Lett. A **374**(8), 1001 (2010)
38. S. Toyabe, T. Sagawa, M. Ueda, E. Muneyuki, M. Sano, Nature Physics **6**(12), 988 (2010)
39. B. Altaner, J. Phys. A: Mathematical and Theoretical **50**(45), 454001 (2017)
40. F. Wang, D.P. Landau, Phys. Rev. Lett. **86**, 2050 (2001)
41. C. Zhou, T.C. Schulthess, S. Torbrügge, D.P. Landau, Phys. Rev. Lett. **96**, 120201 (2006)
42. M. Palassini, F. Ritort, Phys. Rev. Lett. **107**(3), 060601 (2011)
43. T.M. Birshtein, V.A. Pryamitsyn, Macromolecules **24**(7), 1554 (1991)
44. A.Y. Grosberg, D.V. Kuznetsov, Macromolecules **25**(7), 1970 (1992)
45. D. Yang, Q. Wang, ACS Macro Letters **2**(11), 952 (2013)
46. C.F. Abrams, N.K. Lee, S. Obukhov, EPL (Europhysics Letters) **59**(3), 391 (2002)

A Appendix: Simulation Methods

We first obtain the equilibrium free energy map using density of state (DoS) method in the scheme of MC simulations [40]. The details of simulation method are similar to

Ref. [29]. We have estimated DoS by using flat histogram MC scheme first introduced by Wang and Landau [41]. To compare with non-equilibrium pulling experiments, we evaluate DoS as a function of the total twist number Tw and the control parameter, $\lambda = d_s/b$. Once DoS is obtained, thermodynamic function X can be obtained from the density of state $\Omega(Tw, \lambda, E)$,

$$\begin{aligned} \langle X(Tw, \lambda) \rangle &= \Sigma_{EP}(Tw, \lambda, E)X(Tw, \lambda, E) \\ p(Tw, \lambda, E) &= \frac{\Omega(Tw, \lambda, E)\exp(-\beta E)}{Z(Tw, \lambda)} \end{aligned} \quad (\text{A1})$$

with $Z(Tw, \lambda) = \Sigma_E \Omega(Tw, \lambda, E)\exp(-\beta E)$. The free energy $F(Tw, \lambda)$ can be obtained as $F(Tw, \lambda) = -k_B T \log Z(Tw, \lambda)$. As shown in Fig. 2(a), the free energy map is obtained in the phase space of total twist Tw of range $[-1.2 - 1.2]$ and extension λ of range $[1.4 - 21.4]$ (Binning sizes are $\delta E = 1.0, \delta D = 0.1$ and $\delta Tw = 0.1$). It turns out that the limited phase space of MC simulations results in some systematic yet moderate overestimate of the free energy mainly at the large extension regime where states with unsampled structures (e.g., more twist-kinks) may contribute with some statistical weights.

The non-equilibrium pulling experiments are realized by Brownian dynamics simulations for a filament of $S = Nb$, described by the early introduced Hamiltonian, Eq. (17), complemented by the bead-spring Hamiltonian (including the handles). **In the simulations, the discretized version of the Hamiltonian, Eq. (17), is applied to a bead spring model. The bead-bead interactions are modeled by the fully repulsive Lennard-Jones(LJ) potential: $U_{LJ}(r) = 4\epsilon[(\sigma/r)^{12} - (\sigma/r)^6 + 1/4]$ for $r < 2^{1/6}\sigma$ and 0 elsewhere. In addition to LJ potential, we also apply FENE potential so that the bond length b is nearly fixed as $b = 0.99\sigma$ throughout.**

In order to describe the motion of the filament, we integrate Langevin equation with the total energy U ,

$$\zeta \frac{\partial \mathbf{r}_i}{\partial t} = -\frac{dU}{d\mathbf{r}_i} + \boldsymbol{\eta}(t), \quad (\text{A2})$$

together with the Gaussian random force $\langle \boldsymbol{\eta}(t) \rangle = 0$ and $\langle \boldsymbol{\eta}(t)\boldsymbol{\eta}(t') \rangle = 6k_B T \zeta \delta(t-t')$, where ζ is the frictional coefficient. The random forces and diffusion constant $D_0 = k_B T / \zeta$ include temperature dependence. In the simulation, temperature is measured in units of $T_0 = \epsilon/k_B$, time is measured in units of molecular diffusion time $t_0 = \sigma^2 \zeta / k_B T_0$, and lengths are measured in units of σ . Thus, energy and force are measured in units of $k_B T_0$ and $k_B T_0 / \sigma$, respectively.

We perform pulling/releasing under the control of the *single* macroscopic variable d_s , which is the distance between two attached beads parameterized by $\lambda = d_s/b$. The initial conformations prior to pulling are sampled under the constraint that the distance between beads is set to be $\lambda_0 = 4$. Each initial conformation is equilibrated prior to the pulling process. The control parameter λ varies from the initial state λ_0 (λ_f) to the final state λ_f (λ_0) at a given pulling(releasing) rate for each trajectory. With one bead fixed, the second bead is moved along the initial end-to-end vector (which is designated as $+\hat{x}$ -axis). During pulling/releasing, the position of the second bead is

updated by $\pm 0.01 b$ at every time interval of $\delta t = 0.25, 0.5, 1.0, 5.0, 25.0, 50.0 t_0$. These give pulling speeds: $v = 0.0002, 0.0004, 0.002, 0.01, 0.02$ and $0.04 b/t_0$. About 1000 independent experiments are performed for various pulling speeds. At given λ , the time-dependent forces \mathbf{f}_0 and \mathbf{f}_1 are measured by monitoring the extended lengths of the two springs from their preferred length r_0 , where the subscripts 0 and 1 stand for the fixed bead and the moving bead, respectively. The system being out of equilibrium, these forces do not cancel, as can be easily seen at high speed from Fig. 3(a).

B Appendix: Alternative evaluation of information

We define the logarithm of the ratio $i_{o,n} = \ln(p(\theta_n)/p_{eq}(\theta_n))$ as the information associated with orientational degrees of freedom of n -th bond link for a given trajectory. We also define the logarithmic ratio of probabilities to have specific twist rate as compared to the equilibrium value, $i_{tw,n} = \ln(p(\psi_n)/p_{eq}(\psi_n))$, as information associated with twist angle fluctuation of the link n . If all twist angles and bond angles fluctuate independently, total information stored in the filament would be the sum of informations of all bond angles and twist angles.

The sum should be evaluated by run over the independent blocks.

$$\begin{aligned} D_i &= \sum_{\{r\}} k_B \ln\{p(\theta_r, \psi_r)/p^{eq}(\theta_r, \psi_r)\} \\ &= \sum_{\{r\}} k_B \left\{ \ln \frac{p(\theta_r)}{p^{eq}(\theta_r)} + \ln \frac{p(\psi_r)}{p^{eq}(\psi_r)} \right\}. \end{aligned} \quad (\text{A3})$$

The probability to visit a point in $\sim 2N$ dimensional space $\{\theta_j, \psi_j\}$ is nearly impossible to measure as the volume of phase space increases exponentially with the number of degrees of freedom. To avoid such complexity, we group $p(\{r_n, \psi_n\})$ into several (independent) blocks. We then identify the block size n_b from the probability distribution for consecutive monomers for the pulling speed under consideration.

$$D_i = \frac{1}{n_b} (i_{o,n} + i_{tw,n}). \quad (\text{A4})$$

In Fig. 7, we demonstrate the distributions for all monomer positions, all bond angles and twist increments at the pulling speed $v = 0.01 b/t_0$ when they are brought at $\lambda^\dagger = 9.2$ by finite speed pulling procedures. Each distribution obtained from non-equilibrium experiment deviates from the relaxed distribution signifying reduction of phase space. The local twist angle fluctuations are fast relaxing, there are little difference with equilibrium distributions. The main contribution to the information comes from the slowest degrees of freedom. We seek for the number of independent blocks contributing to the information (Eq. A3) by matching the information of all local variables to that of slowly relaxing variables. **The monomer positions and**

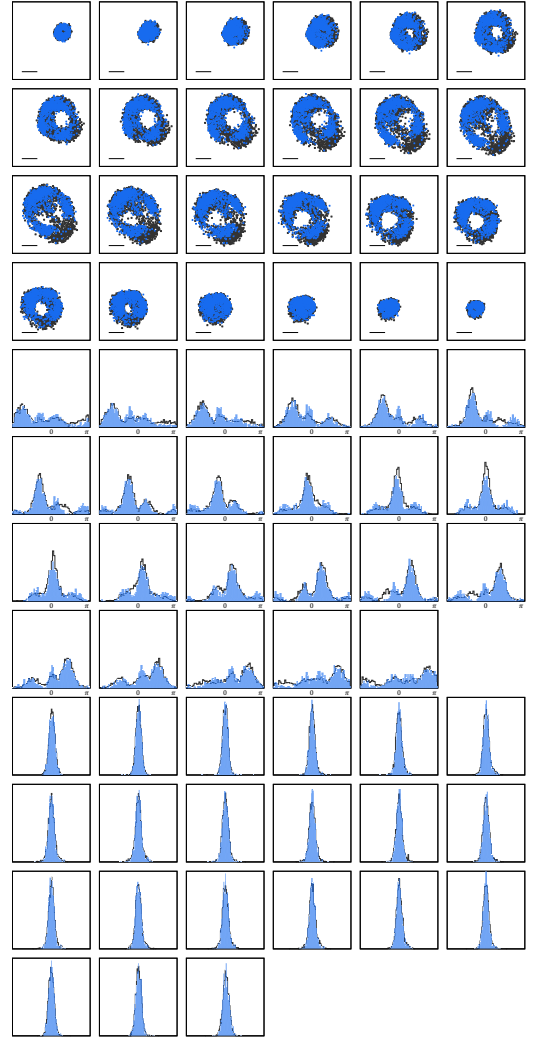


Fig. 7. Distributions of (a) monomer displacements, (b) orientational link angles, for all monomers and links of the model filament of $N = 24$. (c) Distributions of local twist angles of selected links. Coordinates are obtained from 1000 trajectories at the midpoint extension $\lambda^\dagger = \lambda_M = 9.2$. The blue dots represent non-equilibrium distributions with $v_s = 0.002 b/t_0$ and gray points represent the relaxed distributions at the given extension.

bond vector orientations are clearly correlated over some finite lengths. We may determine the effective block size n_b by inspection of the characteristics distributions. Far before the transition point, $\langle n_k \rangle \ll 1$, we found the correlated length is as large as $N = 24$. At transition point, $\langle n_k \rangle \sim 0.5$, the average correlated length is measured to be ~ 6 (See, for example Fig. fig:app). At the final position where $\langle n_k \rangle \sim 1$, we find the average correlated length is ~ 12 . The effective block size is $n_b = N/2$ implies that "most" of conformations are $n_k = 1$ state with strongly localized kink. After determining n_b at the given position, the obtained information (Eq. A4) is used for the evaluation of free energy through GJE or GCFT. We found the results consistent with the method using slow variables.