

**Reducing the variance in the translocation times by prestretching the polymer**Hendrick W. de Haan,<sup>1</sup> David Sean,<sup>2,3</sup> and Gary W. Slater<sup>2</sup><sup>1</sup>*Faculty of Science, University of Ontario Institute of Technology, Oshawa, Ontario, Canada, L1H 7K4*<sup>2</sup>*Physics Department, University of Ottawa, Ottawa, Ontario, Canada, K1N 6N5*<sup>3</sup>*Institut für Computerphysik, Universität Stuttgart, 70569 Stuttgart, Germany*

(Received 8 February 2018; published 3 August 2018)

Langevin dynamics simulations of polymer translocation are performed where the polymer is stretched via two opposing forces applied on the first and last monomer before and during translocation. In this setup, polymer translocation is achieved by imposing a bias between the two pulling forces such that there is net displacement towards the *trans* side. Under the influence of stretching forces, the elongated polymer ensemble contains less variations in conformations compared to an unstretched ensemble. Simulations demonstrate that this reduced spread in initial conformations yields a reduced variation in translocation times relative to the mean translocation time. This effect is explored for different ratios of the amplitude of thermal fluctuations to driving forces to control for the relative influence of the thermal path sampled by the polymer. Since the variance in translocation times is due to contributions coming from sampling both thermal noise and initial conformations, our simulations offer independent control over the two main sources of noise and allow us to shed light on how they both contribute to translocation dynamics. Simulation parameter space corresponding to experimentally relevant conditions is highlighted and shown to correspond to a significant decrease in the spread of translocation times, thus indicating that stretching DNA prior to translocation could assist nanopore-based sequencing and sizing applications.

DOI: [10.1103/PhysRevE.98.022501](https://doi.org/10.1103/PhysRevE.98.022501)**I. INTRODUCTION**

The translocation of polymers across membranes through nanopores is central in biological processes at the cellular level [1] and in the development of DNA analysis techniques [2]. Applications such as nanopore-based sequencing technologies have led to numerous experimental [3–8], theoretical [9–17], and computational [18–32] studies of polymer translocation. Experimental histogram distributions of the translocation times  $\tau$  exhibit surprisingly large variances—particularly so when considering the perfect monodispersity of DNA. This large variance in the translocation time is a prominent source of uncertainty for applications such as using nanopores for DNA size determination. It also introduces complications for sequencing applications where a consistent and orderly passage of bases is desirable. Hence, techniques to reduce this variation are of great interest in the development of nanopore technologies. In this work we propose and explore a methodology for reducing this variation by stretching the polymer prior to translocation.

In an idealized translocation setup that does not include effects such as polymer-wall interactions, the degradation of the nanopore or interactions with ions or impurities in the solvent, the variance in translocation time is known to come from two different contributions: (1) Brownian noise arising from thermal fluctuations (i.e., the *stochastic path* that a polymer follows in a particular translocation event) and (2) conformational noise arising from otherwise identical polymers starting translocation in different conformations [11,33–36]. Hence, the variance in the translocation time can be reduced by either reducing the Brownian noise (e.g., turning down the temperature) or by reducing the variation between the different initial conformations (i.e.,

starting all translocation events from very similar initial states).

In this work, we narrow the distribution of initial conformations by stretching the polymer. Prior to the translocation, the polymer is stretched by the application of a force to the first and last monomer in opposite directions. As the stretching force is increased, conformations in which the polymer starts compressed near the pore or even in a relaxed state become less and less likely. In the limit of a very strong stretching force, all events would start with the polymer in a linear conformation along the pore axis. Reducing conformational noise by confining the polymer in a long cylinder [35] or a cylindrical cavity [36] prior to translocation has been investigated previously. However, the use of this dual-force stretching setup is different than geometrical confinement in that the pathological conformations such as knots and hairpins are nearly impossible to realize. Moreover, the translocation is driven by pulling on the polymer end as opposed to the monomers in the nanopore. This simplifies the translocation dynamics and as well as significantly reduces the effect of monomer crowding on the *trans* side.

The stretching force,  $F_s$ , thus enables us to implicitly control the range of initial conformations while Brownian fluctuations can be controlled via the simulation variable  $k_B T$ . Once the polymer is relaxed in a stretched state, the translocation process is initiated by increasing the force applied on the end of the polymer initially located in the pore by an amount  $F_d$  such that there is a net displacement towards the *trans* side through the nanopore. Driving the translocation via end-pulling eliminates the need to consider the effects of monomer crowding on the *trans* side of the membrane, yielding a clean setup to investigate translocation dynamics.

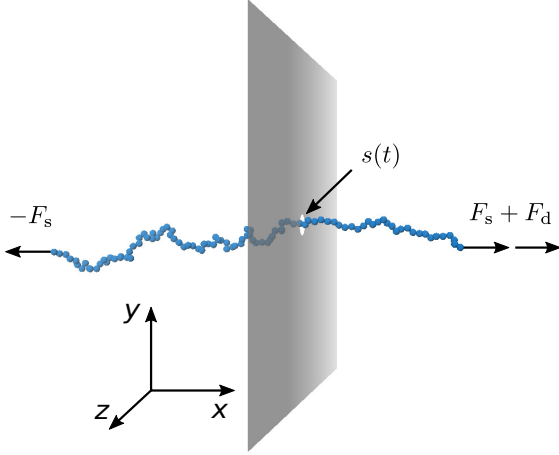


FIG. 1. In this force configuration, the polymer is stretched by applying two equal and opposite forces  $\pm F_s \hat{x}$  on the polymer ends. An additional force  $F_d \hat{x}$  is applied to the first monomer to drive the polymer to the *trans* side. The translocation coordinate  $s(t)$  is defined as the monomer index inside the pore.

## II. THE STRETCHING-PULLING FORCE SETUP

Figure 1 shows a schematic of how both a stretching force and a driving force are implemented in our simulations. A stretching force  $F_s$  is applied to the first monomer ( $i = 1$ ) perpendicular to the plane of the membrane in the *trans* direction while the same force is applied to the last monomer ( $i = N$ ) in the opposite direction. Although the force  $F_s$  stretches out the polymer, the net force on the polymer remains zero. Hence, to drive the translocation process, an additional pulling force  $F_d$  is applied to the first monomer. The DNA is prepared by stretching using  $F_s$  on opposite ends; these stretching forces continue to keep the DNA extended during the course of the translocation process. This pulling force setup corresponds to translocation as controlled by optical tweezers or attached magnetic beads [37–39]. While this has been previously studied via simulations [22,29], here we feature the combination of the stretching and pulling forces. A previous simulation study also employed a double force arrangement, but in that work, a force in the pore opposed the pulling force [40].

Lehtola *et al.* [41] investigated how the initial polymer configuration affects the scaling law behaviors by simulating a polymer chain with an initial configuration of monomers in a straight line. They observed a scaling  $\tau \sim N^\alpha$  with an exponent of  $\alpha = 2$ , which is the same as the driven limit of Sung and Park [9] where the monomer-fluid friction dominates over the monomer-nanopore friction. These findings indicate that not only the applied force and the length of the chain but also the initial polymer conformation affects the scaling exponent strongly.

### Polymer simulations

A setup that is very close to a standard coarse-grained Langevin dynamics (LD) simulation approach is used to model the system [42]. The excluded volume interactions between the  $N = 100$  monomer beads are implemented via the WCA

potential [43]:

$$U_{\text{WCA}}(r) = \begin{cases} 3 \left\{ 4\epsilon_{\text{LJ}} \left[ \left( \frac{\sigma}{r} \right)^{12} - \left( \frac{\sigma}{r} \right)^6 \right] + \epsilon \right\} & \text{for } r < r_c \\ 0 & \text{for } r \geq r_c. \end{cases} \quad (1)$$

Here the nominal well depth  $\epsilon_{\text{LJ}}$  serves as a fundamental energy scale, and the nominal bead diameter  $\sigma$  as the fundamental length scale. Unless otherwise noted, we report length, energy, and force variables in units of  $\sigma$ ,  $\epsilon_{\text{LJ}}$ , and  $\epsilon_{\text{LJ}}/\sigma$  respectively. The cutoff distance is set to  $r_c = 2^{1/6}\sigma$  such that this potential is purely repulsive.

To connect monomers along the contour length of the polymer, the finitely extensible nonlinear elastic (FENE) potential is used. Note that for this work, as relatively high stretching forces are used, we multiply these two standard potentials (WCA and FENE) by a factor of three in order to reduce bond-stretching artifacts:

$$U_{\text{FENE}}(r) = -3 \left[ \frac{1}{2} k r_0^2 \ln \left( 1 - \frac{r^2}{r_0^2} \right) \right]. \quad (2)$$

To prevent bond crossing [44] the parameters are chosen to be  $k = 30\epsilon_{\text{LJ}}/\sigma^2$  and  $r_0 = 1.5\sigma$ .

Monomer positions are integrated in time using the Langevin equation of motion [42]. Monomers are subject to random thermal kicks which have a variance  $2\zeta k_B T / \Delta t$  in each dimension, where  $\Delta t = 0.01$  is the integration time step. Since we will vary the magnitude of thermal fluctuations via the variable of  $k_B T$ , we report temporal quantities in units of  $\zeta \sigma^2 / \epsilon_{\text{LJ}}$ .

The membrane is modeled as a mathematical surface with a pore of radius  $1.5\sigma$ . The beads and the membrane interact via the WCA potential. The nanopore thus has an effective diameter of  $2\sigma$  and a length of  $\sigma$ .

The external stretching force  $\pm F_s \hat{x}$  is applied on the first and last monomer (respectively) and the polymer is equilibrated. The warmup procedure consists of prestretching the polymer with  $-F_s \hat{x}$  while having the first monomer confined in the pore. After the warmup, the monomer held in the pore is released and pulled by additional external force  $(F_s + F_d) \hat{x}$ . However, since the tension blob size scales as  $\sim k_B T / F_s$ , decreasing (increasing) the value of  $k_B T$  for a given stretching force  $F_s$  will yield an increased (decreased) end-to-end distance. To decouple this change in the amount of prestretching upon variations in the control parameter  $k_B T$ , the polymer is stretched by pulling on both ends with a force  $F_s = \hat{F}_s k_B T / \sigma$ . This ensures that the dimensionless stretching force  $\hat{F}_s = F_s \sigma / k_B T$  yields the same degree of deformation. Thus, when we report a particular value for  $\hat{F}_s$ , this will always amount to the same degree of polymer deformation, irrespective of the value of  $k_B T$  used in the simulations.

The simulation is restarted with a new initial conformation whenever the polymer is found to be completely on the *cis* side of the membrane; this is considered to be a failed translocation attempt. A total of 1000 successful translocations are generated for each data point.

In a recent paper, we demonstrated that the dynamics of polymer translocation through a nanopore are significantly

affected by the balance between the magnitude of the force driving the polymer towards the *trans* side and the diffusive aspects arising from thermal noise [45]. To quantify this relationship, one can define the *translocation Péclet* number as the ratio between the polymer relaxation time and its translocation time  $\tau$  [11,45,46]:

$$P_t = \frac{1}{\tau} \frac{R_{g0}^2}{D_0} \approx \frac{F_d}{k_B T}, \quad (3)$$

where  $R_{g0}$  and  $D_0$  are the free solution radius of gyration and diffusion coefficient of the polymer, respectively. In that work, we found that typical coarse-grained (CG) simulation setups, where the force and diffusion coefficient are on the order of unity in LD units, yield a  $P_t$  that is too low for modeling realistic dsDNA translocation events [45]. The two relevant energy scales are thus the thermal energy  $k_B T$  and the potential energy associated with monomers crossing the nanopore,  $F_d \sigma$ , where  $\sigma$  is the effective width of the membrane. In the current simulations, we map the effect of prestretching over a wide range of  $F_d \sigma / k_B T$  ratios—from 1 to 100. Values of interest are thus  $F_d \sigma / k_B T \approx 1-10$ , which corresponds to typical CG setups;  $F_d \sigma / k_B T \approx 50$ , which corresponds to a good (albeit too flexible) model of dsDNA; and  $F_d \sigma / k_B T \geq 100$ , which corresponds to even more CG models.

In order to cover a range where  $F_d \in (1-100)k_B T / \sigma$  while remaining both in the overdamped limit ( $F_d$  cannot be arbitrarily high) and being able to obtain events in a reasonable simulation time ( $F_d$  cannot be arbitrarily low), both  $F_d$  and  $k_B T$  are varied. Simulations at driving forces of  $F_d \sigma / \epsilon_{LJ} = 0.1, 0.2, 0.5$ , and  $1.0$  are performed for two values of the thermal energy:  $k_B T = 0.01 \epsilon_{LJ}$  and  $k_B T = 0.10 \epsilon_{LJ}$ . To compare across cases, we report a dimensionless driving force  $\hat{F}_d = F_d \sigma / k_B T$ . Note that the cases overlap at  $\hat{F}_d = 10$ —attained using the two combinations: (1)  $F_d = 0.1 \epsilon_{LJ} / \sigma$  with  $k_B T = 0.01 \epsilon_{LJ}$  and (2)  $F_d = 1 \epsilon_{LJ} / \sigma$  with  $k_B T = 0.1 \epsilon_{LJ}$ —which allows us to verify that there are only minimal differences between these cases and that the effects overwhelmingly arise from the value of  $\hat{F}_d$ .

### III. RESULTS

#### A. Translocation times $\tau$

Figure 2 shows the translocation time distributions at different stretching forces  $\hat{F}_s = 0, 1, 5, 25$  for three driving forces  $\hat{F}_d$ . Note that for these plots as well as others in the present paper, the translocation time  $\tau$  is scaled by the driving force  $F_d$ . The absolute translocation time is trivially shorter for cases with a higher driving force; the rescaling used here assumes that  $\langle \tau \rangle \sim F_d^{-1}$ , which is not exactly the case, as we shall see later. Examining the qualitative differences between the  $\hat{F}_d = 1, 10, 100$  panels, one notes that the overlap between the distributions diminish with increasing driving magnitude  $\hat{F}_d$ . Recall that the source of variance has contributions from both thermal noise and initial conformations; increasing  $\hat{F}_d$  mostly reduces the contributions from the thermal noise, whereas increasing  $\hat{F}_s$  reduces the contributions from the initial conformations. Thus the segregation between the  $\hat{F}_s$  histograms as  $\hat{F}_d$  is increased arises from reducing the effect of thermal noise. Conversely, the  $\hat{F}_d = 100$  panel clearly showcases how conformational noise

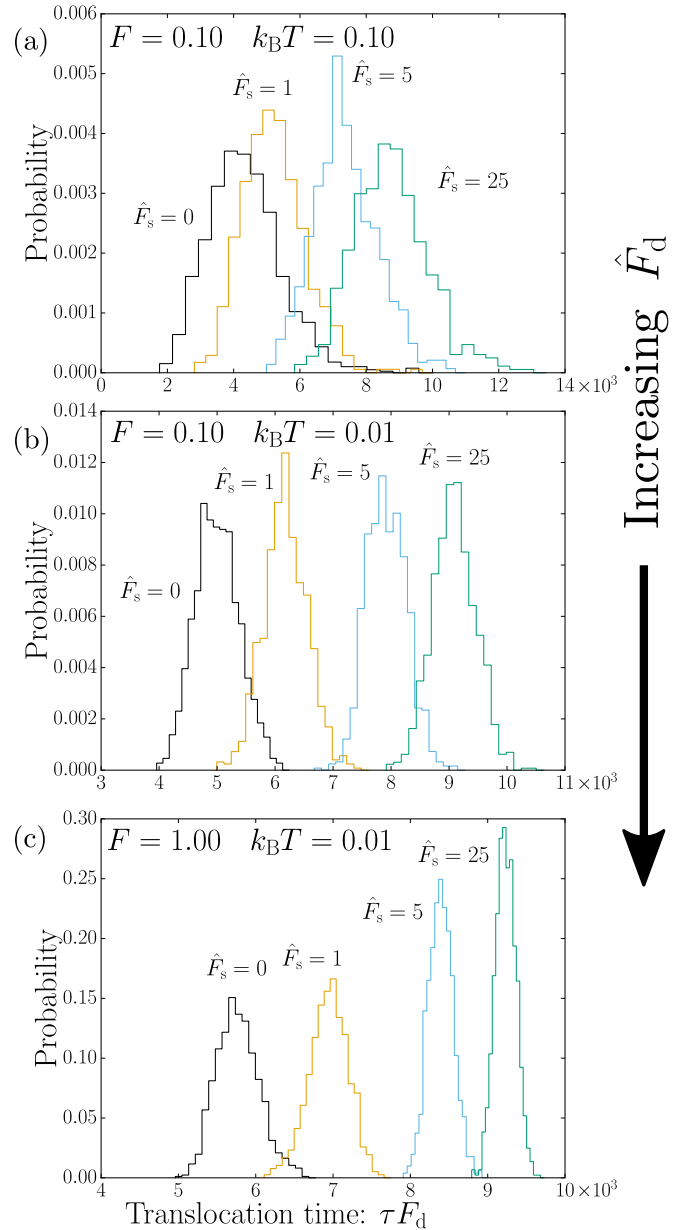


FIG. 2. Histogram distributions of the translocation time with different stretching forces  $\hat{F}_s$ . Three driving regimes are shown: (a)  $\hat{F}_d = 1$ , (b)  $\hat{F}_d = 10$ , and (c)  $\hat{F}_d = 100$ .

is suppressed with prestretching. A striking feature here (and at all simulated  $\hat{F}_d$  values) is that the mean translocation time  $\langle \tau \rangle$  increases considerably with increasing  $\hat{F}_s$ .

This is shown more clearly in Fig. 3 for all  $\hat{F}_d$  values. Again, the mean translocation time  $\langle \tau \rangle$  is rescaled by the force  $F_d$ . The data in Fig. 3 exhibit two plateaus: one at low  $\hat{F}_s$  with no (or little) prestretching and one at  $\hat{F}_s \rightarrow \infty$  where the initial conformations are rodlike. At  $\hat{F}_s = 0$ , the slight increase of  $\langle \tau \rangle F_d$  with increasing  $\hat{F}_d$  highlights how the driving regime affects the translocation time. Both data sets corresponding to  $\hat{F}_d = 10$  (with different combinations of  $k_B T$  and  $F_d$ ) give the same result, demonstrating that these simulations are indeed following the same physical process.

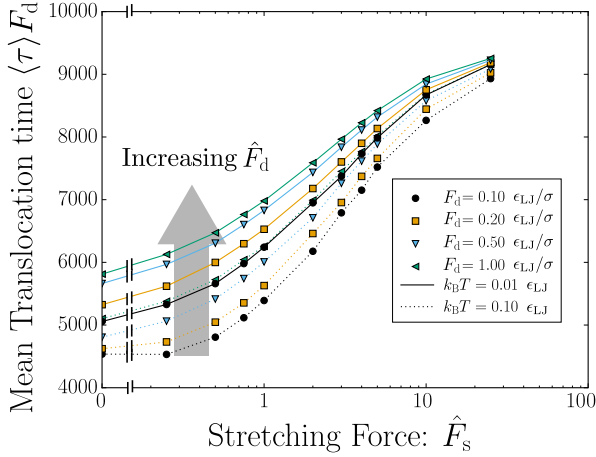


FIG. 3. The average translocation  $\langle \tau \rangle$  time scaled by the driving force  $F_d$  as a function of the scaled stretching force  $\hat{F}_s$ .

In the high prestretching limit  $\hat{F}_s \rightarrow \infty$  the conformations all approach a rodlike initial state. We expect the data to asymptotically collapse for the different  $\hat{F}_d$  values here since the  $\langle \tau \rangle \sim 1/F_d$  should be strictly valid for rods, as confirmed in Fig. 3.

Two effects need to be highlighted in order to understand the increase in  $\langle \tau \rangle$  with  $\hat{F}_s$ . First, we need to consider the tension-propagation (TP) translocation dynamics while taking into account the fact that tension propagates faster along a stretched polymer. As the polymer is pulled through the pore the net force is simply the pulling force  $F_d$  applied to the first monomer. However, the net drag on the polymer is proportional to the number of monomers that have been set into motion; monomers outside the range of the tension front do not contribute to the drag [11–15]. When  $\hat{F}_s$  is low and the polymer is not in a stretched conformation, the number of monomers being dragged increases as translocation proceeds, but it is initially quite low as the pulling force is pulling out the slack in the coil rather than dragging the entire coil. Conversely, when the stretching force is very high, the polymer starts nearly fully extended, and thus more monomers contribute to the drag earlier; the tension front then quickly reaches the end of the polymer at which point all monomers contribute to the drag. This picture is consistent with the TP model and has recently been explicitly tested for the case of driven polymer translocation from a tube [35,36]. The plateau observed at high  $\hat{F}_s$  indicates the saturation of the underlying polymer deformation.

For large  $\hat{F}_d$  values, the driving force dominates over thermal forces, hence the evolution of the polymer is largely deterministic: the relative unimportance of thermal fluctuations means that the polymer does not relax as translocation progresses, i.e., the process is highly out of equilibrium.

Conversely, at small  $\hat{F}_d$  values, the polymer is allowed to partially adapt as translocation proceeds. The driving force deforms the polymer via tension propagation, which effectively moves the center of mass of the *trans* monomers away from the nanopore. If the polymer relaxes somewhat from this nonequilibrium state, it does so by primarily having monomers move *towards* the pore. Hence, as the thermal energy is

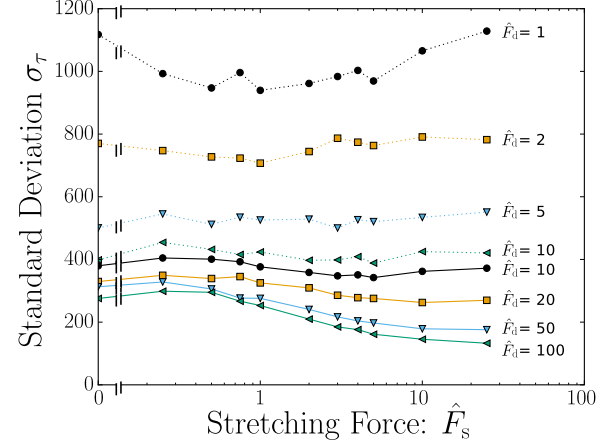


FIG. 4. Standard deviation  $\sigma_\tau$  plotted as a function of the stretching force  $\hat{F}_s$  for several values of the driving force  $\hat{F}_d$ .

increased, the effective drag is reduced and the translocation time decreases. This is seen in Fig. 3 at low  $\hat{F}_s$  where the higher  $k_B T$  curves are below the lower  $k_B T$  curves, and within each  $k_B T$  case, higher driving forces yield larger translocation times. As the stretching force  $\hat{F}_s$  increases, the impact of the nonequilibrium effects is diminished. At high  $\hat{F}_s$  values, the chain is nearly completely stretched and acts almost as a single body for the entire translocation process regardless of the balance between  $F_d$  and  $k_B T$ . Consequently, very little deviation of  $\tau$  between  $\hat{F}_d$  values is observed in this limit and the data converge. Note that the  $\hat{F}_s = 0$  case is studied in detail in Ref. [45], where we demonstrate that this additional increase in  $\tau$  at low  $\hat{F}_s$  reflects varying degrees of nonequilibrium effects.

### B. Translocation time standard deviations $\sigma_\tau$

Returning to Fig. 2, the variation in the widths  $\sigma_\tau$  of the distributions depend not only on  $\hat{F}_s$  but also on  $\hat{F}_d$ . In the last panel with  $\hat{F}_d = 100$ , corresponding to the driving force dominating over thermal forces, the width  $\sigma_\tau$  decreases with increasing  $\hat{F}_s$ . For  $\hat{F}_d = 1$  (first panel) and  $\hat{F}_d = 10$  (middle panel), the behavior is less obvious and the width is only weakly dependent upon  $\hat{F}_s$ . Comparing across panels,  $\sigma_\tau$  decreases with increasing  $\hat{F}_d$  at any particular  $\hat{F}_s$  value. This reflects the suppression of diffusion resulting in reduced variation of the stochastic paths and narrower distributions.

The value of  $\sigma_\tau$  is plotted against  $\hat{F}_s$  for different  $\hat{F}_d$  values in Fig. 4. As expected,  $\sigma_\tau$  decreases with  $\hat{F}_s$  at large  $\hat{F}_d$  values. At  $\hat{F}_d = 1$ , the width initially decreases but subsequently increases, while at  $\hat{F}_d = 2$  and 5,  $\sigma_\tau$  is essentially independent of  $\hat{F}_s$ .

As discussed, there are two contributing factors to the width of the distribution of translocation times: the ensemble of initial conformations and the effects of thermal noise during the translocation process. One expects that the latter mechanism, the variation due to diffusion, to grow with time. Recalling that the translocation time increases significantly as the stretching force increases, this means that there are two competing effects for the data shown in Fig. 4. Due to the fact that the mean translocation time  $\langle \tau \rangle$  increases with  $\hat{F}_s$ , diffusion will



obviously cause  $\sigma_\tau$  to increase with  $\hat{F}_s$ . On the other hand, increasing  $\hat{F}_s$  reduces the ensemble of initial conformations and this causes a reduction in  $\sigma_\tau$  with increasing  $\hat{F}_s$ .

The competition between these effects yields the behavior seen in Fig. 4. At low  $\hat{F}_d$  and low  $\hat{F}_s$ ,  $\sigma_\tau$  is relatively independent of  $\hat{F}_s$  (excluding the nonmonotonic case of  $\hat{F}_d = 1$ ) indicating that these two effects almost balance out. This is largely true also for the case of high  $\hat{F}_d$  and low  $\hat{F}_s$ .

However, the behavior at high  $\hat{F}_s$  depends on  $\hat{F}_d$ . At low  $\hat{F}_d$ ,  $\sigma_\tau$  increases with increasing  $\hat{F}_s$  indicating that the diffusive effects dominate over the further reduction in the range of initial conformations. On the other hand, at high  $\hat{F}_d$ ,  $\sigma_\tau$  decreases with increasing  $\hat{F}_s$ , and hence diffusive effects are marginal compared to the narrowing of the distributions that arises from the reduced conformational phase space. Low  $\hat{F}_d$  thus corresponds to diffusion dominated dynamics, while for high  $\hat{F}_d$  the process is primarily driven, as one would expect.

The effect of increased fluctuations due to increased translocation time can be accounted for by normalizing  $\sigma_\tau$  by  $\tau$ . This is discussed in Sec. III C and shown in Fig. 6, where  $\sigma_\tau/\tau$  decreases with increasing  $F_d$  and  $F_s$ .

Up to this point, we have been interested in the fluctuations at a specific time, namely, the end of the translocation process where the translocation coordinate  $s(t)$  equals  $N$  and the time  $t$  equals  $\tau$ ,  $s(\tau) = N$ . Thus, the focus has been on the accumulated noise  $\sigma_\tau$ . However, the physical picture put forth by the TP theory suggests that when an end-pulling process is considered [11], the noise arising from the initial conformations plateaus when the tension front reaches the last monomer. After this stage the  $N$  monomers of the polymer move in unison with a constant velocity, and the source of fluctuations is solely Brownian. Thus the way by which the fluctuations grow should exhibit the characteristic two-steps of the underlying dynamics. To explore this, we examine statistics of the translocation coordinate  $s$  as a function of time  $t$ .

Log-log plots of the variance  $\sigma_s^2 = \langle (s(t) - \langle s(t) \rangle)^2 \rangle$  as a function of the scaled time  $t/\langle \tau \rangle$  are shown in Fig. 5. In Fig. 5(a) we show a selected  $\hat{F}_s = 5$  case with two additional curves chosen to highlight the contrast between fluctuations arising from Brownian noise versus those due to initial conformations. To do this, we perform a set of deterministic simulations where the polymer is initiated with conformations according to  $\hat{F}_s = 5$  but  $k_B T$  is set to zero such that there is no thermal noise. Hence, the resulting fluctuations arise solely due to variations in the initial conditions. These deterministic data are shown as a labeled dashed line. To contrast this, we also plot the expected fluctuations arising from Brownian motion, modeled as a line with slope unity ( $\sim Dt^1$ ) with an arbitrary prefactor chosen to match the data at the end of the process. This is shown as the dotted line in Fig. 5(a) (note that this line should be taken to guide the eye, as the instantaneous diffusion coefficient of the polymer depends upon the fractional amount of monomers set in motion by the applied force). Examining Fig. 5(a) using these two additional curves, the fluctuations from the selected  $\hat{F}_s = 5$  case start very near the deterministic result. However, the deterministic line saturates near  $t/\langle \tau \rangle \approx 0.08$ . This corresponds to the time where the tension front reaches the last monomer. Correspondingly, the fluctuations in the simulations begin to level off—but cannot completely flatten due to diffusion effects. In general, it is assumed that the motion

of a translocating polymer follows a power law  $\langle \Delta s(t)^2 \rangle \sim t^\beta$  with  $\beta$  close to unity [47]. Fluctuations arising from thermal noise should eventually dominate over the effect from initial conformations which saturate before the translocation time ( $\tau$ ). The proposed linear  $\langle \Delta(t)^2 \rangle \sim t^1$  line closely follows the data for long times.

The family of curves presented in Fig. 5(b) represents data for varying stretching forces. As expected, polymers with higher prestretching exhibit a lower amount of fluctuations: as  $\hat{F}_s$  is increased the curves shift downwards. The curves differ not only in the absolute amount of fluctuations but also in the rate at which these fluctuations grow. As expected in the final—Brownian noise only—stage of translocation, the fluctuations appear to follow the expected diffusive power law. In situations of low prestretching, this diffusive regime is short-lived since the onset occurs very close to the end of the translocation process. In the opposing limit of high prestretching, the diffusive regime completely dominates. In early times, however, fluctuations arise from both the conformational noise (molecular individualism) and Brownian diffusion. Hence, prestretched polymers start with a low variance in  $s(t)$  that increases approximately linearly with  $t$  while polymers at  $\hat{F}_s = 0$  start with a large variance in  $s(t)$  and essentially do not exhibit the diffusive regime. Between these extremes, a crossover from initial conformation to thermal noise can be observed. The crossover (close to the TP time) depends upon the amount of prestretching, consistent with Refs. [17,35].

### C. Scaled variations

It is convenient to calculate the standard deviation  $\sigma_\tau$  normalized by the mean translocation time  $\langle \tau \rangle$ . This quantity, which reflects the relative width of the distributions, is known as the coefficient of variation and is given by

$$c_v = \frac{\sigma_\tau}{\langle \tau \rangle}. \quad (4)$$

These values are plotted in Fig. 6(a). Now, almost without exception, the normalized distribution width decreases with increasing  $\hat{F}_s$ , thus indicating that reducing the range of initial conformations does reduce the uncertainty in the measured translocation time, even though translocation takes longer.

To quantify these effects further, we plot in Fig. 6(b) the percentage difference in the coefficient of variation  $\chi_v$ . It is calculated as the difference between  $c_v$  with no stretching force and  $c_v$  at the highest stretching force normalized by  $c_v$  at  $\hat{F}_s = 0$ :

$$\chi_v \equiv \frac{c_v(\hat{F}_s = 0) - c_v(\hat{F}_s = 25)}{c_v(\hat{F}_s = 0)} \times 100. \quad (5)$$

The decrease in the variance of the translocation times is strongly dependent on the driving regime  $\hat{F}_d$ . For  $\hat{F}_d = 1-10$ , for instance, the percentage relative decrease lies between 40% and 50%. At these low driving forces, the relative decrease actually decreases slightly with increasing driving force. Note that there is a slight discrepancy between the  $k_B T = 0.1$  and  $k_B T = 0.01$  cases, which do not quite overlap at  $\hat{F}_d = 10$ . Above  $\hat{F}_d = 10$ , the relative decrease increases significantly with increased driving force. For  $\hat{F}_d = 100$ , the

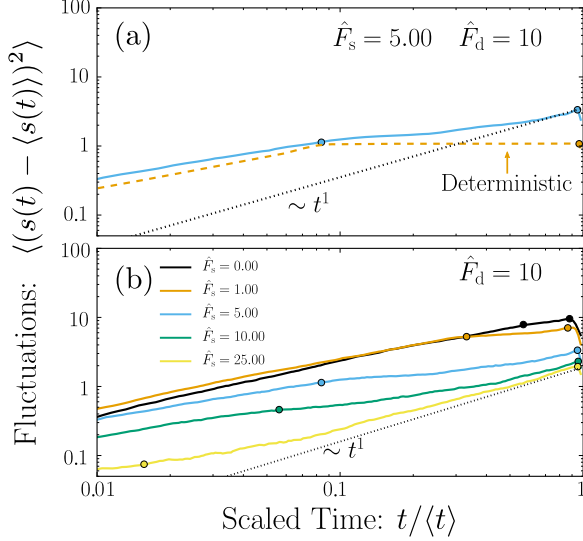


FIG. 5. Fluctuations of  $s(t)$  vs scaled time  $t/\langle t \rangle$ . (a) A single stretching force together with a deterministic simulation and a  $t^1$  line to guide the eye. (b) Different stretching situations as different curves. The curves contain a pair of circular markers. The first indicates the time where the variance contribution from the initial conformations disappears (obtained from the plateau of the deterministic simulation; see top panel). The second marker near  $t = \langle t \rangle$  indicates the time for the first translocation of the ensemble, i.e., the ensemble population decreases from this point onwards.

relative width of the distribution decreases by 70% between  $\hat{F}_s = 0$  and  $\hat{F}_s = 25$ . The data are saturating at large stretching forces.

Figure 7 displays this same data as a phase diagram in  $\hat{F}_s$  and  $\hat{F}_d$ . One can see that both increasing the effective driving force (by increasing  $F_d$  or decreasing the temperature) and prestretching the polymer yield improved (lower) coefficients of variation. The largest effect is observed when both of these factors are employed simultaneously. However, the region near the lowest  $c_v$  is quite flat, indicating that both effects saturate, and thus nearly optimal results are obtained once a sufficient

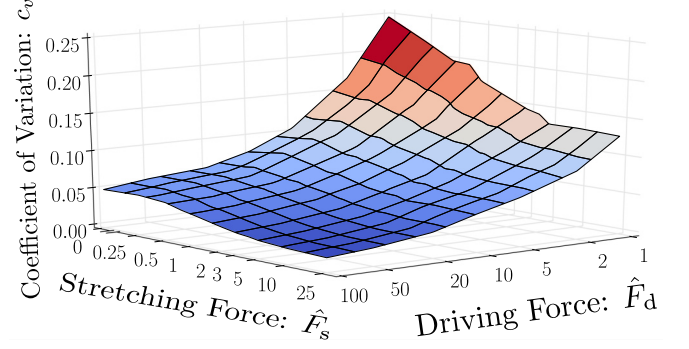


FIG. 7. Phase diagram of the coefficient of variation  $c_v$  as a function of the two variables that control the two noise sources: the stretching force  $\hat{F}_s$  and the driving force  $\hat{F}_d$ . Note the nonlinear layout of both force axes.

strong stretching force is coupled with a sufficiently large Péclet number.

#### IV. CONCLUSIONS

In this paper, we constructed a simulation model to study translocation such that two sources of noise (conformational and Brownian) can be modulated via two independent control parameters. This was achieved by a stretching-pulling force scenario in which a force of  $F_s$  is applied to the first monomer and  $-F_s$  is applied to the last monomer to stretch out the polymer both before translocation begins. These two forces also keep the polymer stretched during translocation, while an additional force  $F_d$  is applied to the first monomer, which pulls the polymer through pore into the *trans* region.

We used  $k_B T$  to control for Brownian noise and  $F_s$  for initial conformations across a range of translocation Péclet numbers by varying  $F_d$  and  $k_B T$ .

We found that prestretching the polymer has two significant benefits: not only does the variance of the translocation time  $\tau$  decrease, but the mean value of  $\tau$  increases. Hence, the process is slowed down and the accuracy of the measured times is increased; both of these are beneficial for DNA sizing

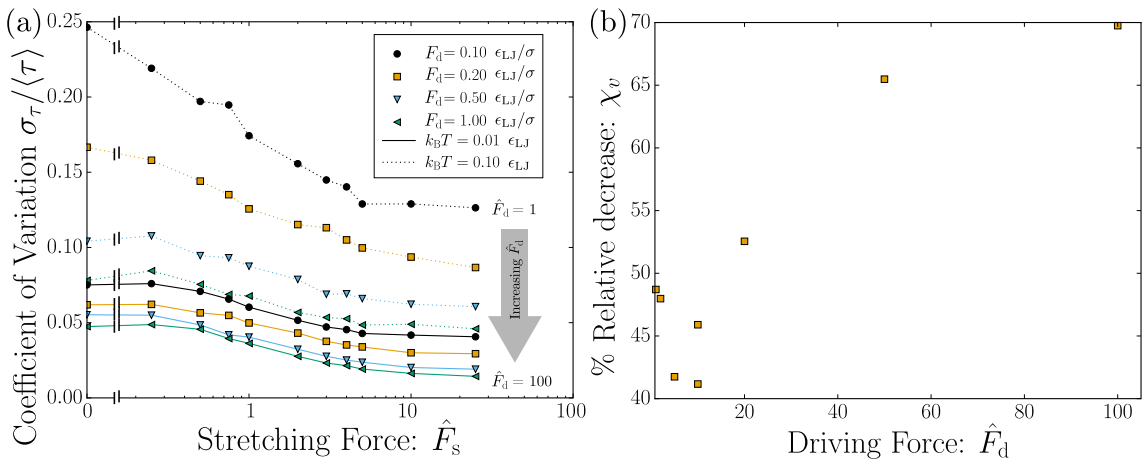


FIG. 6. (a) Coefficient of variation  $c_v = \sigma_\tau / \langle \tau \rangle$  plotted as a function of the stretching force  $\hat{F}_s$  for several values of the driving force  $\hat{F}_d$ . (b) Plot of the percentage decrease of the coefficient of variation  $c_v$  as a function of the driving force  $\hat{F}_d$ .

technologies. The measured benefits are found to strongly depend on the ratio  $F_d\sigma/k_B T$  since high amounts of diffusion can overwhelm the benefits of limiting the range of initial polymer conformations.

It is again worth noting that the current simulations do not incorporate hydrodynamic interactions. While hydrodynamic interactions could change the details of the dynamics, we do not expect the fundamental conclusions to change in that limiting the ensemble of initial conformations can significantly reduce the variations in translocation times for primarily driven processes.

Although the present implementation is difficult to realize experimentally, these results suggest that stretching the polymer prior to and during translocation via some mechanism

could be very beneficial for nanopore-based sorting and sequencing applications. Current work focuses on experimentally viable methods of prestretching DNA that preserve these beneficial effects.

## ACKNOWLEDGMENTS

Simulations were performed using the ESPResSo package [48] on the SHARCNET computer system using VMD [49] for visualization. This work was funded by NSERC (RGPIN/046434-2013) and the University of Ottawa. H.W.d.H. gratefully acknowledges funding from NSERC via the Discovery Grant 2014-06091.

- 
- [1] B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter, *Molecular Biology of the Cell*, 4th ed. (Garland Science, New York, 2002).
- [2] M. Muthukumar, *Polymer Translocation* (CRC Press, Boca Raton, FL, 2011).
- [3] A. J. Storm, C. Storm, J. Chen, H. Zandbergen, J.-F. Joanny, and C. Dekker, *Nano Lett.* **5**, 1193 (2005).
- [4] P. Chen, J. Gu, E. Brandin, Y.-R. Kim, Q. Wang, and D. Branton, *Nano Lett.* **4**, 2293 (2004).
- [5] A. Ivankin, R. Y. Henley, J. Larkin, S. Carson, M. L. Toscano, and M. Wanunu, *ACS Nano* **8**, 10774 (2014).
- [6] C. A. Merchant, K. Healy, M. Wanunu, V. Ray, N. Peterman, J. Bartel, M. D. Fischbein, K. Venta, Z. Luo, A. C. Johnson *et al.*, *Nano Lett.* **10**, 2915 (2010).
- [7] K.-H. Paik, Y. Liu, V. Tabard-Cossa, M. J. Waugh, D. E. Huber, J. Provine, R. T. Howe, R. W. Dutton, and R. W. Davis, *ACS Nano* **6**, 6767 (2012).
- [8] R. Tahvildari, E. Beamish, V. Tabard-Cossa, and M. Godin, *Lab Chip* **15**, 1407 (2015).
- [9] W. Sung and P. J. Park, *Phys. Rev. Lett.* **77**, 783 (1996).
- [10] D. Mondal and M. Muthukumar, *J. Chem. Phys.* **144**, 144901 (2016).
- [11] T. Saito and T. Sakaue, *Phys. Rev. E* **85**, 061803 (2012).
- [12] T. Saito and T. Sakaue, *Eur. Phys. J. E* **34**, 1 (2011).
- [13] T. Sakaue, *Phys. Rev. E* **76**, 021803 (2007).
- [14] T. Sakaue, *Phys. Rev. E* **81**, 041808 (2010).
- [15] T. Saito and T. Sakaue, *Phys. Rev. E* **88**, 042606 (2013).
- [16] P. Rowghanian and A. Grosberg, *J. Phys. Chem. B* **115**, 14127 (2011).
- [17] P. Rowghanian and A. Y. Grosberg, *Phys. Rev. E* **86**, 011803 (2012).
- [18] R. Adhikari and A. Bhattacharya, *Phys. Rev. E* **92**, 032711 (2015).
- [19] P. M. Suhonen and R. P. Linna, *Phys. Rev. E* **93**, 012406 (2016).
- [20] T. Menais, S. Mossa, and A. Buhot, *Sci. Rep.* **6**, 38558 (2016).
- [21] T. Menais, *Phys. Rev. E* **97**, 022501 (2018).
- [22] J. Sarabadani, B. Ghosh, S. Chaudhury, and T. Ala-Nissila, *Eur. Phys. Lett.* **120**, 38004 (2017).
- [23] J. Sarabadani, T. Ikonen, H. Mökkönen, T. Ala-Nissila, S. Carson, and M. Wanunu, *Sci. Rep.* **7**, 7423 (2017).
- [24] P. M. Suhonen, J. Piili, and R. P. Linna, *Phys. Rev. E* **96**, 062401 (2017).
- [25] V. V. Palyulin, T. Ala-Nissila, and R. Metzler, *Soft Matter* **10**, 9016 (2014).
- [26] T. Ikonen, A. Bhattacharya, T. Ala-Nissila, and W. Sung, *Phys. Rev. E* **85**, 051803 (2012).
- [27] T. Ikonen, A. Bhattacharya, T. Ala-Nissila, and W. Sung, *J. Chem. Phys.* **137**, 085101 (2012).
- [28] I. Huopaniemi, K. Luo, T. Ala-Nissila, and S.-C. Ying, *J. Chem. Phys.* **125**, 124901 (2006).
- [29] I. Huopaniemi, K. Luo, T. Ala-Nissila, and S.-C. Ying, *Phys. Rev. E* **75**, 061912 (2007).
- [30] C. Forrey and M. Muthukumar, *J. Chem. Phys.* **127**, 015102 (2007).
- [31] H. H. Katkar and M. Muthukumar, *J. Chem. Phys.* **140**, 135102 (2014).
- [32] J. M. Polson, *J. Chem. Phys.* **142**, 174903 (2015).
- [33] B. Lu, F. Albertorio, D. P. Hoogerheide, and J. A. Golovchenko, *Biophys. J.* **101**, 70 (2011).
- [34] J. Sarabadani, T. Ikonen, and T. Ala-Nissila, *J. Chem. Phys.* **141**, 214907 (2014).
- [35] D. Sean, H. W. de Haan, and G. W. Slater, *Electrophoresis* **36**, 682 (2015).
- [36] D. Sean and G. W. Slater, *J. Chem. Phys.* **146**, 054903 (2017).
- [37] U. F. Keyser, J. van der Does, C. Dekker, and N. H. Dekker, *Rev. Sci. Instrum.* **77**, 105105 (2006).
- [38] A. Sischka, A. Spiering, M. Khaksar, M. Laxa, J. König, K.-J. Dietz, and D. Anselmetti, *J. Phys.: Condens. Matter* **22**, 454121 (2010).
- [39] V. Tabard-Cossa, M. Wiggin, D. Trivedi, N. N. Jetha, J. R. Dwyer, and A. Marziali, *ACS Nano* **3**, 3009 (2009).
- [40] S. Ollila, K. Luo, T. Ala-Nissila, and S.-C. Ying, *Eur. Phys. J. E* **28**, 385 (2009).
- [41] V. V. Lehtola, R. P. Linna, and K. Kaski, *Phys. Rev. E* **78**, 061803 (2008).
- [42] G. Slater, C. Holm, M. Chubynsky, H. De Haan, A. Dubé, K. Grass, O. Hickey, C. Kingsbury, D. Sean, T. Shendruk *et al.*, *Electrophoresis* **30**, 792 (2009).
- [43] J. D. Weeks, D. Chandler, and H. C. Andersen, *J. Chem. Phys.* **54**, 5237 (1971).

- [44] G. S. Grest and K. Kremer, *Phys. Rev. A* **33**, 3628(R) (1986).
- [45] H. W. de Haan, D. Sean, and G. W. Slater, *Phys. Rev. E* **91**, 022601 (2015).
- [46] J. L. A. Dubbeldam, V. G. Rostiashvili, A. Milchev, and T. A. Vilgis, *Phys. Rev. E* **87**, 032147 (2013).
- [47] J. Dubbeldam, A. Milchev, V. Rostiashvili, and T. A. Vilgis, *Eur. Phys. Lett.* **79**, 18002 (2007).
- [48] H.-J. Limbach, A. Arnold, B. A. Mann, and C. Holm, *Comput. Phys. Commun.* **174**, 704 (2006).
- [49] W. Humphrey, A. Dalke, and K. Schulten, *J. Molec. Graphics* **14**, 33 (1996).