

## Resolution limit for DNA barcodes in the Odijk regime

Yanwei Wang,<sup>1,2</sup> Wes F. Reinhart,<sup>2</sup> Douglas R. Tree,<sup>2</sup>  
and Kevin D. Dorfman<sup>2,a)</sup>

<sup>1</sup>*Jiangsu Key Laboratory of Advanced Functional Polymer Design and Application,  
Department of Polymer Science and Engineering, College of Chemistry,  
Chemical Engineering and Materials Science, Soochow University, 199 Ren-ai Road,  
Suzhou 215123, People's Republic of China*

<sup>2</sup>*Department of Chemical Engineering and Materials Science, University of  
Minnesota—Twin Cities, 421 Washington Ave. SE, Minneapolis, Minnesota 55455, USA*

(Received 9 November 2011; accepted 6 December 2011; published online 3 January 2012)

We develop an approximation for the probability of optically resolving two fluorescent labels on the backbone of a DNA molecule confined in a nanochannel in the Odijk regime as a function of the fluorescence wavelength, channel size, and the properties of the DNA (persistence length and effective width). The theoretical predictions agree well with equivalent data produced by Monte Carlo simulations of a touching wormlike bead model of DNA in a high ionic strength buffer. Although the theory is only strictly valid in the limit where the effective width of the nanochannel is small compared with the persistence length of the DNA, simulations indicate that the theoretical predictions are reasonably accurate for channel widths up to two-thirds of the persistence length. Our results quantify the conjecture that DNA barcoding has kilobase pair resolution—provided the nanochannel lies in the Odijk regime. © 2012 American Institute of Physics. [doi:10.1063/1.3672691]

### I. INTRODUCTION

DNA barcoding is emerging as a key tool for high-throughput, single-molecule analysis of the genome at the kilobase pair level, representing an important complement to DNA sequencing.<sup>1</sup> As illustrated in Fig. 1(a), specific sequences (the DNA barcodes) are fluorescently labeled by either binding the probes to the chain<sup>2–5</sup> or inserting labeled nucleotides via nick extension.<sup>6,7</sup> The backbone is fluorescently labeled with a second color through an intercalating dye such as YOYO-1. To determine the genomic distance between barcode markers, the DNA needs to be stretched from its equilibrium, coiled conformation. Several competing technologies have arisen to accomplish this task: using a receding contact line and subsequent binding to the surface (molecular combing),<sup>8–10</sup> extensional flow (direct linear analysis),<sup>11–13</sup> and nanochannel confinement.<sup>5–7,14,15</sup> The genomic distance between neighboring barcodes is then determined by first locating the centers of the sequence-specific labels and then integrating the total backbone fluorescence intensity between these two locations. Nanochannels provide a particularly attractive approach to make such measurements since a confined chain fluctuates about its equilibrium extension; making multiple, statistically independent measurements of the distance reduces the sampling error.<sup>5,16</sup>

A fundamental question related to DNA barcoding in nanochannels is the ability to resolve two nearby barcodes on the DNA backbone. If we consider a sequence-specific probe that emits light with wavelength  $\lambda = 573$  nm (Ref. 14) and a completely extended DNA molecule with a rise of 0.34 nm per base pair (bp), then two probes would have to be separated by at least 843 bp if we take  $\lambda/2$  as an estimate of the resolution limit for diffraction limited optics. However, as we can see from simulation data in Fig. 1(b), the semi-flexible nature of the DNA leads to incomplete stretching, even in very small channels lying in the Odijk regime.<sup>21</sup> In this paper,

<sup>a)</sup>Electronic mail: dorfman@umn.edu.

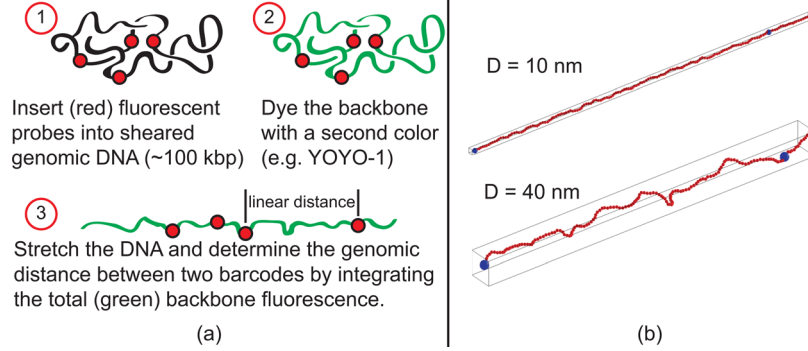


FIG. 1. (a) Principle of DNA barcoding. If the linear distance in the focal plane between barcodes does not exceed the diffraction limit,  $\lambda/2$ , for light of wavelength  $\lambda$ , then the two fluorophores cannot be resolved. (b) Illustration of the chain extension in different regimes of confinement. The configurations were generated using the simulation method in Sec. III with a chain with an approximate size of 27.7 kilobase pair (kbp). Only part of the chain is shown in the figure. The persistence length is  $l_p = 53$  nm (Ref. 17) and the effective width<sup>18</sup> is  $w = 4.6$  nm, appropriate for TBE 5x buffer.<sup>19,20</sup> The two blue spheres in each snapshot are shown to illustrate barcodes with one of the barcodes located at the end of the chain and the other barcode at 2 kbp.

we develop an approximation for the probability of resolving two nearby barcodes as a function of the nanochannel size and the properties of the DNA (persistence length and effective width), which themselves depend intimately on the ionic strength of the buffer.<sup>18,22–24</sup> Our results apply to the Odijk regime, which is the most desirable regime for DNA barcoding.<sup>6,15</sup>

## II. THEORY

The equilibrium two-point distribution function  $\Psi_{ij}(r_{ij})$  quantifies the probability density for segment  $i$  and  $j$ , corresponding to a genomic distance  $|j - i|$ , being separated by a linear distance  $r_{ij} = |\vec{r}_i - \vec{r}_j|$  in the focal plane of the image. In our theoretical analysis, we envision the “segments” to be base pairs of the chain, whereupon the contour length per segment is given by the rise of DNA,  $r$ . In our simulations, we will consider a slightly coarser model where the characteristic length scale is the effective width of the chain,  $w$ . Owing to the symmetry of the square channel, we only need to assume that focal plane is parallel to one of the walls of the channel but we do not need to specify which wall.<sup>25</sup> While  $\Psi_{ij}$  is known for a confined Gaussian chain,<sup>26</sup> there are no results for more complicated cases,<sup>27</sup> such as the wormlike chain in a good solvent we need here. However, we do not need to compute this distribution function to determine the probability of resolving two barcodes. Rather, we can define the probability

$$\psi(j|i) = \int_{\lambda/2}^{\infty} \Psi_{ij}(r_{ij}) dr_{ij} \quad (1)$$

that, given a barcode marker at some point  $i$  along the chain, we are able to resolve a second marker at some different point  $j$ . The average probability of resolving a pair of barcode markers on a chain of  $N$  base pairs separated by  $n = |j - i|$  base pairs is then

$$p(n) = \frac{1}{N - n} \sum_{i=1}^{N-n} \psi(i + n|i). \quad (2)$$

We are interested in computing  $p(n)$  for the Odijk regime,<sup>21</sup> which applies to nanochannel widths much smaller than the chain persistence length,  $l_p$ . The width of the channel accessible to the centerline of the DNA backbone is  $D_{\text{eff}} = D - w$  in the limiting case of steric interactions only, where  $w$  is the effective width of the DNA.<sup>18</sup> While the inequality  $D_{\text{eff}} \ll l_p$  may appear to be so restrictive as to exclude all reasonable nanochannel sizes that allow easy insertion of

DNA,<sup>28</sup> we will see that our final result is a reasonable approximation out to  $D_{\text{eff}} \cong l_p$ . In the Odijk regime, a chain of total length  $L$  consists of a series of deflection segments with the mean span

$$\langle R_{\parallel} \rangle = L \left[ 1 - 2\alpha (D_{\text{eff}}/l_p)^{2/3} \right], \quad (3)$$

where  $\alpha = 0.09137 \pm 0.00007$ .<sup>25</sup> The variance of the extension in this regime is

$$\langle R_{\parallel}^2 \rangle - \langle R_{\parallel} \rangle^2 = 2\beta L D_{\text{eff}}^2 l_p^{-1}, \quad (4)$$

where  $\beta = 0.00478 \pm 0.00010$ .<sup>25</sup>

We can construct an approximate model for  $p(n)$  in the Odijk regime through the following ansatz. We assume the chain is long compared with the distance  $n$  between barcodes and that the barcodes are not located too close to the end of the chain, whereupon  $\Psi_{ij} \approx \Psi_n$  and  $\psi(i+n|i) \approx p(n)$ . If we further neglect the contribution of lateral fluctuations between the two barcode points,  $\Psi_{ij}(r_{ij}) \approx \Psi_n(x)$ , where  $x$  is the distance down the channel axis and  $\Psi_n(x)$  is interpreted as the distribution function for the distance in the  $x$ -direction between segments on the chain separated by a genomic length  $n$ . Equation (1) then reduces to

$$p(n) \approx \int_{\lambda/2}^{\infty} \Psi_n(x) dx. \quad (5)$$

In a high ionic strength buffer, the diffraction limit is long compared with the size of a deflection segment. Neglecting prefactors of order unity, but retaining the definition of the diffraction limit, this corresponds to<sup>21</sup>

$$\lambda/2 \gg D_{\text{eff}}^{2/3} l_p^{1/3}. \quad (6)$$

Provided we satisfy Eq. (6), then the integral in Eq. (5) always involves many deflection segments. It is thus reasonable to assume that the average distance,  $\bar{x}$ , and variance,  $\sigma^2$ , between the two barcodes separated by a contour length  $L = r|i-j| > \lambda/2$ , where  $r$  is the rise of double stranded DNA, are given by Eqs. (3) and (4), respectively. If we further assume a normal distribution for the distance between these barcodes,

$$\Psi_n(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left[ -\frac{(x-\bar{x})^2}{2\sigma^2} \right], \quad (7)$$

then Eq. (5) leads to

$$p(n) = \frac{1}{2} \text{erfc} \left( \frac{\lambda - 2\bar{x}}{2^{3/2}\sigma} \right), \quad (8)$$

where  $\text{erfc}(x) = (2/\pi^{1/2}) \int_x^{\infty} \exp(-t^2) dt$  is the complementary error function. Equation (8) is the key result of our analysis.

### III. SIMULATION METHOD

The logic leading to Eq. (8) involves a number of assumptions. We tested these assumptions by using Monte Carlo simulations to directly compute  $p(n)$  from Eq. (2) in square nanochannels of width  $D$ . The ‘‘touching bead’’ simulation model<sup>29</sup> we used here, illustrated in Fig. 2, is a modification of the model we used previously<sup>20</sup> to study the different regimes of confined DNA. The chain consists of  $N = 2048$  beads of size  $w = 4.6$  nm. All beads interact by a hardcore excluded volume

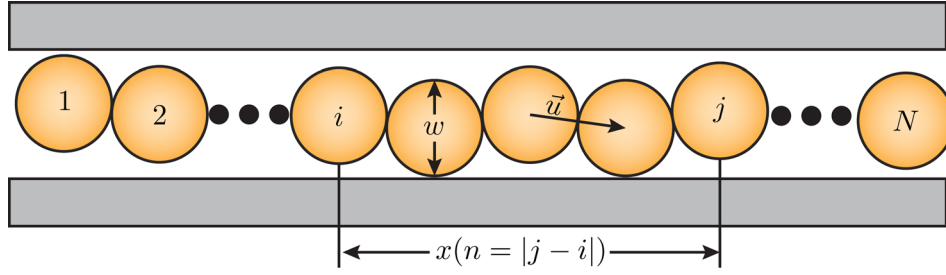


FIG. 2. Schematic illustration of the touching bead model. The “segments” of the chain, numbered from  $k=1$  to  $k=N$ , are spherical beads of size  $w$  that experience hardcore bead-bead and bead-wall interactions. The bending energy is computed from the discrete wormlike chain model in Eq. (9) using the vectors  $\vec{u}$  pointing between bead centers. For calculations of  $p(n)$ , we compute the distance  $x$  between each pair  $(i, j)$  in the direction of the channel axis.

with one another and with the walls, whereupon we identify  $w$  as both the coarse-graining size (in the direction along the backbone) and the effective width of the DNA chain (in the direction perpendicular to the backbone). The value of 4.6 nm is the Stigter effective width<sup>18,30</sup> for DNA in TBE 5x buffer.<sup>19,20</sup> To provide a connection with experiments, we will ultimately convert the simulation data (in terms of the beads) into the experimentally relevant genomic distances using the conversion factor  $w/r$ . We take  $r=0.34$  nm/bp as the rise of DNA. Since intercalating dyes can introduce an increase in the rise of DNA,<sup>31</sup> this conversion should be considered approximate and our results for the resolution limit are likely conservative estimates. Neighboring beads also experience a bending potential given by the discretized wormlike chain model<sup>32</sup>

$$U_{\text{bend}} \approx k_B T (l_p/w) \sum_{k=1}^{N-2} (1 - \vec{u}_k \cdot \vec{u}_{k+1}), \quad (9)$$

where  $l_p=53$  nm is the persistence length<sup>17</sup> and  $\vec{u}_j = (\vec{r}_{j+1} - \vec{r}_j)/l$ ,  $j \in [1, N-1]$  is the unit vector between beads  $j$  and  $j+1$ .

Chain configurations were generated using reptation and crankshaft moves in the Metropolis scheme. Simulations for each channel size were performed in 12 replicates. The simulations for a channel of size  $D$  were initialized using a configuration generated from preceding simulations of a slightly smaller channel to guarantee acceptance of the initial configuration. Upon increasing the channel size, we first waited for the mean extension to stabilize around the average value in the larger channel, and then added an additional  $2.048 \times 10^7$  equilibration steps before starting the production run. This approach introduces a slight bias against hairpin configurations, which are difficult to form in the smallest channels. However, the long equilibration time and the small change in channel width between simulations minimizes this bias.

To determine an appropriate sampling time for computing the chain statistics, we calculated the auto-covariance of the extension,<sup>33</sup>

$$C_{R_{\parallel}}(n) = \frac{\langle R_{\parallel}(m)R_{\parallel}(m+n) \rangle_m - \langle R_{\parallel} \rangle^2}{\langle R_{\parallel}^2 \rangle - \langle R_{\parallel} \rangle^2}. \quad (10)$$

Figure 3 shows the decay of the auto-covariance with the number of Monte Carlo steps. On average, the auto-covariance is sufficiently attenuated after approximately  $2.048 \times 10^8$  trial moves. In what follows, we used sampling rates of  $4.096 \times 10^8$  trial moves for the 10 nm and 20 nm channels and  $2.048 \times 10^8$  trial moves for the 30 nm and 40 nm channels, which produces 10 statistically independent configurations per replicate.

In our prior work,<sup>20</sup> we used a bead-rod model with a Weeks-Chandler Anderson (WCA) potential<sup>34</sup> to provide bead-bead excluded volume. With our code for this model, we were able to reach a long chain (2048 beads, 9.42  $\mu\text{m}$  contour length, 27.7 kbp) without sacrificing the spatial resolution. In free solution, this model produces a root-mean-square end-to-end distance

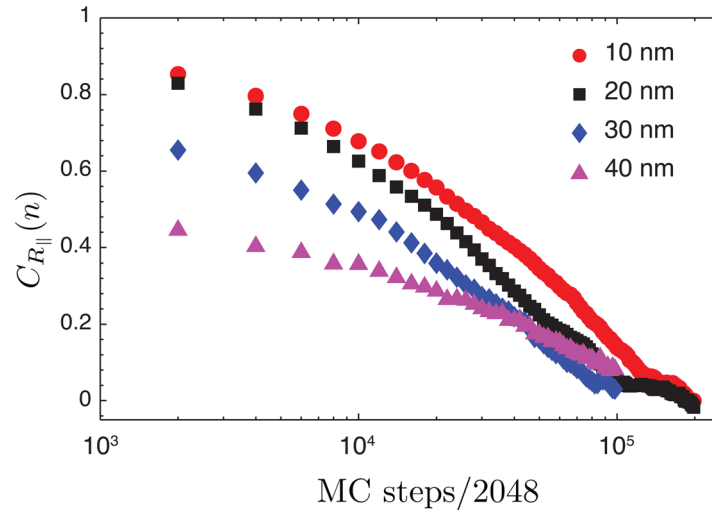


FIG. 3. Auto-covariance of the chain extension as a function of the number of Monte Carlo steps normalized with the size of the chain. Symbols are averaged results over the 12 independent runs.

of  $1.06 \mu\text{m}$ , a root-mean-square radius of gyration of  $0.43 \mu\text{m}$ , and a mean span dimension of  $0.88 \mu\text{m}$ . All of these values agree well with our prior simulations,<sup>20</sup> which themselves match experimental results. Likewise, the results for the fractional extension of this chain exhibit the same scaling laws as simulations of our prior model.<sup>20</sup>

#### IV. RESULTS AND DISCUSSION

In Eq. (7), we assumed that the distance between barcodes is normally distributed, with the mean and variance given by those for the mean span of a chain in the Odijk regime.<sup>21,25</sup> In Fig. 4, we compare this assumption to the simulation data for two barcodes separated by 1001 base pairs. Using a rise of  $0.34 \text{ nm/bp}$  and a persistence length of  $53 \text{ nm}$ , these data correspond to barcodes separated by 6.4 persistence lengths. The normal distribution appears to be an excellent model for the simulation data in Fig. 4. Nevertheless, there is reason to be cautious about assuming a normal distribution in Eq. (7), since it permits unphysical stretching of the chain.

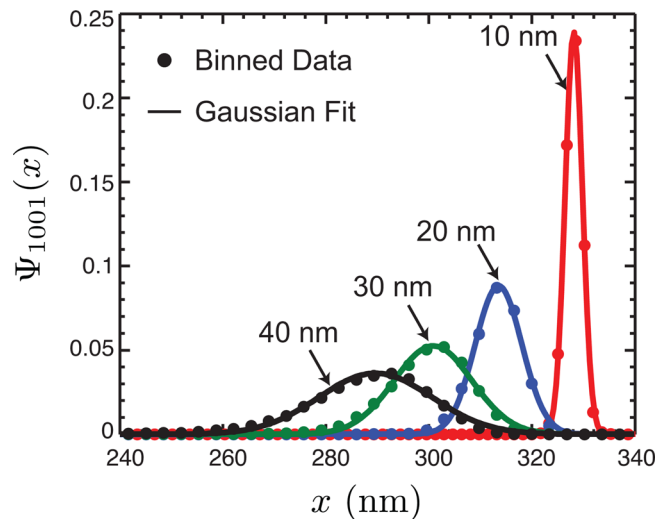


FIG. 4. Comparison between Monte Carlo simulation data for the probability distribution  $\Psi_n$  and the assumed form in Eq. (7) for two barcodes separated by  $n = 1001$  base pairs for  $\lambda = 573 \text{ nm}$ ,  $l_p = 53 \text{ nm}$ , and  $w = 4.6 \text{ nm}$ . The different curves correspond to Eq. (7) for channel sizes  $D$  between 10 and 40 nm, and the symbols are the values of the binned data.

Explicitly, the normal distribution is supported over  $x \in (-\infty, +\infty)$ , so there is a finite probability that two barcodes separated by a contour length  $l$  will be separated by a linear distance  $x > l$ . For  $n = 1001$  bp, the maximum distance between barcodes is approximately 340 nm. Fortunately, for all of the cases plotted in Fig. 4, the relatively small variance in the chain extension given by Eq. (4) leads to the normal distribution being strongly attenuated near  $x = l$ . As a result,  $\Psi_{ij}(x > l)$  makes a negligible contribution to the integral in Eq. (5). The normal distribution also permits a nonsensical probability that  $x < 0$ . Since the probability distribution decays quickly enough at the upper bound  $x = l$  and the mean of the distribution is closer to  $x = l$  than  $x = 0$ , then the probability distribution is exponentially small at the other physical limit,  $x = 0$ . While the data in Fig. 4 give us confidence in the validity of using a normal distribution, we will return to this issue after evaluating the corresponding theoretical predictions for  $p(n)$ .

We thus turn our attention to the second key assumption of our theory, namely, that the distance between barcodes on the interior of the chain obeys the statistics for a long chain in the Odijk regime. Naturally, this can only be the case if the chain as a whole, which contains a large number of deflection segments, also obeys the statistics for the Odijk regime. As we can see in Fig. 5, the prediction of Eq. (3), which contains no adjustable parameters,<sup>25</sup> is an excellent fit for our smallest channel. As the effective channel size increases, there is a systematic deviation between the predictions for the Odijk regime and the simulation data. The origin of the deviation is apparent in Fig. 1(b), where the chain is beginning to form hairpins.<sup>19</sup> However, the deviation from the theoretical prediction is quite small, ranging from 0.1% to 3.2% over the range  $D = 10$  nm to  $D = 40$  nm.

Figure 6 compares the predictions of Eq. (8) to our simulation data. Since we already confirmed in Fig. 4 that the probability distribution  $p(n)$  is reasonably Gaussian for all of these channel sizes, we can conclude that the deviation between the theory and simulation in Fig. 6 is due to a gradual breakdown in the Odijk regime (Fig. 5) as the chain begins to form hairpins (Fig. 1). Indeed, the agreement between the theoretical prediction and the simulation data for  $p(n)$  in Fig. 6 mirrors the phenomenon we observed in Fig. 5. Nevertheless, the approximate theoretical result given by Eq. (8) appears to be a quite good description for all of these channels. Although the numerical results for the mean span and its variations in Eqs. (3) and (4) are only valid in the limit  $D_{\text{eff}} \ll l_p$ , it seems that they provide a reasonable approximation for the configuration of an internal segment of the chain up to  $D_{\text{eff}} \approx 2l_p/3$ .

We also investigated the effect of relaxing the assumption of a normal distribution in Eq. (7) by considering other ‘‘Gaussian-like’’ distributions. A natural alternative is the inverse Gaussian distribution (also called the Wald distribution).<sup>35</sup> In the large  $L$  limit, the inverse

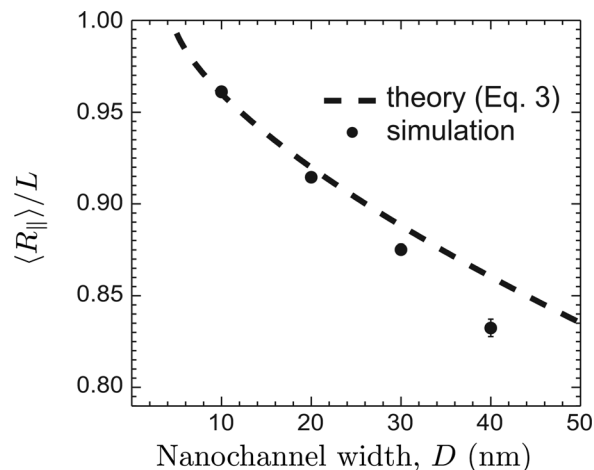


FIG. 5. Comparison between the extension of the chain and the predictions from the Odijk regime with no adjustable parameters given by Eq. (3). The simulations correspond to 2048 touching wormlike beads with  $l_p = 53$  nm and  $w = 4.6$  nm. The channel width accessible to the chain is  $D_{\text{eff}} = D - w$ . The error bars correspond to the standard deviation over the 12 independent replicas.

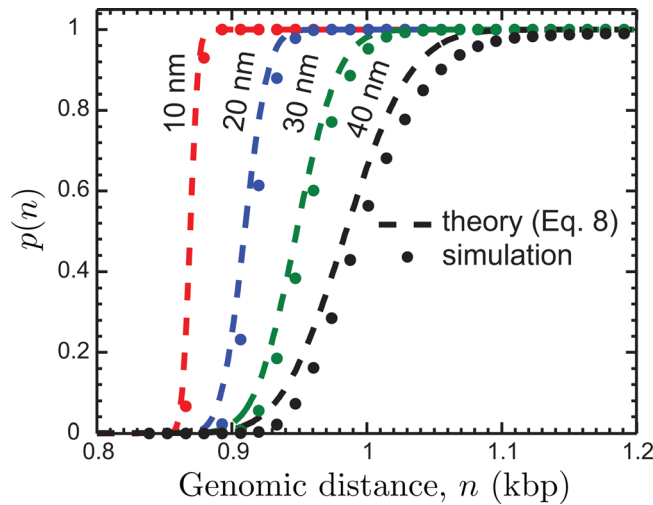


FIG. 6. Comparison between Monte Carlo simulation data for the probability of resolving two barcodes separated by a genomic distance  $n$ ,  $p(n)$ , and the prediction from Eq. (8) for  $\lambda = 573$  nm,  $l_p = 53$  nm and  $w = 4.6$  nm. The different curves correspond to Eq. (8) for channel sizes  $D$  between 10 and 40 nm, and the symbols are the values of the simulation data.

Gaussian distribution limits to the normal distribution, making it an appropriate choice for proximate barcodes separated by a handful of deflection segments. We also considered distributions that would remove the unphysical features of the normal distribution. The log-normal distribution is supported on  $x \in (0, \infty)$ , removing the possibility for a “negative” extension, and the raised cosine distribution has a finite support, which further removes the possibility of overextension. As we can see in Fig. 7, the results for the resolution limit in a 40 nm channel, which has the broadest distribution for  $\Psi_m$ , are effectively unchanged by the choice of any of these “Gaussian-like” probability distributions. The results in Fig. 7 support our choice of the normal distribution to describe the probability distribution even though its limits are not physical; the normal distribution gives essentially the same results for  $p(n)$  as a more realistic model, such as the raised cosine distribution, while permitting the simple expression for  $p(n)$  given by Eq. (8).

All of our simulation data and discussion thus far focused on the configurations of confined DNA in a high ionic strength buffer, where DNA-DNA and DNA-wall electrostatic interactions

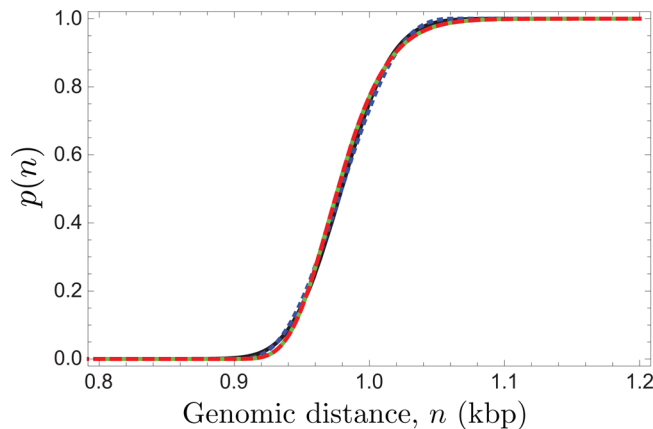


FIG. 7. Comparison of the predictions for  $p(n)$  in the 40 nm channel corresponding to different probability distributions for  $\Psi_n(x)$ , using the statistics for a chain in the Odijk regime given by Eqs. (3) and (4). The various curves correspond to the normal distribution for  $R_{\parallel}$  (black, solid), a raised cosine distribution for  $R_{\parallel}$  (blue, dots), a log-normal distribution for  $L - R_{\parallel}$  (red, dashed) and an inverse Gaussian distribution for  $L - R_{\parallel}$  (green, solid).

are screened. In this case, the inequality given by Eq. (6) holds and the diffraction limit corresponds to many deflection segments. As the ionic strength of the solution is decreased, the electrostatic interactions become sensible and lead to an increase in both the persistence length of the DNA<sup>22–24</sup> and the effective width.<sup>18</sup> Schwartz and coworkers<sup>6,15</sup> have advocated the use of low ionic strength buffers to permit DNA barcoding in relatively wide nanochannels (nanoslits). By increasing the persistence length, these nanochannels still satisfy the criteria for the Odijk regime,  $D_{\text{eff}} \ll l_p$ , especially when we recall that  $D_{\text{eff}} = D - w$  and the effective width  $w$  also increases as the ionic strength decreases. We can make a crude estimate for the bounds of our theory by assuming that these systems operate in a regime  $D_{\text{eff}} \approx l_p$ , corresponding to the edge of the Odijk regime in Fig. 5. In such circumstances, our theory would only be valid for  $\lambda/2 \gg l_p$ , corresponding to ionic strengths down to around 0.1–1 mM.<sup>36</sup> This value is similar to the most recent low ionic strength experiments aimed towards DNA barcoding,<sup>15</sup> so our results represent a conservative estimate for the maximum barcode density for these experiments. However, we should use caution in applying the results for square nanochannels to nanoslits.<sup>37</sup>

## V. CONCLUDING REMARKS

In the present contribution, we have developed an approximation for the probability of resolving two nearby barcodes on a DNA molecule confined in a nanochannel as a function of the channel size, the wavelength of the fluorescence, and the properties of the DNA (persistence length and effective width). Although the final result in Eq. (8) required assuming that (i) the linear distance between barcodes is normally distributed and (ii) that the statistics for these short, internal sections of the chain are equivalent to an infinitely long chain, we found that the theory leads to a reasonable description of the simulation data. While we chose to use a normal distribution for algebraic simplicity, the results for the probability of resolving two barcodes are essentially unchanged when we use more realistic probability distributions (and the concomitantly more difficult algebra.) Our results provide a theoretical basis for claims that the resolution of DNA barcoding in nanochannels is around 1 kbp.<sup>16,38</sup>

Our conclusions are based exclusively on the equilibrium chain statistics. Such an analysis tacitly assumes that we can actually measure the instantaneous chain configuration or, in the case of barcoding, the instantaneous location of the two barcodes. While the fluorophores used for DNA barcodes are bright enough to image before they photobleach, the rate of photon emission relative to the fluctuations in the chain extension is probably low. As a result, the photons need to be collected over a period of time to increase the signal-to-noise ratio to an acceptable level. The imaged fluorophore “position” is then a smeared value of the instantaneous positions corresponding to each photon emission. Our results likely correspond to a theoretical lower bound on the minimum distance between fluorophores.

While we had success in determining the resolution limits in the Odijk regime, we have less confidence that this approach can be extended easily into other regimes of confinement.<sup>20,37,39,40</sup> As the channel size increases, the size of the statistical segment increases as well. For example, in the so-called “extended de Gennes” regime,<sup>20</sup> which is more representative of the confinement regime in recent DNA barcoding experiments,<sup>5,7</sup> the chain consists of anisometric compression blobs with a characteristic size  $H \cong (D_{\text{eff}} l_p)^{2/3} w^{-1/3}$ .<sup>39</sup> For a 250 nm nanochannel with  $l_p = 53$  nm and  $w = 4.6$  nm, the blob size (neglecting prefactors of order unity) is commensurate with the diffraction limit and it is no longer appropriate to adopt a resolution model that requires an assumption similar to Eq. (6). While we can certainly compute the probability of resolving two nearby barcodes using Monte Carlo simulations of the type employed here, developing the corresponding theoretical model remains challenging.

## ACKNOWLEDGMENTS

This work was supported by the NIH (R01-HG005216 and R21-RR031230) and was carried out in part using computing resources at the University of Minnesota Supercomputing Institute.



- <sup>1</sup>T. P. Niedringhaus, D. Milanova, M. B. Kerby, M. P. Snyder, and A. E. Barron, *Anal. Chem.* **83**, 4327 (2011).
- <sup>2</sup>Y. Ebinstein, N. Gassman, S. Kim, J. Antelman, Y. Kim, S. Ho, R. Samuel, X. Michalet, and S. Weiss, *Nano Lett.* **9**, 1598 (2009).
- <sup>3</sup>R. K. Neely, P. Dedecker, J. Hotta, G. Urbanaviciute, S. Klimasauskas, and J. Hofkens, *Chem. Sci.* **1**, 453 (2010).
- <sup>4</sup>B. R. Cipriani, R. Zhao, P. J. Murphy, S. L. Levy, C. P. Tan, H. G. Craighead, and P. D. Soloway, *Anal. Chem.* **82**, 2480 (2010).
- <sup>5</sup>S. F. Lim, A. Karpusenko, J. J. Sakon, J. A. Hook, T. A. Lamar, and R. Riehn, *Biomicrofluidics* **5**, 034106 (2011).
- <sup>6</sup>K. Jo, D. M. Dhingra, T. Odijk, J. J. de Pablo, M. D. Graham, R. Runnheim, D. Forrest, and D. C. Schwartz, *Proc. Natl. Acad. Sci. USA* **104**, 2673 (2007).
- <sup>7</sup>S. K. Das, M. D. Austin, M. C. Akana, P. Deshpande, H. Cao, and M. Xiao, *Nucleic Acids Res.* **38**, e177 (2010).
- <sup>8</sup>D. C. Schwartz, X. Li, L. I. Hernandez, S. P. Ramnarain, E. J. Huff, and Y.-K. Wang, *Science* **262**, 110 (1993).
- <sup>9</sup>A. Bensimon, A. Simon, A. Chiffaudel, V. Croquette, F. Heslot, and D. Bensimon, *Science* **265**, 2096 (1994).
- <sup>10</sup>E. T. Dimalanta, A. Lim, R. Runnheim, C. Lamers, C. Churas, D. K. Forrest, J. J. de Pablo, M. D. Graham, S. N. Copper-smith, S. Goldstein, and D. C. Schwartz, *Anal. Chem.* **76**, 5293 (2004).
- <sup>11</sup>E. Y. Chan, N. M. Goncalves, R. A. Haeusler, A. J. Hatch, J. W. Larson, A. M. Maletta, G. R. Yantz, E. D. Carstea, M. Fuchs, G. G. Wong, S. R. Gullans, and R. Gilmanshin, *Genome Res.* **14**, 1137 (2004).
- <sup>12</sup>R. Dylla-Spears, J. E. Townsend, L. Jen-Jacobson, L. L. Sohn, and S. J. Muller, *Lab Chip* **10**, 1543 (2010).
- <sup>13</sup>R. H. Meltzer, J. R. Krogmeier, L. W. Kwok, R. Allen, B. Crane, J. W. Griffis, L. Knaian, N. Kojanian, G. Malkin, M. K. Nahas, V. Papkov, S. Shaikh, K. Vyavahare, Q. Zhong, Y. Zhou, J. W. Larson, and R. Gilmanshin, *Lab Chip* **11**, 863 (2011).
- <sup>14</sup>T. Su, S. K. Das, M. Xiao, and P. K. Purohit, *PLoS ONE* **6**, e16890 (2011).
- <sup>15</sup>Y. Kim, K. S. Kim, K. L. Kounovsky, R. Chang, G. Y. Jung, J. J. de Pablo, K. Jo, and D. C. Schwartz, *Lab Chip* **11**, 1721 (2011).
- <sup>16</sup>R. Riehn, M. Lu, Y.-M. Wang, S. F. Lim, E. C. Cox, and R. H. Austin, *Proc. Natl. Acad. Sci. USA* **102**, 10012 (2005).
- <sup>17</sup>C. Bustamante, J. F. Marko, E. D. Siggia, and S. Smith, *Science* **265**, 1599 (1994).
- <sup>18</sup>D. Stigter, *Biopolymers* **16**, 1435 (1977).
- <sup>19</sup>T. Odijk, *J. Chem. Phys.* **125**, 204904 (2006).
- <sup>20</sup>Y. Wang, D. R. Tree, and K. D. Dorfman, *Macromolecules* **44**, 6594 (2011).
- <sup>21</sup>T. Odijk, *Macromolecules* **16**, 1340 (1983).
- <sup>22</sup>T. Odijk, *J. Polym. Sci., Part B: Polym. Phys.* **15**, 477 (1977).
- <sup>23</sup>J. Skolnick and M. Fixman, *Macromolecules* **10**, 944 (1977).
- <sup>24</sup>A. V. Dobrynin, *Macromolecules* **38**, 9304 (2005).
- <sup>25</sup>T. W. Burkhardt, Y. Yang, and G. Gompper, *Phys. Rev. E* **82**, 041801 (2010).
- <sup>26</sup>M. Doi and S. F. Edwards, *The Theory of Polymer Dynamics* (Oxford University Press, Oxford, U.K., 1986).
- <sup>27</sup>J. L. Harden and M. Doi, *J. Phys. Chem.* **96**, 4046 (1992).
- <sup>28</sup>H. Cao, Z. N. Yu, J. Wang, J. O. Tegenfeldt, R. H. Austin, E. Chen, W. Wu, and S. Y. Chou, *Appl. Phys. Lett.* **81**, 174 (2002).
- <sup>29</sup>P. J. Hagerman and B. H. Zimm, *Biopolymers* **20**, 1481 (1981).
- <sup>30</sup>D. Stigter, *J. Colloid Interface Sci.* **53**, 296 (1975).
- <sup>31</sup>T. T. Perkins, D. E. Smith, R. G. Larson, and S. Chu, *Science* **268**, 83 (1995).
- <sup>32</sup>J. Wang and H. Gao, *J. Chem. Phys.* **123**, 084906 (2005).
- <sup>33</sup>W. Reisner, K. J. Morton, R. Riehn, Y. M. Wang, Z. Yu, M. Rosen, J. C. Sturm, S. Y. Chou, E. Frey, and R. H. Austin, *Phys. Rev. Lett.* **94**, 196101 (2005).
- <sup>34</sup>J. D. Weeks, D. Chandler, and H. C. Andersen, *J. Chem. Phys.* **54**, 5237 (1971).
- <sup>35</sup>In addition to computing the values of  $\alpha$  and  $\beta$  appearing in Eqs. (3) and (4), Burkhardt et al.<sup>25</sup> showed that an analytical result for the chain extension (in the form of an inverse Gaussian distribution) can be derived by replacing the hardcore repulsion by the walls with a softer, parabolic potential.
- <sup>36</sup>C. C. Hsieh, A. G. Balducci, and P. S. Doyle, *Nano Lett.* **8**, 1683 (2008).
- <sup>37</sup>T. Odijk, *Phys. Rev. E* **77**, 060901(R) (2008).
- <sup>38</sup>W. Reisner, N. B. Larsen, A. Silahtaroglu, A. Kristensen, N. Tommerup, J. O. Tegenfeldt, and H. Flyvbjerg, *Proc. Natl. Acad. Sci. USA* **107**, 13294 (2010).
- <sup>39</sup>F. Brochard-Wyart, T. Tanaka, N. Borghi, and P. G. de Gennes, *Langmuir* **21**, 4144 (2005).
- <sup>40</sup>P. Cifra, *J. Chem. Phys.* **131**, 224903 (2009).