Progress towards solving the vanilla loop extrusion model through a low-density expansion

Henrik Dahl Pinholt
(Dated: May 18, 2023)

The compaction of DNA by active loop extrusion has been found to explain several key features of genome folding. While simple to state, the vanilla loop extrusion model has proven rather difficult to analyze analytically and most current progress has been done with simulations. Here I extend the current models with a Fokker-Planck based formulation which allows for a low-density expansion with a closed form solution. This formulation may provide a novel basis for investigation of several phenomena of relevance to gene regulation and chromosome folding in the future.

I. INTRODUCTION

Nuclear DNA folds into loops with a distribution that is determined by the steady state ensemble of loop extruding factors which bind and unbind to the polymer [1]. Simulations using the interplay of extrusion barriers and motor-motor collisions have provided explanations for both the kilobase loop structure for chromosomes in interphase [2] and the megabase loops for mitotic chromosomes during cell division [3, 4].

Despite its predictive power in simulations, theoretical studies of the steady state ensemble has remained limited. A related process called the totally asymmetric simple exclusion processes (TASEP) has been studied extensively [5]. However, a key difference is that the dispersion in loop extrusion seems to be primarily driven by unbinding and rebinding rather than dispersion in hopping events as is the case for the TASEP [6]. To my knowledge, the TASEP with binding and rebinding has not been investigated.

Direct study of loop extrusion with binding and rebinding has so far relied on simple approximations to investigate the emerging polymer structure. For example, quenched polymer averaging has lead to a contact probability distribution and mean square separation in models with non-interacting loop extruders and no barriers or assumptions of complete jamming and no barriers [6–8]. Similarly, studying a single motor lead to derivable statistics even for the non-equillibrium polymer dynamics regime [9].

Most work in the field seem to rely on the vanilla LEF model [10]. This one dimensional model posits that LEFs extrude bidirectionally and deterministically, with two heads acting independently to reel in DNA (fig. 1A). This means that the dispersion ariseses solely from unbinding and collisions. Unbinding occurs after an exponentially distributed time where both heads dissociate from the DNA into the solvent (fig. 1B). Upon hitting a barrier—or another extruder—the head which collided stalls while the other head keeps extruding (fig. 1C). Further extensions include asymmetric extrusion [11]—where only one head is extruding at a given moment— and bypassing/moving barriers [12].

While conceptually simple, the vanilla LEF model has provided surprisingly difficult to solve. Past analytical studies have relied on combinations of dilute/dense regime limits to arrive at simple emergent properties of the system. In the dense regime, Goloborodko et al. showed that the dynamics reduce to the stochastic growth, splitting, and decay of loop anchors and were able to arrive at range of interesting results [10]. Under the simplifying assumption of no nested loops—with no assumptions on the density—Banigan and Mirny were able to arrive at a simple first-order kinetic approximation for the number of extruders in loops, child loops, and unbound LEFs to estimate compaction fractions. Assuming no LEF-LEF collisions, exact results for the average position of extruder heads has been derived for the vanilla LEF model on a 1D lattice [13] and for a diffusive-type LEF on a continous domain [14, 15].

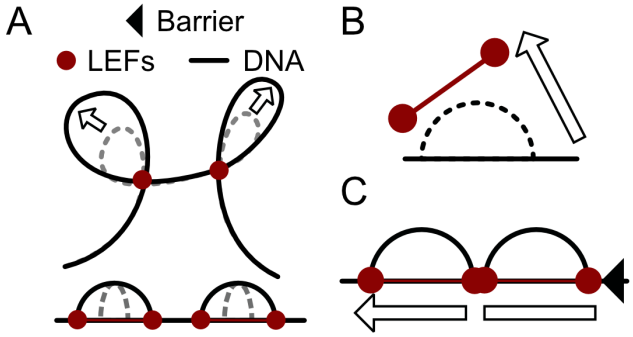

FIG. 1. Graphical depiction of the vanilla loop extrusion model and its three assumptions. A: Diagram of extruding factors (LEFs) bound to DNA which processively extrude loops. The bottom row depicts the loops in an arch-diagram where each arch represents the two heads of a single extruder. The first assumption is that, upon binding, loop extruding factors (LEFs) extrude deterministically at a fixed rate. Dispersion therefore arises soley from binding/unbinding and collisions. B: The second assumption of the model is that unbinding occurs after an exponentially distributed time. C: The third assumption is that collisions with other LEF heads or barriers stalls the heads independently.

Here I formulate a novel perspective on the vanilla LEF model based on a Fokker-Planck formalism. Instead of modelling the whole system of extruders, I focus on calculating the position of a single head relative to the loading site. Collisions are included to lowest order in the

barrier density. This reduced description allows for the derivation of the full time-dependent distribution of head positions, providing a novel approach to calculate various single-LEF and multi-LEF observables with collisions included.

II. FOKKER-PLANCK BASED FORMULATION OF THE VANILLA LOOP EXTRUSION MODEL

We seek to calculate the density $p(x,t)$ of heads which make it to position x at time t conditioned on them loading at time $t=0$ at position $x=0$. The extruders are assumed to unbind at rate α and we shall normalize the distribution to 1 at $t=0$ such that the total probability of being a given distance from the origin is exponentially decaying with time due to unbinding

$$\int_0^\infty p(x,t)\mathrm{d}x = e^{-\alpha t}\,. \tag{1}$$

We assume a deterministic propagation at velocity v which together with the unbinding process yields the Fokker-Planck equation

$$(\partial_t + v\partial_x)\,p(x,t) = -\alpha p\,, \tag{2}$$

wich just propagates the initial condition $p(x,t)=\delta(t)$ forward along the curve $x=vt$, leading to an exponentially decaying delta function in the absence of collisions

$$p(x,t) = \delta(t-x/v)e^{-\alpha t}\,. \tag{3}$$

Collisions with other extruders and barriers alter this picture, making p a sum over densities with various numbers of barriers in front of them.

In steady state, we expect the rate of collisions from extruders to be uniform and equal to the extrusion speed times the extruder density on the DNA $v\rho$. We introduce the density $s_n(x,t)$ to denote the heads with n extruders right in front of it. All states can unbind from the DNA but only the s_0 state can extrude. Collisions occur at the same rate for all states and lead to a degree of stalling one higher $s_n(x,t)\to s_{n+1}(x,t)$. The degree is lowered by unbinding of opposing heads with the unbinding rate $n\alpha$.

Putting all this together leads to a coupled set of Fokker-Planck equations

$$\begin{aligned}(\partial_t + v\partial_x)s_0(x,t) &= -(\alpha+v\rho)s_0(x,t) + \alpha s_1(x,t)\,,\\ \partial_t s_n(x,t) &= -\left((n+1)\alpha + v\rho\right)s_n(x,t)\\ &\quad + v\rho s_{n-1}(x,t) + (n+1)\alpha s_{n+1}(x,t)\,,\end{aligned} \tag{4}$$

with $n=\{1,\dots\}$. From this one can systematically expand in the barrier density $v\rho$ to include higher and higher degrees of stalled states. While this system of equations may perhaps be solvable in full, here I was able to derive useful results by focusing on only the first two terms in this hierarchy of stalled states.

III. LOW DENSITY APPROXIMATION WITH ARBITRARY BARRIER DISTRIBUTION

We assume that stalled states with more than one barrier in front are neglilible. This is exact if collisions are with only static barriers as they can only lead to collisions for the moving s_0 state. If moving barriers are present it would amount to a low-density approximation for extruder-extruder collisions. To include both barrier and extruder collisions we allow for an arbitrary rate of stalling $\gamma b(x)$ where $b(x)$ is a dimensionless function which can include both other extruders and static barriers. Upon stalling, the head may recover at rate u which may be different from the unbinding rate α. Denoting the stalled density by $s_1(x,t)\equiv s(x,t)$ and the free-moving density by $s_0(x,t)\equiv f(x,t)$, the Fokker-Planck equations for deterministic propagation and conversion between them becomes

$$(\partial_t + v\partial_x)\,f(x,t) = -\left(\alpha+\gamma b(x)\right)f(x,t) + us(x,t)\,, \tag{5}$$

$$\partial_t s(x,t) = -\left(\alpha+u\right)s(x,t) + \gamma b(x) f(x,t)\,, \tag{6}$$

These systems of equations may be solved by taking a Laplace transform $\partial_t\to -z$ in which case eq. (6) becomes

$$s(x,z) = \frac{\gamma b(x)}{\alpha+u+z}f(x,s)\,, \tag{7}$$

which may be substituted into eq. (5) to obtain a first order differential equation for f

$$\partial_x f(x,z) = -\left(\frac{\alpha+z}{v} + \frac{\alpha+z}{\alpha+u+z}\frac{\gamma}{v}b(x)\right)f(x,z)\,. \tag{8}$$

The solution to the full system of equations is obtained by solving eq. (8) for $f(x,t)$ and substituting it into eq. (7)

$$\begin{aligned}f(x,z) &= \exp\left[-\frac{\alpha+z}{v}\left(x + \frac{\gamma B(x)}{\alpha+u+z}\right)\right]\,,\\ s(x,z) &= \frac{\gamma b(x)}{\alpha+u+z}\exp\left[-\frac{\alpha+z}{v}\left(x + \frac{\gamma B(x)}{\alpha+u+z}\right)\right]\,,\end{aligned} \tag{9}$$

where I defined the cumulative barrier function $B(x)\equiv\int_0^x b(x)$. The full density is the sum of s and f

$$\begin{aligned}p(x,z) &= s(x,z) + f(x,z)\,,\\ p(x,z) &= \left(1+\frac{\gamma b(x)}{\alpha+u+z}\right)\exp\left[-\frac{\alpha+z}{v}\left(x + \frac{\gamma B(x)}{\alpha+u+z}\right)\right]\,.\end{aligned} \tag{10}$$

To invert the Laplace transform, we first focus on inverting f which may be expressed as a convolution in the time coordinate

$$\mathcal{L}^{-1}[f] = \mathcal{L}^{-1}[e^{-(\alpha+z)\frac{x}{v}}](t) \overset{t}{*} \mathcal{L}^{-1}[e^{-\gamma\frac{\alpha+z}{\alpha+u+z}\frac{B(x)}{v}}](t)\,, \tag{11}$$

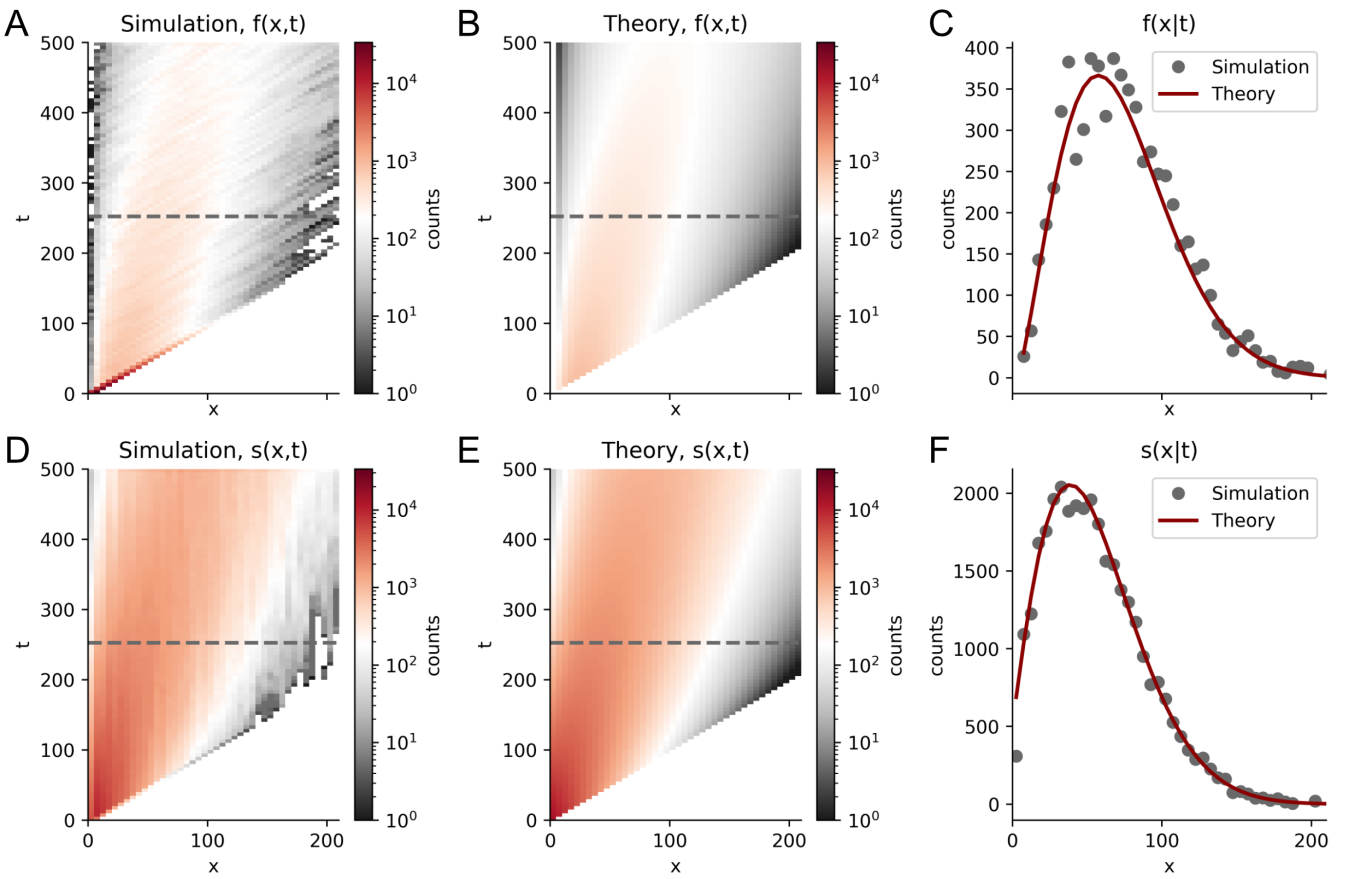

FIG. 2. Comparison of theoretically calculated densities with simulation. A: Simulated free density, B: Theoretically calculated free density using eq. (14) C: Conditional density for the horizontal section marked by a dotted line in A,B. D-F: Same as A-C but for the stalled density using eq. (15) for the theoretical calculation. Length are in units of the lattice spacing and time is in units of the lattice spacing divided by v. Parameters were $p_{\text{stall}} = 0.05 \to \gamma = 0.02$, $p_{\text{unstall}} = 0.01 \to u = 0.004$, $\alpha = 1/1000$.

where the involved transforms are

$$\mathcal{L}^{-1}[e^{-\frac{(\alpha+z)x}{v}}](t) = \delta\left(t - \frac{x}{v}\right) e^{-\alpha t}\,, \tag{12}$$

$$\mathcal{L}^{-1}[e^{-\gamma\frac{\alpha+z}{\alpha+u+z}\frac{B(x)}{v}}](t) = e^{-(u+\alpha)t-\frac{\gamma B(x)}{v}}\left[\delta(t)+\right.$$
$$\left.\theta\left(t-\frac{x}{v}\right)\sqrt{\frac{u\gamma B(x)}{vt}}I_1\left[2\sqrt{\frac{u\gamma t B(x)}{v}}\right]\right], \tag{13}$$

and I_1 is the modified Bessel function of the first kind. Taking the convolution in eq. (11) leads to the inverse Laplace transform

$$\begin{aligned} f(x,t) =& \delta\left(t-\frac{x}{v}\right)e^{-\alpha t-\frac{\gamma B(x)}{v}} + e^{-(u+\alpha)(t-\frac{x}{v})-\alpha\frac{x}{v}-\frac{\gamma B(x)}{v}} \\ & \theta\left(t-\frac{x}{v}\right)\sqrt{\frac{u\gamma B(x)}{v(t-\frac{x}{v})}}I_1\left(2\sqrt{\frac{u(t-\frac{x}{v})\gamma B(x)}{v}}\right) \\ f(x,t) =& e^{-\alpha t-\frac{\gamma B(x)}{v}}\delta\left(t-\frac{x}{v}\right) + e^{-(u+\alpha)(t-\frac{x}{v})-\alpha\frac{x}{v}-\frac{\gamma B(x)}{v}} \\ & \frac{u\gamma B(x)}{v}\theta\left(t-\frac{x}{v}\right){}_0F_2\left[\left(t-\frac{x}{v}\right)\frac{u\gamma B(x)}{v}\right], \end{aligned} \tag{14}$$

where ${}_0F_2$ is the confluent hypergeometric function. From eq. (7) we may now obtain $s(x,t)$ from the convolution

$$\begin{aligned} s(x,t) =& \mathcal{L}^{-1}\left[\frac{\gamma b(x)}{\alpha+u+z}\right](t) \overset{t}{*} f(x,t)\,, \\ s(x,t) =& \gamma b(x)\theta\left(t-\frac{x}{v}\right)e^{-(u+\alpha)(t-\frac{x}{v})-\alpha\frac{x}{v}-\frac{\gamma B(x)}{v}} \\ & {}_0F_1\left[\left(t-\frac{x}{v}\right)\frac{u\gamma B(x)}{v}\right]. \end{aligned} \tag{15}$$

The full distribution for the density of extruders now follows by adding eq. (14) and eq. (15)

$$\begin{aligned} p(x,t) =& \delta\left(t-\frac{x}{v}\right)e^{-\alpha t-\frac{\gamma B(x)}{v}} + \gamma e^{-(u+\alpha)(t-\frac{x}{v})-\alpha\frac{x}{v}-\frac{\gamma B(x)}{v}} \\ & \theta\left(t-\frac{x}{v}\right)\left(b(x)\,{}_0F_1\left[\left(t-\frac{x}{v}\right)\frac{u\gamma B(x)}{v}\right]+\right. \\ & \left.\frac{uB(x)}{v}\,{}_0F_2\left[\left(t-\frac{x}{v}\right)\frac{u\gamma B(x)}{v}\right]\right). \end{aligned} \tag{16}$$

Figure 2 shows eq. (14) and eq. (15) plotted against Monte Carlo simulations. We see that the model fol-

lows the simulations very well. To compare against simulations, I assumed a simple system of uniform barrier density $b(x) = 1$. The simulations were performed following only the right head of an extruder loaded on a lattice at the origin. Each timestep, the head could stall with probability p_{stall} if it was freely moving or it could unstall with a probability p_{unstall} if it was stalled. The head was allowed to move forward every timestep if it was unstalled.

IV. CONCLUSION

The calculations presented in this short communication constitutes a possible avenue to extend the currently available analytical results for the vanilla LEF model. Here I demonstrated how the single head distributions may be obtained, but from a same-time convolution of eq. (16) with itself, the loop size distribution would follow. Furthermore, various moments may be derived from the Laplace transform in eq. (10) and the corresponding solution with x also transformed. Setting $z = 0$ in eq. (10) would give the time-averaged density of extruders which may be of interest in calculating extruder head positions.

These are interesting and useful quantities which may help provide new mechanistic understanding across systems. For example, the LEF cohesin has been hypothesized to act as an agent bringing genes and enhancer regions together to activate transcription [16, 17]. If this was the case, arrival times of cohesins between the enhancer locus and the gene would be directly related to the frequency of activation of the gene these would be possible to calculate from the moments of some of the quantities presented in this communication. Another interesting use is in the modeling of cohesin loading experiments where it is possible to measure the 1D position of extrusion heads arising from a loading site [18] or from Hi-C contact maps around loading sites which has revealed unique features of contacts which might be predictable from this theory [19]. These interesting future avenues may now be possible to investigate with the Fokker-Planck based formalism presented here.

[1] I. F. Davidson and J.-M. Peters, Genome folding through loop extrusion by SMC complexes | Nature Reviews Molecular Cell Biology, Nature Reviews Molecular Cell Biology **22**, 445 (2021).

[2] P. Mach, P. I. Kos, Y. Zhan, J. Cramard, S. Gaudin, J. Tünnermann, E. Marchi, J. Eglinger, J. Zuin, M. Kryzhanovska, S. Smallwood, L. Gelman, G. Roth, E. P. Nora, G. Tiana, and L. Giorgetti, Cohesin and CTCF control the dynamics of chromosome folding | Nature Genetics, Nature Genetics **54**, 1907 (2022).

[3] J. H. Gibcus, K. Samejima, A. Goloborodko, I. Samejima, N. Naumova, J. Nuebler, M. T. Kanemaki, L. Xie, J. R. Paulson, W. C. Earnshaw, L. A. Mirny, and J. Dekker, A pathway for mitotic chromosome formation, Science **359**, eaao6135 (2018).

[4] A. Goloborodko, M. V. Imakaev, J. F. Marko, and L. Mirny, Compaction and segregation of sister chromatids via active loop extrusion, eLife **5**, e14864 (2016).

[5] R. K. P. Zia, J. J. Dong, and B. Schmittmann, Modeling Translation in Protein Synthesis with TASEP: A Tutorial and Recent Developments, J Stat Phys **144**, 405 (2011).

[6] A. Goloborodko, J. F. Marko, and L. A. Mirny, Chromosome Compaction by Active Loop Extrusion, Biophysical Journal **110**, 2162 (2016).

[7] K. Polovnikov, S. Belan, M. Imakaev, H. B. Brandão, and L. A. Mirny, Fractal polymer with loops recapitulates key features of chromosome organization (2022).

[8] S. Belan and V. Parfenyev, Footprints of loop extrusion in statistics of intra-chromosomal distances: An analytically solvable model (2023), arxiv:2301.03856 [cond-mat, physics:physics].

[9] D. Starkov, V. Parfenyev, and S. Belan, Conformational statistics of non-equilibrium polymer loops in Rouse model with active loop extrusion, J. Chem. Phys. **154**, 164106 (2021).

[10] A. Goloborodko, J. F. Marko, and L. A. Mirny, Chromosome Compaction by Active Loop Extrusion, Biophysical Journal **110**, 2162 (2016).

[11] E. J. Banigan and L. A. Mirny, The interplay between asymmetric and symmetric DNA loop extrusion., Elife **9**, 10.7554/eLife.63528 (2020).

[12] H. B. Brandão, P. Paul, A. A. van den Berg, D. Z. Rudner, X. Wang, and L. A. Mirny, RNA polymerases as moving barriers to condensin loop extrusion., Proc Natl Acad Sci U S A **116**, 20489 (2019).

[13] M. Crippa, Y. Zhan, and G. Tiana, Effective model of loop extrusion predicts chromosomal domains, Phys. Rev. E **102**, 032414 (2020).

[14] C. A. Brackley, J. Johnson, D. Michieletto, A. N. Morozov, M. Nicodemi, P. R. Cook, and D. Marenduzzo, Nonequilibrium Chromosome Looping via Molecular Slip Links, Phys. Rev. Lett. **119**, 138101 (2017).

[15] T. Yamamoto and H. Schiessel, Osmotic mechanism of the loop extrusion process, Phys. Rev. E **96**, 030402 (2017).

[16] J. M. Luppino, A. Field, S. C. Nguyen, D. S. Park, P. P. Shah, R. J. Abdill, Y. Lan, R. Yunker, R. Jain, K. Adelman, and E. F. Joyce, Co-depletion of NIPBL and WAPL balance cohesin activity to correct gene misexpression, PLOS Genetics **18**, e1010528 (2022).

[17] M. Gabriele, H. B. Brandão, S. Grosse-Holz, A. Jha, G. M. Dailey, C. Cattoglio, T.-H. S. Hsieh, L. Mirny, C. Zechner, and A. S. Hansen, Dynamics of CTCF and cohesin mediated chromatin looping revealed by live-cell imaging (2021).

[18] R. Han, Y. Huang, I. Vaandrager, A. Allahyar, M. Magnitov, M. J. A. M. Verstegen, E. de Wit, P. H. L. Krijger, and W. de Laat, Targeted cohesin loading characterizes the entry and exit sites of loop extrusion trajectories (2023).

[19] Y. Guo, E. Al-Jibury, R. Garcia-Millan, K. Ntagiantas, J. W. D. King, A. J. Nash, N. Galjart, B. Lenhard, D. Rueckert, A. G. Fisher, G. Pruessner, and M. Merkenschlager, Chromatin jets define the properties of cohesin-driven in vivo loop extrusion, Molecular Cell **82**, 3769 (2022).