An Approach to Polymer Translocation via a Sequential Algorithm

Zbigniew Domański and Andrzej Grzybowski

***Abstract*—The translocation of a polymer belongs to a class of important bio-chemical processes. We propose a sequential algorithm designed to reduce the complexity involved in the dynamics of polymer molecular transport. The key concept behind our algorithm is the sequentialization of a polymer movement between its consecutive conformations into a sequence of steps, a picture borrowed from the definition of an optimal strategy within the theory of games played sequentially. As an example we apply our algorithm to study, in two dimensions, the driven translocation of a polymer-like structure with the length of N monomers through a flat membrane containing two holes separated by the distance Δ. We study the statistics of the translocation time τ computed as the time consumed by the polymer to pass from one side of the membrane to another one as a function of N and Δ. The presence of two close lying holes frustrates the passing polymer and we observe that the average value of τ oscillates around the scaling function $(N/\Delta)^{1.81}$.**

***Index Terms*— multi-hole membrane, passage time, polymer translocation, sequential algorithm.**

I. INTRODUCTION

The transfer of a polymer across a membrane is the subject of extensive theoretical and experimental works. This nano-scale transport primarily exists in the world of biology with such prominent examples as DNA/RNA translocation through nuclear pores, virus injection into a host cell and genetic therapy. The experimental approaches to polymer science allow to determine the physical, bio-chemical and technological properties of polymers [1]-[3]. For example the translocation of polymers through nanometer-scale pores is an experimental technique to analyze physical properties of biomolecules. The results of experiments supply more accuracy in the theoretical models which are formulated on an analogy between a polymer and a type of random walk [4]-[8].

Typically, a flat membrane with an opening is considered with a chain-like sequence of N monomers passing from so-called *cis* side of the membrane to its another side called *trans*. An important theoretical question in this field is to provide the correct estimation of the translocation time τ on the polymer length N [9], [10]. One of the objectives of this work is the numerical analysis of the polymer translocation in terms of polymer length and membrane holes arrangements.

From the physical point of view, polymers, being classical systems, can be studied by computer simulation. However, the complexity of possible polymer conformations along with the volume of a solvent make the search space for a hypothetical algorithm so huge that real polymers have to be mapped onto significantly simplified models [4], [6], [8]. On the other hand, the simplifications of a polymer model can be less harmful with respect to the real polymer if the search space viewed by an algorithm, appropriate to this given model, may be quickly sampled in a reliable manner.

Usually the algorithms which are used for the simulation of polymer behavior consider all future states which are reachable at a given moment [11]-[14]. Such an approach results in computer procedures that are rather inefficient and lead to a very time consuming simulation. Instead we propose here another approach to modeling polymer behaviour. Its main feature is the sequentialization of the polymer move. Amazingly, the idea was inspired by the theory of games: a sequence of choices can be viewed as a sequence of moves in some sequential game. Analogously to a play in a game, a polymer translocation (with rules imposed by Nature) can be viewed as a sequence of choices that lead to some outcome. When one defines the strategy for a game which is played sequentially, then there is no need to define the player decisions at every possible state of the game. To obtain a well defined strategy it is absolutely enough to determine the player decisions only at these stages which may be expected in the course of the game, given earlier decisions made by the player. One can easily notice that after each decision made in the sequential game, the number of its possible future states dramatically decreases. This can seriously simplify the situation of the player and we expect that adopting this idea we obtain more efficient algorithms for the simulation of polymer transport.

The aim of this work is to present a numerical tool for the quick analyzing of the kinetics of chain-like structures from the statistical point of view. We do that by introducing a sequential algorithm built up around a sequential game optimal strategy concept and probabilistic-cellular-automata methodology.

II. THE SEQUENTIAL ALGORITHM FOR MODELING RANDOM MOVEMENT OF POLYMER-LIKE STRUCTURES

In this section we present the algorithm we use to calculate the passage time through a flat membrane. We define our algorithm in a formal way with the help of the following terms and assumptions.

A. Terms and assumptions

An abstract *2D polymer position* is a finite sequence

Manuscript received July 18, 2012. Zbigniew Domański and Andrzej Grzybowski are with the Czestochowa University of Technology, Dabrowskiego 69, PL-42201 Czestochowa, Poland. (corresponding author: Z..D; e-mail: zbigniew.domanski@im.pcz.pl).

$\mathbf{c} = \{c_1, c_2, \ldots c_n\}$ of *2D* points $c_i = (x_i, y_i)$ such that the distance *d* between any pair of its successive elements is bounded by given limits $g : d(c_i, c_{i+1}) < g$, $i = 1, 2, \ldots, n-1$. The elements of the sequence **c** are called *segments* of the polymer. The segment c_n is called the *head* of the polymer, while c_1 is called the *tail*. The number *n* is called the *length* of the chain.

Assumption 1 (discretization of motion space): The polymer moves along the integer lattice nodes, i.e. coordinates (x_i, y_i) are integer.

The *movement trajectory* is a sequence of consecutive chain positions stored in the matrix $\hat{\mathbf{C}}$ whose *i*-th row is interpreted as a polymer position at the moment *i*. Thus the element c_{ij} denotes the position of the segment *j* at the moment *i*.

The relocation of the polymer consists of a sequence of *moves* which transform the polymer from one position to another.

Assumption 2 (sequentialization of the move): Every single move of a chain can be sequentialized into a sequence of *steps*, i.e. moves made sequentially by polymer segments. A single step may transform the given segment c_j only to one of its neighbouring nodes, i.e. $d(c_{ij}, c_{i\pm1,j}) < g$.

The *first to move* (FTM) segment is the segment which in a given move is chosen by the algorithm to make the step as the first from all the polymer segments. The choice is realized according a given probability distribution defined on the polymer segments. The distribution will be denoted as FTMD.

The steps may be influenced by an *outer law* given by a probability distribution defined on the neighbouring nodes. The *outer law probability distribution* (OLDP) may depend on the position of the segment in the motion space. The OLDP reflects the existence of restrictions and constraints imposed on the system from the environment and resulting from the laws of physics.

The step made by the FTM segment as well as all the following steps may be subject to some *additional restrictions connected with the assumed features* of the polymer. These restrictions will be denoted as AFR. For example, one of such restrictions is the upper limit for the distance between successive segments. This restriction assures the continuity of a polymer, and – by assumption – it does not concern the FTM segment. Another example of the AFR restrictions may be the requirement that in a given node, not more than a given number of segments can be placed (e.g. the repton model [5]).

Assumption 3 (polymer nature of the move): Every first move of a polymer is started by only one segment, chosen according to the OLPD and the AFR.

The above assumption allows us to simulate the movement of the described polymer structure effectively and efficiently. However, in many practical problems, such as biopolymer behaviour inside a living tissue, one should also take into account some additional constraints connected with the biochemical nature of the system. Thus we define additionally the cost connected with the polymer structure. The *structure* of the polymer is defined by the relative mutual positions and interactions between the segments. The *cost* of the polymer structure and its location in the motion space is the function *F* representing its *fitness* connected with its conformation and/or other external (e.g. environmental) properties. The lower cost, the better fitness of the polymer structure and position.

Assumption 4 (acceptance of new polymer position): The new position of the polymer is accepted (by Nature) with a probability depending on its cost.

B. Algorithm

The above four assumptions and ideas are implemented in the following sequential algorithm for a polymer movement simulation:

Step 0. (*Initialization*) Set the initial (current) *polymer position* $\mathbf{c}_{curr}$ and evaluate its *current cost* function value F_{curr}.

Step 1. (*FTM segment selection*) According to the given FTMD, select FTM segment $c_{curr,f}$, $f = 1, 2, \ldots, n$.

Step 2. (*Step choice for FTM segment*) According to the given OLPD and AFR, select the neighbouring node for the next position of the segment $c_{new,f}$.

Step 3. (*Successive steps of remaining segments*) To obtain a new polymer position $\mathbf{c}_{new}$ sequentially choose the segments $c_{new,i}$; $i = f-1, \ldots, 1$ and draw the neighbouring nodes for their subsequent positions according the OLPD and AFR. This process is terminated for the first segment *k*, $f-1 \ge k \ge 1$ for which the following condition holds: $d(c_{curr,k}, c_{new,k+1}) < g$. If $k > 1$ then $c_{new,i} = c_{curr,i}$ for $i = 1, \ldots, k$. Next sequentially choose the segments c_i, $i = f+1, \ldots, n$ and draw the neighbouring nodes for their next positions according to the OLDP and AFR. This process terminates for the first *k*: $f+1 \le k \le n$, for which the polymer continuity condition $d(c_{curr,k}, c_{new,k-1}) < g$ is fulfilled. If $k < n$ then, for $i = k, \ldots, n$ we assume $c_{new,i} = c_{curr,i}$.

Step 4. (*Acceptance of new position*) Compute the cost of the new position F_{new}. If $F_{new} < F_{curr}$ accept $\mathbf{c}_{new}$. Alternatively, accept $\mathbf{c}_{new}$ only if the random variable *U* having a uniform probability distribution on interval [0,1] satisfies $U \le \psi(F_{new} - F_{curr})$, with ψ being a given nondecreasing function. If $\mathbf{c}_{new}$ is accepted then $\mathbf{c}_{curr}$ is replaced by $\mathbf{c}_{new}$; else $\mathbf{c}_{curr}$ remains as it is.

Step 5. Terminate the algorithm if the stopping criterion is met; else return to Step 1.

Step 6. Return the final position of the chain, its cost and various statistics connected with the simulated movements.

The nondecreasing function ψ that appears in Step 4 of the above algorithm represents the attitude of Nature towards the acceptance of worse states. If Nature accepts all states, one may assume that $\psi \equiv 1$. Otherwise, similarly as in the famous Metropolis algorithm, we propose to use function $\psi(z) = \exp[-z/T]$, where *T* is a parameter which can be

ISBN: 978-988-19252-4-4
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

additionally subject to change during the movement process, see e.g. [15].

III. Driven Polymer Translocation - Simulation Results

We illustrate the possible usage of the introduced algorithm by considering a simple example of a chain-like structure passing through a membrane with two openings within it. Here we consider the translocation process affected by a homogenous constant force, e.g. an electric drive applied from the outside (assuming that the monomers are similarly charged). Fig. 1 shows the layout for our translocation study.

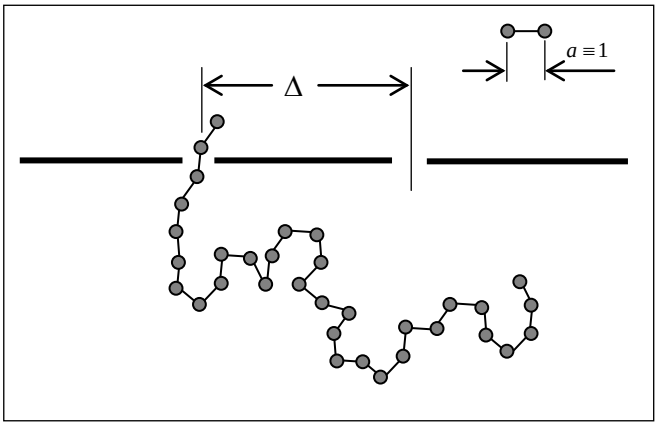

Fig. 1. Schematic representation of 32-segment length chain-like structure translocation through a flat membrane with two holes separated by a distance $\Delta = 6$. All distances are scaled by the factor $1/a$, where a is the length of the segment (monomer).

The simulations are started from a configuration with a polymer placed far away from the membrane. It is in order to ensure that just before the polymer touches the membrane the polymer configuration is not influenced by its initial conformation. We also assume that the hole is large enough to allow the translocation of polymers with folded configurations across the membrane [16].

Below we present the results of the numerical simulations of the above specified setup. For each analyzed length of the polymer we performed 10^4 simulations and then we have built empirical distributions of the translocation time τ.

Our simulations yield a common observation: the average translocation time $\overline{\tau}$ oscillates around the scaling function

$$\overline{\tau} \propto \left(N/\Delta\right)^{1.81\pm0.04}, \qquad (1)$$

see Fig. 2. This dependence is clearly seen for relatively short polymers, i.e. polymers with the length $N < 3\Delta$, see Fig. 3. When only one hole in the membrane is opened this oscillatory variation of $\overline{\tau}$ disappears. Thus, the presence of two holes alters the polymer passage. For some relatively short polymers the translocation process is slowed down. Hence, we conclude that two holes, operating within a distance comparable to the length of polymer may frustrate the system.

Oscillation excepted, our finding is in close correspondence with the scaling function $\left(N/\Delta\right)^{1+\nu_{2D}}$, with $\nu_{2D} = 0.75$ being the Flory exponent in two dimensions. The exponent $1+\nu_{2D}$, connected with the so-called Rouse dynamics of polymers in stationary solvent, has been numerically confirmed [7] and also reported in experiments on long double-strand DNA molecules pushed through a siliconoxide nanopores [2].

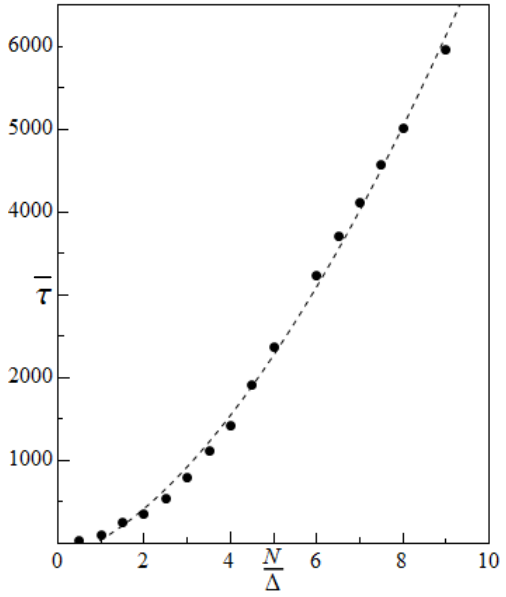

Fig. 2. Scaled mean translocation time of polymers with length $N/\Delta < 10$ averaged over 10^4 cases. The holes are separated by distance $\Delta = 20$. The dashed line is drawn using (1) and it is only a visual guide.

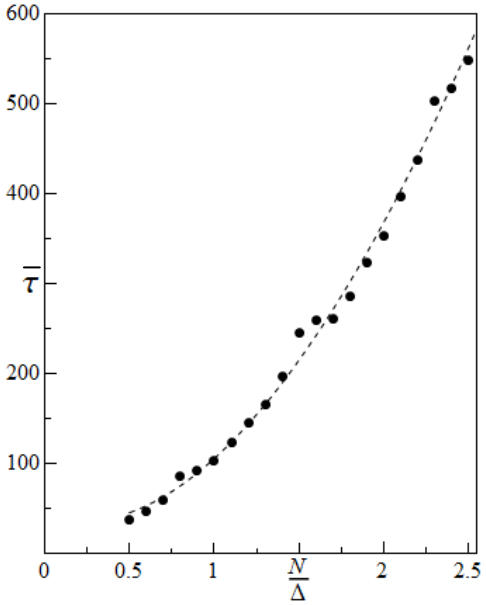

Fig. 3. Same as in Figure 2 but for scaled polymers length $N/\Delta < 3$.

IV. Summary

As biopolymer translocations play an increasing role in the understanding of basic cell biology, there is an increasing need to understand the interplay between the geometry and the transport properties of a long polymer traversing a punched membrane. In this paper we have studied a minimalist model of a polymer, i.e. we have applied our sequential algorithm to a so-called freely-joined model of a polymer being aware of its serious shortcoming. For example

ISBN: 978-988-19252-4-4
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

we did not take into account the interactions among the segments and we also neglected the interactions between the segments and solvent molecules. We have chosen such a simple model only because of the transparency of the algorithm presentation, e.g. we have not used Assumption 4 and Step 4.

Although we have limited our work to the translocation time issue, many other statistical characteristics of the movement process as well as the influence of various parameters describing both the chain structure itself and the distributions involved in the process can be easily examined with the help of the presented algorithm. It is worth emphasizing that the algorithm is very efficient. We have obtained our results with relatively little numerical effort, modest memory and CPU resources.

REFERENCES

[1] A. Meller, L. Nivon, and D. Branton, "Voltage-driven DNA translocations through a nanopores", Phys. Rev. Letters, vol. 86, pp. 3435-3438, April 2001.
[2] A.J. Storm *et all*., "Fast DNA Translocation through a Solid-State Nanopore", Nano Letters, vol. 5, no. 7, pp. 1193-1197, Jan. 2005.
[3] M. Greshow, J.A. Golovchenko, "Recapturing and trapping single molecules with a solid-state nanopore", Nature Technology, vol.2, pp. 775-779, Dec. 2007.
[4] J.M. van Leeuwen and A. Drzewiński, "Stochastic lattice models for the dynamics of linear polymers", Phys. Reports, vol. 475, pp. 53-90, May 2009.
[5] M. Rubinstein, "Discretized model of entangled-polymer dynamics", Phys. Rev. Letters, vol. 59, pp 1946-1949, Oct. 1987.
[6] A. Drzewiński, J.M. van Leeuwen, „Crossover from reptation to Rouse dynamics in the extended Rubinstein-Duke model", Phys. Rev. E, vol. 77, p. 31802, March 2007.
[7] Y. Kantor, M. Kardar, "Anomalous dynamics of forced translocation", Phys. Rev. E, vol. 69, p. 21806, Feb. 2004.
[8] A. Meller, "Dynamics of polynucleotide transport through nanometer-scale pores", J. Phys.: Cond. Matter, vol. 15, pp. R581-R607, April 2003.
[9] A.B. Kolomeisky, "How Polymers Translocate Through Pores: Memory is Important", Biophysical Journal, vol. 94, pp. 1547-1548, March 2008.
[10] R. Metzler, J. Klafter, "When Translocation Dynamics Becomes Anomalous", Biophysical Journal, vol. 85, pp. 2776-2779, Oct. 2005.
[11] A. Milchev, K. Binder, and A. Bhattacharya, "Polymer translocation through a nanopores induced by absorption: Monte Carlo simulation of a coarse-grained model", J. Chem. Phys., vol. 121, p. 34714, Jan. 2006.
[12] K. Luo, T. Ala-Nissila, and See-Chen Ying, "Polymer translocation through a nanopores: Monte Carlo study", J. Chem. Phys., vol. 124, pp.6042-6051, Sept. 2004.
[13] D. Panja and G.T. Barkema, "Simulations of two-dimensional unbiased polymer translocation using the bond fluctuation mode", J. Chem. Phys., vol. 132, p. 14902, Jan. 2010.
[14] C. Chatelain, Y. Kantor, and M. Kardar, "Probability distributions for polymer translocation" , Phys. Rev. E, vol. 78, p. 21129, Aug. 2008.
[15] J.C. Spall, "Introduction to Stochastic Search and Optimization: Estimation, Simulation and Control", John Wiley & Sons Inc. Publication, 2003.
[16] S. Kotsev, A.B. Kolomeisky, "Translocation of polymer with folded configurations across nanopores", J. Chem. Phys., vol. 127, p. 185103, Nov. 2007.

ISBN: 978-988-19252-4-4
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)