

PROBABILISTIC MODEL OF THE CHAIN-LIKE STRUCTURES' TRANSLOCATION THROUGH THE PORE IN MEMBRANE

Kamila Bartłomiejczyk

Institute of Mathematics, Czestochowa University of Technology

Częstochowa, Poland

kamila.bartlomiejczyk@im.pcz.pl

Abstract. The paper is devoted to the analysis of the translocation of the chain-like structures (CLS) through the pore in membrane. It focuses on the so-called *passage time* and the principal aim is to propose a proper probabilistic model for its distribution. The article starts with some preliminary results concerning stochastic processes. Next, an approximate analytic solution of the considered problem that was received in literature is presented. It is known that the resulting probability distribution of the passage time manifests some important shortcomings. Thus it is important to find a reasonable alternative. It is argued here that the beta Moyal probability distribution can be a good candidate for approximation of the theoretical distribution in various interesting situations. In this paper, two different problems connected with CLS translocation are considered. For both problems the theoretical distributions are known from literature and in both situations the beta Moyal approximations turn out to be very satisfactory.

Keywords: *Fokker-Planck equation, translocation of chain-like structure, stochastic process, beta Moyal probability distribution*

1. Introduction: Formal stochastic description of the translocation time

Translocation of the chain-like structures (CLS) through the pore in membrane occurs in many fields of science and it is widely described in literature (e.g. [1-5]). This phenomenon plays a key role, inter alia, in: RNA and DNA sequencing technique, transport of the molecules across membranes in cells, phage infection, bacterial conjugation, gene therapy, and many others. One of the principal research tasks is to study the probability distribution of the CLS passage time.

The translocation through the pore phenomenon can be considered as the Markov process. This class of stochastic processes plays a very important role in description of numerous different physical phenomena [6]. Any stochastic process is the process which is described formally as a directed set of random variables $\{X_t\}_{t \in T}$, where $t \in T$ may be interpreted as the time (continuous if T is uncountable, or a discrete one if T is at most countable). The stochastic process is called

a Markov one, when for any $t \in T$, its next state - at time $t+1$ or $t + \Delta t$ - depends solely on its current state X_t . Thus such processes are generally defined by some initial conditions and the transition rules. Consequently, to define any Markov process, it is sufficient to know two probability functions: $P(x, t_0)$ (initial probability distribution), $P(x, t | x', t')$, $t > t'$ (transition probability). These functions satisfy the following equations:

$$P(x_2, t_2) = \int dx_1 P(x_2, t_2 | x_1, t_1) P(x_1, t_1) \quad (1)$$

and

$$P(x_3, t_3 | x_1, t_1) = \int dx_2 P(x_3, t_3 | x_2, t_2) P(x_2, t_2 | x_1, t_1), \quad t_1 < t_2 < t_3. \quad (2)$$

Formula (2) is known as the integral Chapman-Kolmogorov equation.

Under appropriate assumptions (see [6]) from the integral Chapman-Kolmogorov equation one can obtain the differential Chapman-Kolmogorov equation and its particular variant - the Fokker-Planck equation which describes the diffusion with drift:

$$\frac{\partial P}{\partial t} = -v \frac{\partial P}{\partial x} + D \frac{\partial^2 P}{\partial x^2}. \quad (3)$$

In paper [2] the equation (3) is proposed as a macroscopic description of the CLS translocation process. This description is valid under some assumptions about the considered phenomenon. First, it is assumed that the CLS passes through the pore in the membrane. Next, it is assumed that the pore diameter is comparable to the size of monomers of the CLS in this sense, that no more than one chain can fit in the pore at the same time. In [2] it is also argued that, in the presence of a force driving the CLS through the pore, the system is effectively one-dimensional. Let $P(x, t)$ denotes the probability that a part x of the CLS length has passed through the pore at time t . The variable x is defined in the following way: assuming that CLS has length L , then $x = 0$ when the first segment of CLS reached the pore and $x = L$ when all segments of the structure reached the other side of the membrane. Factors v and D in (3) are, respectively, a drift coefficient and an effective diffusion coefficient. The parameter v is nonzero when the external driving force influences the structure motion. All microscopic details are subsumed in parameters v and D . It is also convenient to consider the related parameter, the so-called diffusive length l_d defined as: $l_d = D / v$.

For fixed v, D and with absorbing boundary conditions $P(0) = P(L) = 0$, the solution of equation (3) on the interval $[0, L]$ gives a distribution of the passage time. Based on the approximate solution found in [2], the probability density function $\psi(t)$ of the passage time in the considered case is given by the formula:

$$\psi(L, \nu, l_d, t) = \frac{\nu}{2} \sqrt{\frac{l_d}{\pi}} \left(\frac{L^2}{l_d (\nu t)^{5/2}} - \frac{2}{(\nu t)^{3/2}} \right) \exp\left(-\frac{(\nu t - L)^2}{4\nu t l_d}\right), \quad \frac{L^2}{\nu t l_d} \gg 1 \quad (4)$$

where, as previously, L is the length of CLS, l_d is the diffusive length, ν is the drift coefficient. Interesting details concerning the solution method can be found in the Appendix of the paper [2]. Simulations of other probability distributions were considered e.g. in [7].

2. The beta Moyal approximation of the passage time distribution

It is worth noticing that expression (4) is not valid for sufficiently large t , because for ts large enough, the function $\psi(t)$ becomes negative. Thus it is an important task to find a *proper* probability density function (pdf) which could be an alternative to (4). Such a pdf should differ as little as possible from ψ , yet it should satisfy the usual conditions characterising any pdf, especially it must be nonnegative for all $t > 0$. Following some ideas from literature, see [8], it seems reasonable to approximate a function given by equation (4) with probability density functions belonging to the beta Moyal family (which is widely discussed, e.g. in [9]). These density functions are given by the following expression:

$$p(a, b, c, d, t) = \frac{\exp\left\{-\frac{1}{2}\left[\frac{t-a}{b} + \exp\left(-\frac{t-a}{b}\right)\right]\right\}}{\sqrt{2\pi b} \Gamma\left(\frac{1}{2}\right)^{d-1} B(c, d)} \left[1 - \frac{\gamma\left\{\frac{1}{2}, \frac{1}{2} \exp\left(-\frac{t-a}{b}\right)\right\}}{\Gamma\left(\frac{1}{2}\right)} \right]^{c-1} \times \left\{ \gamma\left(\frac{1}{2}, \frac{1}{2} \exp\left(-\frac{t-a}{b}\right)\right) \right\}^{d-1}, \quad (5)$$

where $B(c, d) = \Gamma(c)\Gamma(d)/\Gamma(c+d)$ is the beta function, Γ is the Euler's gamma function, $\gamma(\alpha, t)$ is the incomplete gamma function, $-\infty < a < \infty$ is the location parameter, $b > 0$ is the scale parameter and $c, d > 0$ are the shape parameters. When $c = d = 1$, equation (5) reduces to the Moyal density function that was considered in this context in [8]. Since the underlying theoretical distribution of the translocation time is a three-parametric function, and the beta Moyal probability distribution depends on four parameters, while modeling ψ , it can be assumed that $d = f(c)$.

As the criterion for the best fit of distributions the following integral is used:

$$\int_I (p(a, b, c, 2c, t) - \psi(L, \nu, l_d, t))^2 dt \quad (6)$$

where interval $I = \left[0, \frac{L}{2l_d}\right]$. Obviously, the values of the parameters appearing in ψ are exactly the same as in [2]. The aim of our research is to identify such values of a, b, c in (5) for which above criterion takes on the least value. All parameters were looked for in the interval $[0.1, 3]$ (with step $s = 0.01$). In this particular case it was assumed that $f(c) = 2c$. The least value of the integral given by (6) was obtained for $a = 0.33$, $b = 0.09$, $c = 0.51$. It was equal to 0.00147, so the area between both density functions was very small, meaning that the approximation is really good. Moreover, the expected values and the standard deviations of both distributions have almost the same values (they were equal after rounding to one decimal place).

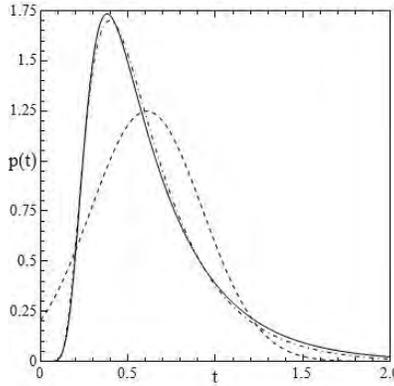


Fig. 1. The probability density functions $p(t)$ of scaled translocation time t . Solid line: the beta Moyal distribution with parameters $a = 0.33$, $b = 0.09$, $c = 0.51$; dash-dotted line: distribution ψ in equation (7). The dashed curve represents a normal distribution. All presented distributions have the same means ($\mu = 0.6$) and standard deviations ($\sigma = 0.3$), which is in agreement with the Figure 2 in [5]

In fact, Figure 1 shows noticeable agreement between the probability density function of the three-parametric beta Moyal probability distribution (with the identified parameters) and the probability density function ψ which has been derived from the coarse-grained equation for chain-like structures' translocation in [2]. A normal probability distribution (suggested by some early works, see [2, 8]) with the same expected value and standard deviation is also included for comparison.

In the remaining part of this section, the beta Moyal approximation of the theoretical translocation time curves obtained in [4] is presented. In [4] the theoretical translocation into and out of an ellipsoidal cavity is studied. It is assumed that the CLS is translocated through the narrow pore in the membrane between an ellipsoidal cavity and a semi-infinite space on one side of the barrier. The diameter of the pore is comparable to the size of the segments forming the CLS. So, similarly as in the previous example, just one chain segment can be placed in the pore at a given time. With the help of the Fokker-Planck formalism, the authors of [4]

were able to calculate translocation times pdfs for both chain insertion and ejection, for more details see [4]. In our studies, the results obtained in [4] are used to verify whether or not a beta Moyal distribution can be a good probabilistic model of the passage time in this, a bit different, case. For this purpose the translocation time curves presented in [4] are analysed (Fig. 6b in there). To obtain the validation data set consisting of the “theoretical” values of the pdf, the coordinates of different points lying on the graphs were read out. Consequently, to obtain the best beta Moyal approximation of the theoretical curve, the following discrete version of the criterion (6) is considered:

$$\sum_{i=1}^n \left(p(a, b, c, \frac{c}{2}, x_i) - y_i \right)^2, \quad (7)$$

where (x_i, y_i) are coordinates of points read out from the curve presented in [4], and n is the number of points (in this case $n = 420$).

Parameters which minimize the above criterion were searched for in the interval $[0.1, 250]$ (with step $s = 0.1$). As in the previous problem, a three-parametric distribution was considered. Thus a parameter d from equation (5) is again related with c . This time it is assumed that $f(c) = c/2$. The least value of criterion (7) is $3.9122 \cdot 10^{-6}$ and it is achieved for $a = 225.2$, $b = 53.8$, $c = 2.05$.

In Figure 2, one can see that the “curve” consisting of points read out from the Figure 6b in [4] is nearly the same as the beta Moyal density function with the parameters found in our studies. Therefore, it can be concluded that the three-parametric beta Moyal probability distribution is a good approximation for theoretical distribution of the CLS passage times in both problems considered here.

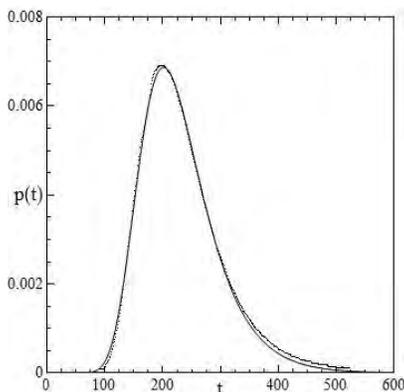


Fig. 2. The probability density functions $p(t)$ of translocation time t . Solid line: the beta Moyal distribution with parameters $a = 225.2$, $b = 53.8$, $c = 2.05$; points: data obtained from the Figure 6b in [4]

3. Final remarks

A major subject of the presented research is the translocation of the CLS through the pore in the membrane. The main conclusion is that the theoretical distribution of the passage time can be well approximated with the three-parametric beta Moyal probability distribution. It is important, because the theoretical probability density function proposed in the literature is not valid for large t (it takes on negative values) while the beta Moyal density function defines a proper probability distribution.

There are several future research directions. First, similar research based on experimental data connected with polymer translocations will be performed. Such data can be found for example in [1, 3].

Another possible area of future studies is connected with the choice of the function f that relates two parameters, c and d , of the beta Moyal distribution. In this paper it was assumed that f is linear in both cases. However, other functions may also be considered and the resulting beta Moyal approximation may appear to be even better.

Yet another subject of future research is to examine data obtained via Monte Carlo studies. In these studies the algorithm described in [8] will be used to analyse the CLS passage times under a number of different assumptions concerning the topology of the pore, its relation to monomers' size, the external force driving the CLS, the algorithm parameters, etc. An important part of this research will be the development of models that relate the algorithm parameters with the parameters of beta-Moyal distribution of the translocation time.

References

- [1] Carson S., Wilson J., Aksimentiev A., Wanunu M., Smooth DNA transport through a narrowed pore geometry, *Biophysical Journal* 2014, 107, 1381-2393.
- [2] Lubensky D.K., Nelson D.R., Driven polymer translocation through a narrow pore, *Biophysical Journal* 1999, 77, 1824-1838.
- [3] Meller A., Branton D., Single molecule measurements of DNA transport through a nanopore, *Electrophoresis* 2002, 23(16), 2583-2591.
- [4] Polson J.M., Polymer translocation into and out of an ellipsoidal cavity, *Journal of Chemical Physics* 2015, 142(17), 174903.
- [5] Żurek S., Kośmider M., Drzewiński A., van Leeuwen J.M.J., Translocation of polymers in a lattice model, *The European Phys. J. E: Soft Matter and Biological Physics* 2012, 35, 47.
- [6] Kosztołowicz T., Lewandowska K.D., Stochastyczny model dyfuzji w układzie membranowym, *Metody Matematyczne w Zastosowaniach* 2014, 2, 153-174.
- [7] Grzybowski A., Puchała P., Remarks about discrete Young measures and their Monte Carlo simulation, *Journal of Applied Mathematics and Computational Mechanics* 2015, 14(2), 13-20.
- [8] Grzybowski A.Z., Domański Z., A sequential algorithm with built in tension-propagation mechanism for modeling the chain-like bodies dynamics, arXiv:1312.4206 [cond-mat.soft].
- [9] Cordeiro G.M., Nobre J.S., Pescim R.R., Ortega E.M.M., The beta Moyal: A useful-skew distribution, *International Journal of Research and Reviews in Applied Sciences* 2012, 10(2), 171-192.