When do diffusion-limited trajectories become memoryless?

Maciej Dobrzyński
CWI (Center for Mathematics and Computer Science)
Kruislaan 413, 1098 SJ Amsterdam, The Netherlands

Abstract

Stochastic description of cellular dynamics by the chemical master equation assumes the exponential distribution of intervals between reaction events. Diffusion-limited reactions violate this assumption. Using the example of the target search we investigate the conditions under which a peaked waiting-time distribution can be approximated by the exponential function. We link the steady-state flux and the dynamic property of the diffusion, the mean first-passage time.

1 Introduction

Cellular regulation involves processes with reactants occurring at low copy numbers per cell (e.g. transcription/translation, signaling). Such processes suffer from thermal noise and diffusion-limitation. The concentration of species involved in such reactions fluctuates significantly. Recent single-cell, single-molecule experiments indicate that fluctuations contribute to heterogeneity of isogenic populations and may be detrimental to cellular information processing [1, 2].

1.1 CME description of biochemical reactions

A conventional modeling approach that accounts for fluctuations in the discrete number of molecules is the chemical master equation (CME) [4, 5]. The equation describes the evolution in time of the probability to occupy

one of the discrete states. State transitions correspond to chemical reactions occurrences. The time to the next transition is drawn from an exponential distribution (CME defines a continuous Markov chain). The mean of the distribution depends on the number of substrates involved in a given reaction and the corresponding macroscopic rate constant [3].

The assumption that waiting times for a transition are distributed exponentially has numerous consequences. For example, if a bimolecular reaction ($A+B \rightarrow C$) is described by a single transition, it implies that binding of the two molecules occurs independently of their relative distance. This applies only if the diffusion is much faster than the time to overcome the potential barrier of a chemical reaction (the "well-mixed" assumption [7]).

1.2 Memorylessness of the waiting time distribution

An important property of the exponential distribution is its memorylessness. In fact, it is the only continuous distribution with that property. Let's denote the time to the first reaction as T - an exponentially distributed random variable. Suppose that the reaction did not occur during past T_1 seconds. The memoryless character of the exponential distribution implies that the probability to wait for another T_2 seconds before the first reaction commences, is the same as the probability to wait T_2 seconds from the initial time T_1. In other words, the random process *does not remember* for how long it failed to produce an event.

In the limit of large numbers of reactants the memorylessness is also reflected in the Law of Mass Action. Consider a simple decay of N molecules of species A ($A \rightarrow \varnothing$). In the stochastic description, the time to decay for a single molecule is distributed exponentially with an average time τ. The time for the first decay event in an ensemble of N molecules, and therefore N reactions, can be computed by multiplying probabilities. The resulting distribution for the first event is also exponentially distributed, but with a smaller average, τ/N. This holds only for the exponentials. If the distribution of waiting times was not memoryless, the dependency of the average first event time (the mean first-passage time) on the number of sources would be more complicated.

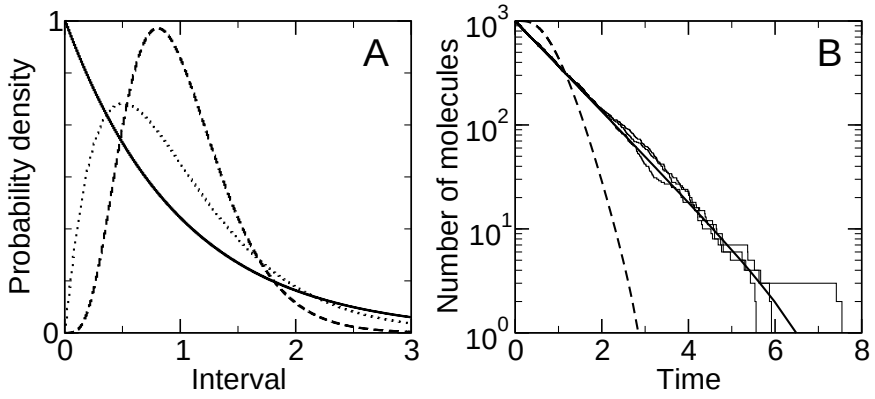

Figure 1: The effect of the peaked interarrival time distribution on a decay reaction. (A) Waiting time distributions for a Poisson process (exponential distribution - solid line), the sum of two exponentially distributed random variables (gamma distribution with $k = 2$ steps - dotted), and the sum of five variables (gamma with $k = 5$ steps - dashed). All distributions have the same mean equal 1. (B) Stochastic dynamics of a decay reaction, $A \rightarrow \varnothing$. Waiting times are drawn from the exponential distribution (solid line) and from a 5-step ($k = 5$) gamma distribution of the same mean equal to 1 (dashed). Averages are taken over 100'000 trajectories. Additionally, three arbitrary trajectories are drawn for the exponential case (light solid).

1.3 Non-exponential waiting time distributions

Under what circumstances does the distribution of interarrival time loose its memoryless property? A non-exponential distribution arises if a reaction results from a sequence of transitions, e.g. sequential degradation of mRNA transcripts [9]. While each of the steps may have exponential waiting times the resulting distribution is peaked. It is a gamma distribution characterized by a shape parameter k, corresponding to the number of steps in a sequence, and the mean waiting time τ for each of the steps. If the mean time to complete the whole sequence is kept fixed, increasing the number of steps results in a distribution with a smaller variance (Fig. 1A). Hence, compared to a single-step process, fluctuations in the interarrival time are lower in a process consisting of steps. As a result, a peaked distribution results in a more effective overall degradation (Fig. 1B).

The computation of stochastic processes with non-exponential waiting times (non-Markovian) becomes more cumbersome. In case of the decay scheme mentioned above, it is enough to model each of the steps explicitly

with exponential waiting times between events. A standard kinetic Monte Carlo simulation technique like Gillespie's direct method can be used [5]. What if the waiting time distribution is not a straightforward result of a sequence of steps? This is the case if a reaction is affected by diffusion. In fact, such a reaction is also a sum of small steps (imagine a lattice and think of diffusion as a first order reaction between two sites of the lattice [5]). Iinstead of a single chain there are infinitely many possible paths of different lengths. Depending on the location of the target, the contribution of short, approximately direct paths, and very long ones may differ.

2 The effect of diffusion on the search time

Diffusion can be a limiting processes for the response time of a signaling network especially when the copy numbers of signaling molecules are low. It is under those conditions that non-exponential waiting times become important. To what extent can they be approximated by exponential distributions? We shall consider a generic model of a prokaryotic two-component signaling [6]. In order for a cell to respond to an external signal, a response regulator (a transcription factor) diffuses to the membrane where it is activated by a receptor. After phosphorylation, it engages in a second random walk and searches for a binding site on the DNA located in the central area of the cell. We shall focus on this second diffusion process. Using the first-passage theory we investigate how diffusion affects the distribution of times to reach the target, the first-passage time.

In 3-D, the time to travel a distance L from the initial position by a single, freely diffusing molecule is $L^2/6D$ time units (D is the diffusion constant). However, this relation cannot estimate a diffusion-limited search for a binding site; it describes the time to reach a sphere of radius L. The model we consider is a single molecule in a sphere of radius R. It initiates at the boundary and searches for a spherical target of radius a in the center. Solving the Smoluchowski diffusion equation with appropriate boundary conditions we obtain the first-passage time (FPT) probability density function (*pdf*) [10]. The FPT *pdf* is a normalized histogram of the time to reach the target for the first time. It is the analog of the exponential next reaction time distribution described in Section 1.2. The mean of this distribution, the

mean first-passage time (MFPT) takes a simple form

$$\tau_{target} = \frac{V}{K_D}\left(1 - \frac{a}{r_0}\right) + \frac{a^2 - r_0^2}{6D}, \tag{1}$$

where $K_D = 4\pi aD$ (units of volume per time, a is the radius of the DNA target) is the Smoluchowski diffusion-limited rate constant, and r_0 is the initial position. In our case, we set $r_0 = R$. If the size of the absorbing target relative to the volume decreases, the first term, $V/K_D \equiv \tau_D$, dominates the mean. Below we explain the origin of this very useful fact.

2.1 Smoluchowski diffusion-limited rate constant

Smoluchowski obtained K_D as the rate of the diffusion-limited annihilation reaction $A + B \rightarrow \varnothing$ [10]. In his model the infinite sea of diffusing A's is fixed at the low concentration to avoid collisions between them. The steady-state flux (the rate) is computed at the fixed absorbing sphere B where molecules annihilate. The ratio V/K_D (units of time) denotes the frequency of arrivals at the absorbing target in an arbitrary volume containing diffusing molecules.

What is the distribution of interarrival times in the Smoluchowski model? In the steady-state, positions of the molecules are random. The average trajectory is so long that the memory of the initial distance is lost entirely. Thus, we expect that arrivals occur at exponentially distributed intervals. If the (imaginary) radius around the target is such that, on average, only one molecule diffuses in V, then the model of Smoluchowski corresponds to our problem where a single molecule starts at a distance R from the absorbing sphere.

2.2 The memoryless character of a diffusion search

Fig. 2 illustrates the dependence of τ_{target} on the target radius. To answer the question why τ_D approximates the MFPT only for small targets, we need to realize under what circumstances the FPT *pdf* of our search model converges to the exponential.

The insets of Fig. 2 contain the FPT *pdfs* for extreme values of the target size. Both functions are peaked, a feature typical for the distribution of the time required to cover a given distance by a diffusing molecule (the most notable example being the inverse Gaussian distribution). Plotting the

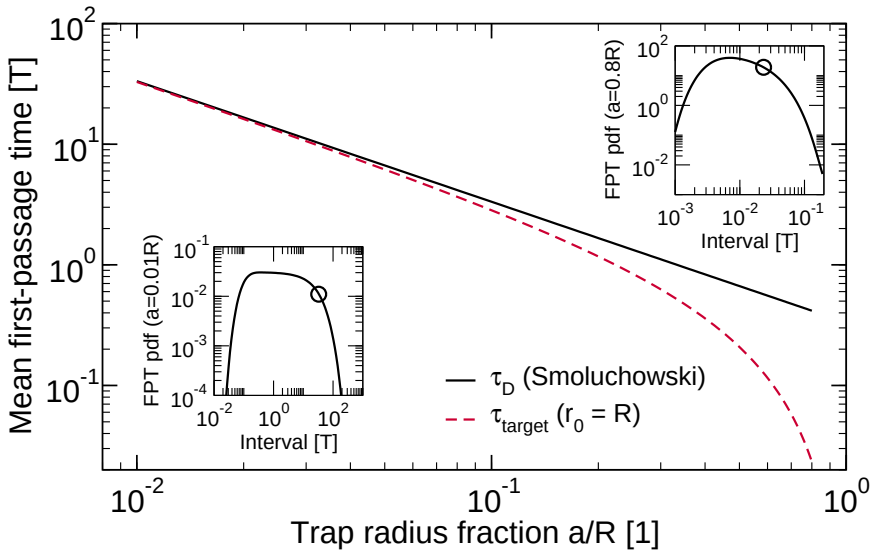

Figure 2: Two estimates of the MFPT to the absorbing sphere as function of the trap radius a: inverse of the Smoluchowski diffusion-limited rate constant τ_D (solid line), exact solution for a molecule starting at the outer boundary, $\tau_{target}(r_0 = R)$ (dashed). A single molecule diffuses in a sphere of radius R (reflective boundary) with diffusion constant $D = 1\,L^2/T$. Insets: FPT *pdf* for two extreme trap sizes, $a = 0.01$ and $0.8\,R$. Curves obtained by numerical inversion of the Laplace transform of the analytical flux at the absorbing target.

distributions in the log-log scale reveals two time scales. The left part of the *pdf* corresponds to the shortest time required to reach the target from R. The right part indicates the duration of the longest search trajectories.

The two *pdfs* differ significantly with respect to their variance. This reflects the obvious fact that the search of the small target involves a significant amount of long trajectories which possibly cover a large volume before reaching the trap. For such trajectories a slight variation in the initial distance from the target does not affect the time to reach it. The situation is different in case of a large trap. The time to reach it is short enough for a trajectory to be heavily dependent on the starting point. One could say that *it is harder to get lost, when the target is close.*

A large variance of the *pdf* for the small target results from long trajectories without the memory of the initial point. This brings us inevitably to the exponential distribution, the only memoryless continuous *pdf*. As shown in Fig. 3A the *pdf* for a single molecule can be very well approximated by the

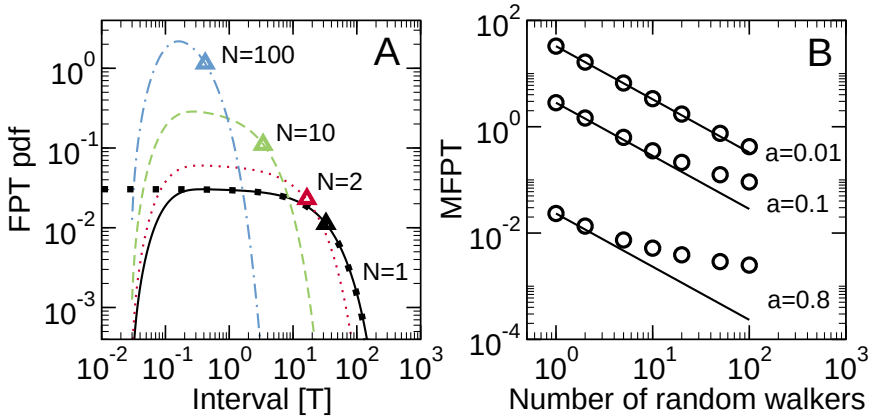

Figure 3: (A) The FPT *pdf* for searching the inner sphere of radius $a = 0.01R$ for a different number N of molecules starting simultaneously at a distance R. For a single walker, $N = 1$, the pdf is approximated by the exponential (dots). Triangles denote the mean of the distributions. (B) The MFPT as function of the number of molecules for different target sizes a. Solid lines denote the exponential-like scaling $1/N$, circles are computed using the order statistics for the analytical result.

exponential function with mean τ_{target} computed from Eq. 1. The difference in the left-most part is not significant for the mean, as it constitutes less than 1% of the total area under the FPT *pdf*.

From the order statistics we obtain the *pdf* for many independent walkers [8]. Increasing the number of molecules enhances the peaked character of the distribution. However, as seen in the panel B, the scaling of the mean with the inverse of the number of walkers holds well up to 100 independent walkers if the target is small.

3 Discussion

The property of memorylessness of the next reaction time in a chemical reaction is a very desired feature for modeling. It allows to construct a Markov chain for a stochastic process and to describe the evolution of the system by the CME. However, waiting times of some processes fail to obey exponential dependence. One of them is a reaction limited by diffusion. Understanding conditions under which waiting times can be approximated by the exponential distribution alleviates the necessity to perform costly spatial simulations.

Using the first-passage theory and the Smoluchowski theory of diffusion-limited kinetics we investigated a random walk towards a spherical target. This process is, for example, part of the response pathway in a two-component signaling network. We find that for a small target (relative to the total volume), the first-passage time probability density function (FPT *pdf*) can be very well approximated by an exponential function. Hence, the scaling of the mean first-passage time (MFPT) for many diffusing molecules as $1/N$ is possible. For a large target, a correction is required for the MFPT approximated by the Smoluchowski diffusion-limited rate constant. This may be relevant in case of cells with large organelles, e.g. vacuoles.

The inverse of the steady-state flux of molecules towards the absorbing target approximates the MFPT of a single diffusing molecule. This holds only if the random walk trajectory is memoryless; well approximated by the exponential FPT *pdf*. Usually the steady-state flux can be computed much easier than a dynamic property as the MFPT.

In this work we apply the first-passage theory to describe the distribution of times to reach the target. A detailed description like this is required when events are rare and involve single molecules, e.g. a binding site on DNA [6]. In a signaling network, the first-passage time estimates the time to respond to the external signal and turn on the gene expression. Variation in this time transmits directly to downstream processes if the lifetime of the cell is too short to average over fluctuations [11, 12]. Mean first passage time theory has an important role to play in cell biology. It allows to understand the design and physical constraints of biological networks in much the same way as the analysis of noise [13] influences our thinking about molecular networks.

4 Acknowledgments

I would like to thank J.G. Blom and F.J. Bruggeman for valuable comments on the manuscript. This research was supported by the Netherlands Organisation for Scientific Research, project NWO-CLS-635.100.007 as part of the Computational Life Science program.

References

[1] Kærn, M., *et al.* (2005) Stochasticity in gene expression: from theories to phenotypes. *Nat Rev Genet* **6**:451-64.

[2] Kaufmann, B.B. & van Oudenaarden, A. (2007) Stochastic gene expression: from single molecules to the proteome. *Curr Opin Genet Dev* **17**:107-12.

[3] Rao, C.V., Wolf, D.M., & Arkin, A.P. (2002) Control, exploitation and tolerance of intracellular noise. *Nature* **420**:231-7.

[4] van Kampen, N.G. (1997) *Stochastic Processes in Physics And Chemistry* (North-Holland).

[5] Dobrzyński, M., *et al.* (2007) Computational methods for diffusion-influenced biochemical reactions. *Bioinformatics* **23**:1969-77.

[6] Vidal Rodríguez, J., Dobrzyński, M. & Bruggeman, F.J. (2008) Optimal prokaryotic two-component signaling at low molecule numbers. *In preparation.*

[7] Gillespie, D.T. (1992) A rigorous derivation of the chemical master equation, *Physica A* **188**:404-25.

[8] Yuste, S.B., Acedo, L., & Lindenberg, K. (2001) Order statistics for d-dimensional diffusion processes *Phys Rev E* **64**:052102-1-052102-4.

[9] Pedraza, J.M. & Paulsson, J. (2008) Effects of molecular memory and bursting on fluctuations in gene expression. *Science* **319**:339-43.

[10] Redner, S. (2001) *A Guide to First-Passage Processes* (Cambridge University Press).

[11] Yu, J., *et al.* (2006) Probing gene expression in live cells, one protein molecule at a time. *Science* **311**:1600-03.

[12] Dobrzyński, M. & Bruggeman, F.J. (2008) Elongation dynamics shape bursty transcription and translation. *Submitted to Proc Natl Acad Sci USA.*

[13] Paulsson, J. (2004) Summing up the noise in gene networks. *Nature* **427**:415-18.