Available online at www.sciencedirect.com

ScienceDirect

Physics of Life Reviews 28 (2019) 88–91

PHYSICS of LIFE reviews

www.elsevier.com/locate/plrev

Comment

Extreme fluctuation dominance in biology: On the usefulness of wastefulness

Comment on “Redundancy principle and the role of extreme statistics in molecular and cellular biology” by Z. Schuss, K. Basnayake and D. Holcman

Igor M. Sokolov

Institute of Physics and IRIS Adlershof, Humboldt-University Berlin, Newtonstr. 15, 12489, Berlin, Germany

Received 1 March 2019; accepted 12 March 2019
Available online 15 March 2019
Communicated by M. Frank-Kamenetskii

There are several reasons which urged me to write this comment to the Review by Z. Schuss, K. Basnayake and D. Holcman [1]. The main one is to put important statements of the review into a broader context, namely to connect them to a concept of fluctuation dominance, once coined in chemical kinetics, and to put them in very simple words, which the authors have only partly done themselves. A personal reason is to give my tribute to Prof. Zeev Schuss who passed away last year, so that the work I comment on is probably his last.

The machinery of the cell is based on complex chemical reactions. This applies to almost all aspects of cell functioning: to its metabolism, to its genetic regulation, and to signaling processes defining cellular responses to changes in their environment. A picture of a cell as a tiny chemical factory is wide-spread and often transmitted to children. The picture of a well-organized and controlled factory production is however misleading to a large extent: At microscopic scales the machinery of the cell has to struggle fluctuations, which are not merely considerable, but huge. The same is also often true at mesoscopic scales of cells. In spite of these fluctuations the functioning of the cell is almost deterministic: each of its parts does its job reliably, and it does it on time.

In a batch production of chemicals, which often takes place in well-stirred reactors, the course of the reaction, and its final outcome may be well-predicted by systems of ordinary differential equations modeling elementary reactions in the reacting mixture. This old approach, called classical kinetics, is based on the law of mass action. This one assumes the reactions to be reversible and connects reaction rates (coefficients in front of combinations of the reactants’ concentrations) with equilibrium properties of the system. This theory, based solely on the means, fundamentally relies on well-mixedness and on the large number of reacting molecules in the system. The reliability of predictions foots on the law of large numbers and on the central limit theorem. Also in molecular and cellular biology one often investigates important cellular processes within simple schemes (“A promotes B and represses C”) which can be combined into complicated integrated networks resembling the logical circuitry of computers [2].

DOI of original article: https://doi.org/10.1016/j.plrev.2019.01.001.

E-mail address: igor.sokolov@physik.hu-berlin.de.

https://doi.org/10.1016/j.plrev.2019.03.003

The important difference of a cell compared to a chemical factory is that the reactions taking place inside the cell do not proceed in well-stirred environments: one cannot grow a chick from a scrambled egg. However, also in these cases the law of large numbers and central limit theorems may lead to the possibility of essentially deterministic description within the reaction-diffusion equations. The main inputs of such theories are, however, the same reaction rates as before.

Elementary acts of simple reactions are either monomolecular (something happens with a particle A even if it is left alone) or bimolecular: something happens only if A meets B. The rate of this elementary reaction derives from the mean time the reactants need to meet and to react. In order to react with B, A has to move in space in all cases when it is not produced on the spot, and to approach B at necessary position and orientation. Already this first, transport, stage preceding the reaction act itself, limits the reaction rate considerably. If the main transport mechanism at small scales is diffusion, the corresponding rates derive from the mean first encounter times of A with B in a diffusive motion. The oldest variant pertinent to three dimensional homogeneous reacting mixture was put forward by Smoluchowski in 1916, which work may be considered as one of the groundbreaking contributions to modern physical chemistry. The discussion of this classical picture in [1] is somewhat incomplete, since it follows immediately after discussing monomolecular conversion and misses to stress the bimolecular nature of the Smoluchowski rate.

Fluctuations around the means may considerably change the kinetics, especially when some amplification (or, speaking chemically, autocatalytic) stages are involved. In such cases the typical outcome of the reaction will be shifted from the predictions of the classical kinetics, but still be captured within the stochastic reaction schemes: the chemical master equations. If analytical treatment gets awkward, stochastic algorithms, like the celebrated Gillespie's algorithm, may do the job [3]. Here, in specific cases, the concepts of large deviations theory are applicable [4] and underline the connection of the processes governing e.g. cell decision making [5] with yet another type of limit theorems, now for large deviations. Also for networks of chemical reactions the limit theorems help, although they do not necessarily lead to simple deterministic limits [6].

For irreversible reactions the spacial aspect may however lead to situations when the continuous deterministic schemes fail badly. In the absence of mixing particle number fluctuations always present on the microscopic scales may get of imminent importance when amplified and carried over to macroscopic scales. This is, for example, the case at late stages of irreversible reactions in homogeneous systems [7] when the reacting background (representing the mean behavior) fades away, or in leading edges of the reaction fronts, where the amount of the invading species is necessarily small by itself [8]. In these cases small microscopic fluctuations lead to kinetics which differ drastically from what can be predicted based on the means. Such regimes were termed "fluctuation-dominated" [9], but, in the domain of chemical kinetics, still constitute a relatively exotic class of processes. In some (but not all) such cases the situations can be described within the systems of reaction-diffusion equations with random initial conditions and noise, and the space-resolved stochastic schemes may still provide reasonable simulation tools. The rate picture may fail fully, or may still be qualitatively correct, but the values of the rates do not follow from equilibrium considerations or form a more detailed picture of molecular encounters discussed above. Note that fluctuations discussed above still fall under the concept of typicality, so that what we called "fluctuation dominated" up to now is essentially a typical behavior when the average background fades away or when these fluctuations are slightly amplified.

At small scales the main process which governs the transport in biological cells is diffusion, sometimes in combination with active transport. In biology one of the reaction partners (a binding site, a receptor) is typically immobile or at least slow. This one will be associated with species B and called a target in what follows. Its mobile reaction partner A will be called a searcher. In many cases there is only one or few targets in the system. The number of searchers may be low or large. We note that the overall reaction picture is not necessarily pertinent to molecular entities, and the target and the searcher may be something else, say an ovum and a spermatozoon.

The distribution of the first reaction times of a diffusing particle with an immobile target may be extremely broad, with its body stretching over several orders of magnitude in time. Moreover, it is strongly skewed, with its mode typically corresponding to much shorter times than the mean [10,11], so that the search for a reaction partner may on the average be quite ineffective. One of the possibilities to cope with search inefficiency when the number of searchers is low, is to start search close to the target, and rely on the typical search time (say, given by the mode rather than by the mean). This corresponds to a so-called rapid search hypothesis [12]. This approach assumes a specific spatial organization of the system, is economical, but is prone to large fluctuations [13]. The noise enhancement in the situation discussed in [13] is connected with a series of binging-unbinding events of a repressor molecule to the DNA promoter site and the competition of this process with the binding of the RNA polymerase. The typical outcome of the reaction may still be described by an effective system of the ordinary reaction equations, but the corresponding rates should be changed considerably compared to the mass action law approximations, which corresponds to a rather moderate case of fluctuation dominance.

The possibility of multiple binding-undinding events is not always disadvantageous, but can lead to positive outcomes, as discussed in [14] on the example of Michaelis-Menten reactions. The mechanism is based on the fact that continuously trying to achieve an effect which is possible but improbable leads at the end to a success. Instead of trying to optimize a single search run, one can simply stop it, and restart from scratch [15]. In many cases, especially when the mean search time is large or formally infinite, this strategy (called "resetting") leads to a considerably faster search, and is essentially economical if the searchers are reused. The resetting strategy implies repeated search, so that new searching attempts are started sequentially.

The number of searchers does not have however to be small. The strategy of coping with fluctuations based on simultaneous use of many searchers in parallel is what the authors call the "redundancy principle". The idea is very simple. Let us assume the probability distribution of reaction times to be proper (each single searcher finds the target at the end of the day) and let $P(t)$ be the (cumulative) probability distribution function of reaction times t, i.e. the probability that the arrival time of a single particle is smaller than t. This function is supported on $t_{\min} \le t < \infty$ and, as we have already stated, may be extremely broad. If the number of searchers N is large, one of them, with a very high probability, will arrive at time quite close to $t_{\min}$, so that the search will be fast. Here we go away from the domain of applicability of a whatever central limit theorem or large deviations theory (which is still about means) into the extreme value statistics and the corresponding limit theorems for extremes, yet the third incarnation of large numbers: If the result of the reaction leads to a decision (a "first takes all" situation), all searchers coming later get redundant, together with all properties of $P(t)$ far from its lower bound. This is, what explains fast responses in the cases when the typical response, as predicted by the mean, should essentially be slow. As an example for such fast response the authors consider calcium release in the endoplasmatic reticulum at neuronal synapses.

Since the outcome of such reaction depends only on the shape of $P(t)$ close to its lower bound, i.e. has nothing to do with its bulk determining the mean arrival time defining the reaction rate, this is only this behavior that has to be calculated to considerable accuracy. This, as the authors show, is often easier than approximating the whole function with a homogeneous precision.

Yet another important statement, reflecting a very different aspect of processes dominated by the first arrival, is as follows. The diffusion (Brownian motion) of particles corresponds to extremely involved tortuous trajectories of formally diverging length. The particles which arrived first, practically do not jiggle, and their trajectories group around the shortest, geodetic paths leading from the particles' release site to their destination. This again shows how the redundancy produces a practically deterministic behavior.

Another example, on quite a different scale, is related to fertilization. The model for a sperm motion adopted by the authors is a simple rectilinear motion in a domain with solid boundaries (modeling the uterus) with random changes in the direction of motion when hitting a wall. The model neglects noise and spiraling produced by the flagella and may seem oversimplified, but taking into account further details of motion would not change considerably the main, qualitative, conclusions of the discussion. One spermatozoon is enough to fertilize the egg cell, and the enormous overweight in the number of produced spermatozoa compared to that of ova might seem to be a waist. This number is necessary to make a process fast and reliable enough in view of the relatively short lifetime of spermatozoa, and their complicated way through a viscous environment. Nature is here even not too much on the safe side: decreasing the amount of sperm by a factor of four might already cause infertility.

Above, we considered the cases when there is only one possible outcome of the competition. There are also situations when several different outcomes are possible. In this case it is not only the speed of response, but the variability of responses which is of interest. This plays a crucial role in the selection process of antibodies, where the work shows how the selection of a unique sequence of gene segments among many copies governed by the first encounter defines the building of a unique antigen receptor.

The principle of redundancy illustrates a general way in which life struggles with strong fluctuations which are present on all scales: Instead of using complicated (and costly) control mechanisms to suppress unfavorable noise, it might simply neglect the losses (if not too costly) and canalize favorable fluctuations. Interestingly enough, this fluctuation-dominated mechanism leads to the overall functioning modus which is extremely reliable and practically deterministic. This mechanism of relying on large numbers is very different form the one following from the central limit considerations, where the deterministic behavior appears via averaging: Here it ensues due to extreme events, which were rare if the number of searchers would be small. This is a principle a researcher should think about when trying to face the problem of robustness of results in spite of large fluctuations. Hence, it seems that the biology is full of examples of extreme fluctuation dominance, which the "normal" chemistry lacks.

There is one more example of the role of extreme value statistics in biology: this is about the time dependence of the length of the shortest telomere, which controls senescence. The authors conclude that the length of telomers, in particular, the shortest ones, does not reflect the age of the cell line, so that the number of cell divisions may be considerably larger than intuitively expected. This example is somewhat away from the main paradigm of the redundancy principle which, as the authors state, is that "many copies of a single object ... are not a waist, but have a specific function in living systems". Here the number of telomers is fixed by the genomic organization, none of them is redundant, and the cell dies when the first one gets too short to perform its job. To my taste, this example falls into the usual concept of fluctuation dominance: the theory taking into account fluctuations leads to longer typical lifetimes than a deterministic theory would predict.

The review finishes with a list of to dos and dont's. The important don't concerns the applicability of reaction-diffusion equations and Gillespie-like algorithms in simulations of activation processes in context of transient responses, and states that a careful discussion of Brownian paths and extreme events is necessary. The authors clearly underline the ways for analytical solutions in the case of simple geometries, where these are feasible, and rely on straightforward Brownian (or detailed stochastic) simulations both to verify them or to obtain the corresponding behaviors when the analytical solution gets impossible to find. Such simulations are definitely costly, which the authors do not discuss. I have a feeling that here much work should be done to provide effective sampling strategies for the reaction time distribution's tails, which might be an interesting direction for the future work.

References

[1] Schuss Z, Basnayake K, Holcman D. Redundancy principle and the role of extreme statistics in molecular and cellular biology. Phys Life Rev 2019;28:52–79. https://doi.org/10.1016/j.plrev.2019.01.001 [in this issue].

[2] McAdams HH, Shapiro L. Circuit simulation of genetic networks. Science 1995;269:650.

[3] McAdams HH, Arkin A. Stochastic mechanisms in gene expression. Proc Natl Acad Sci USA 1997;94:814.

[4] Agazzi A, Dembo A, Eckmann J-P. Large deviations theory for Markov jump models of chemical reaction networks. Ann Appl Probab 2018;28:1821.

[5] Balázsi G, van Oudenaarden A, Collins JJ. Cellular decision making and biological noise: from microbes to mammals. Cell 2011;144:910.

[6] Anderson DF, Kurtz TG. Continuous time Markov chain models for chemical reaction networks. In: Koeppl H, Setti G, di Bernardo M, Densmore D, editors. Design and analysis of biomolecular circuits. New York, NY: Springer; 2011.

[7] Ovchinnikov AA, Zeldovich YaB. Role of density fluctuations in bimolecular reaction kinetics. Chem Phys 1978;28:215.

[8] Kessler DA, Levine H. Fluctuation-induced diffusive instabilities. Nature 1998;394:556.

[9] Kang K, Redner S. Fluctuation-dominated kinetics in diffusion-controlled reactions. Phys Rev A 1985;32:435.

[10] Godec A, Metzler R. Universal proximity effect in target search kinetics in the few-encounter limit. Rhys Rev X 2016;6:041037.

[11] Grebenkov DS, Metzler R, Oshanin G. Towards a full quantitative description of single-molecule reaction kinetics in biological cells. Phys Chem Chem Phys 2018;20:16393.

[12] Kolesov G, Wunderlich Z, Laikova ON, Gelfand MS, Mirny LA. How gene order is influenced by the biophysics of transcription regulation. Proc Natl Acad Sci USA 2007;104:13949.

[13] van Zon JS, Morelli MJ, Tănase-Nicola S, ten Wolde PR. Diffusion of transcription factors can drastically enhance the noise in gene expression. Biophys J 2006;91:4350.

[14] Reuveni S, Urbakh M, Klafter J. Role of substrate unbinding in Michaelis-Menten enzymatic reactions. Proc Natl Acad Sci USA 2014;111:4391.

[15] Pal A, Reuveni S. First passage under restart. Phys Rev Lett 2017;118:030603.