(Article begins on next page)

# Simulating Populations of Protocells
# with Uneven Division

Martina Musa[1], Marco Villani[1,2(✉)], and Roberto Serra[1,2]

[1] Department of Physics, Informatics and Mathematics,
University of Modena and Reggio Emilia, Modena, Italy
{marco.villani, rserra}@unimore.it
[2] European Centre for Living Technology, Ca' Foscari University, Venice, Italy

**Abstract.** Protocells should be similar to present-day biological cells, but much simpler. They are believed to have played a key role in the origin of life, and they may also be the basis of a new technology with tremendous opportunities. In this work we study the effect of uneven division processes on the synchronization of the duplication rates of protocells' membrane and internal materials.

**Keywords:** Protocell · Protocell populations · Models · Synchronization
Replicators

## 1 Introduction

Protocells should be similar to present-day biological cells, but much simpler [1, 2]. They are believed to have played a key role in the origin of life, and they may also be the basis of a new technology with tremendous opportunities (see e.g. the books [1, 3] and further references quoted therein). Among various candidate protocell architectures, those based on lipid vesicles are particularly promising since they can spontaneously undergo fission, giving rise to two daughter protocells. Protocells should also contain a self-replicating set of molecules (the "replicators"): their composition should also affect the growth and replication rate of the container, so that some kind of competition can take place between cells with different chemical compositions [4].

The simplest case is that of even division, where a vesicle splits into two identical daughter cells. In order to assure sustainable growth of a population of protocells, it is necessary that the duplication rate of the replicators be equal to that of the lipid container. It has been shown in a series of papers [5–10] that this synchronization takes spontaneously place, generation after generation, under a wide set of hypotheses concerning the protocell architecture and the kinetic equations for the replicators, and that it is robust with respect to random fluctuations. This is indeed a beautiful example of dynamical self-organization.

It has however also been observed that there are other ways in which lipid vesicles can divide. In this paper we consider a case (inspired by the "budding" processes [11]) where the vesicle splits in two daughter vesicles of different size, a "large" one and a "small" one. Other types of division, like e.g. those due to extrusion processes, might also be investigated. We suppose that, when a critical size has been reached, the protocell

splits into two: the daughters will inherit a fraction of both the lipid container and the replicators. In this work also consider the important situations of (noisy) uneven division and of not constant splitting thresholds.

The paper is organized as follows. In Sect. 2 we introduce the investigated protocell models and discuss the uneven division process. In Sect. 3 we present the experimental setting and the simulations, and discuss the simulations main results. Finally, in Sect. 4 we summarize the main paper results.

## 2   The Protocell Model

### 2.1   The Protocell Model

As anticipated, several different protocell "architectures" have been suggested [1, 12–14]. Many architectures are based upon lipid vesicles, where an aqueous internal environment is separated from the external water phase by a lipid bilayer, similar to those of existing biological cells. Vesicles form spontaneously under appropriate conditions, and it is known that they are able to split giving rise to two (or more) daughter cells [15, 16]. The different architectures are based on different hypotheses about the chemical composition of the protogenetic material (e.g., nucleic acids, or polypeptides, or even lipids themselves) and about the place where the action, i.e., duplication of genetic molecules and growth of the lipid container, takes place (in the internal environment, in the membrane, at the interface, or some combinations of the two) [13].

One might therefore be tempted to guess that no unified treatment is possible, however this turns out not to be the case: indeed it has been shown that at least the problem of synchronization lends itself to be dealt with using abstract models of quite broad applicability [3].

In this paper we consider the case where there is a single replicator, represented by some chemical species, and let its quantity (i.e., number of moles) be denoted by $X$. Let also $C$ be the total quantity of "container" (i.e., lipid membrane forming vesicles or micelles) and $V$ its volume, which is equal to $C/\rho$ (where $\rho$ is the density, assumed to be constant).

We assume that the $X$-molecule favors the formation of the container materials (as for example in [17, 18]). Some models imagine that only the $X$-molecule fraction near the external surface is effective in doing so, since the container precursors are found outside the protocell; other models envisage that these materials could pass through the membrane allowing in such a way an active role also to the internal $X$-molecule fraction. In [8] we demonstrate that these frameworks actually show equivalent behaviors: so, in this paper we follow the design that suppose permeable membranes and inner $X$-molecule materials.[1]

So the catalytic activity of the $X$-molecules favor the growth of the lipid container, which provides in turn the physical conditions appropriate for the replication of the $X$-

---

[1] Note that even in this case it is possible that not all the internal X-molecules be active in supporting the container building; however, in [19] we show that also this difference does not significantly affect the process leading to the synchronization of the X-molecules and container reproduction rates.

molecules, without being however a proper catalyst. Because of its effect on the lipid container and of its ability in maintaining its presence during generations, the chemical species X loosely acts as a sort of "genetic material", and we could refer to it with the term "genetic memory molecule", or GMM for short.

If we follow the assumptions already used and discussed in [10], namely:

1. spontaneous container formation is negligible, so that only the catalyzed term matters
2. the precursors (both of container and X-molecule) are buffered
3. the membrane vesicle is thin, so the volume of the lipid membrane (and as consequence the amount of container C) is approximately proportional to its surface
4. diffusion is very fast, so in each phase the concentrations can be assumed to be homogeneous
5. the protocell breaks into two identical daughter units when its container mass reaches a certain threshold[2]
6. the shape of the mother protocell, as well as those of her daughters, are all spherical[3]
7. the rate limiting step which may appear in the replicator kinetic equations does not play a significant role when the protocell is smaller than the division threshold.

The simplified equations for the total quantities of lipid container C and replicator X during the continuous growth phase from an initial condition to the critical lipid container mass $\theta$ become:

$$\begin{cases} \dfrac{dC}{dt} = \alpha C^{\beta-1} X \\ \dfrac{dX}{dt} = \eta C^{\beta-1} X \end{cases} \tag{1}$$

where $\alpha$ and $\eta$ are two positive constants denoting respectively the rate of self-replication of genetic molecules and the container growth, and the shape factor $\beta$ ranges between 2/3 for a micelle and 1 for a very thin vesicle [8].

As it was proved in [10], in order to determine whether there is a synchronization in the asymptotic time limit, one can limit oneself to consider the $\beta = 1$ case. The final result does not depend on $\beta$, while of course this parameter affects the speed with which it is approached: this is essentially a non-linear rescaling of time, useful to simplify the analysis. With this simplification, the basic equations (which are valid between two successive divisions) are then:

---

[2] The dropping of this hypothesis is one of the topics of this paper.

[3] This assumption is reasonable if we suppose that the flow of water is "fast" enough to allow us to consider the protocell as turgid, on the time scale of interest [20]. This implies that we do not describe here in detail the breakup of a vesicle into two, which certainly requires consideration of shape changes – that are supposed to be fast and to fall below the time scale of the relevant phenomena that the model describes. Moreover, we do not take explicitly into account osmotic effects (as for example in [21]) that might be relevant in the case of hypertonic or hypotonic environments.

$$\begin{cases} \dfrac{dC}{dt} = \alpha X \\ \dfrac{dX}{dt} = \eta X \end{cases} \tag{2}$$

## 2.2   Uneven Division

As anticipated, we assume that a protocell splits into two daughters when its membrane reaches a certain critical mass, in the following indicated as $\theta$. After splitting, one of the daughter cells inherits a fraction $\omega$ of the lipid container, while the other one inherits $1 - \omega$; the same happens to the GMM chemical species that are diluted in the membrane, if any.

In this paper we assume that the shape of the mother protocell, as well as those of her daughters, are all spherical. As it has been discussed elsewhere [3], this implies that some part of the mother's internal aqueous environment is lost in fission - and the same holds for the GMMs that are there diluted. So, only a part of these GMMs part is shared between the two daughters, which inherit L$\omega$ and L$(1 - \omega)$ respectively, L being the fraction of GMMs that is not lost in the bulk.[4]

This fraction can be calculated by simple geometric reasoning, if the concentration of the replicators is uniform in the internal water phase. Indeed, if $r$ is the protocell radius, the surface and the volume of a protocell are respectively $S = 4\pi r^2$ and $V = 4\pi r^3/3$, and the dependence of volume from the surface is:

$$V = \frac{4}{3}\pi r^3 = \frac{4}{3}\pi \left(\frac{S}{4\pi}\right)^{\frac{3}{2}} = \frac{4}{3 \cdot 4\pi \cdot 2\sqrt{\pi}}\pi S^{\frac{3}{2}} = \frac{S^{\frac{3}{2}}}{6\sqrt{\pi}} \tag{3}$$

The sum of the volumes of the two daughters $V_F$ therefore is:

$$V_F = \frac{\omega^{\frac{3}{2}}S^{\frac{3}{2}}}{6\sqrt{\pi}} + \frac{(1-\omega)^{\frac{3}{2}}S^{\frac{3}{2}}}{6\sqrt{\pi}} = \frac{S^{\frac{3}{2}}}{6\sqrt{\pi}}\left(\omega^{\frac{3}{2}} + (1-\omega)^{\frac{3}{2}}\right) = V_{mother}\left(\omega^{\frac{3}{2}} + (1-\omega)^{\frac{3}{2}}\right) \tag{4}$$

Consequently, the fraction of the mother protocell volume that is not lost in the bulk is:

$$L = \omega^{\frac{3}{2}} + (1-\omega)^{\frac{3}{2}} \tag{5}$$

This is also the fraction of GMMs of the initial protocell shared between the two daughters, in case of these GMMs are diluted in the mother's internal aqueous environments.

These rules determine the initial conditions of the two daughter cells at the next generation. The small one will need a longer time to reach the critical size and to undergo fission, while the larger one will be faster.

---

[4] Obviously, $L = 1$ in case of the GMMs are diluted in the membrane.

The continuous growth described by Eq. 2, starting from an initial condition where $C = C_i$ and $X = x_i$ up to the time $T = T_{div}$ when $C = \theta$ (i.e. when splitting takes place) can be analytically determined to be

$$
\begin{cases}
T_{div} = \dfrac{\ln\left(\frac{\eta(\theta - C_i)}{\alpha x_i} + 1\right)}{\eta} \\[3mm]
x_{div} = \dfrac{\eta(\theta - C_i)}{\alpha} + x_i
\end{cases}
\tag{6}
$$

where $x_{div}$ is the quantity of replicator at splitting time $T_{div}$.

Of course, in case of uniform and regular process of division (each progenitor regularly dividing into two descendants of size $\omega$ and $1-\omega$) a large daughter cell will also give rise to another large one, etc. This "pure lineage" of large cells will tend to synchronize, in a way similar to the case of even division. The same will happen for the pure lineage of small cells, although with a different division frequency. Therefore, for a pure $\omega$ lineage the following rule holds:

$$
\begin{cases}
T_{div} = \dfrac{\ln\left(\frac{1}{L\omega}\right)}{\eta} \\[3mm]
x_{div} = \dfrac{\eta\theta}{\alpha}\dfrac{1-\omega}{1-L\omega}
\end{cases}
\tag{7}
$$

A similar rule holds for the pure "$1-\omega$" lineage, by substituting $\omega$ with "$1-\omega$". Finally, we can compute the difference between the asymptotic division times and the GMMs' quantities at division time of these pure lineages:

$$
\begin{cases}
\Delta T_{div} = \dfrac{\ln\left(\frac{(1-\omega)}{\omega}\right)}{\eta} \\[3mm]
\Delta x_{div} = \dfrac{\eta\theta}{\alpha}\left(\dfrac{(L-1)(2\omega-1)}{1-L+L^2\omega-L^2\omega^2}\right)
\end{cases}
\tag{8}
$$

Note that (i) the difference between the asymptotic division times does not depend upon $L$ and that (ii) in case of even division these differences are equal to zero (only one lineage is present).
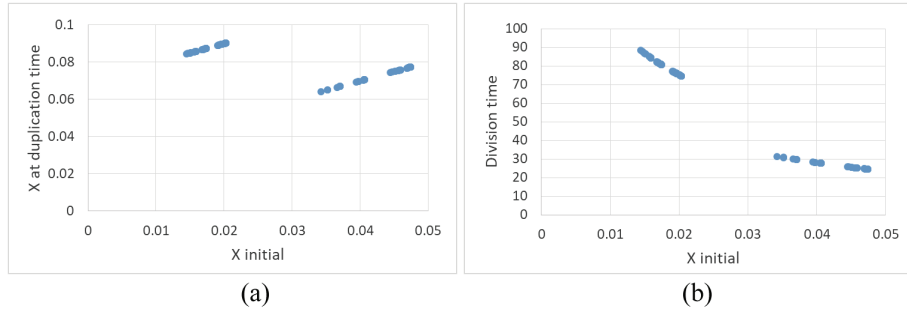
## 3   Population of Protocells

In the previous section, we derived the rules to compute the asymptotic division time and the GMMs' quantity at division time of the protocell pure lineages.

However, as generations increase, the fraction of cells belonging to the pure lineages declines, and most cells have both large and small cells among their ancestors. Indeed, after k generations the large cells will be k, while the total number of cells will be $2^k$, so the fraction of pure lineage cells will vanish in the long k limit.

An interesting question then concerns the distribution of division times after several generations: will there be, on average, a uniform distribution of fission events in time, or will there be some pace, at population level, in the fission processes?

We have therefore simulated the growth of such populations of protocells, originated from a single protocell. Because of computational limits (which anyway mimic real physical constraints) when the population size reach its maximum value each division implies the substitution of an already "born" protocell (stable population phase).

These substitutions are random, so the data are noisy. However, an interesting observation is that the data concerning both the fission intervals and the values of the replicators before fission are divided in two groups and they do not become homogeneous (Fig. 1). As expected, the division time is smaller in the case of the larger initial vesicles with a larger initial quantity of replicators (Fig. 1b). On the other hand, the smaller vesicles, with longer division times, synthesize a larger final quantity of



**Fig. 1.** (a) Quantity of replicators before division and (b) division time (i.e. the interval between two successive divisions) vs initial quantity of replicators.
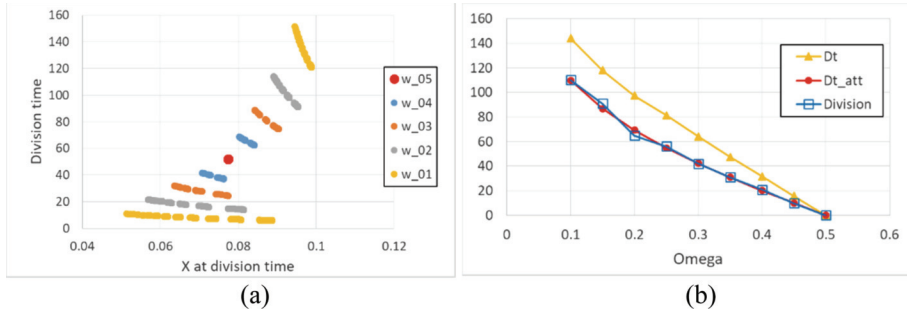


**Fig. 2.** (a) Populations of protocells with uneven fission exhibit two stable subgroups: the more uneven is the fission, the more distant are the characteristics of the subgroups. (b) Difference between the observed and the "theoretical" values for the maximum difference between division times, vs $\omega$. Dt is the measured maximum distance in division times, $Dt\_att$ is the "theoretical distance" defined in the text, Division is the width of the zero-level plateau, whose some instances are visible in Fig. 3a and c.

replicators (Fig. 1a). The bimodality of the distribution of division times and of the quantity of replicators at division time can also be directly observed in Fig. 3, which shows their (stable) probability distribution for different $\omega$. Therefore, after the growth phase the population of protocells shows the presence of two subgroups, composed respectively by protocells with relatively long lifespan that divide with high concentrations of GMMs, and protocells with relatively short lifespan that divide with low concentrations of GMMs. Each protocell divides into two (possibly uneven) daughters, but the two subpopulation are stable – in the sense that the cardinality of each group does not change in time.

It is also possible to analyze the difference between division times as a function of $\omega$ (remember that, given the geometrical hypotheses, $\omega$ determines also the fraction L of replicators that are not lost). Let us define the "theoretical distance" between division times as the difference between the division times of the two pure lineages, determined
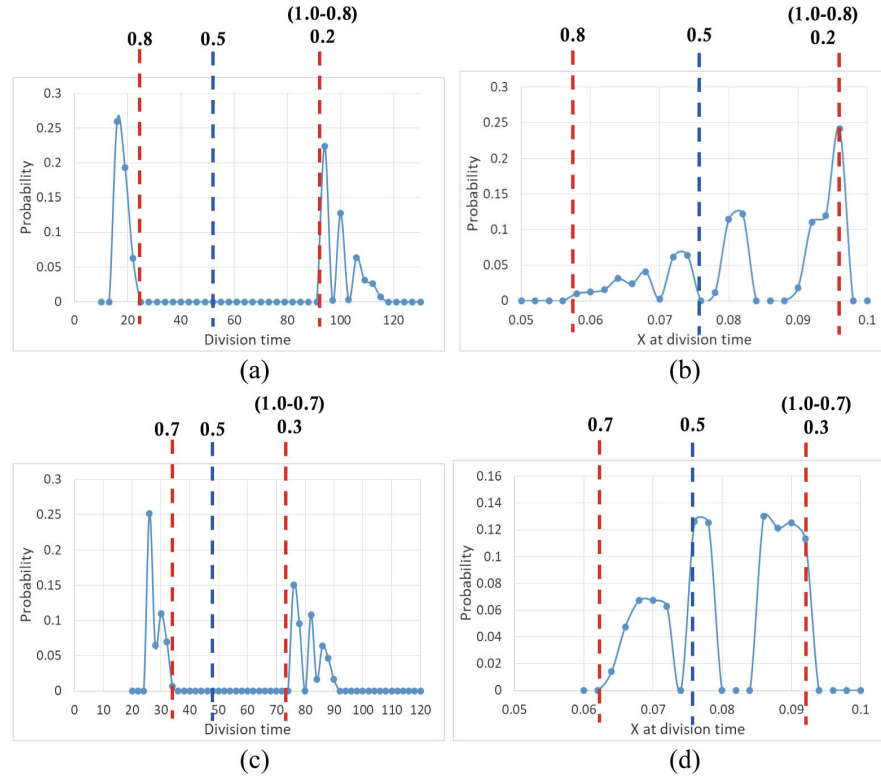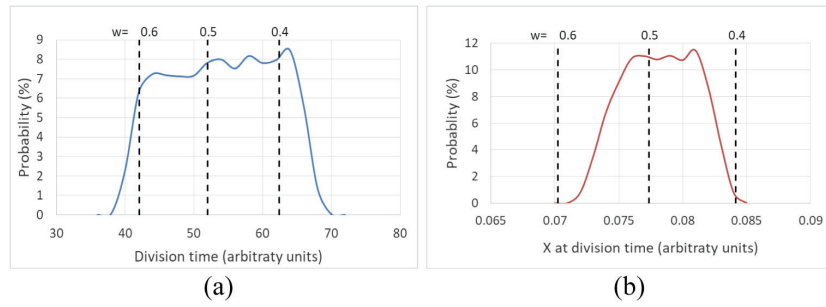


**Fig. 3.** Probability distribution of division times (a), (c) and probability distribution of $X$ quantities at division time (b), (d) observed in simulations; the first row refers to $\omega = 0.2$, whereas the second row refers to $\omega = 0.3$. Red vertical lines indicate the division times and the $X$ quantities at division time of the asymptotic pure lineages, blue vertical lines indicate the same for even division. Note that the division times of the pure lineages precisely individuates the extremities of the empty space dividing the two protocell subgroups, whereas their $X$ quantities at division time embrace the distribution of the protocell populations. (Color figure online)

analytically in Eq. 8. Surprisingly enough, in the case of uneven division one sometimes observes that the maximum of the actual difference between division times typically exceeds this value, as shown in Fig. 2a. Indeed, the theoretical distance closely approximates the zero-level plateau (some instances are visible in Fig. 3a and c).

Till now protocells divide in two "new" individual each one inheriting respectively a fraction equal to Lω and L·(1−ω) of the progenitor's materials. However, this is a very idealized situation: real splitting processes are noisy (or even very noisy), and each



(a)    (b)

**Fig. 4.** Probability distribution of division times (a) and probability distribution of $X$ quantities at division time (b) observed in simulations, when at division time each protocell divides in two "new" protocells inheriting respectively a fraction equal to Lω' and L·(1-ω') of the progenitor's materials, ω' being randomly extracted from a uniform distribution spanning within the range [0.4, 0.5] (a process that simulates a noisy even division). The vertical lines indicate the division times and the $X$ quantities at division time of pure lineages with ω respectively equal to 0.4 and 0.6 (the complementary size of a daughter protocell with ω = 0.4).
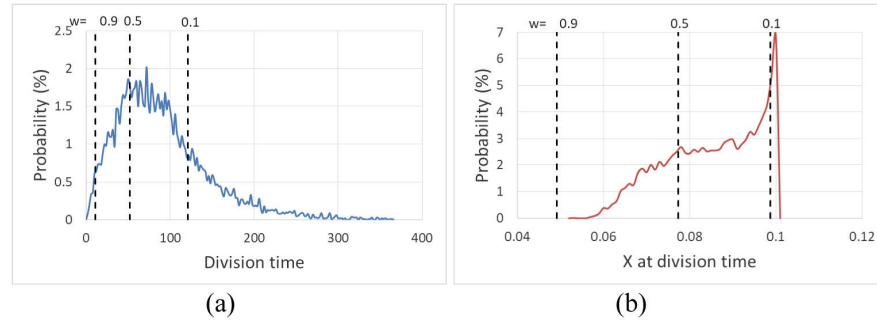


(a)    (b)
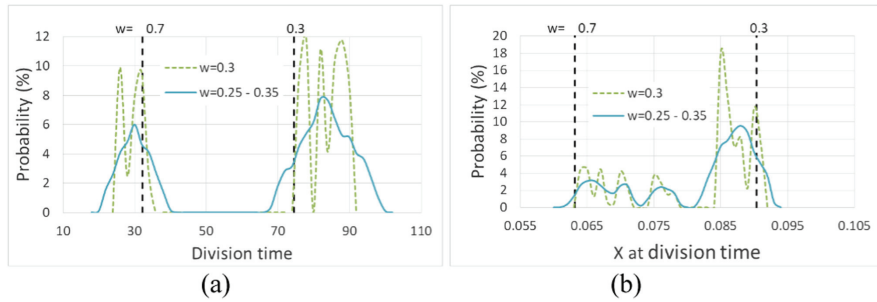
**Fig. 5.** Probability distribution of division times (a) and probability distribution of $X$ quantities at division time (b) observed in simulations, when at division time each protocell divides in two "new" protocells inheriting respectively a fraction equal to Lω' and L·(1-ω') of the progenitor's materials, ω' being randomly extracted from a uniform distribution spanning within the range [0.25, 0.3] (a process that simulates a noisy uneven division with ω = 0.3). The vertical lines indicate the division times and the $X$ quantities at division time of pure lineages with ω respectively equal to 0.3 and 0.7 (the complementary size of a daughter protocell with ω = 0.3).

division could (i) happen at thresholds more or less distant from θ and/or (ii) giving birth to slightly different descendants.

The effect of adding noise to even division is that of "blurring" the delta distributions of the asymptotic division times and of the $X$ quantities at division time (Fig. 4); a similar effect is observed in uneven divisions, where it is still possible however to re-observe the presence of the two previously individuated protocell subpopulations (Fig. 5).

Interestingly, a very "disordered" division process leads to an asymmetric distribution of division times (with a small fraction of protocell owning very long lifespans), that corresponds to a very sparse $X$ quantities at division times distribution. Note that the presence of very long lifespans allows the formation of protocells with relatively high final concentrations (Fig. 6).



(a)                                    (b)

**Fig. 6.** Probability distribution of division times (a) and probability distribution of $X$ quantities at division time (b) observed in simulations, when at division time each protocell divides in two "new" protocells inheriting respectively a fraction equal to $L\omega'$ and $L\cdot(1-\omega')$ of the progenitor's materials, $\omega'$ being randomly extracted from a uniform distribution spanning within the range [0.01, 0.5] (a process that simulates a very noisy uneven division). The vertical lines indicate the division times and the $X$ quantities at division time of pure lineages with $\omega$ respectively equal to 0.5, 0.9 and 0.1 (the complementary size of a daughter protocell with $\omega = 0.9$).

The effect of splitting events occurring at not constant protocell size (simulated by adding noise to θ, the membrane threshold quantity that induces the protocells instabilities leading to the protocell division) is similar to the ones just shown for noisy uneven division. A random size of splitting makes more "fuzzy" the delta distributions of the asymptotic division times and of the $X$ quantities at division time (Fig. 7); very high noise levels possibly override the presence of the two previously individuated protocell subpopulations, at least of one of the observed variables (Fig. 7d).
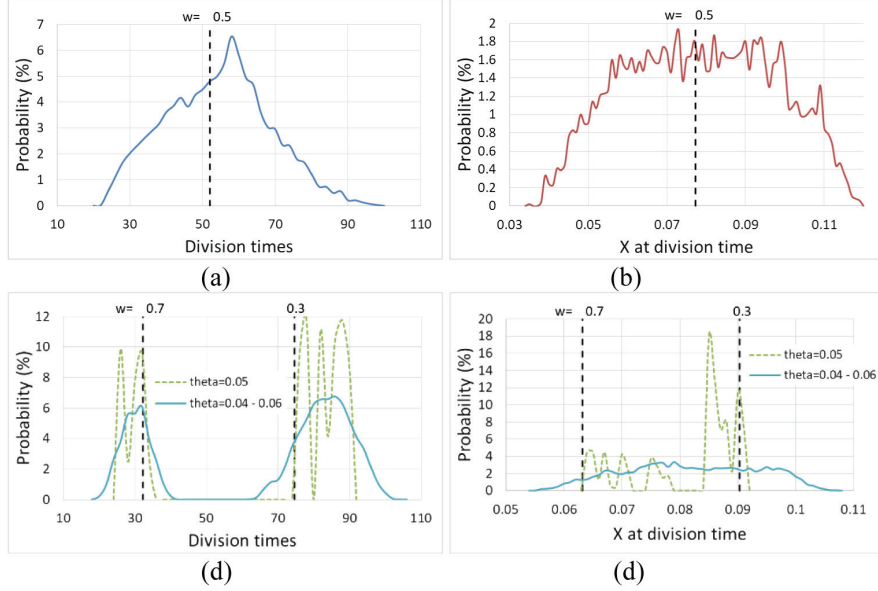
**Fig. 7.** Probability distribution of division times (a), (c) and probability distribution of *X* quantities at division time (b), (d) observed in simulations, for even division (first row) and uneven division (second row, uneven division with $\omega = 0.3$), in case of the membrane threshold quantity that induces the protocells instabilities leading to the protocell division is not constant. The vertical lines indicate the division times and the *X* quantities at division time of pure lineages with $\omega$ respectively equal to 0.5, 0.7 and 0.3 (the complementary size of a daughter protocell with $\omega = 0.7$).

## 4   Conclusions

Protocells should be similar to, but much simpler than biological cells. Protocell populations do not yet exist and mathematical models are therefore extremely important to address the key questions concerning their synthesis and behavior. Different protocell architectures have been proposed so, due to uncertainties about the details, high-level abstract models like those that are presented in this paper are particularly relevant. In this context, the problem of synchronization plays a particularly relevant role: indeed, growth and evolution of a population of protocells require that reproduction of the whole protocell and replication of its "genetic memory molecules" take place at the same pace.

Despite the fact that only "pure lineage" streams of protocells can rigorously synchronize (that is, reproduce the same amount of materials at precise and regular time intervals), in this paper we show that the macroscopic output of the random super-position of thousands of these processes is the presence within the protocell population of stable distributions of the relevant protocell variables. In case of uneven division

these distributions become bimodal, highlighting in such a way the presence of two stable subpopulations, the macroscopic consequence of the fact that protocells, when divide, split into two not symmetric descendants.

Further works will explore the effects of changing the protocell architecture on the regularity or on the shape of this very interesting macroscopic output.

# References

1. Rasmussen, S., Bedau, M.A., Chen, L., Deamer, D., Krakauer, D.C., Packard, N.H., Stadler, P.F. (eds.): Protocells. The MIT Press, Cambridge (2008)
2. Schrum, J.P., Zhu, T.F., Szostak, J.W.: The origins of cellular life. Cold Spring Harb. Perspect. Biol. **2**, a002212 (2010)
3. Serra, R., Villani, M.: A stochastic model of growing and dividing protocells. Modelling Protocells. UCS, pp. 105–147. Springer, Dordrecht (2017). https://doi.org/10.1007/978-94-024-1160-7_5
4. Smith, J.M., Szathmáry, E.: The Major Transitions in Evolution. W.H. Freeman Spektrum, Oxford (1995)
5. Serra, R.: The complex systems approach to protocells. In: Pizzuti, C., Spezzano, G. (eds.) Advances in Artificial Life and Evolutionary Computation, WIVACE 2014. Communications in Computer and Information Science, vol. 445, pp. 201–211. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-12745-3_16
6. Villani, M., Filisetti, A., Graudenzi, A., Damiani, C., Carletti, T., Serra, R.: Growth and division in a dynamic protocell model. Life **4**, 837–864 (2014)
7. Filisetti, A., Serra, R., Carletti, T., Villani, M., Poli, I.: Non-linear protocell models: synchronization and chaos. Europhys. J. B **77**, 249–256 (2010)
8. Carletti, T., Serra, R., Poli, I., Villani, M., Filisetti, A.: Sufficient conditions for emergent synchronization in protocell models. J. Theor. Biol. **254**, 741–751 (2008)
9. Filisetti, A., Serra, R., Carletti, T., Poli, I., Villani, M.: Synchronization phenomena in protocell models. BRL. Biophys. Rev. Lett. **3**(1/2), 325–342 (2008)
10. Serra, R., Carletti, T., Poli, I.: Synchronization phenomena in surface reaction models of protocells. Artif. Life **13**, 1–16 (2007)
11. Svetina, S.: Vesicle budding and the origin of cellular life. ChemPhysChem **10**, 2769–2776 (2009)
12. Solé, R.V., Macía, J., Fellermann, H., Munteanu, A., Sardanyés, J., Valverde, S.: Models of protocell replication. In: Rasmussen, S., Bedau, M.A., Chen, L., Deamer, D., Krakauer, D.C., Packard, N.H., Stadler, P.F. (eds.) Protocells, pp. 213–231. The MIT Press, Cambridge (2008)
13. Ruiz-Mirazo, K., Briones, C., de la Escosura, A.: Prebiotic systems chemistry: new perspectives for the origins of life. Chem. Rev. **114**, 285–366 (2014)
14. Luisi, P.L., Ferri, F., Stano, P.: Approaches to semi-synthetic minimal cells: a review. Naturwissenschaften **93**, 1–13 (2006)
15. Luisi, P.L.: The Emergence of Life: From Chemical Origins to Synthetic Biology. Cambridge University Press, New York (2007)
16. Terasawa, H., Nishimura, K., Suzuki, H., Matsuura, T., Yomo, T.: Coupling of the fusion and budding of giant phospholipid vesicles containing macromolecules. Proc. Natl. Acad. Sci. **109**, 5942–5947 (2012)

17. Rasmussen, S., Chen, L., Stadler, B.M.R., Stadler, P.F.: Photo-organism kinetics: Evolutionary dynamics of lipid aggregates with genes and metabolism. Orig. Life Evol. Biosph. **34**, 171–180 (2004)
18. Rocheleau, T., Rasmussen, S., Nielsen, P.E., Jacobi, M.N., Ziock, H.: Emergence of protocellular growth laws. Philos. Trans. R. Soc. Lond. B Biol. Sci. **362**, 1841–1845 (2007)
19. Calvanese, G., Villani, M., Serra, R.: Synchronization in near-membrane reaction models of protocells. In: Rossi, F., Piotto, S., Concilio, S. (eds.) WIVACE 2016. CCIS, vol. 708, pp. 167–178. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-57711-1_15
20. Sacerdote, M.G., Szostak, J.W.: Semipermeable lipid bilayers exhibit diastereoselectivity favoring ribose. Proc. Natl. Acad. Sci. **102**, 6004–6008 (2005)
21. Mavelli, F., Ruiz-Mirazo, K.: Theoretical conditions for the stationary reproduction of model protocells. Integr. Biol. **5**, 324–341 (2013)