# Giant fluctuations in logistic growth

## Bahram Houchmandzadeh

# Giant fluctuations in logistic growth

Bahram Houchmandzadeh
*CNRS, LIPHY, F-38000 Grenoble, France*
*Univ. Grenoble Alpes, LIPHY,*
*F-38000 Grenoble, France*

We analyze the fluctuation of the number of individuals when two competing species, beginning with a few initial individuals, are submitted to a logistic growth. We show that when the total number of individuals reaches the carrying capacity, the number of each species is subject to giant fluctuations (variance $\sim$ mean$^2$) if the two species have similar growth rate. We show that the deterministic logistic equation can be used only when the growth rates are significantly different, otherwise such growth has to be investigated by stochastic processes tools. These results generalize to a wide class of growth law.

## I. INTRODUCTION.

In many chemical or biological systems, fluctuations can be large and drastically modify the results expected from a mean field approximation[1]. A famous early example was investigated by Delbrück [2] for the unbounded autocatalytic chemical reaction $A \to 2A$ where he showed that the number $n(t)$ of $A$ molecules at time $t$ displays giant fluctuations: the variance $V(t)$ is of the order of the *square* of the mean $V(t) = \langle n(t) \rangle^2 / n_0$, where $n_0$ is the initial number of $A$ molecules. It can be shown that spatial diffusion is not fast enough to dilute these local fluctuations and this phenomenon can lead to spatial clustering for example of organisms in ecological systems[3, 4] or of neutrons in nuclear reactors[5].

The unbounded autocatalytic reaction captures the initial growth period, but may seem unrealistic for systems where resources are limited[6]. More realistic scenarios are captured by a logistic growth where the reaction constant tends toward zero as the number of replicating agents increases. If only one species is subject to such a growth, fluctuations will become negligible when the number of replicating agents reaches the carrying capacity of the system. On the other hand, as we show below, if different species are competing for the same resources, the number of each species can display large fluctuations similar to the above example. This situation is relevant for example when independent cellular pathways compete for the same resources[7], when a cell is infected initially by a few bacteria or viruses carrying different mutations or when different mutants of cancerous cells compete with each other in the organism[8, 9]. Another important example is chemical/biological reactions in small compartments such as droplets[10, 11] which can be used for example for high throughput directed evolution[12].

Consider the simple competition of two species of autoreplicators $A$ and $B$ subject to a logistic growth where their deterministic evolution equation is given by

$$\frac{dn}{dt} = an(N_s - n - m) \tag{1}$$

$$\frac{dm}{dt} = bm(N_s - n - m) \tag{2}$$

where $n$, $m$ are the (continuous) number of each species, $a$, $b$, their respective growth rate at small concentration and $N_s$ the carrying capacity of the system. The solution of the above equations is given by

$$\frac{n}{n_0} = \left( \frac{m}{m_0} \right)^r \tag{3}$$

where $n_0$ and $m_0$ are the initial number of each species and $r = a/b$ is the relative growth rate of $A$ in respect to $B$ species. The final number of each species is found by solving $n_\infty + m_\infty = N_s$ in combination with relation (3). In particular, for the neutral case $r = 1$, the final number of each species is explicitly given by $n_\infty = pN_s$ and $m_\infty = (1 - p)N_s$ where $p = n_0/(n_0 + m_0)$ is the initial proportion of $A$ species.

Equations (1-2) are mean field approximations of the discrete stochastic process given by the following rates:

$$W(n, m \to n + 1, m) = an(N_s - N) \tag{4}$$

$$W(n, m \to n, m + 1) = bm(N_s - N) \tag{5}$$

where $N = n + m$ is the total number of individuals at time $t$. Figure (1) displays the stochastic behavior of the logistic growth (4,5) for $r = 1$. We observe that as expected[6], fluctuations in the total number of individuals $N = n+m$ disappear as $N$ reaches the carrying capacity $N_s$ (Fig. 1a). However, the number of individuals of each species is extremely variable (Fig. 1b). In fact, as we will show below, the probability of finding $n$ individuals of type $A$ when the system reaches saturation ($N = N_s$) is *uniform* in this case $P(n, N = N_s) = 1/(N_s - 1)$. For such giant fluctuations, the deterministic solution $n_\infty = pN$ is devoid of information and we have as much chance of finding one $A$ individual as finding $pN$ individuals!

In this article, we investigate analytically and numerically the stochastic equations (4,5) in general and discuss the origin of such large fluctuations when $r \approx 1$. The following section is devoted to the transformation of equations (4,5) ; section III investigates the problem for the neutral case $r = 1$ ; section IV generalizes the solution to $r \neq 1$. The last section is devoted to discussion and concluding remarks. Details of some computations are given in the appendices.

Figure 2. Mapping of the logistic growth into a flow problem in the $(N, n)$ plane.

$$P(n, N+1) = \alpha_N^{n-1} P(n-1, N) + (1 - \alpha_N^n) P(n, N) \quad (6)$$

at the initial time, the system contains $N_0$ individuals, $n_0$ of which are of type $A$ ; the initial condition for the Master equation (6) is

$$P(n, N_0) = \delta_{n_0}^n$$

where $\delta$ designates the Kronecker delta. The Master equation (6) is the mapping of the logistic growth into a flow problem in the $(N, n)$ plane, where each node distributes its content $P(n, N)$ to the adjacent ones $(N+1, n+1)$ and $(N+1, n)$ with proportion $\alpha_N^n$ and $\beta_N^n$ (figure 2).

Because of the form of the flow, the number of $A$ individuals $n$ is bounded by $n_0$ and $N - N_0 + n_0$ (figure 2). More over, on the two boundaries, the Master equation (6) reduces to a one term recurrence relation. For example, on the lower boundary,

$$P(n_0, N+1) = (1 - \alpha_N^{n_0}) P(n_0, N) \quad (7)$$

The probability is found to be

$$P(n_0, N) = \frac{(N_0 - n_0)_{N-N_0}}{(N_0 + sn_0)_{N-N_0}} \quad (8)$$

where $s = r - 1$ is the excess relative fitness of species $A$. $(x)_p$ designates the Pochhammer symbol (raising factorial) :

$$(x)_p = x(x+1)...(x+p-1) \quad (9)$$

Similarly, on the higher boundary,

$$P(N - N_0 + n_0, N) = \frac{(n_0)_{N-N_0}}{\left(\frac{N_0 + sn_0}{r}\right)_{N-N_0}} \quad (10)$$

Relation (10) can also be deduced from (8) by exchanging the role of $A$ and $B$ individuals.

The mean of various quantities can be computed theoretically from the Master equation (6). Let $f(.)$ be an arbitrary function and define
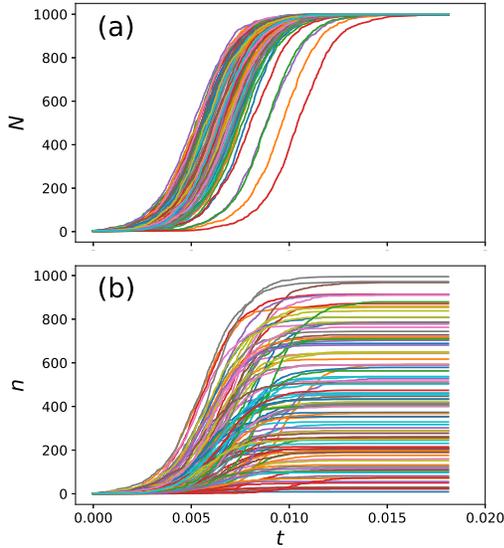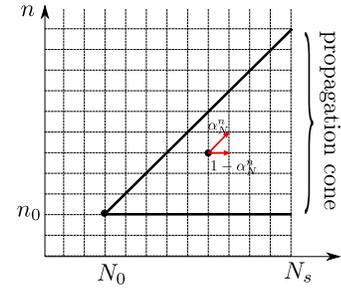
$$\langle f(n)(N) \rangle = \sum_n f(n) P(n, N)$$

Figure 1. Neutral logistic growth of two competing species. 100 individual based numerical simulation of equations (4,5) with $a = b = 1$ and $N_s = 1000$ are displayed. The initial number of each species is $n_0 = m_0 = 1$. (a) Total number of individuals $N$ ; (b) number of individuals of species $A$.

## II. MAPPING TO A SIMPLE PROBLEM.

Equations (4,5) represent a 2+1 dimensional system where because of the non-linearities, moment closure is lost and no closed form solution can be obtained. However, if we change the *independent* variable from time $t$ to the total number of individuals, the problem is mapped to a much simpler, one dimensional one : instead of computing the probability $P(n, t)$ of finding $n$ individuals of type $A$ at time $t$, we compute the probability $P(n, N)$ of finding $n$ individuals of type $A$ when the total number of individuals is $N$. For long times, $N$ reaches the carrying capacity $N_s$ and therefore, $P(n, t = \infty)$ and $P(n, N = N_s)$ contain the same information. A similar transformation was recently used to compute the Luria-Delbrück distribution of the number of mutants for a general growth curve[13].

The Master equation governing $P(n, N)$ is simple. *Once* a replication event happens $(N \to N + 1)$, the probability that it was an $A$ replicating $(n \to n + 1)$ is

$$\alpha_N^n = \frac{W(n, m \to n+1, m)}{W(n, m \to n+1, m) + W(n, m \to n, m+1)}$$
$$= \frac{rn}{N + (r-1)n}$$

The probability that it was a $B$ replicating ($n$ remains constant) is

$$\beta_N^n = 1 - \alpha_N^n = \frac{N - n}{N + (r-1)n}$$

The master equation for $P(n, N)$ is therefore

then

$$\langle f(n)(N+1)\rangle = \langle f(n)(N)\rangle + \langle \alpha_N^n \left( f(n+1)(N) - f(n)(N) \right)\rangle \tag{11}$$

For example, for $f(n) = n$, we have

$$\langle n(N+1)\rangle - \langle n(N)\rangle = \langle \alpha_N^n \rangle$$

the mean field, continuous approximation of the above expression leads to

$$\frac{d\langle n\rangle}{dN} = \alpha_N^{\langle n\rangle} \tag{12}$$

which is the equation deduced from the deterministic evolution (relation 1-2).

Finally, note that it is very simple to compute numerically the probabilities obeying the Master equation (6): The right-hand side of the equation (6) is the product of a bi-diagonal $(N+1) \times N$ matrix by an $N-$ column vector.

The next two sections are devoted to the computation of the means and probabilities for the neutral and non-neutral case.

## III. SOLUTION FOR THE NEUTRAL CASE.

In the neutral case $r = 1$, $\alpha_N^n = n/N$ ; the linearity of $\alpha$ in $n$ allows for moment closure and efficient computation of moments and probabilities. In particular, using relation (11), the mean $\langle n(N)\rangle$ and variance $\sigma^2(N)$ are found to obey the recurrence equation

$$\langle n(N+1)\rangle = \left(1 + \frac{1}{N}\right)\langle n(N)\rangle \tag{13}$$

$$\sigma^2(N+1) = \left(1 + \frac{2}{N}\right)\sigma^2(N) + p(1-p) \tag{14}$$

where $p = n_0/N_0$ is the initial proportion of the $A$ type. The two first moments are then found to be

$$\langle n(N)\rangle = pN \tag{15}$$

$$\sigma^2(N) = \frac{p(1-p)}{N_0+1}N(N-N_0) \tag{16}$$

We observe that regardless of the population size $N$ of the system, the fluctuations are of the same magnitude as the mean ($\sigma \sim \langle n\rangle$) if the initial population size is small:

$$\mathrm{cv} = \frac{\sigma(N)}{\langle n(N)\rangle} \approx \sqrt{\frac{1-p}{p(N_0+1)}} \tag{17}$$

Figure (3a) shows the perfect agreement between stochastic numerical simulations (equations 4,5) and the above results on the moments.

Using expression (11), it can be shown (see appendix A 1) that the raising factorial moments obey a simple relation:

$$\langle (n)_k\rangle = \langle n(n+1)...(n+k-1)\rangle = \frac{(n_0)_k}{(N_0)_k}(N)_k \tag{18}$$
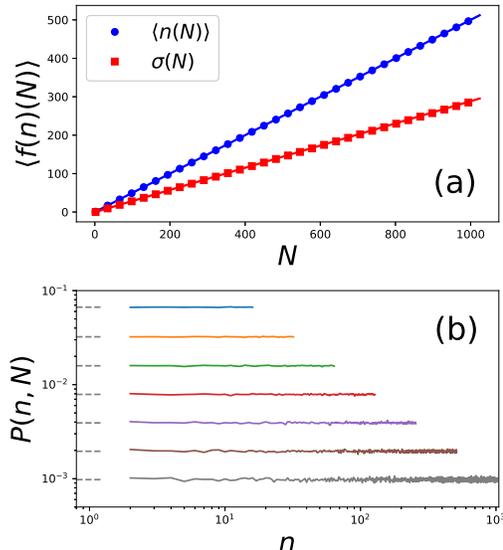


Figure 3. Numerical stochastic simulations of equations with rates (4,5) and comparison to theoretical values for the neutral case $r = 1$ and initial values $N_0 = 2$ and $n_0 = 1$. (a) Evolution of the two first moments $\langle n\rangle$ (circle) and $\sigma$ (squares) as a function of the number of individuals $N$. Symbols : numerical stochastic simulations ; solid lines : theoretical values given by relations (15,16). The moments were computed from $M = 5000$ realizations. (b)Solid lines: $P(n,N)$ as a function of the number of $A$ individuals $n$ for various values of $N = 2^k$, $k = 4, 8, ...10$. The gray dashed lines on the left designate the theoretical value $P(n,N) = 1/(N-1)$ (relation 20). The probabilities were computed from $M = 10^6$ realizations

In the neutral case, we can go beyond moments computation and solve the Master equation (6) for $P(n,N)$. In general, $P(n,N|n_0,N_0)$ is a polynomial of $n$ of degree $N_0 - 2$, where $n_0$, $N_0$ are the initial conditions for the number of $A$ individuals and all individuals. It is straightforward to check that (see appendix A 2)

$$P(n,N|n_0,N_0) = A\frac{(n-n_0+1)_{n_0-1}(m-m_0+1)_{m_0-1}}{(N-N_0+1)_{N_0-1}} \tag{19}$$

where $m = N - n$, and by convention, $(x)_0 = 1$. The normalization constant is found to be

$$A = \frac{(N_0-1)!}{(n_0-1)!(m_0-1)!}$$

In particular,

$$P(n,N|1,2) = \frac{1}{N-1} \tag{20}$$

$$P(n,N|2,3) = \frac{2(n-1)}{(N-1)(N-2)} \tag{21}$$

The initial condition $n_0 = 1$, $N_0 = 2$ was used in numerical simulations of figures 1,3.

The solution (19) is in perfect agreement with the numerical solution of the Master equation (6) (figure 4 ).
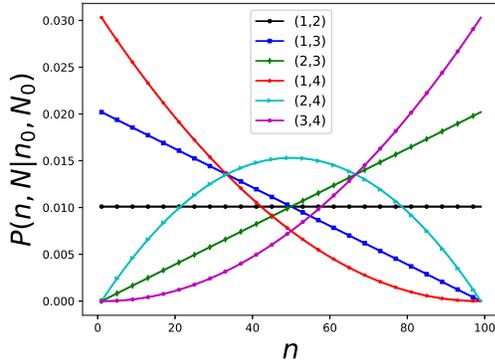
Figure 4. The probability $P(n, N|n_0, N_0)$ as a function of $n$ for $N = 100$ and various initial conditions $(n_0, N_0)$ indicated in the legend. Solid line : theoretical solution (19) ; symbols : numerical solutions of the master equation (6).



Figure 5. Function $n(\tilde{n})$ obtained by numerically solving the algebraic equation (24) for $N_0/N = 10^{-3}$ and various values of $r$.

## IV. SOLUTION FOR $r > 1$.

For the non-neutral case $r > 1$,

$$\alpha_N^n = \frac{rn}{N + (r-1)n}$$

is not anymore linear in $n$ and an exact solution for $P(n, N)$ becomes hard to obtain. However, as we are interested in the solution for large $N$, we can treat $n$ and $N$ as *continuous* variables and approximate the Master equation (6) by a partial differential equation (PDE). The Master equation (6) has indeed a simple structure and can be set into

$$\partial_N P(n, N) + \partial_n \left[\alpha_N^n P(n, N)\right] = 0 \qquad (22)$$

Equation (22) is a first order PDE and can be solved by the methods of characteristics[14]. Its general solution is found to be (see appendix B)

$$P(n, N) = \frac{\partial}{\partial n} f\left(\frac{(N-n)^r}{n}\right) \qquad (23)$$

where $f(.)$ is an arbitrary function to be determined from the initial condition. The implicit function $(N - n)^r/n = $ Cte is the solution of the mean field equation (12) $dn/dN = \alpha_N^n$.

Let us define $\tilde{n}$ such that (figure 5)

$$\frac{(N-n)^r}{n} = \frac{(N_0 - \tilde{n})^r}{\tilde{n}} \qquad (24)$$

Then for the initial condition $P(n, N_0) = \phi_0(n)$, the complete solution of equation (22) is given by (see appendix B)

$$P(n, N) = \frac{\partial \tilde{n}}{\partial n} \phi_0(\tilde{n}) \qquad (25)$$

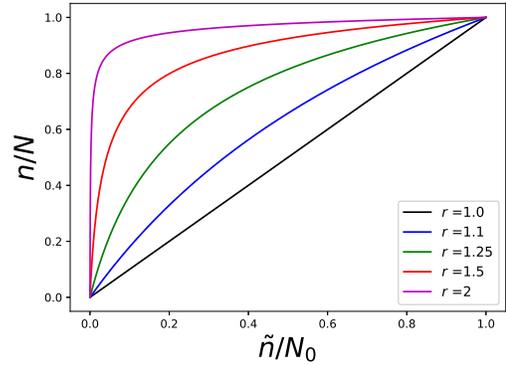$$= \frac{\tilde{n}(N_0 - \tilde{n})}{N_0 + (r-1)\tilde{n}} \frac{N + (r-1)n}{n(N-n)} \phi_0(\tilde{n}) \qquad (26)$$
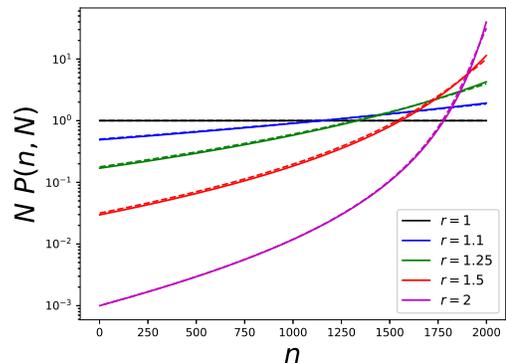


Figure 6. Solution (30) of the continuous Master equation (22) (continuous lines) compared to numerical solutions of the discrete Master equation (6) (dashed lines) for $N = 2000$, $N_0 = 2$, $n_0 = 1$ and various values of $r$. The solution (30) is obtained by numerically solving equation (24) and then using relation (26).

No special function is defined in the mathematical literature to deal with equations of type $x^r + ux - u = 0$ ; however, it is straightforward to find the numerical solution of equation (24) and use expression (26) to compute $P(n, N)$.

To make it more concrete, let us consider in some details the neutral case $r = 1$, and compare the exact known solution (19) to the solution (25) of the PDE approach. In this case, relation (23) transforms into the explicit form $\tilde{n} = (N_0/N)n$. The initial condition has to be chosen in order to match the known solution (19) ; once it has been fixed for $r = 1$, it will be used for all $r > 1$. The initial condition corresponding to the discrete case $n_0 = 1$, $N_0 = 2$ (relation 20) is

$$\phi_0(n) = \Pi(n - 1) \qquad (27)$$

where the gate function is defined as $\Pi(x) = 1/2$ for

$|x| < 1$ and is zero outside this domain. Therefore,

$$P(n, N) = \frac{2}{N}\Pi\left(\frac{2}{N}n - 1\right) \quad (28)$$

$$= \frac{1}{N} \quad n \in ]0, N[ \quad (29)$$

which approximates the exact solution (20) to $O(1/N)$.

The general solution for arbitrary $r$ corresponding to initial condition $n_0 = 1$, $N_0 = 2$ is then simply

$$P(n, N) = \frac{1}{2}\frac{\partial\tilde{n}}{\partial n} \quad n \in ]0, N[ \quad (30)$$

Figure 6 shows the excellent agreement between expression (30) and the numerical solution obtained from the exact discrete Master equation (6).

Various moments can be extracted from solution (25):

$$\left\langle n^k(N)\right\rangle_r = \int_0^N n^k P(n, N)dn = \int_0^{N_0} n^k \phi_0(\tilde{n})d\tilde{n} \quad (31)$$

where $n$ inside the integrand on the right-hand side of eq. (31) is a function of $\tilde{n}$ through relation (24). For the neutral case $r = 1$, $n/N = \tilde{n}/N_0$ and therefore

$$\int_0^{N_0} \frac{\tilde{n}^k}{N_0^k}\phi_0(\tilde{n})d\tilde{n} = \frac{\left\langle n^k(N)\right\rangle_1}{N^k} = \frac{(n_0)_k}{(N_0)_k} + O(1/N) \quad (32)$$

We can obtain an explicit form of $n$ as a function of $\tilde{n}$ for various conditions. If $s = r - 1 \ll 1$, we can obtain a perturbative solution of equation (24) in powers of $s$. On the other hand, for high values of integer $r$ such as $r = 2, 3, 4$, we can exactly solve the algebraic equation (24). These two cases constitute the near neutral and highly non-neutral situations and allows us to understand the general behavior of the system.

### A. Perturbative solution.

Let us first consider the case $s = r - 1 \ll 1$. Setting $\kappa = \log(N/N_0)$, we have, to the second order in $s$ :

$$x = \tilde{x} + \kappa\tilde{x}(1 - \tilde{x})s - \kappa\tilde{x}(1 - \tilde{x})\left((\kappa + 1)\tilde{x} - \kappa/2\right)s^2 \quad (33)$$

where $\tilde{x} = \tilde{n}/N_0$, $x = n/N$. The symmetry of equation (24) implies that $\tilde{x}$ can be expressed as a function of $x$ by simply replacing $\kappa$ by $-\kappa$ in expression (33). Using expression (31,32) for the initial conditions $N_0, n_0$, , to the first order perturbations, the moments are found to be

$$\left\langle n(N)\right\rangle_r = \left\langle n(N)\right\rangle_1\left\{1 + \kappa s\frac{(N_0 - n_0)}{N_0 + 1}\right\} \quad (34)$$

$$\sigma_r^2(N) = \sigma_1^2(N)\left\{1 + 2\kappa s\frac{N_0 - 2n_0}{N_0 + 2}\right\} \quad (35)$$

$$\frac{\sigma_r(N)}{\left\langle n(N)\right\rangle_r} = cv_1\left\{1 - \kappa s\frac{N_0(n_0 + 1)}{(N_0 + 1)(N_0 + 2)}\right\} \quad (36)$$
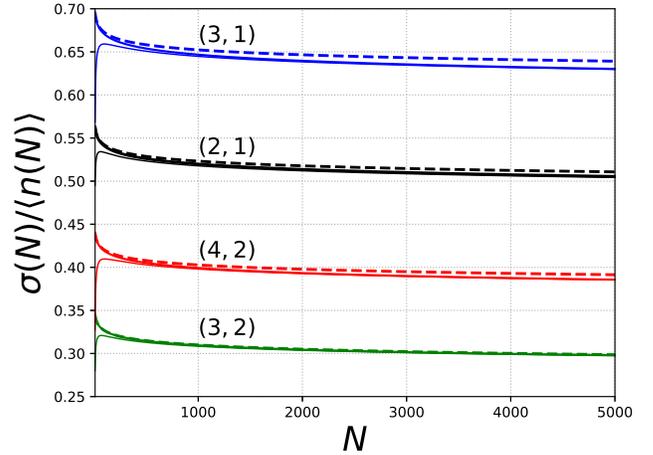


Figure 7. Coefficient of variation $\sigma(N)/\left\langle n(N)\right\rangle$ for $s = 0.05$ ($r = 1 + s$) and various initial conditions $(N_0, n_0)$. Thin solid lines : exact values obtained from direct numerical resolution of the Master equation (6) ; Dashed lines : first order perturbations given by expression (36) ; thick solid lines : second order perturbations. The initial condition $(N_0, n_0)$ of each curve is displayed above it.

Where the subscript 1 refers to the neutral expressions (15-17). Figure 7 shows the comparison of the above expressions to exact values obtained from numerical solutions of the exact Master equation (6).

We observe that the correction of the above expressions compared to neutral values (equation 15-17) are logarithmic and of the order of $s\kappa = s\log(N/N_0)$ : the fluctuations amplitude $\sigma$ is still large and of the order of the mean $\left\langle n\right\rangle$. The perturbative approach is valid for $\kappa s \ll 1$ ; the solution for higher values of $s$ can be slightly improved by using higher order perturbations ( figure 7) but the perturbative approach reaches its limit for $\kappa s \lesssim 1$.

### B. High values of $r$.

High values of $r$ can be understood by investigating integer values such as 2,3,4 for which the equation (24) can be exactly solved. For the case $r = 2$

$$x = \frac{\gamma(1 - \tilde{x})^2 + 2\tilde{x} - \sqrt{\gamma^2(1 - \tilde{x})^2 + 4\gamma\tilde{x}}}{2\tilde{x}} \quad (37)$$

where $\gamma = N_0/N \ll 1$ and as before, $x = n/N$ and $\tilde{x} = \tilde{n}/N_0$. We will investigate the simplest case corresponding to the initial condition $N_0 = 2$, $n_0 = 1$ where $\phi_0(u) = \Pi(u - 1)$ (relation 27). For this initial condition, the moments equation (31) is greatly simplified :

$$\frac{\left\langle n^k\right\rangle}{N^k} = \frac{N_0}{2}\int_0^1 x^k d\tilde{x} \quad (38)$$

Using expression (37), performing the integrations involved by equation (38) and keeping only the leading
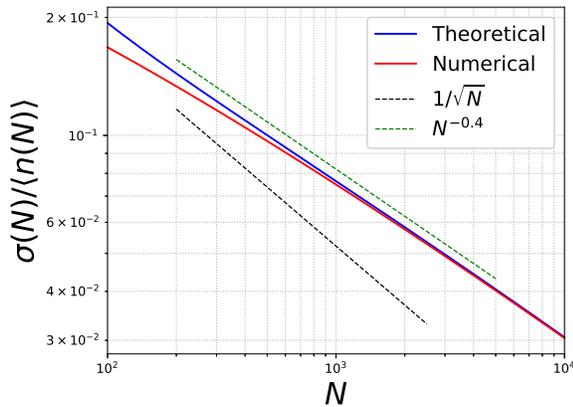
Figure 8. Coefficient of variation $\sigma(N)/\langle n \rangle_N$ for $r = 2$ with initial condition $N_0 = 2$, $n_0 = 1$. The theoretical value is obtained from expressions (39,??) ; the exact, numerical value is obtained by numerical resolution of the Master equation (6). As a guide for the eye, $N^{-0.5}$ and $N^{-0.4}$ are also displayed.

orders of $\gamma$, we find that

$$\frac{\langle n(N) \rangle}{N} = 1 - \frac{4}{3}\sqrt{\gamma} + \frac{\gamma}{4}\left(1 - 2\log\gamma\right) + O(\gamma^{3/2}) \quad (39)$$

$$\frac{\sigma^2(N)}{N^2} = \gamma\left(-\log\gamma - \frac{77}{18}\right) +$$

$$\gamma^{3/2}\left(-\frac{4}{3}\log\gamma + \frac{106}{15}\right) + O(\gamma^2) \quad (40)$$

Expression (40) is valid for $N/N_0 \gtrsim 72$ which is indeed the regime of interest (figure 8) . We see that for $r = 2$, the variance increases only as $N\log N$ and not $N^2$ as in the neutral case. Therefore, for high values of $N$, the coefficient of variation $\sigma/\langle n \rangle$ decreases as $(\log N/N)^{1/2}$ . In this regime, fluctuations become negligible and the deterministic approach is valid.

## V. DISCUSSION AND CONCLUSION.

In this article, we have investigated the distribution of the number of individuals $n$ and $m$ of two species $A$,$B$ during a logistic growth. We have shown that the investigation is greatly simplified if instead of time $t$, the independent variable is chosen to be the total number of individuals $N = n + m$. This paper was focused on the well known logistic growth, but the method and conclusions are valid for any stochastic growth of the form

$$W(n, m \to n + 1, m) = anf(n, m) \quad (41)$$

$$W(n, m \to n, m + 1) = bmf(n, m) \quad (42)$$

where $f(n, m)$ is an arbitrary function not necessarily symmetric in $m$ and $n$.

The most interesting feature of the investigated system is the large amplitude of fluctuations in the neutral case $r = a/b = 1$, where both species have similar growth

rate. Suppose that we draw (and replace) $N_s$ individuals at random from a pool of $N_0$ individuals when $n_0$ are of the $A$ type. The distribution of the number of $A$ type in the $N_s$ sample is a binomial one with parameter $p = n_0/N_0$ ; the fluctuation amplitude of this experiment $\sigma/\langle n \rangle \sim 1/\sqrt{N_s}$ is small if $N_s \gg 1$. One could naively suppose that a logistic growth when two types $A$ and $B$ individuals are competing and the system expands from $N_0$ to $N_s$ individuals ($N_s \gg N_0$) is similar to the above drawing experiment : each individual in the final pool draws at random its ancestor from the initial pool. This is however not the case and we have shown that contrary to the binomial case, the fluctuation amplitude $\sigma/\langle n \rangle \sim 1/\sqrt{N_0}$ is always large and independent of the final system size.

Various experiments can be devised to test the relevance of the above computations. For example, a phage such as $\lambda$ can be modified into few different mutants, each expressing a different fluorescent proteins (such as GFP, RFP, YFP,...) ; the mutants can then be used to co-infect a bacterial culture. The distribution of the colors in the culture after some time can be related to the probabilities we have computed through a convolution by a Poisson-Binomial distribution to account for variation in the initial number of co-infectors. A similar experiment can be performed using PCR amplification of few similar DNA strands[15] of the same length and characteristics and then analyze the number of strands copy in each droplets.

The problem we have investigated can also be used to extend the Wright-Fisher (WF) model of population genetics to variable size population (see for example[16–18]). In the WF model with fixed population size $N_0$ and two mutant types $A$ and $B$, each generation is formed by selecting randomly $N_0$ individuals among the progeny of generation $i$ to form generation $i + 1$. If $x$ is the proportion of the $A$ type with reproductive advantage $r = 1 + s$, then a diffusion (Kimura) equation can be derived for the evolution of the population ([19, 20]) where the drift and diffusion coefficient are $a(x) = sx(1 - x)$ and $b(x) = x(1 - x)/(2N)$.

We can generalize the WF model by allowing, at each generation $i$, the population to expand to size $N_s$ and then select $N_0$ individuals among them to form the new generation $i + 1$. By using the result of subsection IV A, it is straightforward to show that the diffusion equation governing this system is the same as before except that the relative excess fitness is now renormalized to $s' = s\log(N_s/N_0)$. The fact that the effective fitness increases in a growing population was already noted by Ewens ([16]), although the amplifying factor in the problem investigated by him was proportional to the harmonic mean $N_s$ and $N_0$ rather than their logarithmic difference as here.

In summary, we have shown that populations subjects to logistic-like growth such as equation (41,42) can be modeled by deterministic equations only if there is significant difference ($r \gtrsim 2$) between their growth rates.

If they have similar growth rate, the deterministic equation must be abandoned and a stochastic treatment used instead.

## Appendix A: Various neutral computations

### 1. Factorial moments.

Consider the function $f(n) = (n)_k = n(n+1)...(n+k-1)$ ; then

$$f(n+1) - f(n) = (n+1)_{k-1}(n+k-n) = k(n+1)_{k-1}$$

and therefore

$$n\left(f(n+1) - f(n)\right) = k(n)_k = kf(n)$$

Therefore, using the general expression (11), we find the one term recurrence relation

$$\langle f(n)(N+1)\rangle = \left(1 + \frac{k}{N}\right)\langle f(n)(N)\rangle$$

which is trivially solved and leads to expression (18).

### 2. Expression of the probability.

To shorten the notations, we use $m = N - n$ whenever needed. The Master equation in the neutral case is

$$P(n, N+1) = \frac{n-1}{N}P(n-1, N) + \frac{m}{N}P(n, N) \quad \text{(A1)}$$

Consider

$$P(n, N) = \frac{(n-n_0+1)_{n_0-1}(m-m_0+1)_{m_0-1}}{(N-N_0+1)_{N_0-1}} \quad \text{(A2)}$$

Pochhammer manipulation is similar to factorial manipulation. In particular,

$$N(N-N_0+1)_{N_0-1} = (N-N_0+1)_{N_0}$$
$$(n-1)(n-n_0)_{n_0-1} = (n-n_0)_{n_0}$$
$$m(m-m_0+1)_{m_0-1} = (m-m_0+1)_{m_0}$$

and therefore, the right hand side of relation (A1) is found to be

$$\frac{(n-n_0+1)_{n_0-1}(m-m_0+2)_{m_0-1}}{(N-N_0+1)_{N_0}}(n-n_0+m-m_0+1)$$

As

$$n - n_0 + m - m_0 + 1 = N - N_0 + 1$$

and

$$\frac{N-N_0+1}{(N-N_0+1)_{N_0}} = \frac{1}{(N+1-N_0+1)_{N_0-1}}$$

expression (A2) is indeed a solution of the Master equation, up to a multiplicative constant. The constant is found by stating $P(n_0, N_0) = 1$. As the master equation conserves the probability, the constant is valid for all $N$.

## Appendix B: Solving the PDE

Consider a first order partial differential equation (PDE) of first order for the function $P(x, t)$ of type

$$\partial_t P + \partial_x(\alpha P) = 0 \quad \text{(B1)}$$

where $\alpha = \alpha(x, t)$ is a known function. Let us call $R(x, t) = \text{Cte}$ the solution of the characteristic equation

$$\frac{dx}{dt} = \alpha(x, t)$$

Then by definition,

$$\partial_t R + \alpha \partial_x R = 0$$

Consider the function

$$Q(x, t) = \frac{\partial}{\partial x}f\left(R(x, t)\right) \quad \text{(B2)}$$

where $f()$ is an arbitrary function. Then

$$\partial_t Q + \partial_x(\alpha Q) = \partial_x\left\{(\partial_t R + \alpha \partial_x R)f'(R)\right\} = 0$$

and therefore $Q(x, t)$ is a solution of equation (B1). For example, for $\alpha = c$, the solution is the trivial propagation $P(x, t) = f(x - ct)$.

The function $f(.)$ has to be determined from the initial condition $P(x, t_0) = \phi_0(x)$. Consider two points $(t_0, \tilde{x})$ and $(t, x)$ in the plane, related through $R(x, t) = R(\tilde{x}, t_0)$, i.e. they belong to the same characteristic curve. Obviously, we can reverse this relation as $\tilde{x} = g\left(R(x, t), t_0\right)$ and therefore write the general solution (B2) as $P(x, t) = \partial_x f(\tilde{x}) = (\partial \tilde{x}/\partial x)f'(\tilde{x})$. On the other hand, at the initial time $t_0$, $x = \tilde{x}$, $\partial \tilde{x}/\partial x = 1$ and therefore $f'() = \phi_0()$. The solution of the PDE (B2) with the initial condition $\phi_0(x)$ is then

$$P(x, t) = \frac{\partial \tilde{x}}{\partial x}\phi_0(\tilde{x})$$

$P(., t)$ can be seen as a transformation, i.e. scaling and deformation of the initial condition $\phi_0(.)$. An initial Dirac distribution however propagates without deformation along a characteristic curve because $f(x)\delta(x) = \delta(x)$

: in this case, the PDE is reduced to the deterministic equations $dx/dt = \alpha$.

Let us precise the function $\phi_0(.)$ used in this article for the PDE (22) . The true probability $P_d(n, N)$ is function of *discrete* variables $n$ and $N$. In order to estimate this probability, we have used the probability density $P_c(n, N)$ of *continuous* variable $n, N$. $P_c$ must approximate $P_d$ for *large* $N$. $\phi_0()$ has to be chosen to make this approximation as precise as possible. However, we cannot use the discrete initial condition $P(n, N_0) = \delta_{n_0}^n$,

because the continuous PDE will be reduced to a deterministic equation. We make the assumption that the choice of $\phi_0(.)$ is independent of $r$ and therefore can be deduced from the known expression of neutral probability. For $r = 1$, $\tilde{n} = (N_0/N)n$, and therefore we have

$$\phi_0(\tilde{n}) = \frac{N}{N_0} P_1(\frac{N}{N_0}\tilde{n}, N)$$

where $P_1(,)$ is the neutral probabilities but the arguments are continuous.

[1] Lev S. Tsimring. Noise in biology. *Reports on Progress in Physics*, 77(2):026601, 2014.

[2] Max Delbrück. Statistical Fluctuations in Autocatalytic Reactions. *The Journal of Chemical Physics*, 8(1):120–124, January 1940.

[3] B Houchmandzadeh. Neutral clustering in a simple experimental ecological community. *Phys Rev Lett*, 101(7):78103, 2008.

[4] Bahram Houchmandzadeh. Theory of neutral clustering for growing populations. *Physical Review E*, 80(5):051920, November 2009.

[5] Eric Dumonteil, Fausto Malvagi, Andrea Zoia, Alain Mazzolo, Davide Artusio, Cyril Dieudonné, and Clélia De Mulatier. Particle clustering in Monte Carlo criticality simulations. *Annals of Nuclear Energy*, 63:612–618, January 2014.

[6] Dipjyoti Das, Dibyendu Das, and Ashok Prasad. Giant number fluctuations in microbial ecologies. *Journal of Theoretical Biology*, 308:96–104, September 2012.

[7] Anthony J. Genot, Teruo Fujii, and Yannick Rondelez. Computing with Competition in Biochemical Networks. *Physical Review Letters*, 109(20):208102, November 2012.

[8] John A. Spratt, D. von Fournier, John S. Spratt, and Ernst E. Weber. Decelerating growth and human breast cancer. *Cancer*, 71(6):2013–2019.

[9] Nkiruka C. Atuegwu, Lori R. Arlinghaus, Xia Li, A. Bapsi Chakravarthy, Vandana G. Abramson, Melinda E. Sanders, and Thomas E. Yankeelov. Parameterizing the Logistic Model of Tumor Growth by DW-MRI and DCE-MRI Data to Predict Treatment Response and Changes in Breast Cancer Cellularity during Neoadjuvant Chemotherapy. *Translational Oncology*, 6(3):256–264, June 2013.

[10] Helen Song, Delai L. Chen, and Rustem F. Ismagilov. Reactions in Droplets in Microfluidic Channels. *Angewandte Chemie International Edition*, 45(44):7336–7356, November 2006.

[11] Alexandre Baccouche, Shu Okumura, Rémi Sieskind, Elia Henry, Nathanaël Aubert-Kato, Nicolas Bredeche, Jean-François Bartolo, Valérie Taly, Yannick Rondelez, Teruo Fujii, and Anthony J. Genot. Massively parallel and multiparameter titration of biochemical assays with droplet microfluidics. *Nature Protocols*, 12(9):1912–1932, September 2017.

[12] Jeremy J. Agresti, Eugene Antipov, Adam R. Abate, Keunho Ahn, Amy C. Rowat, Jean-Christophe Baret, Manuel Marquez, Alexander M. Klibanov, Andrew D. Griffiths, and David A. Weitz. Ultrahigh-throughput screening in drop-based microfluidics for directed evolution. *Proceedings of the National Academy of Sciences*, 107(9):4004–4009, March 2010.

[13] Bahram Houchmandzadeh. General formulation of Luria-Delbr\"uck distribution of the number of mutants. *Physical Review E*, 92(1):012719, July 2015.

[14] Andrei D. Polyanin, Valentin F. Zaitsev, and Alain Moussiaux. *Handbook of First-Order Partial Differential Equations*. CRC Press, London ; New York, 1 edition edition, November 2001.

[15] Yolanda Schaerli, Robert C. Wootton, Tom Robinson, Viktor Stein, Christopher Dunsby, Mark A. A. Neil, Paul M. W. French, Andrew J. deMello, Chris Abell, and Florian Hollfelder. Continuous-Flow Polymerase Chain Reaction of Single-Copy DNA in Microfluidic Microdroplets. *Analytical Chemistry*, 81(1):302–306, January 2009.

[16] W. J. Ewens. The probability of survival of a new mutant in a fluctuating environment. *Heredity*, 22(3):438–443, August 1967.

[17] Sarah P. Otto and Michael C. Whitlock. The Probability of Fixation in Populations of Changing Size. *Genetics*, 146(2):723–733, June 1997.

[18] Karl Wienand, Erwin Frey, and Mauro Mobilia. Evolution of a Fluctuating Population in a Randomly Switching Environment. *Physical Review Letters*, 119(15):158301, October 2017. WOS:000412751700014.

[19] M. Kimura. Solution of a process of random genetic drift with a continuous model. *Proc. Nat. Ac. Sci. (USA)*, 41:144, 1955.

[20] W. J. Ewens. *Mathematical Population Genetics*. Springer-Verlag, 2004.