

Large population solution of the stochastic Luria–Delbrück evolution model

David A. Kessler^{a,1} and Herbert Levine^{b,1}

^aDepartment of Physics, Bar-Ilan University, Ramat-Gan IL52900, Israel; and ^bCenter for Theoretical Biological Physics, Rice University, Houston, TX 77005

Contributed by Herbert Levine, May 22, 2013 (sent for review April 30, 2013)

Luria and Delbrück introduced a very useful and subsequently widely adopted framework for quantitatively understanding the emergence of new cellular lineages. Here, we provide an analytical treatment of the fully stochastic version of the model, enabled by the fact that population sizes at the time of measurement are invariably very large and mutation rates are low. We show that the Lea–Coulson generating function describes the “inner solution,” where the number of mutants is much smaller than the total population. We find that the corresponding distribution function interpolates between a monotonic $1/m(m+1)$ decrease at relatively small populations, (compared with the inverse of the mutation probability), whereas it goes over to a Lévy α -stable distribution in the very large population limit. The moments are completely determined by the outer solution, and so are devoid of practical significance. The key to our solution is focusing on the fixed population size ensemble, which we show is very different from the fixed time ensemble due to the extreme variability in the evolutionary process.

Landau distribution | mutant distribution

In 1943, Luria and Delbrück (LD) (1) used a simple model to argue that data regarding bacterial resistance were consistent with selection of preexisting mutations. The model postulated a constant probability at which dividing normal cells would have a single mutated, resistant daughter; these mutations were assumed not to affect the growth rate until the bacteria were challenged by an infectious virus and were assumed never to revert to the sensitive state. In recent years, the LD model has been used extensively to help understand the emergence of antibiotic-resistant microbes (2, 3) as well as drug therapy-resistant cancer cells (4–8). The basic approach is also used for understanding stem cell lineage data (9) and to help estimate mutation rates (10, 11). These applications require generalizations of the LD model to include differential fitness effects and cell death (12, 13), but the basic notion of the systematic accumulation of resistant clones remains unchanged.

Despite extensive analyses by the applied mathematics, statistical physics, and evolutionary biology communities (14), a tractable analytical expression for the distribution of the number of mutants in an overall population of size N for the general stochastic problem has remained elusive (15). In this work, we develop a unique approach that directly solves the master equation in the limit of (eventually) large population sizes and a small mutation rate. We will find an explicit expression for the mutant number probability distribution. Remarkably, in the further limit wherein the mutation rate μN is large, the distribution takes the form of a Landau distribution (16), the one-sided Lévy α -stable distribution (with $\alpha=1$) originally obtained (17) in the context of energy loss due to ionization. (The Landau distribution should not be confused with the Cauchy distribution, which is the Lévy $\alpha=1$ two-sided stable distribution.) For pedagogical purposes, we will start with the original pseudodeterministic version of the model,

wherein our results agree with those obtained by earlier researchers (2, 18).

Original LD Model

The initial LD approach assumes that all populations grow with a deterministic exponential growth rate β starting from an initial sensitive population of size N_0 . The only sources of stochasticity are the number and timing of the mutational events. At each moment in time, there is a fixed probability per individual of generating a new mutant, corresponding to an average fraction μ of birth events, leading to an overall rate of $N_0\mu\beta e^{\beta t}$. The overall probability of having m mutants can be broken down into a sum over the number of mutational events and integrals over the possible times at which those events occurred, so that

$$P(m) = \sum_{k=0}^{\infty} e^{-\mu(N-N_0)} \frac{\mu^k \beta^k}{k!} \times \prod_{i=1}^k \int_0^t dt_i e^{\beta t_i} \delta\left(m - \sum_k e^{\beta(t-t_i)}\right), \quad [1]$$

where the factor $e^{-\mu(N-N_0)}$ comes from an integral over the times when no mutation occurred. Changing variables to

$$z_k \equiv e^{\beta(t-t_k)} = N e^{-\beta t_k} / N_0,$$

we obtain

$$P(m) = \sum_{k=0}^{\infty} e^{-\mu(N-N_0)} \frac{\mu^k N^k}{k!} \prod_{i=1}^k \int_1^{N/N_0} \frac{dz_i}{z_i^2} \delta\left(m - \sum z_i\right). \quad [2]$$

We already see from here why $P(m)$ contains features of the Landau distribution. The random variable m is the sum over a Poisson-distributed [with mean $\mu(N-N_0)$] number of random variables, each distributed as $1/z^2$, and if this were a sum over a large definite number of z 's with no upper cutoff on the distribution, by the generalized central-limit theorem (19), m would be distributed as a one-sided Lévy α -stable distribution, with $\alpha=1$ (i.e., the Landau distribution). Here, things are a bit more subtle; nevertheless, as N increases, the number of terms becomes large and relatively more definite, and the cutoff on the distribution goes off to infinity. Expanding the exponential $e^{i\alpha z}$, carrying out the integrals over the z_i , and rescaling $\tilde{\alpha} = \alpha N / N_0$ gives

Author contributions: D.A.K. and H.L. designed research, performed research, and wrote the paper.

The authors declare no conflict of interest.

¹To whom correspondence may be addressed. E-mail: kessler@dave.ph.biu.ac.il or herbert.levine@rice.edu.

$$\begin{aligned}
 P(m) &= e^{-\mu(N-N_0)} \int_{-\infty}^{\infty} \frac{d\alpha}{2\pi} \exp\left(i\alpha m + \mu N \int_1^{N/N_0} \frac{dz}{z^2} e^{-i\alpha z} \right) \\
 &= \frac{N_0}{2N\pi} \int_{-\infty}^{\infty} d\tilde{\alpha} \exp\left(i \frac{\tilde{\alpha} m N_0}{N} \right) \\
 &\quad \times \exp\left(\mu N \left[-i \frac{\tilde{\alpha} N_0}{N} \ln \frac{N}{N_0} + \frac{N_0}{N} \sum_{k=2}^{\infty} \frac{(-i\tilde{\alpha})^k}{k!(k-1)} \right] \right) \\
 &= \frac{N_0}{2N\pi} \int_{-\infty}^{\infty} d\tilde{\alpha} \exp\left(i\tilde{\alpha} N_0 \frac{m - \mu N \ln N/N_0}{N} + \mu N_0 \sum_{k=2}^{\infty} \frac{(-i\tilde{\alpha})^k}{k!(k-1)} \right). \tag{3}
 \end{aligned}$$

This result is in accord with the Lea–Coulson scaling ansatz (2) that the N dependence of $P(m)$ is captured by the variable

$$\delta = \frac{m - \mu N \ln(N/N_0)}{N},$$

where it is easy to show that $\mu N \ln(N/N_0)$ is the mean of m to lowest order in μ .

For small μ , there is an inner region where $\delta \sim O(\mu)$ and an outer region where $\delta \sim O(1)$. In the inner region, we rescale $\delta = \mu \tilde{\delta}$ and note that we can approximate the sum by its large argument limit:

$$\sum_{k=2}^{\infty} \frac{(-ix)^k}{k!(k-1)} = 1 - \cos x - x \text{Si}(x) + i[x(\gamma - 1 + \ln x - \text{Ci}(x)) + \sin x],$$

where $\gamma \approx 0.5772$ is the Euler–Mascheroni constant and Si and Ci are the sine and cosine integrals, respectively:

$$\text{Si}(x) \equiv \int_0^x \frac{\sin t}{t} dt; \quad \text{Ci}(x) \equiv \int_0^x \frac{\cos t - 1}{t} dt + \gamma + \ln x. \tag{4}$$

Using $\text{Si}(x) \sim \pi/2 - \cos(x)/x$, $\text{Ci}(x) \sim \sin(x)/x$ as $x \rightarrow \infty$, we have

$$\sum_{k=2}^{\infty} \frac{(-ix)^k}{k!(k-1)} \approx -x\pi/2 + 1 + i[x(\gamma - 1 + \ln x)]. \tag{5}$$

Thus,

$$\begin{aligned}
 P(\tilde{\delta}) &= \frac{N_0}{2\pi\mu N} \int_{-\infty}^{\infty} d\alpha \exp\left[N_0 \left(i\tilde{\delta}\alpha - \frac{\alpha\pi}{2} + i\alpha \left(\gamma - 1 + \ln \frac{\alpha}{\mu} \right) \right) \right] \\
 &= \frac{1}{2\pi\mu N} \int_{-\infty}^{\infty} d\alpha \exp\left(i\tilde{\delta}\alpha - \frac{\alpha\pi}{2} + i\alpha \left(\gamma - 1 + \ln \frac{\alpha}{\mu N_0} \right) \right). \tag{6}
 \end{aligned}$$

This is nothing other than the Landau distribution:

$$P(m) = \frac{1}{\mu N} P_{\text{Landau}}\left(\frac{m}{\mu N} - \ln \mu N + \gamma - 1 \right). \tag{7}$$

The asymptotic behavior of $P_{\text{Landau}}(x)$ is

$$P_{\text{Landau}}(x) \approx \frac{1}{x^2}, \tag{8}$$

as expected. Notice that this inner formula is independent of N_0 . This is because small clones originate near the end of the process, and so are independent of the initial conditions.

Finally, for $\delta \sim O(\mu)$, which means $m \sim O(N)$, we can expand Eq. 3 in μ and find the outer solution:

$$\begin{aligned}
 P(m) &= \frac{N_0}{2N\pi} \int_{-\infty}^{\infty} d\alpha e^{i\alpha\delta N_0} \mu N_0 \sum_{k=2}^{\infty} \frac{(-i\alpha)^k}{k!(k-1)} \\
 &= -\frac{\mu}{2N\pi\delta^2} \int_{-\infty}^{\infty} d\alpha e^{i\alpha\delta} \sum_{k=2}^{\infty} \frac{(-i\alpha/N_0)^{k-2}}{(k-1)!} \\
 &= -\frac{\mu}{2N\pi\delta^2} \int_{-\infty}^{\infty} d\alpha e^{i\alpha\delta} \frac{e^{-i\alpha/N_0} - 1}{-i\alpha} \\
 &= \frac{\mu}{N\delta^2} [\theta(\delta) - \theta(\delta - 1/N_0)],
 \end{aligned} \tag{9}$$

which cuts off the $1/m^2$ tail at $m = N/N_0$.

To check the validity of this result, we have directly simulated the sum in Eq. 2. The results are shown in Fig. 1, demonstrating the direct relevance of the Landau distribution as well as the sharp cutoff at $m = N/N_0$.

Stochastic LD Model: Outer Solution

An obvious refinement of the original LD model is to allow for the fact that birth events, for both the sensitive and mutant subpopulations, are also stochastic (20). Again, we assume that development of resistance is irreversible and mutants grow with no fitness disadvantage compared with wild-type (WT). As we shall see, the question of conditioning becomes crucial here, an issue that obviously does not arise in the quasideterministic original model, or in the Lea–Coulson (2) variant, where the WT grows deterministically. We start by conditioning on the total population, N , which matches most closely the quasideterministic earlier formulations. We will later see that conditioning on time (20, 21) yields very different answers in general.

The system, conditioned on N , is fully described by the master equation governing the probability $P_N(m)$ of having m mutants among the N total members of the population, [i.e., that has undergone $N - N_0$ birth events (22)]:

$$\begin{aligned}
 P_{N+1}(m) &= \frac{1}{N} \left[P_N(m-1)\mu(N-m+1) \right. \\
 &\quad \left. + P_N(m)(1-\mu)(N-m) + P_N(m-1)(m-1) \right], \tag{10}
 \end{aligned}$$

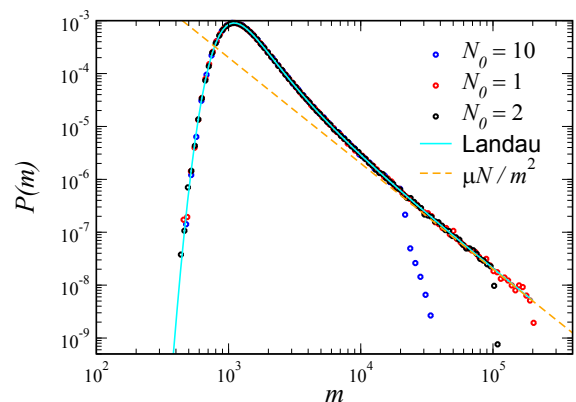


Fig. 1. $P(m)$ for the original LD model from a simulation of 10^6 realizations of the stochastic process in Eq. 2, together with the analytical Landau formula, for $\mu = 0.001$; $N = 2 \cdot 10^5$; and $N_0 = 1, 2$, and 10 .

where $N > N_0$ and $m > 0$. As before, μ is the mutation rate. The probability that there are no mutants is clearly

$$P_N(0) = (1 - \mu)^{N - N_0}. \quad [11]$$

As in the original model, things are tremendously simplified if $\mu N \ll 1$, and we work to linear order in μ . Then, the master equation takes the form

$$P_{N+1}(m) = \frac{1}{N} [\delta_{m,1} \mu N + P_N(m)(N - m) + P_N(m - 1)(m - 1)]. \quad [12]$$

The last two terms are just birth without mutation, and the first term reflects the uniform rate of mutant creation, which arises from using the fact that $P_N(0) = 1 + O(\mu)$, whereas $P_N(m)$ is $O(\mu)$ for all $m \neq 0$. The solution of this recursion relation can be seen to be

$$P_N(m) = \mu \frac{(N - N_0)!(N - 1 - m)!(N - m + mN_0)}{m(m + 1)(N - 1)!(N - N_0 - m)}. \quad [13]$$

This is particularly simple for $N_0 = 1$, where $P_N(m) = \mu N / (m(m + 1))$ for $1 < m < N$, so that the power-law tail is sharply cut off past $m = N - 1$. In general, for large N ,

$$P_N(m) \simeq \frac{\mu N}{m(m + 1)} \left(1 + \frac{m(N_0 - 1)}{N}\right) \left(1 - \frac{m}{N}\right)^{N_0 - 1}, \quad [14]$$

and we again have a monotonic $1/m(m + 1)$ distribution, now with a smooth cutoff of the $1/m^2$ power-law tail as m approaches N . It is important to note that the presence of the power-law tail implies that all moments of the distribution are dominated by rare early birth, large m events, and thus are extremely difficult to measure in practice and strongly dependent on the initial conditions. As in the original model, this first order in μ solution breaks down for large enough N , such that μN is not small. Nonetheless, it is always valid (for small μ) for $m \sim O(N)$, which does, in fact, arise from early birth events that are always rare and so are treatable to linear order in μ . The vanishing of $P_N(m)$ for $m > N - N_0$ is an immediate consequence of the fixed N ensemble in which we are working. In the fixed time ensemble, the nature of the outer behavior is exponential, because there is no limit to the total population at any fixed time. We next turn to the problem of general μN for $m \ll N$, i.e., in the inner region.

Stochastic Model: Inner Solution

To derive the inner solution valid for large N and small μ (but with no constraint on μN), we have to allow for terms in the distribution that are of arbitrary order in μ . However, to each order, we just need to keep the leading order in N . It is easy to see that terms of order μ^k have leading order dependence N^k . If we denote the coefficient of this term in $P_N(m)$ as A_m^k , we find on substitution into the master equation (Eq. 12):

$$(k + m)A_m^k = (m - 1)A_{m-1}^k - A_m^{k-1} + A_{m-1}^{k-1}. \quad [15]$$

To derive this result, we used the fact that the coefficient of the term highest order in N (i.e., the term proportional to N^k) automatically vanishes, and the resultant equation emerges from the requirement that the coefficient of the order N^{k-1} term vanishes as well. This equation is supplemented with $A_m^0 = 0$ for $m > 1$ (because only the no-mutant sector is populated if $\mu = 0$) and with $A_0^k = (-1)^k / k!$, which follows from Eq. 11.

To solve this equation, we use a transform method. We multiply Eq. 15 by $x^m \cdot y^k$ and sum from $k = m = 1$, and we denote the resulting function as $F(x, y)$. The distribution we want will

be found from F via setting $y = \mu N$ and identifying $P_N(m)$ as the coefficient of x^m in F . Proceeding, we obtain

$$\left[y \frac{\partial}{\partial y} + x(1 - x) \frac{\partial}{\partial x}\right] F(x, y) = -y(1 - x)F + xye^{-y}. \quad [16]$$

We can convert this to a homogeneous equation, writing $F = G - e^{-y}$, so that

$$\left[y \frac{\partial}{\partial y} + x(1 - x) \frac{\partial}{\partial x}\right] G(x, y) = -y(1 - x)G. \quad [17]$$

This homogeneous equation can be solved by the method of characteristics. We write $G(x, y) = G(x(y), y) \equiv \hat{G}(y)$ and choose $dx(y)/dy = x(1 - x)/y$, whose solution is

$$x = \frac{y}{\alpha + y}, \quad [18]$$

where α labels the characteristic. Then, the equation for \hat{G} is the ordinary differential equation:

$$y \frac{d\hat{G}}{dy} = -\frac{\alpha y}{\alpha + y} \hat{G}. \quad [19]$$

Solving with the boundary condition $G(y = 0) = 1$ yields

$$G = \left(1 + \frac{y}{\alpha}\right)^{-\alpha}. \quad [20]$$

Finally, we restore the original functional dependence on x and y using (from Eq. 18) $\alpha = y(1 - x)/x$ to obtain

$$F = (1 - x)^{y(1-x)/x} - e^{-y}. \quad [21]$$

This result was originally derived by Lea–Coulson (2), but for the model where the WT population has deterministic growth dynamics. Note that the singular behavior at $x = 1$ is ultimately responsible for the $1/m^2$ behavior at large m . Explicitly, for small $y = \mu N$:

$$\begin{aligned} F &\approx \mu N \frac{1 - x}{x} \ln(1 - x) = -\mu N \sum_k \left[\frac{x^{k-1}}{k} - \frac{x^k}{k}\right] \\ &= \mu N \sum_{m=1}^{\infty} \frac{x^m}{m(m + 1)}, \end{aligned} \quad [22]$$

which is, of course, the lowest order solution away from the cutoff region. It should be noted that our scaling solution cannot be used to determine the exact answers to higher than linear order in μ , however, because we have dropped higher order terms in μ , terms that do not scale as y .

As already mentioned, we can recover P_m by extracting the coefficient of the x^m term in an expansion about $x = 0$. This can be accomplished by contour integration in a small circle around the origin:

$$P_m = \frac{1}{2\pi i} \oint \frac{F(x)}{x^{m+1}} dx. \quad [23]$$

The integrand has a branch cut along the positive x axis starting at $x = 1$, associated with the aforementioned singular behavior. We can thus transform the contour integral into an integral of the discontinuity along the branch cut:

$$P_m = \frac{1}{\pi} \int_1^{\infty} dx x^{-m-1} (x - 1)^{-y(x-1)/x} \sin \frac{\pi y(x-1)}{x}. \quad [24]$$

A similar expression based on the Lea–Coulson result appears in a paper by Kepler and Oprea (10).

If we further specialize to the case of large y , large m (because the scale of m is set by y), the integral is dominated by the region near $x = 1$. We thus write $m = y\bar{m}$ and denote $x = 1 + t/y$, giving

$$\begin{aligned}
 P_m &= \frac{1}{\mu N \pi} \int_0^\infty dt e^{-\bar{m}t - t \ln t/y} \sin \pi t \\
 &= \frac{1}{\mu N \pi} \int_0^\infty dt e^{-(\bar{m} - \ln y)t - t \ln t} \sin \pi t.
 \end{aligned}
 \tag{25}$$

This is just the Landau distribution, and our final answer is $P_N(m) = \frac{1}{\mu N} P_{\text{Landau}}(\frac{m}{\mu N} - \ln \mu N)$. Amazingly, this is exactly the same result [aside from a small shift in m of $\mu N(1 - \gamma)$] as we found for the original LD model. Note that this inner solution does not depend on N_0 , as we have come to expect. This lack of dependence of the inner solution on N_0 explains why the Lea–Coulson model, which ignores the stochasticity of the WT, produces the same large N distribution in the inner region as the fully stochastic model. Similarly, this is why the distribution agrees with that provided by Mandelbrot (18), who also used deterministic dynamics for the WT population. Because here there is no N_0 dependence, the effect of stochasticity can be made arbitrarily weak by taking N_0 large, and the models then coincide.

In Fig. 2, we compare this solution with an exact numerical calculation of the master equation with $N_0 = 1$, with the latter not assuming anything regarding μN . The agreement is essentially perfect for the case $\mu N = 200$ but deviates for small μ from the exact answer when μN is reduced by an order of magnitude. The Lea–Coulson distribution, however, still works perfectly. Both approximations work all the way out to N , because the cutoff is sharp for $N_0 = 1$. Finally, in Fig. 3, we multiply this inner solution by the appropriate outer solution to create a composite solution valid for all values of m . This restores the N_0 dependence and shows how the Landau distribution is modified for $m \sim \mathcal{O}(N)$.

Discussion

We have seen that in the fixed N ensemble for large N , small μ , the distribution can be divided into two regions, one an N_0 -independent inner region, where the Lea–Coulson generating function obtains, and the other an outer region, where N_0 is important and determines how the distribution is cut off. The

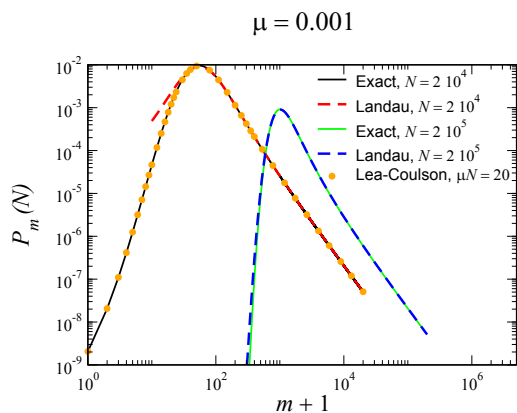


Fig. 2. Comparison of exact integration of the master equation (solid lines) vs. the Landau approximation (dashed lines) and the distribution represented by the Lea–Coulson generating function in Eq. 21 (dots) for $\mu = 0.001$, $N = 2 \cdot 10^4$ and $2 \cdot 10^5$, and $N_0 = 1$.

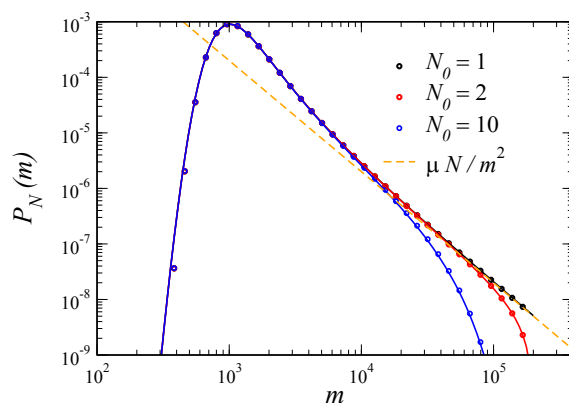


Fig. 3. Comparison of exact integration of the master equation (dots) vs. the Landau approximation (solid lines) with the appropriate $(1 + (N_0 - 1)m/N)/(1 - m/N)^{N_0 - 1}$ factor for $\mu = 0.001$; $N = 2 \cdot 10^5$; and $N_0 = 1, 2$, and 10 . Also shown is the asymptotic $\mu N/m^2$ behavior.

Lea–Coulson distribution interpolates between two simple limiting behaviors: the $\mu N/m(m + 1)$ monotonic decrease for small μN and the Landau distribution with its sharp rise at small m to a peak followed by a slow $1/m^2$ decay for large μN . In the first case, the most likely value of m is, of course, $m = 0$, whereas in the second, the mode is at $m \approx \mu N (\ln \mu N - 0.22)$. This is to be compared with the first moment of m , namely, $\bar{m} = \mu N \ln N$, which we see larger by approximately $\mu N \ln(0.8\mu)$. For a not untypical set of parameters, say $\mu = 10^{-7}$, $N = 10^8$, the mode is at approximately $m = 21$, whereas the mean is ~ 184 . This reiterates how uninformative the mean is as a characterization of the distribution. Similar statements hold for the variance, and using these to estimate parameters of the evolution process by comparison with data is essentially impossible. This is a direct consequence of the (cutoff) power-law nature of the distribution, which is the underlying mechanism of the LD fluctuation experiment. The convergence to the α -stable Landau distribution at large μN is another consequence of this. As already mentioned, only one previous work (18) appears to have noticed this fundamental property of this process and that aspect of this work seems to have been completely ignored.

In this paper, we have focused on the original model with deterministic growth as well as the stochastic variation, where birth events occur probabilistically. In extensions to be discussed elsewhere, we show that the results presented here can be generalized to include differential fitness effects as well as effects of death. For completeness, we note that in each of those cases, the distribution in the large μN limit is a one-sided Lévy α -stable distribution, whose asymptotic power-law is determined solely by the fitness difference of the two subpopulations. The extension of the Lea–Coulson generating function to the differential fitness case is known (18), and it is again the inner solution to the fixed N stochastic problem. No similar results exist, however, for the case in which death is included.

It should be pointed out that the stochastic LD model is exactly solvable, and a formal but not very tractable explicit expression for the probability distribution has appeared previously (22). This was shown to reproduce the Lea–Coulson generating function in the inner region. The outer region was not analyzed. More crucially, the connection to power-law tails and, in particular, the Landau distribution seems not to have been appreciated in this work. This exact solution also applies only for equal birth rates and in the absence of death; thus, our approximate technique is more informative in the general case. It also should be noted that the analytical exact solution is, as a numerical tool, less effective than a direct solution of the master

equation, because the number of operations is comparable and the exact solution involves terms of both signs, so that round off becomes an issue.

For the general case, Antal and Krapivsky (AK) (21) have recently derived a rather complex set of formulae embodying the solution for the generating function of the probability distribution in the fixed time ensemble, generalizing the results of Bartlett (20) for the neutral pure birth stochastic model. Bartlett succeeded in recovering the Lea–Coulson distribution in the limit of large N_0 , because the WT dynamics become deterministic there and the two ensembles coincide. To appreciate the problem, we first discuss the results of the time ensemble, working from the AK solution for the neutral pure birth model. The generating function of AK for $N_0 = 1$, which is called $A(y, x)$, where y stands for WT and x for mutants, satisfies

$$\dot{A} = A^2 + \nu B - (1 + \nu)A, \quad [26]$$

where B is the generating function for the pure mutant growth problem, starting from a single mutant, which is given by

$$B = -\frac{z}{1-z}, \quad z = -\frac{x}{1-x}e^{-\lambda_2 t}, \quad [27]$$

and $\nu = \mu/(1-\mu)$, $\lambda_2 = 1/(1-\mu)$. The solution for A is given by

$$A = -\frac{Cz}{(1-z)^\mu + C - Cz}, \quad [28]$$

which can be seen, using the definition of z , to satisfy Eq. 26. The constant C is determined by the condition that $A(t=0) = y$, so that

$$C = -\frac{y(1-x)^{1-\mu}}{y-x}. \quad [29]$$

The generating function for the mutants, independent of the number of WT cells, is given by $F = A(1, x)$. It is also convenient to work in terms of \bar{N} , the expected total population size, rather than t :

$$\bar{N} \equiv e^{\lambda_2 t}. \quad [30]$$

We thus obtain

$$F(x) = \frac{x}{x + \bar{N}(1-x) \left[1 - \left(1 - x + x/\bar{N} \right)^\mu \right]}. \quad [31]$$

We see immediately that the singularity at $x = 1$ has been shifted to $x = 1/(1 - 1/\bar{N}) > 1$, so that the asymptotic decay of $P(m)$ is exponential, and there is again a crossover from the inner region when $m \sim \mathcal{O}(N)$.

The limit we wish to consider is the large \bar{N} , small μ limit: $\mu\bar{N} \sim \mathcal{O}(1)$. In the inner region, we can drop the x/\bar{N} term, and we have

$$F(x) \approx \frac{x}{x - (1-x)\mu\bar{N} \ln(1-x)}, \quad [32]$$

so that, again, $\mu\bar{N}$ is the relevant control parameter. However, this is not the Lea–Coulson generating function. In particular, $P_0 = 1/(1 + \mu\bar{N})$, so that it does not decay exponentially with $\mu\bar{N}$ as in the fixed N ensemble. The reason for this is the presence of anomalous members of the fixed time ensemble with small total population sizes. Plotting the distribution that arises from this solution for various values of $\mu\bar{N}$ (Fig. 4), we see that there is no peak, ostensibly for the same reason.

Thus, going to the time ensemble makes a huge difference and, in fact, is not the relevant ensemble for application to cases

where one measures N , for example, by determining the tumor size. We can recover relevant results if we additionally condition on N , that is, pick only those members of the statistical ensemble that have exactly $N = \bar{N}(t)$ individuals. This is gotten by looking at the coefficient of y^N in A . We can do this analytically, because A is of the form $A_1 y / (A_2 y + A_3)$, where $A_1 = z$:

$$A_2 = z - 1 + \frac{[(1-z)(1-x)]^\mu}{1-x}, \quad A_3 = -\frac{x[(1-z)(1-x)]^\mu}{1-x}.$$

Thus, the normalized generating function for the subensemble is explicitly

$$Q(N, y) = \mathcal{N} \frac{A_1}{A_3} \left(\frac{A_2}{A_3} \right)^{N-1}, \quad [33]$$

where \mathcal{N} is the normalization factor, chosen so that $Q(N, 1) = 1$. Now, in the limit $x \rightarrow 1$,

$$\begin{aligned} z &\rightarrow -\frac{1}{N(1-x)} \rightarrow -\infty, \\ A_2 &\rightarrow -\frac{1}{N(1-x)} + \frac{N^{-\mu}}{1-x} = \frac{N^{-\mu}(1-N^{\mu-1})}{1-y}, \\ A_3 &\rightarrow -\frac{N^{-\mu}}{1-x}. \end{aligned} \quad [34]$$

Thus,

$$\begin{aligned} \mathcal{N} &= N^{1-\mu} (1 - N^{\mu-1})^{1-N} = N^{1-\mu} e^{-(N-1)\ln(1-N^{\mu-1})} \\ &\approx N^{1-\mu} e^{1+\mu \ln N} = N e^1. \end{aligned} \quad [35]$$

Given this, we now need to calculate the leading N behavior of $Q(N, y)$. For fixed y , large N , $z \rightarrow 0$, and

$$A_2/A_3 \approx -1 + \frac{1}{N} - \mu \frac{(1-x)\ln(1-x)}{x}, \quad [36]$$

and

$$A_1/A_3 \approx \frac{1}{N}. \quad [37]$$

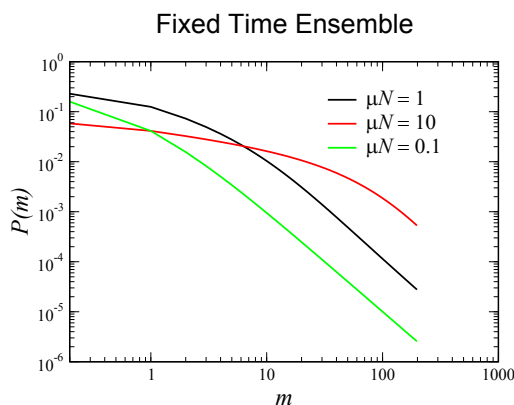


Fig. 4. Distribution for the fixed time ensemble resulting from Eq. 32 for $\mu\bar{N} = 0.1$, $\mu\bar{N} = 1$, and $\mu\bar{N} = 10$. Notice that in all cases, the distribution is monotonic decreasing and not at all like the Lea–Coulson result with its Landau limit at large $\mu\bar{N}$.

Collecting all the pieces, we arrive at the result that $Q(N, y)$ in this limit is equal to

$$Q(N, y) \approx Ne^1 \left(\frac{1}{N} \right) e^{-1 + \mu N \frac{(1-y)\ln(1-y)}{y}} \quad [38]$$

$$= (1-x)^{\mu N(1-x)/x}.$$

This is just the Lea–Coulson result, including the $m=0$ term; thus, it also contains the small μN and Landau limits discussed above. Of course, the true fixed N ensemble contains contributions from different elapsed times, but this has no impact, because conditioning on N forces the time to be close to the nominal time (although conditioning on time does not force N to be close to its mean value). The small variation in these times clearly does not change the distribution, as can be shown explicitly.

Methods

Here, we present more information regarding the generation of the curves presented in the figures. The data for the original LD model was obtained for

1. Luria SE, Delbrück M (1943) Mutation of bacteria from virus sensitivity to virus resistance. *Genetics* 28(6):491–511.
2. Lea DE, Coulson CA (1949) The distribution of the numbers of mutants in bacterial populations. *J Genet* 49(3):264–285.
3. Lenski RE, Slatkin M, Ayala FJ (1989) Mutation and selection in bacterial populations: Alternatives to the hypothesis of directed mutation. *Proc Natl Acad Sci USA* 86(8):2775–2778.
4. Coldman AJ, Goldie JH (1986) A stochastic model for the origin and treatment of tumors containing drug-resistant cells. *Bull Math Biol* 48(3-4):279–292.
5. Moolgavkar SH (1986) Carcinogenesis modeling: From molecular biology to epidemiology. *Annu Rev Public Health* 7(1):151–169.
6. Komarova N (2006) Stochastic modeling of drug resistance in cancer. *J Theor Biol* 239(3):351–366.
7. Tomasetti C, Levy D (2010) An elementary approach to modeling drug resistance in cancer. *Math Biosci Eng* 7(4):905–918.
8. Diaz LA, Jr., et al. (2012) The molecular evolution of acquired resistance to targeted EGFR blockade in colorectal cancers. *Nature* 486(7404):537–540.
9. Clayton E, et al. (2007) A single type of progenitor cell maintains normal epidermis. *Nature* 446(7132):185–189.
10. Kepler TB, Oprea M (2001) Improved inference of mutation rates: I. An integral representation for the Luria–Delbrück distribution. *Theor Popul Biol* 59(1):41–48.
11. Gerrish P (2008) A simple formula for obtaining markedly improved mutation rate estimates. *Genetics* 180(3):1773–1778.

each realization by picking k from a Poisson distribution with mean $\mu(N - N_0)$ and then summing k independent identically distributed random variables drawn from a $1/z^2$ distribution running from $z=1$ to $z=N/N_0$. The individual sums were then logarithmically binned to generate the histogram.

The Landau distribution, $P_{\text{Landau}}(x)$, was computed for $x > -2$ via the integral in Eq. 25. For $x < -2$, the form in Eq. 6 was used, where the region of integration was restricted to $\alpha > 0$ by symmetry, and the contour was then deformed to the steepest descent contour, extending from the critical point on the negative α axis off into the complex plane (23).

The data for the stochastic birth model were obtained via a direct solution of the master equation, Eq. 10. The data for the Lea–Coulson distribution came from a Taylor expansion of the generating function obtained from the symbolic algebra program Maple (Maplesoft). The data for the AK distribution came from a Maple-generated Taylor expansion of the corresponding generating function as well.

ACKNOWLEDGMENTS. This work was supported by the National Science Foundation Center for Theoretical Biological Physics (Grant PHY-1308264). H.L. was also supported by Cancer Prevention and Research Institute of Texas Scholar Program, and D.A.K. was also supported by the Israeli Science Foundation.

12. Dewanji A, Luebeck EG, Moolgavkar SH (2005) A generalized Luria–Delbrück model. *Math Biosci* 197(2):140–152.
13. Iwasa Y, Nowak MA, Michor F (2006) Evolution of resistance during clonal expansion. *Genetics* 172(4):2557–2566.
14. Zheng Q (1999) Progress of a half century in the study of the Luria–Delbrück distribution. *Math Biosci* 162(1-2):1–32.
15. Kashdan E, Pareschi L (2012) Mean field mutation dynamics and the continuous Luria–Delbrück distribution. *Math Biosci* 240(2):223–230.
16. Gentle JE (2003) *Random-Number Generation and Monte Carlo Methods* (Springer, NY), p 196.
17. Landau L (1944) On the energy loss of fast particles by ionization. *J Phys (USSR)* 8:201–205.
18. Mandelbrot B (1974) A population birth-and-mutation process!: Explicit distributions for the number of mutants in an old culture of bacteria. *J Appl Probab* 11(3):437–444.
19. Gnedenko BV, Kolmogorov AN (1954) *Limit Distributions for Sums of Independent Random Variables* (Addison Wesley, Reading, MA).
20. Bartlett MS (1955) *An Introduction to Stochastic Processes* (Cambridge Univ Press, UK).
21. Antal T, Krapivsky PL (2011) Exact solution of a two-type branching process: Models of tumor progression. *J Stat Mech* 2011(8):P08018.
22. Angerer WP (2001) An explicit representation of the Luria–Delbrück distribution. *J Math Biol* 42(2):145–174.
23. Börsch-Supan W (1961) On the evaluation of the function $\varphi(\lambda) = \frac{1}{2\pi} \int_{\sigma-i\infty}^{\sigma+i\infty} e^{u \ln u + \lambda u} du$ for real values of λ . *Journal of Research of the National Bureau of Standards B: Mathematics and Mathematical Physics* 65B(4):245–250.