axioms

MDPI

Article

Reversible Transitions in a Fluctuation Assay Modify the Tail of Luria–Delbrück Distribution

Pavol Bokes [1,*], Anna Hlubinová [1] and Abhyudai Singh [2]

[1] Department of Applied Mathematics and Statistics, Comenius University, 84248 Bratislava, Slovakia
[2] Departments of Electrical and Computer Engineering, Biomedical Engineering, University of Delaware, Newark, DE 19716, USA
* Correspondence: pavol.bokes@fmph.uniba.sk

Abstract: We consider a fluctuation test experiment in which cell colonies were grown from a single cell until they reach a given population size and were then exposed to treatment. While they grow, the cells may, with a low probability, acquire resistance to treatment and pass it on to their offspring. Unlike the classical Luria–Delbrück fluctuation test, and motivated by recent work on drug-resistance acquisition in cancer/microbial cells, we allowed the resistant cell state to switch back to a drug-sensitive state. This modification did not affect the central part of the Luria–Delbrück distribution of the number of resistant survivors: the previously developed approximation by the Landau probability density function applied. However, the right tail of the modified distribution deviated from the power law decay of the Landau distribution. Here, we demonstrate that the correction factor was equal to the Landau cumulative distribution function. We interpreted the appearance of the Landau laws from the standpoint of singular perturbation theory and used the asymptotic matching principle to construct uniformly valid approximations. Additionally, we describe the corrections to the distribution tails in populations initially consisting of multiple sensitive cells, a mixture of sensitive and resistant cells, and a cell with a randomly drawn state.

Keywords: Luria–Delbrück fluctuation test; reversible state switching; back-mutation; mutation rate; Landau distribution; mathematical biology

MSC: 60J80; 92D25; 60E07; 34E20

check for updates

Citation: Bokes, P.; Hlubinová, A.; Singh, A. Reversible Transitions in a Fluctuation Assay Modify the Tail of Luria–Delbrück Distribution. *Axioms* **2023**, *12*, 249. https://doi.org/10.3390/axioms12030249

Academic Editor: Qi Zheng

Received: 26 December 2022
Revised: 16 February 2023
Accepted: 18 February 2023
Published: 1 March 2023

Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

In a fluctuation experiment, a colony of bacterial/cancer cells is grown and eventually exposed to a treatment [1]. Most cells are killed, but some are resistant to the treatment. Repeated experiments each give a different number of resistant cells. The variability in the number of survivors is key in a fluctuation test. If the resistance is acquired when the treatment is administered, a low Poissonian variability is expected; on the other hand, if it can be acquired by a cell at an early stage of colony growth and passed on to its (many) descendants, variability is large. Resistant cells are typically assumed to have developed a genetic mutation and are referred to as "mutants". However, several studies have shown that stochastic expression of specific proteins can lead to resistance to drug therapy arising transiently. In such cases, cells switch to becoming resistant for several generations, but then switch back into a drug-sensitive state [2–6]. Several tools that exploit colony-to-colony variation to estimate the rates of switching into and out of the resistant state have recently been developed [7,8].

Mathematical analysis has been applied for decades to quantify variability in a fluctuation test experiment under different modelling assumptions and parameter regimes [9]. The modelling framework is either semideterministic or fully stochastic [10]. In the former case, proliferation is deterministic, and the acquisition of resistance is stochastic [11–13]; in the latter case, both are stochastic. A fully stochastic framework further differentiates into

fixed-time or fixed-population statistical ensembles. In the former case, one waits until a given end time and then measures the cell population [14]. The final measurement includes the effect of total population noise. In the fixed population case, the measurement is taken when the total cell population reaches a given level [15,16]. Such an approach excludes the population noise and is thus consistent with the semideterministic framework. An alternative approach is to look at the proportion of resistant cells rather than the absolute number [8].

In most models, resistance is never lost once acquired: resistant mothers bear resistant daughters. Such models are asymmetric, whereas ones in which bidirectional change is possible are structurally symmetric. The possibility of a "back-mutation" was considered in [17,18] and modelled deterministically in [19]. Back-mutation was inherently present in a recent innovative model that combined population growth with mutations in a sequence of nucleotides [20]. The loss of resistance through random changes in gene expression was suggested to reduce variability in a fluctuation experiment [2,7,8].

The aim of this paper is to understand the probability distribution of the number of resistant cells in a structurally symmetric, fully stochastic, fixed-population-ensemble fluctuation test. Some explicit results are available in the asymmetric case, but they reveal little [15]. In the asymptotic regime of small probability of resistance acquisition, the Luria–Delbrück distribution of the number of surviving resistors can be approximated through (discrete) Lea–Coulson distribution [10,15,21,22]. Lea–Coulson distribution has a shape parameter that corresponds to the population rate of resistance acquisition. If this is sufficiently large, (discrete) Lea–Coulson distribution simplifies to its coarse-grained continuous counterpart, the Landau distribution [10,23,24]. The Landau distribution is one-sided stable with characteristic exponent 1 [25–27]. Landau approximation is universal (or self-similar) because the resistance rate only affects its position and scale parameters. Both Lea–Coulson and Landau distributions have a heavy tail. Our paper contributes by characterizing the correction to the power-law tail that is due to loss of resistance.

Lea–Coulson and Landau distributions approximate the Luria–Delbrück distribution provided that the overwhelming majority of cells are sensitive, and few are resistant. This represents a small layer at the left boundary of the interval of all possible sizes of the resistant subpopulation. In a nod to the classical asymptotic theory of ordinary differential equations [28], we refer to the Lea–Coulson/Landau approximations as the (left) boundary layer solution. This approximation can be combined with a regular power series solution to construct a composite solution that uniformly approximates the Luria–Delbrück distribution [10]. Here, we argue that the reversal of resistance introduces another (right) boundary layer solution that is valid if the overwhelming majority of cells are resistant. We constructed a uniform solution that combined the regular power series solution with two boundary-layer solutions.

Our results were methodologically based on the matched asymptotic framework. While falling short of providing rigorous estimates, matched asymptotics have been an effective tool throughout applied mathematics [29], including stochastic modelling [30–32].

2. Model Formulation

The growing population of resistant and sensitive cells is modelled with a two-type branching process in continuous time [33]. In the model, each cell has an exponentially distributed cell-cycle duration that is independent of other cells in the population. We assumed that the mean cell cycle was the same for sensitive and resistant cells. Cells divide at the end of the cell cycle; the state of the daughter cell is chosen probabilistically depending on the mother cell, as given in Figure 1.

We consider this branching process at the time points of cell division. If there are N cells immediately before a cell division time point of which m are resistant, then the dividing cell is resistant with probability m/N and sensitive with complementary probability $1 - m/N$. As shown in Figure 1, the daughter cell is, thus, resistant with probability $(1 - \lambda)m/N + \mu(1 - m/N)$ or sensitive with complementary probability $(1 - \mu)(1 - m/N) + \lambda m/N$. As

the total population increases by 1 in the cell division event, the number of resistant cells can remain the same or increase by 1 depending on the state of the new cell (Figure 2a).

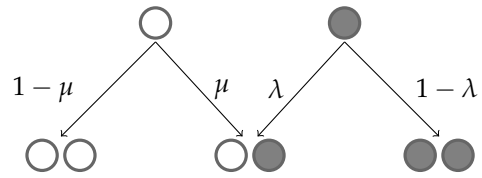


Figure 1. White balls in the diagram represent sensitive cells, while black balls represent resistant cells. Arrows pointing from the top tier balls (mother cells) downwards illustrate the four possible outcomes of cell division: (**left**) Either a sensitive parent cell has a sensitive descendant or its offspring is resistant with probabilities $1-\mu$ and μ, respectively; (**right**) a resistant cell may produce a sensitive or resistant offspring with probabilities λ and $1-\lambda$, respectively.

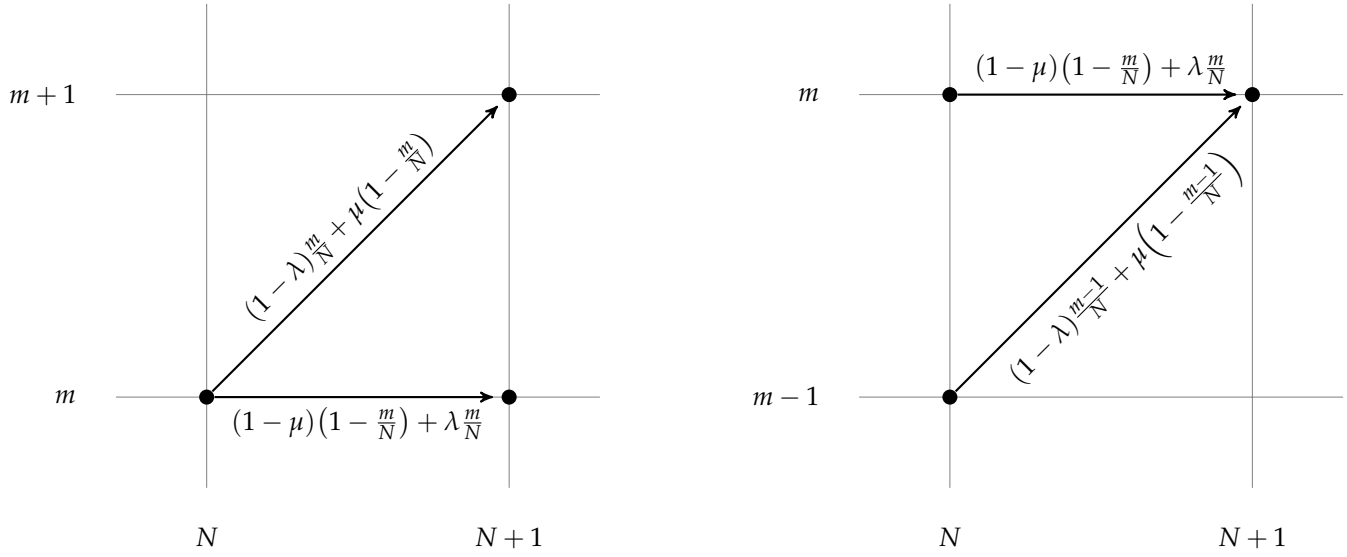


Figure 2. Forward and adjoint schematic illustrations of transitional probabilities in the Markov chain model. Parameters μ and λ are the resistance acquisition and loss probabilities, respectively. (**a**) Forward diagram: given a population of size N containing m resistant cells, the probabilities that one offspring becomes resistant or sensitive are illustrated. (**b**) Adjoint diagram: considering a population of size $N+1$ containing m resistant cells, the probabilities of the two possible ways by which the state could have been reached are illustrated.

The aim of this paper was to characterise the probability $P_N(m)$ of observing m resistant cells as the total population reached N. This probability could be evaluated recursively with a variation in the above argument. In order to have m resistant cells in the population of $N+1$ cells after a cell division event, we must have had m resistant cells before the division and added a sensitive cell or added an extra resistant cell to the $m-1$ resistant cells before the division (Figure 2b). Adding up the probabilities of these two options, $P_N(m)$ satisfies the master equation:

$$\begin{aligned} P_{N+1}(m) = \frac{1}{N}\Big\{ & P_N(m-1)\Big[\mu(N-(m-1))+(1-\lambda)(m-1)\Big] \\ & + P_N(m)\Big[(1-\mu)(N-m)+\lambda m\Big]\Big\}. \end{aligned} \tag{1}$$

Equation (1) was considered in [34]. For $\lambda = 0$, (1) is simplified into the master equation for the irreversible case as previously studied in [10].

Master Equation (1) is studied subject to the following boundary conditions:

$$P_N(m) = 0 \quad \text{if} \quad m < 0 \text{ or } m > N \tag{2}$$

and the following initial condition:

$$P_{N_0}(m) = \begin{cases} 1 & \text{if } m = m_0, \\ 0 & \text{otherwise,} \end{cases} \tag{3}$$

where N_0, the initial population size; m_0, the initial resistant subpopulation size.

3. Methods

Following the previous studies in [10,11] and empirical evidence [2], we assumed that the resistance acquisition and loss probabilities were small perturbation parameters $\mu \ll 1$ and $\lambda \ll 1$. The aim of this paper was to characterise the asymptotic behaviour of the solution to Master Equation (1) subject to (2) and (3) as the perturbation parameters tended to zero. Therefore, we assumed that they tended to zero with the same rate: $\mu/\lambda = O(1)$.

3.1. Overview of Asymptotic Approximations

The problem in question is singularly perturbed, and there are several alternative distinguished approximations that apply to different regions of the definition domain $0 \le m \le N < \infty$ of the solution $P_N(m)$, and intermediate approximations through which the former can be matched. Additionally, the analysis proceeds differently depending on whether we start with a homogeneous population, i.e., with $m_0 = 0$ resistant and $N_0 > 0$ sensitive cells or $m_0 = N_0 > 0$ resistant cells, or we start with a mixed population with $m_0 > 0$ resistant and $N_0 - m_0 > 0$ sensitive cells. Our aim was to treat all the possible combinations exhaustively; to help the reader in navigating through the resulting complexity, Table 1 outlines the region of validity for the individual approximations and gives references to their formulae. Below, we provide a brief commentary on the table. The derivations are found in the subsections that follow.

The first asymptotic approximation was developed under the assumption of bounded population sizes $N = O(1)$ as the perturbation parameters tended to zero. The number of resistant cells was then also bounded via the natural constraint $m \le N$. Mathematically, the solution was developed into a regular power series in a perturbation parameter and approximated with the leading order term. Therefore, it is a regular approximation to the solution or simply regular solution. The approximation occupies the first row of Table 1 and is derived in Section 3.2.

Regular approximation is an example of a distinguished limit [35]: the variables are of definite order $O(1)$. Another distinguished limit occurs if the total population size N is inversely proportional to the small perturbation parameter while the number m of resistant cells is $O(1)$ (Table 1, row 3, column 1). This is the large-population and low-mutation rate scenario discussed in previous works on the unidirectional case [10,11]. Given that $m \ll N$, this scenario represents a left boundary layer of the definition interval $0 \le m \le N$ of the solution $P_N(m)$. Our structurally symmetric model could also develop a right boundary layer in which the number $N - m$ of sensitive cells is $O(1)$ (Table 1, row 3, column 3).

Three intermediate limits are obtained by taking the total large population in the regular solution and considering various scales of resistant and sensitive subpopulations (Table 1, row 2; Sections 3.3–3.5). The intermediate left and right regular solutions are used for the asymptotic matching of the distinguished left and right solutions (Sections 3.6 and 3.7). The coarse-grained regular solution is applicable to the space between the two boundary layers (Table 1, row 3, middle column) and can be combined with the layer solutions into a uniform composite approximation (Section 3.8).

The triangular character of definition domain $0 \le m \le N < \infty$ in Table 1 renders the concept of inner and outer solutions ambiguous. On the one hand, following [10], one can refer to the boundary layer solutions as inner because they focus on small fractions

of the total population. On the other hand, following classical applications of singular perturbations in biology [36], one can also refer to the regular solution as inner because it captures the system's early dynamics. In order to avoid a clash with earlier conventions, we do not use "inner and outer solutions" in this work.

Table 1. Overview of approximate solutions to Equation (1) and their regions of validity within definition domain $0 \leq m \leq N < \infty$. The first reference indicates a formula that applies for a mixed initial population; the second reference indicates an exclusively sensitive initial population. The linked formulae contain population transition Rates (23) and the (partial summations of the) Lea–Coulson probability mass function (33)–(34).

	$m = O(1)$	$m = O(N) = N - m$	$N - m = O(1)$
$N = O(1)$		regular (4), (10)	
$1 \ll N \ll \frac{1}{\mu}$	regular left (13), (14)	regular coarse-grained (20), (21)	regular right (16), (17)
$N = O\left(\frac{1}{\mu}\right)$	left (32), (36)		right (40) (41)

3.2. Regular Solution

Through a regular solution to Master Equation (1), we find a solution that regularly depends on perturbation parameter μ , i.e., is a power series in μ. Our aim was to determine this solution at the leading order, i.e., to provide the first nonzero term of the power series. Cases $m_0 > 0$ and $N_0 - m_0 > 0$ (initially present resistant and sensitive cells) were, perhaps surprisingly, easier than the $m_0 = 0$ and $N_0 > 0$ cases or the symmetrically opposite case of $m_0 = N_0 > 0$, and were treated first.

3.2.1. Case $m_0 > 0$ and $N_0 - m_0 > 0$

Rather than formally solving Difference Equation (1), it is both easier and more instructive to derive the regular solution from probabilistic considerations. The leading-order behaviour of the regular solution was obtained simply by neglecting perturbations μ and λ in (1). The resulting equation describes a growing culture of cells that deterministically pass their state to their progeny. Growth is as follows: one chooses a cell at random and looks at its state; the chosen cell is returned to the culture, and another cell (a daughter) is added that has the same state.

Aside from the terminology (cells = balls; state = colour), this is a classical Polya urn model [37]. A detailed description of this model can be found in Appendix A. As a result, with μ (and λ) approaching zero,

$$P_N(m) \sim \frac{(m-1)!(N-m-1)!(N-N_0)!}{B(m_0, N_0-m_0)(m-m_0)!(N-m-N_0+m_0)!(N-1)!}, \tag{4}$$

where $m_0 \leq m \leq N - N_0$ and the tilde sign are understood in the usual sense of asymptotic expansions [35]. $B(x,y) = \Gamma(x)\Gamma(y)/\Gamma(x+y) = (x-1)!(y-1)!/(x+y-1)!$ is the beta function.

3.2.2. Case $m_0 = 0$, $N_0 > 0$

The case of an entirely sensitive initial population is of prime interest here. It is more difficult than the mixed initial population case in Section 3.2.1. We used the previous case in deriving the new result. First, neglecting perturbations μ and λ in (1) leads to a problem with a trivial solution, namely, one always chooses a sensitive cell and keeps adding new sensitive cells. Mathematically,

$$\begin{aligned} P_N(0) &= 1 + O(\mu), \\ P_N(m) &= O(\mu),\ m > 0. \end{aligned} \tag{5}$$

The second equation in (5) tells us that the probability of finding some sensitive cells is small, but does not describe the leading-order behaviour—here, O(μ). For this, we needed a subtler estimate that we once again obtained from probabilistic considerations rather than formal solution techniques for difference equations.

Let $\tilde{L}$ denote the population size, such that the first $\tilde{L}$ cells are sensitive, and the $(\tilde{L}+1)$-th cell is resistant. Its distribution is shifted geometrically: $\mathrm{Prob}(\tilde{L}=L) = (1-\mu)^{L-N_0}\mu$, $L \geq N_0$. Conditioning the value of $\tilde{L} = L$, we have

$$P_N(m) = \sum_{L=N_0}^{N-m} P(m, N|1, L+1)\ (1-\mu)^{L-N_0}\ \mu, \tag{6}$$

where $P(m, N|1, L+1)$ denotes the probability of observing m resistant cells in a population of size N that was initially of size $L+1$ and contained one resistant cell. This can be estimated by setting $m_0 = 1$ $N_0 = L+1$ into (4), which yields

$$P(m, N|1, L+1) \sim \frac{L\ (N-m-1)!\ (N-L-1)!}{(N-1)!\ (N-m-L)!}. \tag{7}$$

Since the main aim was to derive a leading-order approximation of $P_N(m)$, we could neglect μ in the factors $(1-\mu)$ appearing in (6). As a result, we obtained

$$P_N(m) \sim \mu\ \frac{(N-m-1)!}{(N-1)!} \sum_{L=N_0}^{N-m} L\ \frac{(N-L-1)!}{(N-m-L)!}. \tag{8}$$

Moreover, using symbolic software (details in Appendix B), the sum appearing in the previous equation can be simplified as follows:

$$\sum_{L=N_0}^{N-m} L\ \frac{(N-L-1)!}{(N-L-m)!} = \frac{(N-N_0)!\ (N-m+mN_0)}{m\ (m+1)\ (N-N_0-m)!}, \tag{9}$$

giving the final result

$$P_N(m) \sim \mu\ \frac{(N-N_0)!\ (N-m-1)!\ (N-m+mN_0)}{m(m+1)\ (N-1)!\ (N-N_0-m)!}, \quad m > 0. \tag{10}$$

Formula (10) was derived with other means for the unidirectional model in [10].

3.2.3. Case $m_0 = N_0 > 0$

This case is a mirror reflection of the case discussed in Section 3.2.2. The result, which reads $P_N(N) \sim 1$ and

$$P_N(m) \sim \lambda\ \frac{(N-N_0)!\ (m-1)!\ (m+(N-m)N_0)}{(N-m)(N-m+1)\ (N-1)!\ (m-N_0)!}, \quad m < N, \tag{11}$$

was readily obtained by replacing μ by λ, $N_0 - m_0$ by m_0, $N-m$ by m, and vice versa in (10).

3.3. Regular Left Solution

A regular left solution means the approximate solution that considers a relatively large total population $1 \ll N \ll \frac{1}{\mu}$ that includes only a small number of resistant cells $m = O(1)$. The regular left solution can be derived from regular solution Formulae (4), (10), and (11) by approximating descending factorials by powers as appropriate; for example, one has

$$\frac{(N-N_0)!}{(N-1)!} = \frac{1}{(N-1)(N-2)\ldots(N-N_0+1)} \sim \frac{1}{N^{N_0-1}} \tag{12}$$

for $N \gg 1$.

1. Case $m_0 > 0$ and $N_0 - m_0 > 0$

$$P_N(m) \sim \frac{(m-1)(m-2)\dots(m-m_0+1)}{N^{m_0} B(m_0, N_0 - m_0)}. \tag{13}$$

2. Case $m_0 = 0$, $N_0 > 0$

$$P_N(m) \sim \begin{cases} 1 & \text{if } m = 0, \\ \frac{\mu N}{m(m+1)} & \text{otherwise.} \end{cases} \tag{14}$$

3. Case $m_0 = N_0 > 0$

$$P_N(m) \sim \frac{\lambda N_0 (m-1)(m-2)\dots(m-N_0+1)}{N^{N_0}}. \tag{15}$$

3.4. Regular Right Solution

A regular right solution is the approximate solution that considers a relatively large total population, $1 \ll N \ll \frac{1}{\mu}$ that includes a large number of resistant cells; mathematically, $N - m = O(1)$. These approximations are most conveniently obtained by performing substitutions $m \leftrightarrow N - m$, $m_0 \leftrightarrow N_0 - m_0$, and $\lambda \leftrightarrow \mu$ in (13), (14), and (15).

3.4.1. Case $m_0 > 0$ and $N_0 - m_0 > 0$

The approximation reads

$$P_N(m) \sim \frac{(N-m-1)(N-m-2)\dots(N-m-(N_0-m_0)+1)}{N^{N_0-m_0}\, B(m_0, N_0 - m_0)}. \tag{16}$$

3.4.2. Case $m_0 = 0$, $N_0 > 0$

The approximation reads

$$P_N(m) \sim \frac{\mu\, N_0 (N-m-1)(N-m-2)\dots(N-m-N_0+1)}{N^{N_0}}. \tag{17}$$

3.4.3. Case $m_0 = N_0 > 0$

The approximation reads

$$P_N(m) \sim \begin{cases} \frac{\lambda N}{(N-m)(N-m+1)} & \text{if } m < N, \\ 1 & \text{if } m = N. \end{cases} \tag{18}$$

3.5. Regular Coarse-Grained Solution

A regular coarse-grained solution is an approximate solution that considers a large number of cells in the total population, while the order of the number of resistant and sensitive cells in the population is the same. Mathematically, $m = O(N) = N - m$. The regular coarse-grained solution was obtained, similarly as regular left/right solutions, by approximating descending factorials by appropriate powers in regular solution Formulas (4), (10), and (11). For example,

$$\begin{aligned} \frac{(N-m-1)!}{(N-m-N_0+m_0)!} &= (N-m-1)\,(N-m-2)\dots(N-m-N_0+m_0+1) \\ &\sim (N-m)^{N_0-m_0-1} \end{aligned} \tag{19}$$

holds for m and N both large and of the same order.

3.5.1. Case $m_0 > 0$ and $N_0 - m_0 > 0$

The approximation reads

$$P_N(m) \sim \frac{m^{m_0-1}\,(N-m)^{N_0-m_0-1}}{N^{N_0-1}B(m_0, N_0-m_0)}. \tag{20}$$

3.5.2. Case $m_0 = 0$, $N_0 > 0$

The approximation reads

$$P_N(m) \sim \frac{\mu\,(N-m+mN_0)}{m^2}\left(1-\frac{m}{N}\right)^{N_0-1}. \tag{21}$$

3.5.3. Case $m_0 = N_0 > 0$

The approximation reads

$$P_N(m) \sim \frac{\lambda\,(m+(N-m)N_0)}{(N-m)^2}\left(\frac{m}{N}\right)^{N_0-1}. \tag{22}$$

The (20) relates to the classical convergence result to the beta distribution for the Polya urn [38]; (21) and (22) are symmetric.

3.6. Left Boundary Layer Solution

As in the previous sections, we assumed that the resistance acquisition and loss probabilities were small, i.e., $\mu \ll 1$ and $\lambda \ll 1$ such that $\lambda/\mu = O(1)$, and that the total population size was large, i.e., $N \gg 1$. In contrast with the previous sections, here, we focus on the distinguished large-population low-mutation-rate regime in which the population is so large as to produce the population rates of resistance acquisition and loss

$$y = \mu N, \quad \tilde{y} = \lambda N, \tag{23}$$

of order 1. The distinguished limit was found for $m = O(1)$ resistant cells. This represents a (left) boundary layer of the large interval of admissible values $m \in \{0, 1, \ldots, N\}$. The right boundary layer occurring for $N - m = O(1)$ is treated in the next section with symmetry arguments. Unless stated otherwise, asymptotic equivalence sign $\sim$ is used in this section in the context of limit process $\mu \to 0$, $y = O(1)$, $\tilde{y} = O(1)$, $m = O(1)$.

The boundary-layer solution is found with the method of the generating function, which is defined via

$$G_N(x) = \sum_{m=0}^{N} P_N(m)x^m. \tag{24}$$

The usual rules [39] transform Master Equation (1) into a difference–differential equation:

$$N(G_{N+1}(x) - G_N(x)) = (x-1)(\mu N + (1-\mu-\lambda)x\partial_x)G_N(x) \tag{25}$$

for the generating function.

Guided by Intermediate Asymptotics (13), (14), and (15) of the regular solution, we looked for a generating function with the same kind of asymptotics:

$$G_N(x) \sim N^{-\beta}H(x,y), \quad 0 < x < 1, \tag{26}$$

where β is to be matched to a regular solution.

$$G_{N+1}(x) = (N+1)^{-\beta}H(x, y(1+1/N)) \sim N^{-\beta}H + N^{-\beta-1}(-\beta H + y\partial_y H) + O(N^{-\beta-2}). \tag{27}$$

Inserting (26) and (27) into (25), and collecting leading order terms, we obtained the following partial differential equation:

$$-\beta H + y\partial_y H = (x-1)(y + x\partial_x)H. \tag{28}$$

For $\beta = 0$, (28) is the Lea–Coulson equation [9]. Parameter β can be viewed as an eigenvalue, and solution H as an eigenfunction of the Lea–Coulson differential operator.

Solving (28) with the method of characteristics, one uncovers the following parametric family of solutions:

$$H(x,y) = Cy^{\gamma}\left(\frac{x}{1-x}\right)^{\beta-\gamma}(1-x)^{y(1-x)/x}, \tag{29}$$

where C and γ are arbitrary constants. In what follows, we determine these constants and eigenvalue β by matching them to the regular solution. The distinct previously introduced cases were thereby treated separately.

3.6.1. Case $m_0 > 0$, $N_0 - m_0 > 0$

The limiting operations of infinite summation (to obtain a generating function from a sequence of probabilities) and the asymptotic limit process can be interchanged for $x < 1$ using a dominated convergence-type argument. Specifically, the left–regular approximation (13) can be expressed in terms of the generating function as follows:

$$G_N(x) \sim N^{-m_0}\frac{\Gamma(N_0)}{\Gamma(N_0-m_0)}\left(\frac{x}{1-x}\right)^{m_0}, \quad 0 < x < 1. \tag{30}$$

The approximation is valid in the left boundary layer $m = O(1)$ for an intermediate range of population sizes $1 \ll N \ll 1/\mu$.

Matching (30) with (26) and (29) yields $\gamma = 0$, $\beta = m_0$, and $C = \Gamma(N_0)/\Gamma(N_0 - m_0)$, i.e.,

$$G_N(x) \sim N^{-m_0}\frac{\Gamma(N_0)}{\Gamma(N_0-m_0)}\left(\frac{x}{1-x}\right)^{m_0}(1-x)^{y(1-x)/x}, \quad 0 < x < 1, \tag{31}$$

which is valid in the left boundary layer $m = O(1)$ for population sizes extending up to $N = O(1/\mu)$.

The right-hand side of (31) can be transformed into the probability space by combining the following rules:

- x transforms into the shift operator.
- $(1-x)^{-1}$ transforms into the summation operator (a discrete convolution with a sequence of ones).
- $(1-x)^{-k}$ transforms into the kth power of this operator (a discrete convolution with the sequence $\{\Gamma(k+m)/\Gamma(k)m!\}_{m\geq 0}$).
- $(1-x)^{y(1-x)/x}$ transforms into the Lea–Coulson probability mass function (PMF).

Therefore, Approximation (31), when expressed in terms of the probability mass function, reads

$$P_N(m) \sim N^{-m_0}\frac{\Gamma(N_0)}{\Gamma(N_0-m_0)}P_{L-C}^{(m_0)}(m-m_0;y), \tag{32}$$

where $P_{L-C}^{(k)}(m;y)$ is the kth summation of the Lea–Coulson probability mass function, i.e.,

$$P_{L-C}^{(0)}(m;y) = \left.\frac{\partial_x^m(1-x)^{y(1-x)/x}}{m!}\right|_{x=0}, \tag{33}$$

$$P_{L-C}^{(k)}(m;y) = \sum_{j=0}^{m}P_{L-C}^{(k-1)}(j;y) = \sum_{j=0}^{m}\frac{\Gamma(k+m)}{\Gamma(k)m!}P_{L-C}^{(0)}(m-j;y), \quad k = 1,2,\ldots \tag{34}$$

In particular, $P^{(0)}_{L-C}(m;y)$ is the Lea–Coulson PMF and $P^{(1)}_{L-C}(m;y)$ is the Lea–Coulson cumulative distribution function (CDF).

3.6.2. Case $m_0 = 0$, $N_0 > 0$

The left-regular intermediate approximation (14) implies $G_N(x) \sim 1$ for $1 \ll N \ll 1/\mu$ at leading order. Matching this to (29) and (26) gives $\beta = \gamma = 0$, $C = 1$. The left boundary approximation then reads

$$G_N(x) \sim (1-x)^{y(1-x)/x} \tag{35}$$

in the generating function space and

$$P_N(m) \sim P^{(0)}_{L-C}(m;y) \tag{36}$$

in the probability space. In this case, the matching procedure is equivalent to directly imposing the initial condition (of starting without any resistant cells) on the solution to the Lea–Coulson equation [10].

3.6.3. Case $m_0 = N_0 > 0$

Intermediate Approximation (15) implies

$$G_N(x) \sim \tilde{y} N^{-N_0-1} \Gamma(N_0+1) \left(\frac{x}{1-x}\right)^{N_0}, \quad 0 < x < 1, \tag{37}$$

for $1 \ll N \ll 1/\mu$. Matching with (29) and (26) yields $\gamma = 1$, $\beta = N_0 + 1$, $C = \Gamma(N_0 + 1)\lambda/\mu$, i.e.,

$$G_N(x) \sim \tilde{y} N^{-N_0-1} \Gamma(N_0+1) \left(\frac{x}{1-x}\right)^{N_0} (1-x)^{y(1-x)/x}, \quad 0 < x < 1. \tag{38}$$

This reads

$$P_N(m) \sim \tilde{y} N^{-N_0-1} \Gamma(N_0+1) P^{(N_0)}_{L-C}(m - N_0; y) \tag{39}$$

in the probability space.

3.7. Right Boundary Layer Solution

Symmetry considerations imply that $P_N(m) = \bar{P}_N(N-m)$, where the bar indicates a model in which the substitutions $m_0 \leftrightarrow N_0 - m_0$ and $\lambda \leftrightarrow \mu$ ($\tilde{y} \leftrightarrow y$) are performed. Combining the symmetry result with the previous section's results on the left boundary layer yields the following list of approximations. The equivalence sign is used in the context of $\mu \to 0$, $y = O(1)$, $\tilde{y} = O(1)$, $N - m = O(1)$.

3.7.1. Case $m_0 > 0$, $N_0 - m_0 > 0$

The approximation reads

$$P_N(m) \sim N^{-N_0+m_0} \frac{\Gamma(N_0)}{\Gamma(m_0)} P^{(N_0-m_0)}_{L-C}(N - m - N_0 + m_0; \tilde{y}). \tag{40}$$

3.7.2. Case $m_0 = 0$, $N_0 > 0$

The approximation reads

$$P_N(m) \sim y N^{-N_0-1} \Gamma(N_0+1) P^{(N_0)}_{L-C}(N - m - N_0; \tilde{y}). \tag{41}$$

3.7.3. Case $m_0 = N_0 > 0$

The approximation reads

$$P_N(m) \sim P^{(0)}_{L-C}(N - m; \tilde{y}). \tag{42}$$

3.8. Log-Composite Solution

Composite solutions provide a means of combining multiscale asymptotic approximations into a single uniformly valid approximation. Typically, the composite solution is equal to the sum of the inner and outer solutions minus their common limit value in the overlap region; this construction was applied to, e.g., enzyme kinetics [36] and the mathematical analysis of a hematopoietic genetic switch [40]. In a log-composite solution, addition and subtraction are replaced by multiplication and division.

For $N = O(1/\mu)$, three approximations are available in different sections of the definition domain $m \in \{0, 1, \ldots, N\}$: the left boundary layer solution applies to $m = O(1)$; the regular coarse-grained solution applies to $m = O(N) = N - m$; the right boundary layer solution applies to $N - m = O(1)$. Table 1, row 3, contains references to the specific formulae for the mixed and exclusively sensitive initial conditions. The product of these three approximations provides the numerator of the log-composite solution, which is constructed according to the following scheme:

$$P_N(m) \sim \frac{\text{left} \times \text{regular coarse-grained} \times \text{Right}}{\text{left overlap} \times \text{right overlap}}, \tag{43}$$

in which the 'left', 'right', and 'regular coarse-grained' labels are replaced with the appropriate formulae in the usual cases of different initial conditions. Let us now define the 'left overlap' and 'right overlap' terms in the denominator; for concreteness, consider the case $m_0 = 0$ of a sensitive initial population. The overlap between Left Solution (36) and Regular Coarse-Grained Solution (21) can be obtained by taking $m \to \infty$ in the former or $m \to 0$ in the latter; both procedures yield the same left overlap solution $P_N(m) \sim \mu N/m^2$ [23] because of the asymptotic matching principle [35]. Similarly, taking $N - m \to \infty$ in Right Solution (41) or $N - m \to 0$ in the Regular Coarse-Grained Solution (21) yields right overlap solution $P_N(m) \sim \mu N_0 (1 - m/N)^{N_0 - 1}/N$. For any value of m, one of the three factors in the numerator of (43) provides a valid approximation to the exact distribution, while the two remaining factors reduce to their overlap form and cancel with the denominator. Thus, Formula (43) provides a uniform approximation across the entire interval of admissible values of m.

The application of Pattern (43) in the usual cases leads to the following.

3.8.1. Case $m_0 > 0$, $N_0 - m_0 > 0$

The approximation reads

$$P_N(m) \sim N^{-N_0+1} \Gamma(N_0) P_{L-C}^{(m_0)}(m - m_0; y) P_{L-C}^{(N_0 - m_0)}(N - m - N_0 + m_0; \tilde{y}). \tag{44}$$

3.8.2. Case $m_0 = 0$, $N_0 > 0$

The approximation reads

$$P_N(m) \sim N^{-N_0+1} \Gamma(N_0) P_{L-C}^{(0)}(m; y) \left(1 + \frac{(N_0 - 1)m}{N}\right) P_{L-C}^{(N_0)}(N - m - N_0; \tilde{y}). \tag{45}$$

3.8.3. Case $m_0 = N_0 > 0$

The approximation reads

$$P_N(m) \sim N^{-N_0+1} \Gamma(N_0) P_{L-C}^{(N_0)}(m - N_0; y) \left(N_0 - \frac{(N_0 - 1)m}{N}\right) P_{L-C}^{(0)}(N - m; \tilde{y}). \tag{46}$$

3.9. Landau Distribution

As the shape parameter y of the Lea–Coulson PMF $P_{L-C}^{(0)}(m; y)$ increases, the PMF can be approximated with the probability density function (PDF) of the Landau distribution [10].

Additionally, partial summation $P^{(k)}_{L-C}(m;y)$, $k = 1,2,\ldots$ can be approximated with the partial integration of the PDF. Precisely,

$$P^{(k)}_{L-C}(m-k;y) \sim y^{k-1} F^{(k)}_{\text{Landau}}\left(\frac{m}{y} - \ln y\right), \tag{47}$$

where

$$F^{(0)}_{\text{Landau}}(\xi) = f_{\text{Landau}}(\xi) = \frac{1}{\pi}\int_0^\infty \exp(-v(\xi + \ln v))\sin(\pi v)\mathrm{d}v \tag{48}$$

is the density of Landau distribution [26,41] and

$$F^{(k)}_{\text{Landau}}(\xi) = \int_{-\infty}^{\xi} F^{(k-1)}_{\text{Landau}}(\tilde{\xi})\mathrm{d}\tilde{\xi} = \frac{1}{\Gamma(k)}\int_{-\infty}^{\xi} \tilde{\xi}^{k-1} f_{\text{Landau}}(\xi - \tilde{\xi})\mathrm{d}\tilde{\xi} \tag{49}$$

is the kth partial integration of the density. In particular, $F_{\text{Landau}}(\xi) = F^{(1)}_{\text{Landau}}(\xi)$ is the CDF of the Landau distribution.

4. Results

This paper provides asymptotic approximations to the Luria–Delbrück distribution $P_N(m)$ of observing m resistant cells as the total population reaches size N. Specifically, distribution $P_N(m)$ is defined as the solution to Master Equation (1), subject to boundary and initial Conditions (2)–(3). The input parameters of the distribution were the final population size N and the probabilities μ and λ of resistance acquisition and loss, respectively. In the examples below, we adopt the values $N = 2 \times 10^5$ and $\mu = 10^{-3}$ from the study of the unidirectional model [10]; we compare the unidirectional ($\lambda = 0$) and bidirectional ($\lambda = 10^{-3}$) cases.

The distribution $P_N(m)$ also depends on the initial population size N_0 and the initial number of resistant cells m_0. In the following subsections, we report the distribution asymptotics for initial conditions of increasing complexity. Section 4.1 deals with a population that starts from a single sensitive cell ($N_0 = 1$, $m_0 = 0$). Section 4.2 focuses on a population that begins with two cells of different kinds ($N_0 = 2$, $m_0 = 1$). The case of a general (deterministic) initial condition is outlined in Section 4.3. Section 4.4 briefly considers the case of a nondeterministic initial condition.

4.1. One Sensitive Cell at the Beginning

This type of initial condition has been widely considered in studies concerning the Luria–Delbrück fluctuation test. Approximations of the resistant population distribution by the Lea–Coulson probability mass function and the Landau probability density function were reported. This paper establishes the appearance of partial summations/integrations of these distributions, in particular the cumulative distribution function, in the tail of the population distribution.

For low-to-moderate values m, the probability of seeing m resistant cells among N total cells is approximated with the following (see Figure 3a):

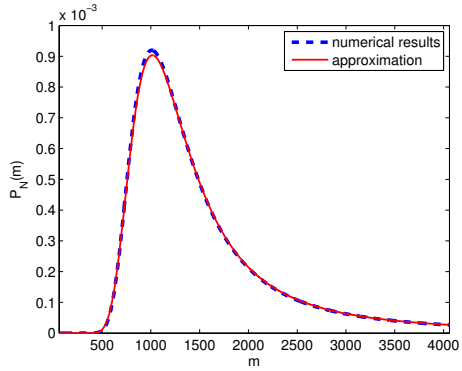
$$P_N(m) \sim \frac{1}{\mu N} f_{\text{Landau}}\left(\frac{m}{\mu N} - \ln \mu N\right), \tag{50}$$

where $f_{\text{Landau}}(\xi)$ is Density (48) of Landau distribution. The approximation follows from (36) and (47). This approximation was previously derived for the unidirectional model in [10].

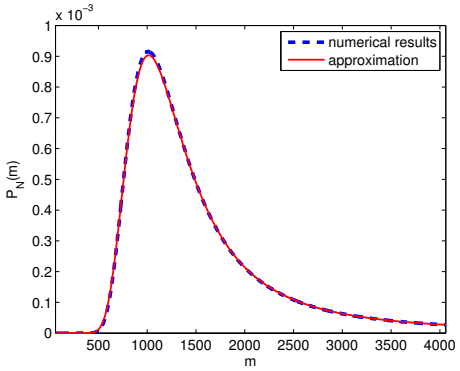
For a large number of resistant cells m in the population, we have the following approximation (see Figure 3b):

$$P_N(m) \sim \frac{\mu}{N} F_{\text{Landau}}\left(\frac{N-m}{\lambda N} - \ln \lambda N\right), \tag{51}$$

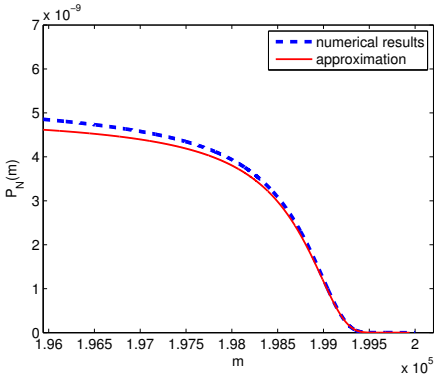
where $F_{\text{Landau}}(x) = \int_{-\infty}^{x} f_{\text{Landau}}(y)\mathrm{d}y$ is the cumulative distribution function of Landau distribution. The approximation follows from (41) for $N_0 = 1$ and (47).



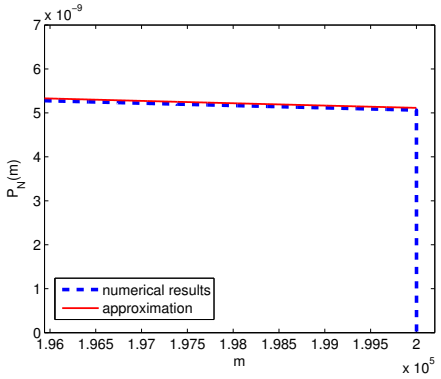
(**a**) Bidirectional case; small m. Left-boundary Landau PDF approximation (50) was used.



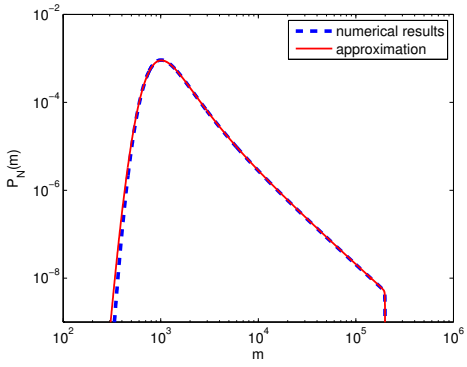
(**d**) Unidirectional case; small m. Landau PDF approximation (50) was used.



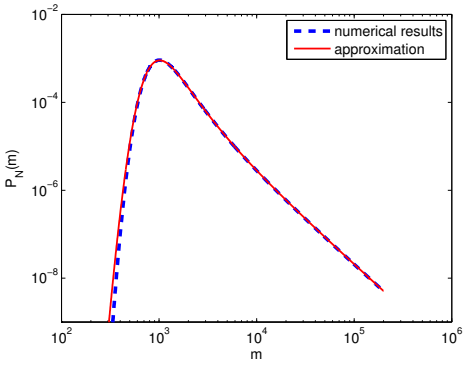
(**b**) Bidirectional case; large m. Right-boundary Landau CDF approximation (51) was used.



(**e**) Unidirectional case; large m. Landau PDF approximation (50) was used.



(**c**) Bidirectional case; all m. Composite Landau PDF/CDF approximation (52) was used.



(**f**) Unidirectional case; all m. Landau PDF approximation (50) was used.

Figure 3. The probability $P_N(m)$ of having m resistant cells in a total population of N cells as calculated numerically (blue dashed lines) and its asymptotic approximations (50), (51), (52) (red solid lines). The population started from a single sensitive cell and was left to grow until it had consisted of $N = 2 \times 10^5$ cells. (**a**–**c**) Bidirectional case: the probabilities of resistance acquisition and loss were $\mu = \lambda = 10^{-3}$. (**d**–**f**) Unidirectional case: $\mu = 10^{-3}$, $\lambda = 0$.

Lastly, the following composite approximation (Figure 3c):

$$P_N(m) \sim \frac{1}{\mu N} f_{\text{Landau}}\left(\frac{m}{\mu N} - \ln \mu N\right) F_{\text{Landau}}\left(\frac{N-m}{\lambda N} - \ln \lambda N\right) \tag{52}$$

works well for all possible numbers m of resistant cells. The approximation follows from (45) for $N_0 = 1$ and (47).

In the irreversible and unidirectional case ($\lambda = 0$), Landau PDF (50) applies uniformly for the entire range of $0 \leq m < N$ of resistant subpopulation sizes (Figure 3d–f). The probability distribution then exhibits a sharp cut-off at $m = N$ (Figure 3e); contrastingly, the presence of a reverse transition smooths the cut-off into a Landau CDF in the right boundary layer (Figure 3b).

4.2. One Sensitive and One Resistant Cell at the Beginning

A mixed model develops a dramatically different dynamics than the homogeneous initial population. In an initial phase, which lasts until the population size reaches the reciprocal of the small resistance acquisition parameter, the dynamics are driven by the competition between the independent progeny of the two parent cells. A classical result implies that this tug of war leads to the uniform distribution of resistant/sensitive cells [38]. In this paper, this classical result was incorporated into Regular Coarse-Grained Approximation (20) with $m_0 = 1$, $N_0 = 2$. After the initial phase, boundary layers appeared at the left and right corners of $m \in \{0, 1, \ldots, N\}$ due to resistance acquisition and loss. This paper provides nontrivial approximations of the boundary behaviour through the Landau cumulative distribution function.

For a small-to-moderate number of resistant cells m in the population, we have the following approximation (Figure 4a):

$$P_N(m) \sim \frac{1}{N} F_{\text{Landau}}\left(\frac{m}{\mu N} - \ln \mu N\right). \tag{53}$$

The approximation was derived from (32) for $m_0 = 1$, $N_0 = 2$ and (47).

For a large number of resistant cells m in the population, we have the following approximation (Figure 4b):

$$P_N(m) \sim \frac{1}{N} F_{\text{Landau}}\left(\frac{N-m}{\lambda N} - \ln \lambda N\right). \tag{54}$$

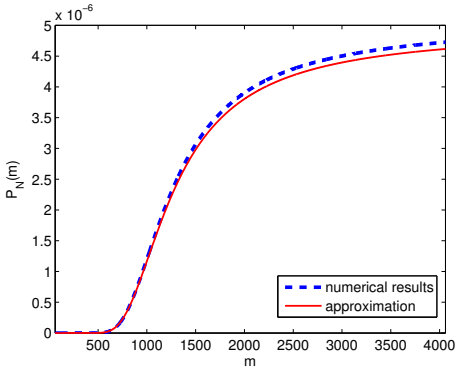
The approximation was derived from (40) for $m_0 = 1$, $N_0 = 2$ and (47).

Lastly, just like in the case of a sensitive initial population, in the case of a mixed initial population, one can find a composite approximation that works very well for all possible m. In this case, it is given by (Figure 4c):

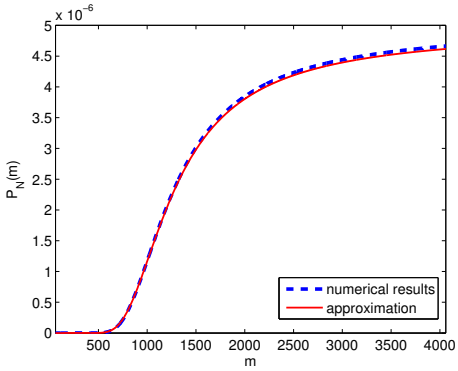
$$P_N(m) \sim \frac{1}{N} F_{\text{Landau}}\left(\frac{m}{\mu N} - \ln \mu N\right) F_{\text{Landau}}\left(\frac{N-m}{\lambda N} - \ln \lambda N\right). \tag{55}$$

The approximation was derived from (44) for $m_0 = 1$, $N_0 = 2$ and (47).

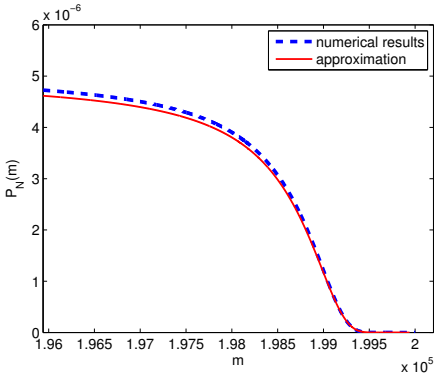
In the unidirectional case ($\lambda = 0$), the Landau CDF approximation (50) applied uniformly to all resistant subpopulation sizes $0 \leq m < N$ (Figure 4d–f). Again, a sharp cut-off was observed at $m = N$ in the unidirectional case (Figure 4e), whereas a Landau CDF in the right boundary layer appeared in the reversible case (Figure 4b).



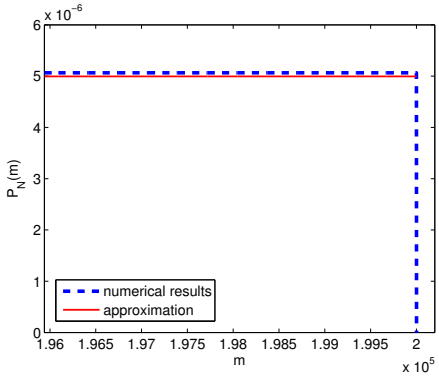
(**a**) Bidirectional case; small m. Left-boundary Landau CDF approximation (53) was used.



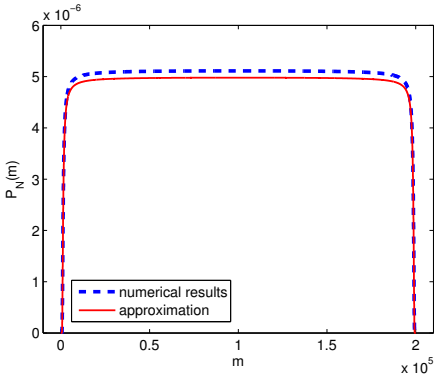
(**d**) Unidirectional case; small m. Landau CDF approximation (53) was used.



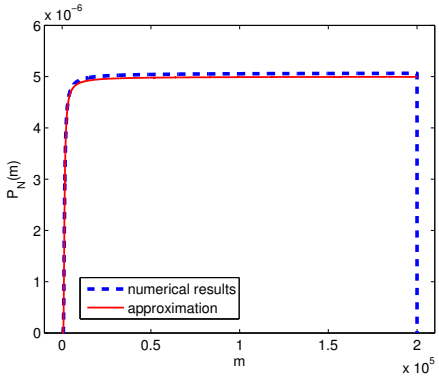
(**b**) Bidirectional case; large m. Right-boundary Landau CDF approximation (54) was used.



(**e**) Unidirectional case; large m. Landau CDF approximation (53) was used.



(**c**) Bidirectional case; all m. Composite Landau CDF approximation (55) was used.



(**f**) Unidirectional case; all m. Landau CDF approximation (53) was used.

Figure 4. The probability $P_N(m)$ of having m resistant cells in a total population of N cells as calculated numerically (blue dashed lines) and its asymptotic approximations (53), (54) and (55) (red solid lines). The population started from two cells, one sensitive and one resistant, and was left to grow until it had consisted of $N = 2 \times 10^5$ cells. (**a**–**c**) Bidirectional case: the probabilities of resistance acquisition and loss are $\mu = \lambda = 10^{-3}$. (**d**–**f**) Unidirectional case: $\mu = 10^{-3}$, $\lambda = 0$.

4.3. More than One Sensitive or Resistant Cells at the Beginning

The cumulative distribution function is a partial integration (up to a given level) of the probability density function. The second, third, etc. partial integrations are defined recursively. The higher-order integrations of Landau density appear in the boundary behaviour of the Luria–Delbrück distribution if the initial condition contains more than one cell of either type (resistant/sensitive).

For the sake of brevity, we only give the composite approximations. If we initially have $m_0 = 0$ resistant and $N_0 > 0$ sensitive cells, then (52) generalises to

$$P_N(m) \sim \frac{\Gamma(N_0)\lambda^{N_0-1}}{\mu N} f_{\text{Landau}}\left(\frac{m}{\mu N} - \ln \mu N\right)\left(1 + \frac{(N_0-1)m}{N}\right) F^{(N_0)}_{\text{Landau}}\left(\frac{N-m}{\lambda N} - \ln \lambda N\right), \tag{56}$$

If we initially have $m_0 > 0$ resistant and $N_0 - m_0 > 0$ sensitive cells, on the other hand, then (55) generalises to

$$P_N(m) \sim \frac{\Gamma(N_0)\mu^{m_0-1}\lambda^{N_0-m_0-1}}{N} F^{(m_0)}_{\text{Landau}}\left(\frac{m}{\mu N} - \ln \mu N\right) F^{(N_0-m_0)}_{\text{Landau}}\left(\frac{N-m}{\lambda N} - \ln \lambda N\right). \tag{57}$$

Equations (56)–(57) were derived by combining (44)–(45) with (47); Γ stands for the gamma function, and $F^{(k)}_{\text{Landau}}$ gives the k-th partial integration (49) of the Landau probability density function.

4.4. Stochastic Initial Conditions

Our analysis so far concerned deterministic initial Condition (3). A previous study of a bidirectional model (based on different assumptions) [8] considered a stochastic initial condition in which the population was derived from a single initial cell, which could be with a certain probability ρ sensitive; with the complementary probability, it is resistant. Via the superposition principle, the solution that is subject to the stochastic initial condition is equal to the weighted sum of the solution starting deterministically with one sensitive parent and the solution starting deterministically with one resistant parent. The Landau cumulative distribution functions in the tails of the two solutions are then negligible. For low-to-moderate values m, the probability of seeing m resistant cells among N total cells is approximated as follows:

$$P_N(m) \sim \frac{\rho}{\mu N} f_{\text{Landau}}\left(\frac{m}{\mu N} - \ln \mu N\right). \tag{58}$$

For a large number of resistant cells m in the population, we have the following approximation:

$$P_N(m) \sim \frac{1-\rho}{\lambda N} f_{\text{Landau}}\left(\frac{N-m}{\lambda N} - \ln \lambda N\right). \tag{59}$$

A uniformly valid solution is obtained by adding up (58) and (59):

$$P_N(m) \sim \frac{\rho}{\mu N} f_{\text{Landau}}\left(\frac{m}{\mu N} - \ln \mu N\right) + \frac{1-\rho}{\lambda N} f_{\text{Landau}}\left(\frac{N-m}{\lambda N} - \ln \lambda N\right). \tag{60}$$

This is a composite rather than log-composite approximation.

A special type of initial condition is if the switching between resistance and sensitivity is in balance: $\rho = \lambda/(\mu+\lambda)$, $1-\rho = \mu/(\mu+\lambda)$. Then, the mean resistant fraction satisfies $\sum_{m=0}^{N} mP_N(m)/N = \mu/(\mu+\lambda)$, i.e., it does not change as the population grows [8].

4.5. Limitations of the Approximations

The left boundary solution has the form of the Landau PDF/CDF located at $m = \mu N \ln \mu N$; see Equations (50), (53) and (58). For self-consistency, we required that the boundary layer location be much lower than the size of the entire definition domain, meaning that $\mu N \ln \mu N \ll N$. This is equivalent to requiring $\ln N \ll 1/\mu$, where $\mu \ll 1$. Thus, N should be large but not exponentially large for the asymptotic approximations to apply.

5. Discussion

The paper focused on a stochastic Luria–Delbrück fluctuation test model with reversible switching between two cellular states (for example, single cells being drug-resistant

or drug-sensitive). We considered a fixed-population ensemble formulation in which the population was expanded until it had reached a given size. The main contribution of the paper is the characterisation of the Luria–Delbrück distribution of resistant cells in the final population. We focused on the parametric regime of low switching probabilities and large populations; we used singular perturbation methods to obtain asymptotic approximations to the desired distribution. The novelty of the work lies in the systematic use of asymptotic matching between the alternate approximations. The approach is applicable to other known approximations, e.g., Fréchet distribution [42].

While our analysis focused on a fixed-population ensemble, the results also have implications in the case of a fixed-time ensemble. In the latter formulation, the population is grown until a fixed end time point, as explained in Section 2. The total population grows exponentially as per the Yule growth process with a reciprocal growth rate to the mean generation time. If one conditions on the Yule process reaching a given value N, one recovers the fixed-ensemble model for the number of resistant cells as studied in this paper.Conversely, drawing N from the state of a Yule growth process at a given time t, the fixed-population ensemble is turned into a fixed-time ensemble formulation. Conditioning a fixed population thus eliminates proliferation noise. Another approach to this consists of calculating the relative fractions of sensitive/resistant cells [8].

Our model assumes that switching between sensitivity and resistance occurs at cell division events. A previous study [8] of a reversible switching case was based on different assumptions that warrant a brief discussion. Like in the fixed time ensemble version of our model, the cell cycle was assumed to be exponentially distributed, and the mean generation time was the same for the two cell types. However, switching was uncoupled from the cell cycle in [8]: a daughter cell is always of the same type as the mother cell, and any cell has a constant propensity to transition at any time to the opposite state.

Importantly, unlike in our model, there could be more than one switching event per cell cycle in the model of [8]. Therefore, although this needs to be supported by further analysis, the difference between the two formulations is expected to be pronounced in the fast switching regime in which multiple switching events per cell cycle are probable. In the model of [8], proliferation events increased the population by one, but the switching events did not; this renders the relationship between the fixed-time and fixed-population ensemble formulations subtler. Further complications in the dynamics arise if the switching rates depend on the presence of drugs [43,44].

The reversible model studied in this paper is a direct extension of the irreversible switching model studied in [10]; indeed, setting probability λ that a resistant mother bears a sensitive daughter cell to zero, one recovers the earlier model. The earlier study provides methods that are specifically applicable to the important regime of a large population subject to low-rate switching. We extended these methods to the current reversible case. The generalisation required nontrivial modifications in the methodology, notably in the systematic use of asymptotic matching between the small and large population solutions.

The new approximation differs from that given in [10] only regarding the low-probability distribution tail, as evidenced in Figure 3. Nevertheless, the appearance of the Landau cumulative distribution function in the tail is widely interesting. Additionally, we established that it also appeared for mixed initial conditions, as illustrated in Figure 4.

6. Conclusions

In this paper, we theoretically showed that reversible switching in a fluctuation test experiment affects the tail of Luria–Delbrück distribution. Methodologically, our results were based on the matched asymptotic framework that allows for alternate approximations to the distribution under different parameter scaling regimes, combining them into a uniform composite approximation. We expect that the approach could be extended, particularly to more complicated (and realistic) descriptions of the fluctuation test experiment, and provide valuable insights into the underlying dynamics.

Author Contributions: Conceptualization, A.S.; methodology, P.B.; analysis, P.B. and A.H.; visualization, A.H.; writing—review and editing, P.B., A.H. and A.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Slovak Research and Development Agency under contract no. APVV-18-0308, and VEGA grants 1/0339/21 and 1/0755/22. A.S. acknowledges support from ARO W911NF-19-1-0243, and NIH grants R01GM124446 and R01GM126557.

Institutional Review Board Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors are grateful to the three anonymous referees whose insightful review reports helped in improving the manuscript.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

PMF	Probability mass function
PDF	Probability density function
CDF	Cumulative distribution function
L–C	Lea–Coulson

Appendix A. Polya urn Models

Suppose that we have an urn that first consists of w white balls and b black balls. Then, we withdraw a ball from the urn one at a time, look at its colour, return the ball to the urn, and add c balls of the same colour. The probability that x white balls are drawn in a sample of n withdrawals is

$$P[X = x] = \frac{\binom{\frac{w}{c} + x - 1}{x} \binom{\frac{b}{c} + n - x - 1}{n - x}}{\binom{\frac{w+b}{c} + n - 1}{n}}.$$

The number of balls of a specific colour in the Polya urn model follows negative hypergeometric distribution [37].

We can adapt the Polya urn model to our situation in the following way. First, $c = 1$ in our case, as in every birth event, and only one new cell is born. Furthermore, w and b are the number of resistant and sensitive cells in the initial population, respectively; n denotes the number of birth events, and x denotes the number of resistant cells among these n births. The following relations with the notation used in the main text hold:

$$\begin{aligned} m_0 &= w, \\ N_0 &= b + w, \\ m &= w + x, \\ N &= w + b + n. \end{aligned}$$

Therefore, we can write that, as μ (and λ) tends to zero,

$$P_N(m) \sim \frac{\binom{m-1}{m-m_0} \binom{N-m-1}{N-m-(N_0-m_0)}}{\binom{N-1}{N-N_0}}, \qquad \text{where } m_0 \le m \le N - N_0, \tag{A1}$$

where the tilde sign is understood in the usual sense of asymptotic expansions. The rearrangement of combinatorial numbers yields (4).

Appendix B. Computations

The symbolic computation of Sum (9) was performed using MATLAB's Symbolic Math Toolbox [45]:

```
syms L N m N_0
A = factorial(N-L-1);
B = factorial(N-L-m);
s = symsum(L*A/B, L, N_0, N-m);
assume(m>0);
simplify(s)
```

References

1. Luria, S.E.; Delbrück, M. Mutations of bacteria from virus sensitivity to virus resistance. *Genetics* **1943**, *28*, 491. [CrossRef] [PubMed]
2. Shaffer, S.M.; Dunagin, M.C.; Torborg, S.R.; Torre, E.A.; Emert, B.; Krepler, C.; Beqiri, M.; Sproesser, K.; Brafford, P.A.; Xiao, M.; et al. Rare cell variability and drug-induced reprogramming as a mode of cancer drug resistance. *Nature* **2017**, *546*, 431–435. [CrossRef] [PubMed]
3. Shaffer, S.M.; Emert, B.L.; Hueros, R.A.R.; Cote, C.; Harmange, G.; Schaff, D.L.; Sizemore, A.E.; Gupte, R.; Torre, E.; Singh, A.; et al. Memory sequencing reveals heritable single-cell gene expression programs associated with distinct cellular behaviors. *Cell* **2020**, *182*, 947–959. [CrossRef] [PubMed]
4. Hossain, T.; Singh, A.; Butzin, N.C. Escherichia coli cells are primed for survival before lethal antibiotic stress. *bioRxiv* **2022**. [CrossRef]
5. Chang, C.A.; Jen, J.; Jiang, S.; Sayad, A.; Mer, A.S.; Brown, K.R.; Nixon, A.M.; Dhabaria, A.; Tang, K.H.; Venet, D.; et al. Ontogeny and vulnerabilities of drug-tolerant persisters in her2+ breast cancer. *Cancer Discov.* **2022**, *12*, 1022–1045. [CrossRef]
6. Harmange, G.; Hueros, R.A.R.; Schaff, D.L.; Emert, B.L.; Saint-Antoine, M.M.; Nellore, S.; Fane, M.E.; Alicea, G.M.; Weeraratna, A.T.; Singh, A.; et al. Disrupting cellular memory to overcome drug resistance. *bioRxiv* **2022**. [CrossRef]
7. Saint-Antoine, M.M.; Grima, R.; Singh, A. A fluctuation-based approach to infer kinetics and topology of cell-state switching. *bioRxiv* **2022**. [CrossRef]
8. Saint-Antoine, M.M.; Singh, A. Moment-based estimation of state-switching rates in cell populations. *bioRxiv* **2022**. [CrossRef]
9. Zheng, Q. Progress of a half century in the study of the Luria–Delbrück distribution. *Math. Biosci.* **1999**, *162*, 1–32. [CrossRef]
10. Kessler, D.A.; Levine, H. Large population solution of the stochastic Luria–Delbrück evolution model. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 11682–11687. [CrossRef]
11. Keller, P.; Antal, T. Mutant number distribution in an exponentially growing population. *J. Stat. Mech. Theory Exp.* **2015**, *2015*, P01011. [CrossRef]
12. Nicholson, M.D.; Antal, T. Universal asymptotic clone size distribution for general population growth. *Bull. Math. Biol.* **2016**, *78*, 2243–2276. [CrossRef]
13. Pakes, A.G. Mutant number laws and infinite divisibility. *Axioms* **2022**, *11*, 584. [CrossRef]
14. Antal, T.; Krapivsky, P. Exact solution of a two-type branching process: Models of tumor progression. *J. Stat. Mech. Theory Exp.* **2011**, *2011*, P08018. [CrossRef]
15. Angerer, W.P. An explicit representation of the Luria–Delbrück distribution. *J. Math. Biol.* **2001**, *42*, 145–174. [CrossRef] [PubMed]
16. Kessler, D.A.; Levine, H. Scaling solution in the large population limit of the general asymmetric stochastic Luria–Delbrück evolution process. *J. Stat. Phys.* **2015**, *158*, 783–805. [CrossRef] [PubMed]
17. Armitage, P. The statistical theory of bacterial populations subject to mutation. *J. R. Stat. Soc. Ser. B (Methodol.)* **1952**, *14*, 1–33. [CrossRef]
18. Jolly, C.; Cook, A.; Raferty, J.; Jones, M. Measuring bidirectional mutation. *J. Theor. Biol.* **2007**, *246*, 269–277. [CrossRef]
19. Sorace, R.; Komarova, N.L. Accumulation of neutral mutations in growing cell colonies with competition. *J. Theor. Biol.* **2012**, *314*, 84–94. [CrossRef]
20. Cheek, D.; Antal, T. Genetic composition of an exponentially growing cell population. *Stoch. Process. Their Appl.* **2020**, *130*, 6580–6624. [CrossRef]
21. Lea, D.E.; Coulson, C.A. The distribution of the numbers of mutants in bacterial populations. *J. Genet.* **1949**, *49*, 264–285. [CrossRef] [PubMed]
22. Zheng, Q. A fresh approach to a special type of the Luria–Delbrück distribution. *Axioms* **2022**, *11*, 730. [CrossRef]
23. Pakes, A.G. Remarks on the Luria–Delbrück distribution. *J. Appl. Probab.* **1993**, *30*, 991–994. [CrossRef]
24. Angerer, W.P. A note on the evaluation of fluctuation experiments. *Mutat. Res. Mol. Mech. Mutagen.* **2001**, *479*, 207–224. [CrossRef] [PubMed]

25. Bulyak, E.; Shul'ga, N. Landau distribution of ionization losses: History, importance, extensions. *arXiv* **2022**, arXiv:2209.06387.
26. Nolan, J.P. *Univariate Stable Distributions*; Springer: Berlin/Heidelberg, Germany, 2020.
27. Rao, C.R.; Shanbhag, D.N.; Sapatinas, T.; Rao, M.B. Some properties of extreme stable laws and related infinitely divisible random variables. *J. Stat. Plan. Inference* **2009**, *139*, 802–813.
28. Bender, C.M.; Orszag, S.A. *Advanced Mathematical Methods for Scientists and Engineers I: Asymptotic Methods and Perturbation Theory*, 1st ed.; Springer Science & Business Media: Berlin/Heidelberg, Germany, 1999.
29. Hinch, E.J. *Perturbation Methods*; Cambridge University Press: Cambridge, UK, 1991.
30. Hinch, R.; Chapman, S.J. Exponentially slow transitions on a Markov chain: The frequency of calcium sparks. *Eur. J. Appl. Math.* **2005**, *16*, 427–446. [CrossRef]
31. Bressloff, P.C. *Stochastic Processes in Cell Biology*; Springer: Berlin/Heidelberg, Germany, 2014.
32. Bokes, P. Heavy-tailed distributions in a stochastic gene autoregulation model. *J. Stat. Mech. Theory Exp.* **2021**, *2021*, 113403. [CrossRef]
33. Athreya, K.B.; Ney, P.E.; Ney, P. *Branching Processes*; Courier Corporation: Chelmsford, CA, USA, 2004.
34. Kepler, T.B.; Oprea, M. Improved inference of mutation rates: I. An integral representation for the Luria–Delbrück distribution. *Theor. Popul. Biol.* **2001**, *59*, 41–48. [CrossRef]
35. Kevorkian, J.; Cole, J.D. *Perturbation Methods in Applied Mathematics*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2013.
36. Murray, J.D. *Mathematical Biology: I. An Introduction*; Springer: Berlin/Heidelberg, Germany, 2002.
37. Johnson, N.L.; Kotz, S.; Kemp, A.W. *Univariate Discrete Distributions*; John Wiley & Sons: Hoboken, NJ, USA, 2005.
38. Mahmoud, H. *Pólya urn Models*; Chapman and Hall/CRC: Boca Raton, FL, USA, 2008.
39. Walczak, A.M.; Mugler, A.; Wiggins, C.H. Analytic methods for modeling stochastic regulatory networks. *Comput. Model. Signal. Netw.* **2012**, *880*, 273–322.
40. Bokes, P.; King, J.R.; Loose, M. A bistable genetic switch which does not require high co-operativity at the promoter: A two-timescale model for the PU. 1–GATA-1 interaction. *Math. Med. Biol. A J. IMA* **2009**, *26*, 117–132. [CrossRef] [PubMed]
41. Landau, L.D. On the energy loss of fast particles by ionization. *J. Phys.* **1944**, *8*, 201–205.
42. Frank, S.A. Numbers of mutations within multicellular bodies: Why it matters. *Axioms* **2022**, *12*, 12. [CrossRef]
43. Su, Y.; Wei, W.; Robert, L.; Xue, M.; Tsoi, J.; Garcia-Diaz, A.; Homet Moreno, B.; Kim, J.; Ng, R.H.; Lee, J.W.; et al. Single-cell analysis resolves the cell state transition and signaling dynamics associated with melanoma drug-induced resistance. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 13679–13684. [CrossRef] [PubMed]
44. Angelini, E.; Wang, Y.; Zhou, J.X.; Qian, H.; Huang, S. A model for the intrinsic limit of cancer therapy: Duality of treatment-induced cell death and treatment-induced stemness. *PLoS Comput. Biol.* **2022**, *18*, e1010319. [CrossRef] [PubMed]
45. Symbolic Math Toolbox. Available online: https://www.mathworks.com/products/symbolic.html (accessed on 16 January 2022).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.