

Prediction coverage in inherited-state branching models: an exact mean-closure counterexample, a repaired count threshold, and the effect of founder number

16 September 2026

Abstract

Low-cell-count proliferation assays are often summarized by a scalar birth–death model whose per-cell rates are read off the expected phenotypic composition of the population. We show that such a mean-composition closure can reproduce the exact expected cell count, and even the exact expected birth and death fluxes, while assigning misleading probabilities to future counts. The witness is an explicit two-type branching process with positive switching in both directions, one founder of random type, and an observation horizon of one resistant doubling time. The scalar closure assigns probability at least $332/343 > 0.9679$ to the counts $\{0, 1, 2\}$, whereas the inherited-state process assigns them probability at most $1 - L < 0.95$ with $L = 0.05044\dots$ an explicit rational number. Both inequalities are proved analytically: the scalar tail is bounded through an explicit box for the mean and the accumulated-birth parameter, certified by one positive polynomial identity, and the actual tail is bounded from below by seven no-switch histories. Widening the region to $\{0, \dots, 3\}$ restores coverage at least $H = 0.9713\dots$, and three is the smallest upper endpoint that reaches 95% for this preparation. A variance comparison ($\text{Var } Z_T \leq 11/16 < 4/5 < \text{Var } N_T$) shows that the discrepancy does not disappear with many independent founders: the limiting actual coverage of the scalar central 95% interval is below 0.9308, and numerically about 0.849. We give a general identity that expresses the variance defect of any mean-composition closure as an integrated covariance between population size and composition-weighted net growth, and closed low-count backward equations that compute corrected thresholds without solving the joint population law. Numerical checks reproduce the witness, scan the switching rates, exhibit the same mechanism with positive birth, death and day-scale switching in both states, and retain a source-linked preparation in which the closure performs well. The tail obstruction, the analytic box, the explicit scalar law and the repaired threshold are verified in Lean 4; the two process-identification bridges are proved conventionally and are identified as such. The witness is a synthetic persistent-state regime, not a fitted experiment.

1 Introduction

1.1 The question

A low-founder proliferation assay produces a random cell count at a fixed time. A model can reproduce the average of that count over many replicate wells and still give the wrong probability for unusually small or unusually large outcomes. When cells inherit a phenotypic state at division, one founder’s state influences the growth of its whole family, so a description that first averages the composition of the population and then computes fluctuations from that average may discard exactly the dependence that governs family sizes.

The specific closure we study is the following. A two-type process (S_t, R_t) of sensitive and resistant cells has constant per-cell rates and exact type means $s(t) = \mathbb{E}S_t$, $r(t) = \mathbb{E}R_t$. The *mean-composition closure* replaces it by a single-type birth–death process Z_t whose per-cell birth and death rates at time t are the population-average rates computed from $(s(t), r(t))$. By construction $\mathbb{E}Z_t = \mathbb{E}(S_t + R_t)$ for all t ; the closure also matches the expected instantaneous birth and death fluxes. Closures of this kind are attractive in low-cell-count settings because they yield a one-dimensional master equation whose solution is available in closed form [3, 19]. Browning et al. [8] develop such an approximate count equation for a stochastic phenotype model and explicitly discuss the limitation introduced by phenotypic inheritance; Iyer et al. [18] show that inheritable states shape drug-persister correlations. The question addressed here is quantitative:

Can mean-composition closure invalidate a count-prediction region even when all rates and the initial preparation are known exactly and the expected count trajectory is matched exactly?

The answer is yes, and the failure can be certified by hand. The witness is deliberately small. All of its probabilities are explicit rational or elementary expressions, the decisive inequalities are strict rational-arithmetic statements, and the argument never computes the full joint population law.

1.2 Main results

Section 2 defines the process: sensitive cells die at rate $d = 3/10$ per day, resistant cells divide at rate $b = 1/10$ per day, sensitive cells switch to resistant at rate $\alpha = 1/1000$ and resistant cells switch to sensitive at rate $\beta = 1/100\,000$ per day; one founder is resistant with probability $5/24$. The horizon is $T = \log 2/(b + \beta)$, one resistant doubling time, and $c = b/(b + \beta) = 10000/10001$.

- (i) *Coverage failure with exact mean matching* (Theorem 3.1). $\mathbb{E}Z_t = \mathbb{E}N_t$ for all t , where $N_t = S_t + R_t$, yet

$$\mathbb{P}(Z_T \leq 2) \geq \frac{332}{343} > 0.9679, \quad \mathbb{P}(N_T \leq 2) \leq 1 - L < 0.95,$$

where $L = \frac{5}{48} \sum_{j=2}^6 (\frac{c}{2})^j = 0.05044 \dots$ is an explicit rational number.

- (ii) *Instance-specific repair and minimality* (Theorem 3.1, Corollary 3.2). $\mathbb{P}(N_T \leq 3) \geq H = 0.97131 \dots$, an explicit rational number, and 3 is the smallest k for which $\{0, \dots, k\}$ has actual coverage at least 95%.
- (iii) *Persistence under many founders* (Proposition 5.1, Theorem 5.2). $\text{Var } Z_T \leq 11/16$ while $\text{Var } N_T > 4/5$. For F independent founders the actual coverage of the scalar central 95% interval converges to $2\Phi(z_{.975}\sqrt{v_Z/v_N}) - 1 < 2\Phi(z_{.975}\sqrt{55/64}) - 1 < 0.9308$.
- (iv) *What the closure discards* (Proposition 5.4). For any finite-type process with constant per-cell rates and count-preserving switching, the variance defect of the mean-composition closure satisfies

$$v_N(t) - v_Z(t) = 2m(t)^2 \int_0^t \frac{\text{Cov}(N_u, G_u - \bar{g}(u)N_u)}{m(u)^2} du,$$

where G is the population net-growth propensity and $\bar{g} = \mathbb{E}G/\mathbb{E}N$. Inheritance therefore inflates variance exactly when population size covaries with composition-weighted excess growth; when all types have the same net growth rate the defect vanishes identically.

- (v) *Computable corrections* (Proposition 6.1). The probabilities $\mathbb{P}(N_T = n)$ for $n \leq k$ satisfy a closed system of $2(k + 1)$ ordinary differential equations obtained from the backward generating-function equations. Corrected thresholds for other preparations, and for F independent or binomially plated founders, follow without solving a two-dimensional master equation. A mean-only Markov bound would certify endpoint 10 for the witness; the type-aware calculation certifies 3.

Section 7 reproduces the witness numerically, tabulates independent-founder coverage, scans both switching rates over five decades, exhibits the mechanism with strictly positive birth, death and switching rates in both states and switching times of ten days, and reports a source-linked preparation from Browning et al. [8] in which the closure covers correctly. Section 8 states what the result does and does not show.

1.3 Relation to existing work

The scalar terminal law used here is the linear-fractional law of the time-dependent birth–death process [3, 19]; nothing about it is new. Two-type Markov branching processes and their moment equations are classical [2, 15, 16], and their use for drug resistance in cancer [5, 10, 17] and for phenotypic switching in bacteria [4, 21] is well established. That inherited traits generate correlated family sizes and excess variance in counts has been understood since Luria and Delbrück [22]. Reversible drug-tolerant states in cancer cell subpopulations [25, 26] motivate the two-type structure, and Iyer et al. [18] model inherited states and persister correlations directly. On the approximation side, moment-closure methods and their accuracy have been reviewed extensively [14, 24, 27], and identifiability from low-cell-count experiments is the subject of Browning et al. [7, 8]. Calibration of prediction intervals as a criterion in its own right is discussed by Gneiting et al. [13].

Against this background the contribution of the present paper is not the existence of inherited-state correlations. It is (a) an explicit, fully certified example in which a mean-composition closure that matches the exact mean and the exact expected fluxes nevertheless fails a stated prediction-coverage requirement, with a small-event proof that avoids the joint population law; (b) the instance-specific corrected threshold and the proof that it is minimal in its family; (c) the founder-number limit, which shows that the failure is a variance-ratio error preserved by the central limit theorem rather than a small-count artefact; (d) the variance-defect identity, which locates the failure in one covariance and gives a sign condition; and (e) the low-count backward equations as a practical route to corrected thresholds in the same model class.

1.4 Scope, methods and verification

All probabilities concern repeated realizations of a process with known, fixed rate parameters. A prediction region covers a future cell count; it is neither a confidence interval for an estimated parameter nor a posterior credible interval. The population is well mixed and non-interacting, with exponential event times, no immigration, no density dependence, no observation error and no shared random environment. The witness is a synthetic persistent-state regime and is not a calibration of any experiment.

Every inequality in Theorems 3.1 and 5.2 and Propositions 5.1, 5.4 and 6.1 is proved in the text. The actual-process tail obstruction, the nonexplosive process law it is stated for, the analytic bounds on the first-moment system, the explicit scalar law with its normalization and its 332/343 coverage, and the process-level repaired threshold are additionally verified in Lean 4 with Mathlib [9, 28]; Appendix B lists the declarations and what they do and do not cover. Two bridges remain conventional: that the unrestricted process means coincide with the ODE solution, and that the time-inhomogeneous scalar jump process has the explicit

terminal law. Both are standard and both are proved in Section 4; we describe the development as a conventional proof with machine-checked components. The script `check_paper.py` distributed with the sources replays every constant printed below in exact rational arithmetic and regenerates the figures and floating-point diagnostics.

2 Model, closure and statistical target

2.1 The inherited-state process

Let $X_t = (S_t, R_t) \in \mathbb{N}_0^2$ be a continuous-time Markov chain with the transitions of Table 1. A division adds one cell of the parent's current type, equivalently replaces the parent by two cells of that type; a switch changes the type of one cell and preserves the total count $N_t = S_t + R_t$. Cells interact only through genealogy: given the current state, every cell acts independently. There is exactly one founder, whose type is drawn once,

$$\mathbb{P}(X_0 = (1, 0)) = \frac{19}{24}, \quad \mathbb{P}(X_0 = (0, 1)) = \frac{5}{24}. \quad (1)$$

The founder's type is not redrawn at later events; the descendants inherit the current type of their parent and switch independently thereafter. This is the mechanism under study.

Table 1: Transitions of the witness. All per-cell rates are in day^{-1} ; sensitive division and resistant death have rate zero in the exact witness (Section 7 relaxes this numerically).

Event	Change in (S, R)	Total intensity	Parameter
Sensitive death	$(-1, 0)$	dS	$d = 3/10$
Resistant division	$(0, +1)$	bR	$b = 1/10$
Sensitive $\rightarrow$ resistant	$(-1, +1)$	αS	$\alpha = 1/1000$
Resistant $\rightarrow$ sensitive	$(+1, -1)$	βR	$\beta = 1/100\,000$

The observation horizon and a recurring constant are

$$T = \frac{\log 2}{b + \beta} = 6.9307787 \dots \text{ days}, \quad c = \frac{b}{b + \beta} = \frac{10000}{10001}. \quad (2)$$

Thus T is one doubling time of an isolated resistant lineage, and c is the probability that the next event of a resistant cell is a division rather than a switch.

2.2 Mean-composition closure

Write $s(t) = \mathbb{E}S_t$, $r(t) = \mathbb{E}R_t$ and $m(t) = s(t) + r(t)$. Lemma 4.1 shows that these are finite, that $m(t) > 0$, and that they solve the linear system (10). The scalar closure is the time-inhomogeneous birth–death process Z_t on $\mathbb{N}_0$ with $Z_0 = 1$, birth intensity $\lambda(t)Z_t$ and death intensity $\delta(t)Z_t$, where

$$\lambda(t) = \frac{b r(t)}{m(t)}, \quad \delta(t) = \frac{d s(t)}{m(t)}. \quad (3)$$

These are the per-cell birth and death rates of a population whose composition is replaced by its expected composition. They are deterministic functions of time computed from the exact means, not from a fitted or estimated trajectory, and they differ in general from rates conditional on the realized count. Once mean matching is established (Lemma 4.1),

$$\lambda(t) \mathbb{E}Z_t = b \mathbb{E}R_t, \quad \delta(t) \mathbb{E}Z_t = d \mathbb{E}S_t,$$

so the closure also reproduces the expected birth and death fluxes at every time. Whatever it gets wrong is therefore not the average flux; it is the distribution and dependence structure that averaging removes.

2.3 Prediction regions

Two constructions are relevant. A one-sided integer region $\{0, \dots, k\}$ is acceptable at level 95% when its probability under the governing law is at least 0.95; Theorem 3.1 treats the regions with $k = 2$ and $k = 3$ analytically. A *central* scalar quantile interval has endpoints at the 2.5% and 97.5% integer quantiles of the scalar law; it is the construction a user of the closure would report. For the witness the scalar central 95% interval is $[0, 2]$ numerically (Section 7). The analytic bound $\mathbb{P}(Z_T \geq 3) \leq 11/343$ alone does not certify that identification, because $11/343 > 0.025$; we keep the analytic statement about $\{0, 1, 2\}$ separate from the numerical quantile calculation.

2.4 Physical scale of the witness

The sensitive death time is $1/d \approx 3.3$ days and the resistant doubling time is $T \approx 6.93$ days. The isolated switching clocks have means $1/\alpha = 1000$ days and $1/\beta = 100\,000$ days, so $\alpha T \approx 0.0069$ and $\beta T \approx 0.0000693$: over the horizon a switch is rare, and the founder's state is effectively inherited by its whole family. This is a persistent-state stress test, and it must not be read as a measured reversible switching rate. Rescaling every rate and the horizon by a common factor changes units but not the ratios that matter. Sections 7.3–7.4 examine numerically what happens when the switching clocks are shortened to days and when every event class has a positive rate.

3 Main theorem

Theorem 3.1 (Coverage failure and its repair). *For the processes of Section 2, $\mathbb{E}Z_t = \mathbb{E}N_t$ for all $t \geq 0$. At the horizon T of (2),*

$$\mathbb{P}(Z_T \leq 2) \geq \frac{332}{343} = 0.96793\dots, \quad (4)$$

$$\mathbb{P}(N_T \leq 2) \leq 1 - L < 0.95, \quad (5)$$

$$\mathbb{P}(N_T \leq 3) \geq H > 0.9713, \quad (6)$$

where

$$L = \frac{5}{48} \sum_{j=2}^6 \left(\frac{c}{2}\right)^j = \frac{151415045117539070312500}{3001800450060004500180003} = 0.050441409293\dots, \quad (7)$$

$$H = \frac{19}{24} \cdot \frac{300}{301} + \frac{5}{48} \left(1 + \frac{c}{2} + \left(\frac{c}{2}\right)^2\right) = \frac{467880212635}{481696324816} = 0.971317796152\dots \quad (8)$$

Corollary 3.2 (Minimal upper endpoint). *Among the regions $\{0, \dots, k\}$, $k \in \mathbb{N}_0$, the smallest k whose coverage under the inherited-state process is at least 95% is $k = 3$. The scalar closure, by contrast, already attributes more than 96.79% to $k = 2$.*

Proof. Coverage is nondecreasing in k ; (5) rules out every $k \leq 2$ and (6) admits $k = 3$. $\square$

Remark 3.3 (What the certified numbers mean). The certified gap between the two models on $\{0, 1, 2\}$ is at least $L - 11/343 = 0.01837\dots$, about 1.84 percentage points. The certified amount by which the actual coverage misses 95% is smaller, at least $L - 1/20 = 0.00044\dots$. Numerically the actual coverage of $\{0, 1, 2\}$ is 0.94760, so the true shortfall is about 0.24 percentage points and the true gap about 2.9 points. Remark 4.5 sharpens the certified shortfall to 0.00207 with a geometric series. The corollary concerns the family $\{0, \dots, k\}$ only; it does not assert that $\{0, \dots, 3\}$ is the shortest prediction set, an equal-tailed interval, or a universal add-one rule.

4 Proof of Theorem 3.1

4.1 Nonexplosion and exact first moments

Lemma 4.1 (Yule domination and mean equations). *The process X is nonexplosive. On a common probability space there is a Yule process Y with per-cell rate b and $Y_0 = 1$ such that $\sup_{u \leq t} N_u \leq Y_t$ almost surely; in particular*

$$\mathbb{E} \sup_{u \leq t} N_u \leq \mathbb{E} Y_t = e^{bt}, \quad \mathbb{E} \sup_{u \leq t} N_u^2 \leq \mathbb{E} Y_t^2 = 2e^{2bt} - e^{bt}. \quad (9)$$

The means s, r are finite, continuously differentiable, and are the unique solution of

$$\begin{pmatrix} s' \\ r' \end{pmatrix} = \begin{pmatrix} -(d + \alpha) & \beta \\ \alpha & b - \beta \end{pmatrix} \begin{pmatrix} s \\ r \end{pmatrix}, \quad \begin{pmatrix} s(0) \\ r(0) \end{pmatrix} = \begin{pmatrix} 19/24 \\ 5/24 \end{pmatrix}, \quad (10)$$

that is, with the witness rates, $s' = -\frac{301}{1000}s + \frac{1}{100000}r$ and $r' = \frac{1}{1000}s + \frac{9999}{100000}r$. Moreover $m(t) > 0$ for every $t \geq 0$. The scalar process Z is well defined and nonexplosive, is dominated by the same kind of Yule process, and $\mathbb{E} Z_t = m(t)$ for all t .

Proof. Embedding. Construct the population inside a Yule tree of per-cell rate b started from one cell, as follows. Every biological cell alive at time t is attached to a distinct Yule cell. Each Yule cell carries an exponential(b) division clock; when the clock of a Yule cell attached to a resistant biological cell rings, both the Yule cell and the biological cell divide, and the two new biological cells are attached to the two new Yule cells. When the clock of a Yule cell attached to a sensitive biological cell rings, or of a Yule cell with no attachment, only the Yule cell divides; its children are unattached. Each biological cell additionally carries independent exponential clocks for death (rate d if sensitive, 0 if resistant) and for switching (rate α if sensitive, β if resistant); death removes the biological cell and leaves its Yule cell unattached, and switching relabels the cell. By the memoryless property this construction realizes the chain of Table 1, the attached cells are always a subset of the Yule cells, and $N_u \leq Y_u \leq Y_t$ for $u \leq t$. The Yule process is nonexplosive [23, Chapter 2], and on any bounded interval the finitely many Yule cells bound the total intensity of all death and switching clocks, so the biological process has finitely many events on bounded intervals as well. The moments in (9) are the classical Yule moments, obtained by applying the pure-birth generator to n and n^2 or from the geometric law of Y_t [19, 30].

Stopped generator identity. Let $\tau_K = \inf\{t : N_t > K\}$. Every transition changes N by at most one, so the stopped process $X_{t \wedge \tau_K}$ takes values in the finite set $\{N \leq K+1\}$ and Dynkin's formula for it is a finite-state identity requiring no interchange of limits [11, Chapter 4]:

$$\mathbb{E} S_{t \wedge \tau_K} = \frac{19}{24} + \mathbb{E} \int_0^{t \wedge \tau_K} (-(d + \alpha)S_u + \beta R_u) du, \quad \mathbb{E} R_{t \wedge \tau_K} = \frac{5}{24} + \mathbb{E} \int_0^{t \wedge \tau_K} (\alpha S_u + (b - \beta)R_u) du.$$

As $K \rightarrow \infty$, $\tau_K \rightarrow \infty$ by nonexplosion; the stopped coordinates are bounded by Y_t and the integrands by $(d + \alpha + \beta + b)Y_t$, both integrable by (9). Dominated convergence on the endpoint and on the time integral removes the stopping and gives the integral form of (10). The integrands are bounded by an integrable function uniformly in $u \leq t$, so s and r are continuous, hence the integral equations make them continuously differentiable, and the linear system (10) has a unique solution. Positivity of m : with probability at least $\frac{5}{24}e^{-(b+\beta)t} > 0$ a resistant founder has no event before t , so $\mathbb{P}(N_t \geq 1) > 0$ and $m(t) > 0$.

Scalar process. By Lemma 4.3 below, or simply by nonnegativity of s, r and $m > 0$, the rates (3) are continuous with $0 \leq \lambda \leq b$ and $0 \leq \delta \leq d$. The same embedding (attach each scalar cell to a Yule cell, accept a Yule division at time u with probability $\lambda(u)/b$ independently, and run independent death clocks of rate $\delta(u)$) produces a nonexplosive process with the prescribed

intensities, dominated by Y . The stopped identity gives $\mathbb{E}Z_{t \wedge \tau_K} = 1 + \mathbb{E} \int_0^{t \wedge \tau_K} (\lambda(u) - \delta(u)) Z_u \, du$, and dominated convergence yields $(\mathbb{E}Z)' = (\lambda - \delta)\mathbb{E}Z$ with $\mathbb{E}Z_0 = 1$. From (10), $m' = br - ds = (\lambda - \delta)m$ with $m(0) = 1$, so $\mathbb{E}Z_t = m(t)$ by uniqueness for this scalar linear equation. $\square$

4.2 The scalar terminal law

Define the accumulated-birth parameter

$$A(t) = m(t) \int_0^t \frac{\lambda(u)}{m(u)} \, du. \quad (11)$$

Since $m' = (\lambda - \delta)m$, $A' = \lambda + (\lambda - \delta)A$, and differentiating $1/m$ gives the normalization identity

$$1 + A(t) - m(t) = m(t) \int_0^t \frac{\delta(u)}{m(u)} \, du \geq 0. \quad (12)$$

Lemma 4.2 (Kendall's law for the closure). *Write $m = m(T)$ and $A = A(T)$. For $0 \leq z \leq 1$,*

$$\mathbb{E}z^{Z_T} = 1 - \frac{m(1-z)}{1+A(1-z)}, \quad (13)$$

so that

$$\mathbb{P}(Z_T = 0) = w_0 = 1 - \frac{m}{1+A}, \quad \mathbb{P}(Z_T = n) = w_n = \frac{m}{(1+A)^2} \left(\frac{A}{1+A} \right)^{n-1} \quad (n \geq 1), \quad (14)$$

and in particular

$$\mathbb{P}(Z_T \geq 3) = \frac{mA^2}{(1+A)^3}, \quad \text{Var } Z_T = m(1+2A) - m^2. \quad (15)$$

Proof. The law (13) is the classical linear-fractional law of the time-dependent birth–death process [19]. We include a proof that identifies it with the process constructed in Lemma 4.1 without assuming uniqueness of an infinite forward system. Fix $z \in (0, 1)$ and, for $0 \leq u \leq T$, set

$$M_u = \frac{m(T)}{m(u)}, \quad B_u = m(T) \int_u^T \frac{\lambda(v)}{m(v)} \, dv, \quad h(u) = 1 - \frac{M_u(1-z)}{1+B_u(1-z)}.$$

Integrating $(1/m)' = -(\lambda - \delta)/m$ over $[u, T]$ and multiplying by $m(T)$ gives $1 + B_u - M_u = m(T) \int_u^T \delta/m \, dv \geq 0$, so $M_u(1-z) \leq (1+B_u)(1-z) \leq 1+B_u(1-z)$ and $0 \leq h(u) \leq 1$. Direct differentiation with $M'_u = -(\lambda - \delta)M_u$ and $B'_u = -\lambda M_u$ gives

$$h'(u) = (\lambda(u)h(u) - \delta(u))(1 - h(u)), \quad h(T) = z. \quad (16)$$

Let $f(u, n) = h(u)^n$ with $f(u, 0) = 1$. The scalar generator at time u acts by $\mathcal{L}_u f(u, n) = \lambda(u)n(h^{n+1} - h^n) + \delta(u)n(h^{n-1} - h^n) = nh^{n-1}(1-h)(\delta - \lambda h)$, so by (16) $(\partial_u + \mathcal{L}_u)f = 0$. Stop Z at $\tau_K = \inf\{u : Z_u > K\}$; the stopped process lives on $\{0, \dots, K+1\}$, the rates are continuous, and f is C^1 in u , so the finite-state time-inhomogeneous Dynkin formula gives $\mathbb{E}f(T \wedge \tau_K, Z_{T \wedge \tau_K}) = f(0, 1) = h(0)$. Since $0 \leq f \leq 1$ and $\tau_K \rightarrow \infty$, bounded convergence gives $\mathbb{E}z^{Z_T} = h(0)$. At $u = 0$, $M_0 = m(T)$ and $B_0 = A(T)$, which is (13) for $z \in (0, 1)$; the endpoints $z = 0, 1$ follow by monotone convergence, avoiding any convention at 0^0 . Expanding (13) in powers of z gives (14); the geometric series sums the weights to one, and (12) makes $w_0 \geq 0$. Finally $1 - w_0 - w_1 - w_2 = mA^2/(1+A)^3$ and $\text{Var } Z_T = G''(1) + G'(1) - G'(1)^2 = 2mA + m - m^2$ with G the right-hand side of (13). $\square$

4.3 An explicit analytic box

Lemma 4.3 (Box for the mean and the birth parameter). $0 < m(T) \leq 11/20$ and $0 \leq A(T) \leq 2/5$.

Proof. Let $m_0(t) = \frac{19}{24}e^{-3t/10} + \frac{5}{24}e^{t/10}$ be the total mean of the same system with $\alpha = \beta = 0$. Integrating factors in (10) give, for $0 \leq t \leq 7$,

$$m(t) \leq e^{t/10}, \quad r(t) \leq \frac{26}{25} \cdot \frac{5}{24} e^{t/10}, \quad (17)$$

$$s(t) \leq \frac{19}{24}e^{-3t/10} + \frac{1}{40000}e^{t/10}, \quad m(t) \geq \frac{993}{1000} m_0(t). \quad (18)$$

Indeed $m' = br - ds \leq bm$ gives the first bound. Then $r' = \alpha s + (b - \beta)r \leq \alpha e^{bt} + br$, so $(e^{-bt}r)' \leq \alpha$ and $e^{-bt}r \leq \frac{5}{24} + 7\alpha = \frac{5}{24} + \frac{7}{1000} < \frac{26}{25} \cdot \frac{5}{24}$. Next $s' \leq -ds + \beta r \leq -ds + \beta e^{bt}$, so $(e^{dt}s)' \leq \beta e^{(b+d)t}$ and $s \leq \frac{19}{24}e^{-dt} + \frac{\beta}{b+d}e^{bt}$ with $\beta/(b+d) = 1/40000$. For the lower bound keep only the diagonal terms, $s \geq \frac{19}{24}e^{-(d+\alpha)t}$ and $r \geq \frac{5}{24}e^{(b-\beta)t}$, and use $e^{-\beta t} \geq e^{-\alpha t} \geq 1 - \alpha t \geq 993/1000$.

Mean. Since $T < 7$ we have $e^{T/10} = 2^c \leq 2$, and $e^{-3T/10} = 2^{-3c} = \frac{1}{8}e^{3\beta T} \leq \frac{1}{8} \cdot \frac{1001}{1000}$ because $3\beta T < 21/100000$ and $e^x \leq (1-x)^{-1} < 1001/1000$ there. Adding the bounds on $s(T)$ and $r(T)$,

$$m(T) \leq \frac{19}{24} \cdot \frac{1001}{8000} + \frac{1}{20000} + \frac{26}{25} \cdot \frac{5}{24} \cdot 2 = 0.53244 \dots < \frac{11}{20}.$$

Birth parameter. By (11) and (3), $A(T) = m(T) \int_0^T br(u)/m(u)^2 du$. Substitute $x = e^{u/10} - 1$, so $du = 10 dx/(1+x)$ and $x \leq e^{T/10} - 1 \leq 1$, and write $D(x) = 5(1+x)^4 + 19$, so that $m_0(10 \log(1+x)) = D(x)/(24(1+x)^3)$. Inserting (17)–(18) and $\frac{5}{24} \cdot 24^2 = 120$,

$$A(T) = m(T) \int_0^{e^{T/10}-1} \frac{r(10 \log(1+x))}{(1+x)m(10 \log(1+x))^2} dx \leq \frac{11}{20} \cdot \frac{26/25}{(993/1000)^2} \int_0^1 \frac{120(1+x)^6}{D(x)^2} dx. \quad (19)$$

Let $I = \int_0^1 [120(1+x)^6 + 1368(1+x)^2]/D(x)^2 dx$. With $\mu_0(x) = D(x)/(24(1+x)^3)$ one checks $(1/\mu_0)' = [1368(1+x)^2 - 120(1+x)^6]/D(x)^2$ and $1/\mu_0(1) - 1/\mu_0(0) = 31/33$, hence

$$\int_0^1 \frac{120(1+x)^6}{D(x)^2} dx = \frac{1}{2} \left(I - \frac{31}{33} \right) \leq \frac{173}{264},$$

where the inequality uses $I \leq 9/4$, proved in Appendix A by one positive polynomial identity. Substituting in (19),

$$A(T) \leq \frac{11}{20} \cdot \frac{26}{25} \cdot \left(\frac{1000}{993} \right)^2 \cdot \frac{173}{264} = \frac{1124500}{2958147} = 0.38013 \dots < \frac{2}{5}. \quad \square$$

Remark 4.4. The box is deliberately coarse. The numerical values are $m(T) = 0.51862$ and $A(T) = 0.32163$ (Table 2); the exact value of I is 2.1788, against the certified $132541/60000 = 2.2090$. A sharper box would sharpen (4) but would not change the decision, and the polynomial certificate is the only auxiliary inequality the proof needs.

4.4 The scalar tail

The function $a \mapsto a^2/(1+a)^3$ has derivative $a(2-a)/(1+a)^4 \geq 0$ on $[0, 2]$, so it is nondecreasing on $[0, 2/5]$. By (15) and Lemma 4.3,

$$\mathbb{P}(Z_T \geq 3) = \frac{m(T)A(T)^2}{(1+A(T))^3} \leq \frac{(11/20)(2/5)^2}{(7/5)^3} = \frac{11}{343}, \quad (20)$$

which is (4).

4.5 The actual tail from seven histories

Start from one resistant cell and kill the process at its first switching event. Before killing, the population is a pure-birth process in which each of the n cells divides at rate b and the whole population is killed at rate βn . Let $q_n(t)$ be the probability of being alive with n cells at time t . Then $q'_1 = -(b + \beta)q_1$ and $q'_n = b(n - 1)q_{n-1} - (b + \beta)nq_n$ for $n \geq 2$, with $q_1(0) = 1$, $q_n(0) = 0$, whose solution is

$$q_n(t) = e^{-(b+\beta)t} \left[c(1 - e^{-(b+\beta)t}) \right]^{n-1}, \quad n \geq 1. \quad (21)$$

At the horizon $e^{-(b+\beta)T} = \frac{1}{2}$, so $q_n(T) = \frac{1}{2}(c/2)^{n-1}$. These equations describe a finite family of histories (no switch, n cells); no truncation of the unrestricted law is assumed, and the event “alive with n cells at T ” for the killed process is a sub-event of $\{N_T = n\}$ for the unrestricted process. Restricting to an initially resistant founder (probability $5/24$) and final counts $3, \dots, 7$,

$$\mathbb{P}(N_T \geq 3) \geq \frac{5}{24} \sum_{n=3}^7 q_n(T) = \frac{5}{48} \sum_{j=2}^6 \left(\frac{c}{2}\right)^j = L,$$

and rational arithmetic gives $L > 1/20$, which proves (5). This is the decisive compression: seven live states and one cemetery state certify the failure without any computation of the joint branching law. Every omitted history, sensitive-founder or post-switch, could only increase the tail.

Remark 4.5 (A sharper tail from the full geometric series). Summing all no-switch resistant histories with final count at least three gives

$$\mathbb{P}(N_T \geq 3) \geq \frac{5}{48} \cdot \frac{(c/2)^2}{1 - c/2} = \frac{7812500}{150045003} = 0.052067711978\dots, \quad (22)$$

so $\mathbb{P}(N_T \leq 2) \leq 0.947932\dots$, within 3.4×10^{-4} of the numerical value 0.947598. The seven-history bound is what is formalized (Appendix B); the series bound is a conventional strengthening, useful when a larger certified margin is wanted.

4.6 The repair from selected successful histories

For a sensitive founder, either no event occurs before T , or the death clock rings before both the switching clock and T . Both histories finish with at most one cell. Their combined probability is

$$e^{-(d+\alpha)T} + \int_0^T d e^{-(d+\alpha)t} dt = \frac{d}{d+\alpha} + \frac{\alpha}{d+\alpha} e^{-(d+\alpha)T} \geq \frac{d}{d+\alpha} = \frac{300}{301}. \quad (23)$$

For a resistant founder use the no-switch histories with final counts 1, 2, 3 in (21): probability $\frac{1}{2}(1 + \frac{c}{2} + (\frac{c}{2})^2)$. The disjoint founder types give $\mathbb{P}(N_T \leq 3) \geq \frac{19}{24} \cdot \frac{300}{301} + \frac{5}{48} (1 + \frac{c}{2} + (\frac{c}{2})^2) = H$, and rational arithmetic gives $H > 0.9713$. This proves (6) and completes the proof of Theorem 3.1. The repair is proved directly for the positive-switching process; no no-switching theorem is transferred.

5 Variance separation and the number of founders

5.1 Variance bounds

Proposition 5.1. *At the horizon, $v_N = \text{Var } N_T > 4/5$ and $v_Z = \text{Var } Z_T \leq 11/16$; hence $v_Z/v_N < 55/64$.*

Proof. By (9) both second moments are finite. Retaining only the initially resistant no-switch histories with final counts $1, \dots, 7$ in $\mathbb{E}N_T^2$ and using $m(T) \leq 11/20$,

$$v_N \geq \frac{5}{48} \sum_{n=1}^7 n^2 \left(\frac{c}{2}\right)^{n-1} - \left(\frac{11}{20}\right)^2 = \frac{487556732813619898056609881}{600360090012000900036000600} = 0.81210\dots > \frac{4}{5}.$$

For the closure, (15) and $A \leq 2/5$ give $v_Z = m(1 + 2A) - m^2 \leq \frac{9}{5}m - m^2$, and the quadratic $\frac{9}{5}m - m^2$ is increasing on $[0, 11/20]$ with value $11/16$ at $m = 11/20$. $\square$

5.2 Independent founders

Let $F \geq 1$ founders be prepared independently, each with its own type draw (1), and let $U_F = \sum_{i=1}^F N_T^{(i)}$ be the total count, a sum of independent copies of N_T . The scalar closure of the F -founder assay is the process with $Z_0 = F$ and the same rates (3); by the branching property its terminal count V_F is a sum of F independent copies of Z_T , and $\mathbb{E}U_F = \mathbb{E}V_F = Fm$. For $p \in (0, 1)$ let $q_F(p) = \min\{k \in \mathbb{N}_0 : \mathbb{P}(V_F \leq k) \geq p\}$ be the integer quantile of the closure, so that $[q_F(.025), q_F(.975)]$ is the scalar central 95% interval.

Theorem 5.2 (Founder-number limit). *With Φ the standard normal distribution function and $z_{.975} = \Phi^{-1}(0.975)$,*

$$\begin{aligned} \lim_{F \rightarrow \infty} \mathbb{P}\{q_F(.025) \leq U_F \leq q_F(.975)\} &= 2\Phi\left(z_{.975}\sqrt{v_Z/v_N}\right) - 1 \\ &< 2\Phi\left(z_{.975}\sqrt{55/64}\right) - 1 = 0.930772\dots < 0.95. \end{aligned} \quad (24)$$

Proof. Both terminal laws have finite, strictly positive variance (Proposition 5.1). By the central limit theorem [12, Chapter X], $(V_F - Fm)/\sqrt{Fv_Z}$ converges in distribution to Φ , and because Φ is continuous the convergence of distribution functions is uniform (Pólya's theorem, 29, Lemma 2.11). Fix $p \in (0, 1)$ and $\varepsilon > 0$, and let $k_{\pm} = Fm + (\Phi^{-1}(p) \pm \varepsilon)\sqrt{Fv_Z}$. Then $\mathbb{P}(V_F \leq \lfloor k_+ \rfloor) \rightarrow \Phi(\Phi^{-1}(p) + \varepsilon) > p$ and $\mathbb{P}(V_F \leq \lceil k_- \rceil) \rightarrow \Phi(\Phi^{-1}(p) - \varepsilon) < p$, so for large F , $k_- \leq q_F(p) \leq k_+ + 1$; letting $\varepsilon \downarrow 0$,

$$\frac{q_F(p) - Fm}{\sqrt{Fv_Z}} \longrightarrow \Phi^{-1}(p) \quad (p = 0.025, 0.975).$$

Now $\mathbb{P}\{q_F(.025) \leq U_F \leq q_F(.975)\} = \mathbb{P}(U_F \leq q_F(.975)) - \mathbb{P}(U_F \leq q_F(.025) - 1)$. On the $\sqrt{Fv_N}$ scale the two endpoints converge to $\pm z_{.975}\sqrt{v_Z/v_N}$ (the subtraction of one vanishes after division by $\sqrt{Fv_N}$), and the central limit theorem for U_F with uniform convergence and continuity of the limit gives $\Phi(z_{.975}\sqrt{v_Z/v_N}) - \Phi(-z_{.975}\sqrt{v_Z/v_N})$, which is the first equality. The strict inequalities follow from $v_Z/v_N < 55/64 < 1$ and strict monotonicity of Φ ; they do not depend on the decimal evaluation. $\square$

Remark 5.3. The theorem is an application of the central limit theorem to two different variances, not a failure of it. Relative count fluctuations decrease like $F^{-1/2}$ in both models, but the ratio of the two variances does not change with F , and the closure's central interval is too narrow by the factor $\sqrt{v_Z/v_N}$ in the limit. The theorem supplies no finite- F rate, and it does not say that every finite F undercovers: integer endpoints move in jumps, and Table 3 shows coverage above 95% at $F = 2$ and $F = 3$. It also does not apply unchanged to related founders, to founder counts that are fixed rather than sampled, to shared culture effects or to estimated parameters. Estimating the variance from independent replicate wells is a different procedure from imposing the closure's variance and is not addressed by the theorem.

5.3 What the closure discards: a variance-defect identity

The following identity holds for any finite-type inherited-state process and explains the mechanism of Proposition 5.1 without reference to the witness.

Consider finitely many types $i \in \{1, \dots, r\}$ with counts $X = (X_i)$, total $N = \sum_i X_i$, type-specific per-cell division and death rates $b_i, d_i \geq 0$, and switching events that change one cell's type and preserve N , with constant per-cell rates. Let $B = \sum_i b_i X_i$ and $D = \sum_i d_i X_i$ be the total division and death propensities, $G = B - D$ the net-growth propensity, $m = \mathbb{E}N$, $\bar{g} = \mathbb{E}G/m$, and let Z be the mean-composition closure with per-cell rates $\lambda = \mathbb{E}B/m$ and $\delta = \mathbb{E}D/m$ and $\text{Var } Z_0 = \text{Var } N_0$. Write $v_N = \text{Var } N$, $v_Z = \text{Var } Z$ and $\Delta = v_N - v_Z$.

Proposition 5.4 (Variance defect of the mean-composition closure). *Assume $m(t) > 0$ and that the second moments are finite and satisfy the generator identities for N and N^2 on $[0, t]$ (as they do whenever the population is dominated by a Yule process, by the argument of Lemma 4.1). Then*

$$v'_N = 2 \text{Cov}(N, G) + \mathbb{E}(B + D), \quad v'_Z = 2\bar{g} v_Z + \mathbb{E}(B + D), \quad (25)$$

$$\Delta' = 2\bar{g} \Delta + 2 \text{Cov}(N, G - \bar{g}N), \quad (26)$$

and consequently

$$\Delta(t) = m(t)^2 \frac{\Delta(0)}{m(0)^2} + 2m(t)^2 \int_0^t \frac{\text{Cov}(N_u, G_u - \bar{g}(u)N_u)}{m(u)^2} du. \quad (27)$$

Proof. A division adds one cell and a death removes one, while switches preserve N , so the generator acts by $\mathcal{L}N = B - D = G$ and $\mathcal{L}N^2 = (2N + 1)B + (-2N + 1)D = 2NG + B + D$. Taking expectations, $(\mathbb{E}N)' = \mathbb{E}G$ and $(\mathbb{E}N^2)' = 2\mathbb{E}(NG) + \mathbb{E}(B + D)$, so $v'_N = 2\mathbb{E}(NG) - 2\mathbb{E}N\mathbb{E}G + \mathbb{E}(B + D) = 2\text{Cov}(N, G) + \mathbb{E}(B + D)$. For the closure the same computation with B replaced by λZ and D by δZ gives $v'_Z = 2(\lambda - \delta)\mathbb{E}Z^2 + (\lambda + \delta)\mathbb{E}Z - 2(\lambda - \delta)(\mathbb{E}Z)^2 = 2\bar{g}v_Z + \mathbb{E}(B + D)$, using $\mathbb{E}Z = m$ and $(\lambda + \delta)m = \mathbb{E}(B + D)$. Subtracting, and writing $\text{Cov}(N, G) = \text{Cov}(N, G - \bar{g}N) + \bar{g}v_N$, gives the third identity. Since $(m^2)' = 2\bar{g}m^2$, the integrating factor m^{-2} turns it into $(\Delta/m^2)' = 2\text{Cov}(N, G - \bar{g}N)/m^2$, which integrates to (27). $\square$

Corollary 5.5 (Sign condition). *If $\Delta(0) = 0$ and $\text{Cov}(N_u, G_u - \bar{g}(u)N_u) \geq 0$ for all $u \in [0, t]$, then $v_N(t) \geq v_Z(t)$, with strict inequality if the covariance is positive on a set of positive measure. If all types have the same net growth rate $b_i - d_i = g$, then $G = gN$, the covariance vanishes identically, and $v_N \equiv v_Z$ whenever $\Delta(0) = 0$, even if the turnover rates $b_i + d_i$ differ between types.*

The identity locates the defect precisely: the closure discards the covariance between population size and the composition-weighted excess of net growth over its mean. Inherited states inflate variance when large populations tend to be those enriched in fast-growing types, which is what happens after a resistant founder in the witness. Inheritance alone does not imply a universal strict variance ordering: the equal-net-growth case has correlated family sizes and different turnover by type, yet the count variance of the closure is exact. Equal variances do not imply equal count laws.

For two types the covariance is computable from the closed second-moment system. With column mean $\mu = (s, r)^\top$, drift matrix M as in (10) (with general rates, $M_{11} = b_S - d_S - \alpha$, $M_{22} = b_R - d_R - \beta$, $M_{12} = \beta$, $M_{21} = \alpha$) and $C = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$, the covariance matrix Σ of (S, R) satisfies

$$\mu' = M\mu, \quad \Sigma' = M\Sigma + \Sigma M^\top + \text{diag}((b_i + d_i)\mu_i) + (\alpha\mu_S + \beta\mu_R)C, \quad (28)$$

because each jump with increment v contributes $vv^\top$ times its rate to the second-moment drift and the switching increments $(\mp 1, \pm 1)$ have $vv^\top = C$ [1]. For one founder of random type, $\Sigma(0) = p(1-p)C$ with $p = 5/24$, so $v_N(0) = \mathbf{1}^\top \Sigma(0) \mathbf{1} = 0 = v_Z(0)$; setting $\Sigma(0) = 0$ instead would silently replace the random-type preparation by a deterministic one. For the witness the integrand of (27) is nonnegative on $[0, T]$ with maximum 0.064, and the identity reproduces $v_N(T) - v_Z(T) = 0.5043$ (Figure 5, numerical). These formulas sit within classical multitype moment theory [2, 20]; their role here is to say exactly what the closure loses.

6 Corrected thresholds beyond the witness

Theorem 3.1 certifies $k = 3$ for one preparation and horizon. It is not a universal add-one rule. Within the same constant-rate branching class, however, corrected thresholds can be computed for any rates, founder law and horizon from a small closed system, without the joint population law.

Proposition 6.1 (Closed low-count equations). *Let $u_S(z, t)$ and $u_R(z, t)$ be the generating functions of N_t given one founder of type S , respectively R , for a two-type process with constant per-cell rates b_S, d_S, α and b_R, d_R, β and faithful same-type division. Then*

$$\partial_t u_S = b_S(u_S^2 - u_S) + d_S(1 - u_S) + \alpha(u_R - u_S), \quad \partial_t u_R = b_R(u_R^2 - u_R) + d_R(1 - u_R) + \beta(u_S - u_R), \quad (29)$$

with $u_S(z, 0) = u_R(z, 0) = z$. Writing $u_i(z, t) = \sum_{n \geq 0} a_{i,n}(t) z^n$, the coefficients with $0 \leq n \leq k$ satisfy the closed system

$$a'_{i,n} = b_i \left(\sum_{j=0}^n a_{i,j} a_{i,n-j} - a_{i,n} \right) + d_i (\mathbf{1}_{n=0} - a_{i,n}) + \sigma_i (a_{\bar{i},n} - a_{i,n}), \quad a_{i,n}(0) = \mathbf{1}_{n=1}, \quad (30)$$

where $\sigma_S = \alpha$, $\sigma_R = \beta$ and $\bar{i}$ is the other type. No coefficient of degree above k appears. For founder resistant probability p ,

$$\mathbb{P}(N_T \leq k) = \sum_{n=0}^k ((1-p) a_{S,n}(T) + p a_{R,n}(T)).$$

Proof. (29) is the backward Kolmogorov equation for the generating functions of a two-type Markov branching process with bounded per-cell rates [2, Chapter V]: conditioning on the first event of the founder, a division replaces it by two independent copies (u_i^2), a death by nothing (1), and a switch by a founder of the other type. The degree- n coefficient of u_i^2 is the Cauchy product $\sum_{j \leq n} a_{i,j} a_{i,n-j}$, which involves only degrees at most n , so the system for degrees $0, \dots, k$ closes. The mixture formula is linearity in the founder type. $\square$

Three points about the use of (30). First, it is not a cap on the population: histories that grow beyond k cells and later lose descendants remain represented in the low-degree coefficients, because the closure is a property of the generating-function convolution and not of a truncated state space. Second, the cost is $2(k+1)$ ordinary differential equations with a right-hand side of order k^2 per evaluation; for $k \leq 10$ this is negligible, and the witness values in Section 7 were obtained this way and cross-checked against a truncated joint forward equation. Third, for F independent identically prepared founders the terminal generating function is $g(z)^F$ with $g = (1-p)u_S + pu_R$, and for M potential founders each present independently with probability f it is $(1-f + fg(z))^M$; the required low-order coefficients follow by polynomial multiplication. Related founders and shared environments need a different model. For time-varying rates the backward equation must be posed in the starting time with fixed terminal time; the elapsed-time form (29) is specific to constant rates.

A mean-only bound is available but weak. Markov’s inequality gives $\mathbb{P}(N_T \geq k + 1) \leq m(T)/(k + 1)$; with the certified $m(T) \leq 11/20$ it certifies 95% coverage only for $k \geq 10$. The type-aware calculation certifies $k = 3$. Any method that uses only the mean must accept this loss, because the mean does not determine the tail.

For certification rather than point estimation, the workflow is: declare the source, founder law, horizon and interval convention; compute the closure’s candidate endpoints; evaluate the actual coefficients; widen the upper endpoint (or recompute both endpoints under the declared rule) until the lower coverage bound meets the target; and, to certify minimality, show that the preceding endpoint’s upper bound fails it. A point estimate from an ordinary solver is a numerical recommendation; a certificate needs an enclosure of the finite system by validated numerics or an analytic argument of the kind used in Section 4. Uncertain parameters require coverage valid over the parameter set or a stated inferential procedure.

7 Numerical results

All numbers in this section are floating-point evaluations, separate from the inequalities of Sections 3–6. Terminal laws were computed from the low-count system (30) with $k = 48$ and, independently, from the joint forward equation on $\{s + r \leq K\}$ with an absorbing overflow state that retains every exit, so that in exact arithmetic the retained mass of any count event is a lower bound and adding the overflow gives an upper bound. Means and variances were computed from (10) and (28). Adaptive integration used relative tolerance 2×10^{-12} and absolute tolerance 2×10^{-14} . The two implementations agree on the witness coverage of $\{0, 1, 2\}$ to 3.8×10^{-15} ; the overflow was 4.84×10^{-11} at $K = 32$ and 7.4×10^{-16} at $K = 48$, the coverages at the two caps differ by 2.7×10^{-15} , and mass residuals were below 4×10^{-15} . These are consistency diagnostics, not roundoff bounds. Software: Python 3.11.15, NumPy 2.4.6, SciPy 1.17.1.

7.1 The one-founder witness

Table 2 compares every certified quantity with its numerical value, and Figure 1 shows the coinciding means and the differing terminal laws. The closure’s central 95% interval is $[0, 2]$: its own distribution function is 0.6076 at 0, 0.9045 at 1 and 0.9768 at 2. Discreteness makes the closure conservative under its own law, with 97.68% coverage; that conservatism still fails to guarantee 95% under the inherited-state law, which puts 5.24% on counts of three or more against the closure’s 2.32%.

In an idealized assay with this preparation, accepting counts zero through two as the nominal 95% prediction region misses more than 5% of future outcomes, and raising the upper endpoint to three has a certified coverage above 97.13%. A real assay would additionally need plating, observation and parameter uncertainty accounted for before such a guarantee could be attached to it.

7.2 Independent founders

Table 3 and Figure 3 convolve the one-founder laws for F up to 100; the omitted actual mass at $F = 100$ is below 10^{-13} with the $k = 48$ coefficient system. The scalar central interval and its actual coverage are recomputed at every F . Coverage is not monotone in F , because integer endpoints move in jumps: it exceeds 95% at $F = 2$ and $F = 3$, and rises slightly between $F = 10$ and $F = 20$. At $F = 100$ the actual coverage is 0.8511, consistent with the numerical limit 0.8488 from the variances of Table 2; both are below the certified bound 0.9308 of Theorem 5.2.

Table 2: One-founder witness at $T = 6.9307787$ days. Analytic statements are proved in Sections 4–5; numerical values are floating-point evaluations of the idealized model, not experimental observations.

Quantity	Analytic statement	Numerical value
Common mean $m(T)$	≤ 0.55	0.5186169458
Birth parameter $A(T)$	≤ 0.4	0.3216267862
Closure coverage of $\{0, 1, 2\}$	$\geq 332/343 = 0.967930$	0.9767606151
Actual coverage of $\{0, 1, 2\}$	$\leq 1 - L = 0.949558$	0.9475977178
Actual coverage of $\{0, 1, 2\}$ (Remark 4.5)	≤ 0.947933	0.9475977178
Actual coverage of $\{0, \dots, 3\}$	$\geq H = 0.971317$	0.9738271218
Actual variance v_N	> 0.8	1.0875939888
Closure variance v_Z	≤ 0.6875	0.5832556124
Variance ratio v_Z/v_N	$< 55/64 = 0.859375$	0.5362806511
Limiting founder coverage	< 0.930773	0.8487997562

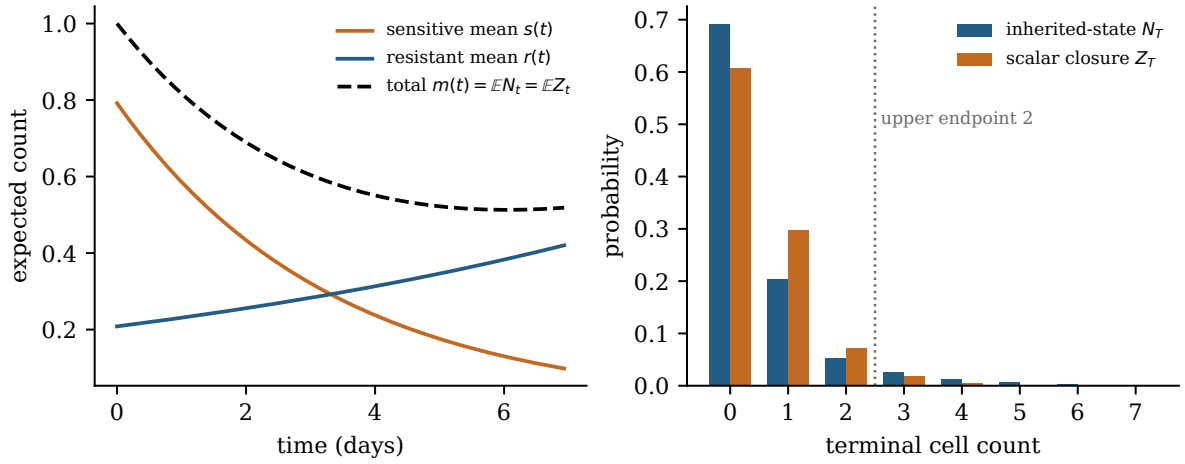


Figure 1: Numerical one-founder witness. Left: exact type-mean trajectories from (10); the total mean coincides with the closure’s mean for all t . Right: terminal count laws at T . The inherited-state process puts more mass both at extinction and at three or more cells while keeping the same mean.

7.3 Scanning both switching rates

The witness’s switching clocks are extreme. Figure 4 and Table 4 vary α over 10^{-4} –1 and β over 10^{-5} –1 per day, holding $b = 0.1$, $d = 0.3$, $b_S = d_R = 0$, $p = 5/24$ and T fixed, and recomputing the scalar central interval at every point. All lower endpoints are zero. Undercoverage of the scalar central interval occurs at seven of the thirty points: the six with $\alpha \leq 10^{-3}$ and $\beta \leq 10^{-3}$, and the point $(\alpha, \beta) = (0.01, 0.01)$, where increasing β from 0.001 to 0.01 moves the scalar upper endpoint from 3 to 2 and drops the actual coverage from 97.28% to 94.79%. The variance ratio v_Z/v_N (right panel, through the limit of Theorem 5.2) varies smoothly and stays below 0.6 over the whole slow-switching corner; the finite-count coverage does not follow it, because the integer quantile transitions dominate at one founder. A smooth variance surface therefore cannot serve as a surrogate for one-founder coverage, whereas it is exactly what governs the many-founder limit. The relevant asymptotic threshold is variance ratio one, not 55/64; the latter is the certified witness value. The grid is synthetic and sparse; it is not interpolated into a certified boundary.

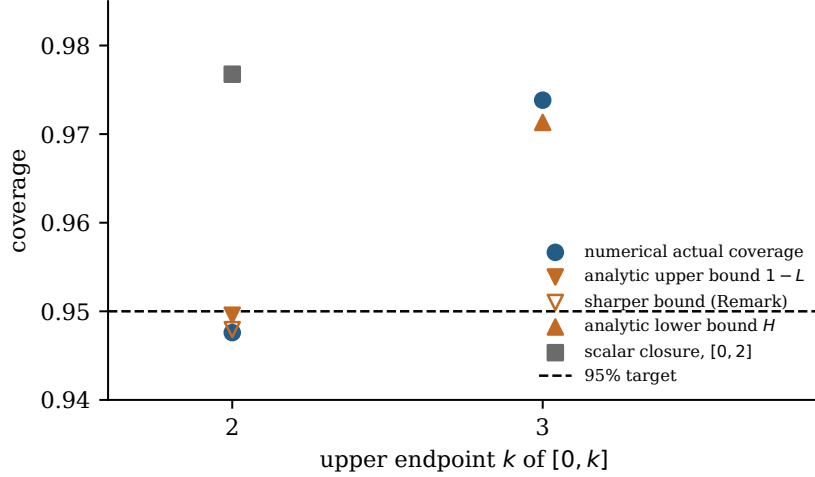


Figure 2: Actual coverage of $\{0, \dots, k\}$ at the two candidate endpoints. Dots: numerical values. Filled triangles: the certified upper bound $1 - L$ at $k = 2$ and lower bound H at $k = 3$; the open triangle is the sharper conventional bound of Remark 4.5. The square is the closure’s own coverage of $\{0, 1, 2\}$.

Table 3: Scalar central 95% interval and its actual coverage for F independent witness founders (numerical; the analytic bound constrains only the limit).

F	Scalar interval	Actual coverage	F	Scalar interval	Actual coverage
1	$[0, 2]$	0.947598	10	$[1, 10]$	0.903909
2	$[0, 4]$	0.962995	20	$[4, 18]$	0.904355
3	$[0, 5]$	0.958707	50	$[16, 37]$	0.868375
5	$[0, 6]$	0.931584	100	$[38, 67]$	0.851146

7.4 Positive rates in both states and day-scale switching

The exact witness converts negative net growth of sensitive cells into pure death and positive net growth of resistant cells into pure division. Table 5 relaxes this. The first two rows add sensitive division and resistant death at a common rate e to the witness: at $e = 0.001$ per day the undercoverage persists (0.9481); at $e = 0.01$ it disappears for this interval criterion (0.9520). The remaining rows change the switching clocks and the zero rates together and follow the assay to $F = 100$ founders. The middle row is the informative one: every event class has a positive rate, both isolated switching clocks have mean ten days, and the autonomous two-state relaxation time $1/(\alpha + \beta)$ is five days, yet the scalar central interval covers only 90.63% at 100 founders and its numerical limit is 90.07%. Its moments are $m(T) = 0.55538$, $v_N = 0.91188$, $v_Z = 0.64471$ and $v_Z/v_N = 0.70702$; the mean exceeds the witness box, so the witness’s certified constants do not transfer, and this row is numerical evidence only. A separate forward calculation at cap 40 agreed with the coefficient system within 3×10^{-14} , with retained overflow at most 1.2×10^{-13} . None of these rows is a fitted biological estimate.

7.5 A source-linked control in which the closure covers

To avoid presenting a counterexample to universal validity as evidence of universal failure, we retain a preparation in which the closure performs well. The pinned Figure 8 code of Browning [6] specifies a two-type preparation with $\alpha = 1$, $\beta = 0.01$, $b_R = 0.1$, $d_S = 0.3$ per day, horizon 7 days, initial count Binomial(1000, f) with $f = 817 \cdot 614 / (\pi \cdot 4500^2)$, and initial resistant fraction

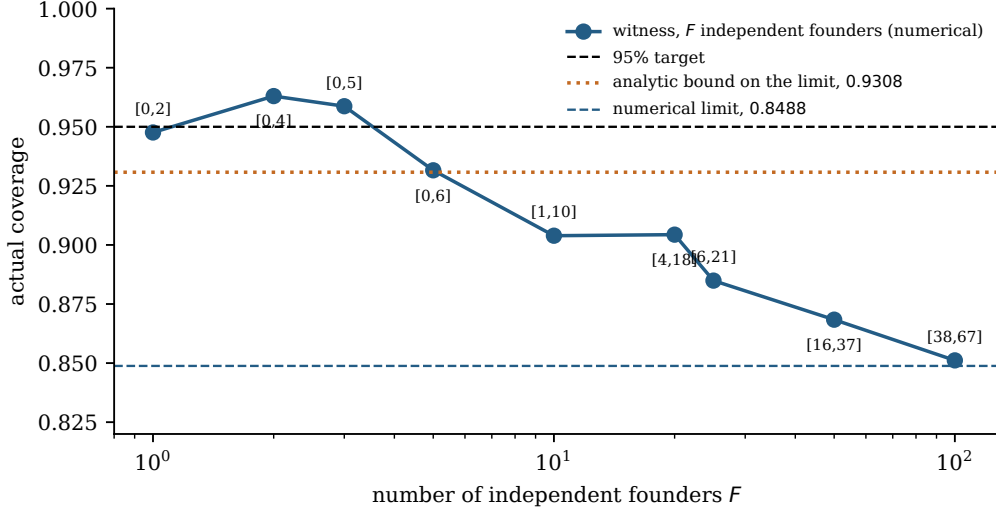


Figure 3: Actual coverage of the scalar central 95% interval against founder number (numerical, labels give the integer endpoints), the numerical limit 0.8488, and the certified upper bound 0.9308 on the limit. The bound constrains the limit, not every finite point.

Table 4: One-founder actual coverage (percent) of the scalar central 95% interval and, in brackets, its upper endpoint, on the switching grid of Figure 4. Rows α , columns β , in day^{-1} . Entries below 95% are in bold.

$\alpha \backslash \beta$	10^{-5}	10^{-4}	10^{-3}	10^{-2}	10^{-1}	1
10^{-4}	94.79 [2]	94.79 [2]	94.82 [2]	95.10 [2]	97.16 [2]	99.82 [1]
10^{-3}	94.76 [2]	94.76 [2]	94.79 [2]	95.07 [2]	97.14 [2]	99.81 [1]
10^{-2}	97.26 [3]	97.26 [3]	97.28 [3]	94.79 [2]	96.94 [2]	99.73 [1]
10^{-1}	96.16 [3]	96.16 [3]	96.18 [3]	96.39 [3]	97.92 [3]	98.93 [1]
1	98.02 [5]	98.02 [5]	98.03 [5]	98.11 [5]	97.14 [4]	97.51 [2]

the smaller root of $0.05\theta^2 - 0.57\theta + 0.02 = 0$, about 0.0352. This is a check of the two-type rate law and its scalar variant, not a reproduction of the full continuous-phenotype inference of Browning et al. [8]. The closure’s central interval [3, 24] has actual coverage 0.95322 (overflow 2.4×10^{-8} at cap 60); the mean is 11.32 and the variance ratio is 1.034. With fast switching, $\alpha T = 7$, the founder state is forgotten within the horizon and the closure is adequate. The control has a different founder-count law and composition from the witness and is not a point on the founder curve of Figure 3.

8 Discussion

Mechanism. The closure fails because of correlated family growth from an inherited founder state. Conditional on a resistant founder, the family grows as a Yule process for the whole horizon with probability close to one; conditional on a sensitive founder, it almost surely dies or stays at one cell. The mixture has the same mean as the closure but more mass at zero and more mass at three or more. Proposition 5.4 makes this quantitative: the closure discards $\text{Cov}(N, G - \bar{g}N)$, which is positive here throughout the horizon. The tail discrepancy is detected by seven explicit histories, the repair by a few more, and the founder-number limit by comparing two variances. None of these steps requires the joint law, which is why the whole argument is certifiable by hand and, in large part, by machine.

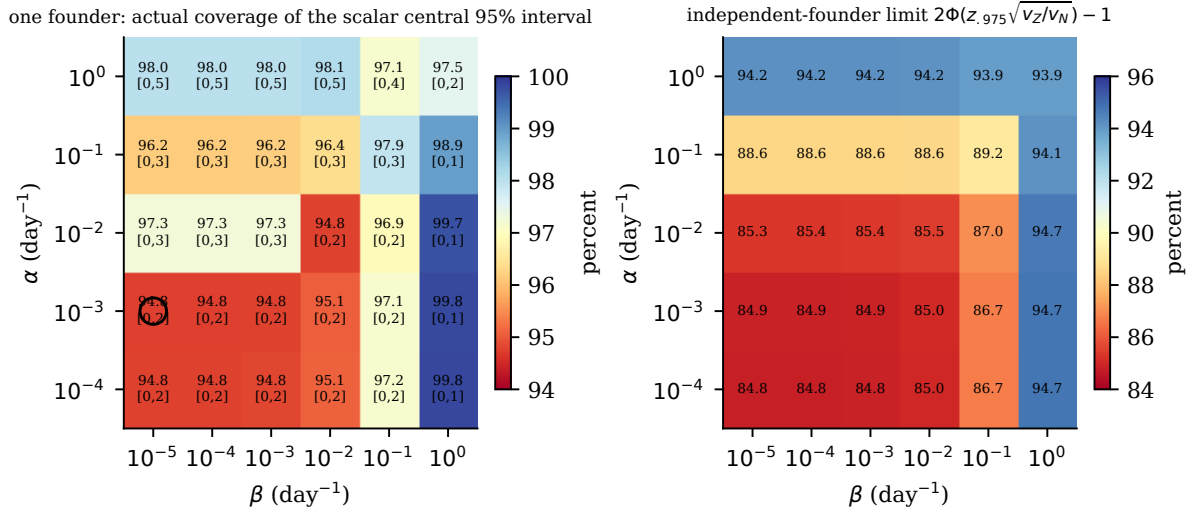


Figure 4: Switching-rate scan at fixed $b = 0.1$, $d = 0.3$, $b_S = d_R = 0$, $p = 5/24$ and T (numerical). Left: one-founder actual coverage of the scalar central 95% interval, with the scalar interval printed in each cell; the circle marks the exact witness. Right: the independent-founder limit $2\Phi(z_{.975}\sqrt{v_Z/v_N}) - 1$ at the same points.

Table 5: Perturbations of the witness with positive sensitive division and resistant death rates $b_S = d_R = e$ (numerical; $b_R = 0.1$, $d_S = 0.3$, $p = 5/24$, horizon T fixed). The limit column is $2\Phi(z_{.975}\sqrt{v_Z/v_N}) - 1$.

α	β	e	$F = 1$ interval	coverage	$F = 100$ interval	coverage	limit
0.001	0.00001	0	[0, 2]	0.947598	[38, 67]	0.851146	0.848800
0.001	0.00001	0.001	[0, 2]	0.948058	—	—	—
0.001	0.00001	0.01	[0, 2]	0.951965	—	—	—
0.01	0.01	0	[0, 2]	0.947846	[38, 68]	0.868361	0.854998
0.1	0.1	0.01	[0, 3]	0.980862	[41, 72]	0.906285	0.900652
1	0.01	0.01	[0, 5]	0.983256	[119, 171]	0.948059	0.943195

What the result does not show. It does not show that mean-composition closures fail in general; Section 7.5 and the fast-switching corner of Table 4 show them working. It does not show that inheritance always inflates the count variance relative to a closure; Corollary 5.5 gives a case with correlated families and no inflation. It does not identify a molecular mechanism, discriminate between forms of phenotypic memory, calibrate any cell line, or prescribe a treatment. The theorem is about prediction regions for a future count under known rates; parameter estimation, posterior prediction, plating randomness, measurement error, density dependence, non-exponential division times and shared environments are outside its assumptions and each can introduce its own calibration problem.

Implications for assay design. Three measurements bear directly on whether a closure of this kind is adequate. Independent well or founder replicates estimate variability across genuinely independent units, whereas descendants within a well are not independent replicates. Founder-state information tests whether high-growth families originate preferentially in one initial state and whether that state persists through division. Perturbations of memory or turnover test whether the predicted count distribution, not just its mean, responds as the model says. Observed overdispersion can also come from variable plating efficiency, uncertain founder number, measurement error or between-well parameter variation, and a model comparison

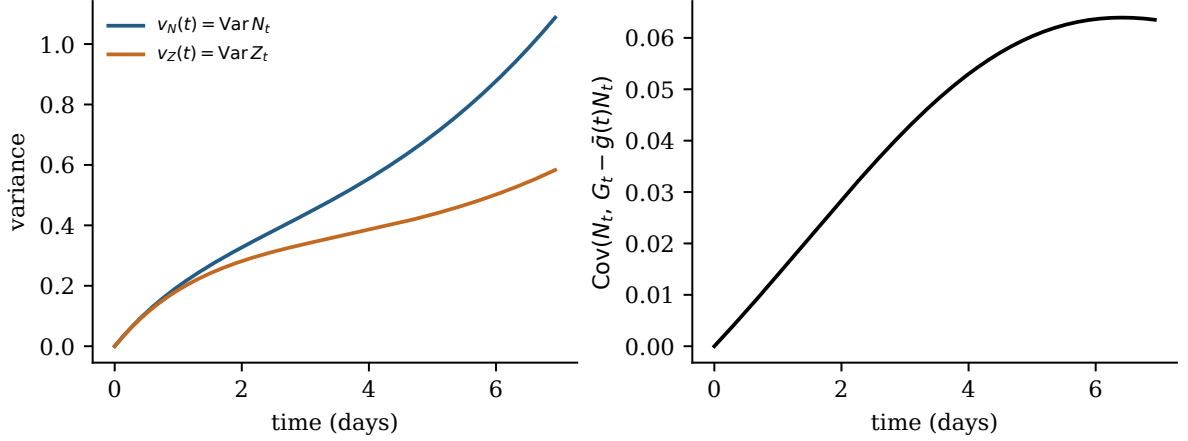


Figure 5: The variance-defect identity for the witness (numerical). Left: $\text{Var } N_t$ and $\text{Var } Z_t$; both start at zero because the random founder type contributes $\Sigma(0) = p(1-p)C$ with $\mathbf{1}^\top C \mathbf{1} = 0$. Right: the integrand $\text{Cov}(N_t, G_t - \bar{g}(t)N_t)$ of (27), nonnegative throughout; the identity reproduces $v_N(T) - v_Z(T) = 0.5043$.

should not attribute all of it to inheritance. Where data permit, birth and death should be distinguished, because a net growth rate alone does not fix the turnover that drives count noise; the positive-rate rows of Table 5 were included for this reason.

Formal status. The stochastic comparison is the principal result and is proved conventionally in full. Lean verifies the actual-process tail obstruction on a normalized, nonexplosive process law, the analytic bounds on the first-moment system, the explicit scalar law with its normalization and 332/343 coverage, and the repaired threshold at the process level (Appendix B). The identification of the unrestricted process means with the ODE solution and of the scalar jump process with Kendall’s law are standard arguments that are proved in Section 4 and are not yet formalized. Completing them would strengthen the verification claim without changing the biological scope.

A The polynomial certificate

Only one auxiliary inequality is needed for Lemma 4.3: $I = \int_0^1 U(x)/D(x)^2 dx \leq 9/4$ with $D(x) = 5(1+x)^4 + 19$ and $U(x) = 120(1+x)^6 + 1368(1+x)^2$. Let

$$P(x) = \frac{25897}{10000} + \frac{3457}{2000}x - \frac{32099}{5000}x^2 + \frac{179}{50}x^3.$$

Expansion gives the exact identity

$$P(x)D(x)^2 - U(x) = \sum_{j=0}^{11} a_j x^j (1-x)^{11-j} \quad (31)$$

with the unnormalized Bernstein coefficients (no additional binomial factor)

$$(a_0, \dots, a_{11}) = \left(\frac{2292}{625}, \frac{41292}{625}, \frac{10987}{625}, \frac{39713}{625}, \frac{12938273}{2500}, \frac{28405139}{1250}, \right. \\ \left. \frac{54190307}{1250}, \frac{26139962}{625}, \frac{174528933}{10000}, \frac{131199}{1250}, \frac{58443}{2000}, \frac{836124}{625} \right).$$

Every coefficient is positive and every basis term is nonnegative on $[0, 1]$, while $D > 0$, so $U/D^2 \leq P$ pointwise on $[0, 1]$ and

$$I \leq \int_0^1 P(x) dx = \frac{132541}{60000} = 2.20901\dots < \frac{9}{4}.$$

The identity (31) is an algebraic certificate, checked by exact expansion of both sides in `check_paper.py`, not a sampled inequality. The exact value of I is 2.17876.

B Lean verification status

The conventional argument of Sections 4–5 is self-contained. Its central components are additionally verified in Lean 4 (toolchain `leanprover/lean4:v4.30.0`) against a frozen Mathlib manifest, with warnings treated as errors, in the namespace `InheritedCellAssay` under `proofs/InheritedCellAssay/`. The verifier exports the elaborated interface of each requested declaration and lists its axioms; the exported root declarations use only `propext`, `Classical.choice` and `Quot.sound`, and no `sorry`, admitted proposition or native numerical certificate enters. The formal clock is $\tau = (b + \beta)t$, so the horizon is $\log 2$ and per-day rates are multiplied by $100000/10001$; the process carries an explicit unit-rate state-preserving event so that the total holding rate stays positive at extinction, which does not alter the observed population. Table 6 maps the mathematical components to their formal status at manuscript build.

Table 6: Claim-to-proof map. “Compiled” means strict verification with a receipt at manuscript build; “conventional” means proved in this paper only.

Mathematical component	Formal status
Nonexplosive actual law; $\mathbb{P}(N_T \geq 3) > 1/20$ (Section 4.5)	Compiled: <code>PositiveLaw.lean</code> , root <code>PositiveTrajectory.actual_count_probability_gt</code> , stated for the normalized chronological measure of the two-type chain and the founder mixture; <code>nonexplosion</code> <code>source_nonexplosion</code> ; killed-history transport <code>finite_le_actual</code> through <code>FiniteClockProjection.clock_finite_projection</code> ; exact seven-history value <code>PositiveKilled.finite_failure_exact</code> . Receipt 61.7 s, 85 source files hash-matched.
Exact seven-history lower bound L as a probability of the actual law	Compiled: <code>PositiveRepair.lean</code> , <code>actual_failure_ge_exact</code> : $\mathbb{P}(N_T \geq 3) \geq L$ with the exact rational L of (7), for the same founder mixture of chronological measures.

Mathematical component	Formal status
Repaired threshold $\mathbb{P}(N_T \leq 3) \geq H$ at the process level (Section 4.6)	Compiled: <code>PositiveRepair.lean</code> , root <code>actual_repair_probability_ge</code> : $\mathbb{P}(N_T \leq 3) \geq 467880212635/481696324816$ for the founder mixture of chronological measures, from a two-live-state killed chain for the sensitive founder (<code>sensitiveSource</code> , <code>finite_sensitive_exact</code> , value $\frac{300}{301} + \frac{1}{301}e^{-(30100/10001)\log 2}$) and the resistant killed chain with payoff “alive with at most three cells” (<code>finite_repair_exact</code>), both transported by <code>stopped_clock_le_unrestricted</code> . Receipt 71.4 s, axioms as above.
Endpoint 2 fails, endpoint 3 succeeds, for the actual law (Corollary 3.2)	Compiled: <code>PositiveRepair.lean</code> , <code>interval_two_undercovers</code> : the mixture probability of $\{N_T \leq 2\}$ is strictly below 19/20, from the probability-measure complement of the tail event; together with the root above this is Corollary 3.2 at the process level.
First-moment ODE bounds (17)–(18); explicit nonnegative ODE solution; $m(T) \leq 11/20$; $A(T) \leq 2/5$ (Lemma 4.3)	Compiled: <code>PositiveMoments.lean</code> , <code>PositiveMomentConstruction.lean</code> (three-state killed chain construction of the solution), <code>PositiveHorizon.lean</code> (<code>endpoint_mean_upper</code>), <code>PositiveBirthIntegral.lean</code> (<code>birth_integral_bound</code>), reusing the Bernstein certificate of Appendix A.
Explicit scalar weights (14): nonnegativity, normalization, coverage $\geq 332/343$ (Lemma 4.2, (20))	Compiled: <code>PositiveScalarLaw.lean</code> (<code>weight_nonneg</code> , <code>weight_pgf</code> , <code>weight_normalized</code> , <code>scalar_coverage_sharp</code>); a distribution-level result about the explicit weights built from the compiled mean and birth integral.
Backward rate identities (16) and rate bounds $0 \leq \lambda \leq b$, $0 \leq \delta \leq d$	Compiled: <code>PositiveScalarRates.lean</code> (<code>backwardMean_derivative</code> , <code>backwardBirth_derivative</code> , <code>scalar_rate_bounds</code>).
Actual unrestricted means equal the ODE solution (Lemma 4.1)	Conventional (Yule embedding, stopped Dynkin identity, dominated convergence); Lean bridge open.
Scalar jump process has the law (13) (Lemma 4.2)	Conventional (bounded backward test, stopped identity, bounded convergence); Lean bridge open.
Variance bounds, founder-number limit, variance-defect identity, low-count equations (Sections 5–6)	Conventional.

Two cautions apply. Compilation of the explicit normalized scalar distribution does not by itself identify that distribution with the time-inhomogeneous jump process; that is the second open bridge. And the earlier no-switching comparison in the same repository

(`CountComparison.lean`, coverage 91/96 and 187/192) is a different theorem and is not used as a substitute for any positive-switching statement.

C Reproducibility

The source package contains `main.tex`, `refs.bib`, the figures, and `check_paper.py`, which (i) replays in exact rational arithmetic every constant in Theorems 3.1 and 5.2, Proposition 5.1, Remark 4.5, Lemma 4.3 and Appendix A, including the expansion of (31) and the symbolic verification of the generator identities behind Proposition 5.4; (ii) recomputes every numerical value of Section 7 from the low-count system (30) and the moment system (28), cross-checks the witness and the positive-rate rows against the truncated joint forward equation, and evaluates the reference control from the pinned source; and (iii) draws Figures 1–5. The whole run takes about two seconds. The Lean sources, receipts and the historical campaign ledger accompany the package separately.

References

- [1] David F. Anderson and Thomas G. Kurtz. *Stochastic Analysis of Biochemical Systems*. Springer, Cham, 2015. doi: 10.1007/978-3-319-16895-1.
- [2] Krishna B. Athreya and Peter E. Ney. *Branching Processes*. Springer, Berlin, 1972. doi: 10.1007/978-3-642-65371-1.
- [3] Norman T. J. Bailey. *The Elements of Stochastic Processes with Applications to the Natural Sciences*. Wiley, New York, 1964.
- [4] Nathalie Q. Balaban, Jack Merrin, Remy Chait, Lukasz Kowalik, and Stanislas Leibler. Bacterial persistence as a phenotypic switch. *Science*, 305(5690):1622–1625, 2004. doi: 10.1126/science.1099390.
- [5] Ivana Bozic, Johannes G. Reiter, Benjamin Allen, Tibor Antal, Krishnendu Chatterjee, Preya Shah, Yo Sup Moon, Amin Yaqubie, Nicole Kelly, Dung T. Le, Evan J. Lipson, Paul B. Chapman, Luis A. Diaz, Bert Vogelstein, and Martin A. Nowak. Evolutionary dynamics of cancer in response to targeted combination therapy. *eLife*, 2:e00747, 2013. doi: 10.7554/eLife.00747.
- [6] Alexander P. Browning. `phenotypic_heterogeneity_ibm`: source code accompanying *Identifiability of phenotypic adaptation from low-cell-count experiments and a stochastic model*. GitHub repository, commit 401c054dcf588262edfa1263767ffa7ce366533b, 2025. URL https://github.com/ap-browning/phenotypic_heterogeneity_ibm.
- [7] Alexander P. Browning, David J. Warne, Kevin Burrage, Ruth E. Baker, and Matthew J. Simpson. Identifiability analysis for stochastic differential equation models in systems biology. *Journal of the Royal Society Interface*, 17(173):20200652, 2020. doi: 10.1098/rsif.2020.0652.
- [8] Alexander P. Browning, Rebecca M. Crossley, Chiara Villa, Philip K. Maini, Adrianne L. Jenner, Tyler Cassidy, and Sara Hamis. Identifiability of phenotypic adaptation from low-cell-count experiments and a stochastic model. *PLOS Computational Biology*, 21(6): e1013202, 2025. doi: 10.1371/journal.pcbi.1013202.
- [9] Leonardo de Moura and Sebastian Ullrich. The Lean 4 theorem prover and programming language. In *Automated Deduction – CCADE 28*, volume 12699 of *Lecture Notes in Computer Science*, pages 625–635. Springer, 2021. doi: 10.1007/978-3-030-79876-5_37.

- [10] Richard Durrett. *Branching Process Models of Cancer*. Springer, Cham, 2015. doi: 10.1007/978-3-319-16065-8.
- [11] Stewart N. Ethier and Thomas G. Kurtz. *Markov Processes: Characterization and Convergence*. Wiley, New York, 1986. doi: 10.1002/9780470316658.
- [12] William Feller. *An Introduction to Probability Theory and Its Applications*, volume I. Wiley, New York, 3rd edition, 1968.
- [13] Tilmann Gneiting, Fadoua Balabdaoui, and Adrian E. Raftery. Probabilistic forecasts, calibration and sharpness. *Journal of the Royal Statistical Society: Series B*, 69(2):243–268, 2007. doi: 10.1111/j.1467-9868.2007.00587.x.
- [14] Ramon Grima. A study of the accuracy of moment-closure approximations for stochastic chemical kinetics. *Journal of Chemical Physics*, 136(15):154105, 2012. doi: 10.1063/1.3702848.
- [15] Patsy Haccou, Peter Jagers, and Vladimir A. Vatutin. *Branching Processes: Variation, Growth, and Extinction of Populations*. Cambridge University Press, Cambridge, 2005. doi: 10.1017/CBO9780511629136.
- [16] Theodore E. Harris. *The Theory of Branching Processes*. Springer, Berlin, 1963.
- [17] Yoh Iwasa, Martin A. Nowak, and Franziska Michor. Evolution of resistance during clonal expansion. *Genetics*, 172(4):2557–2566, 2006. doi: 10.1534/genetics.105.049791.
- [18] Arun Iyer, Arnob Alva, Adrián E. Granada, and Shaon Chakrabarti. Inheritable cell-states shape drug-persister correlations and population dynamics in cancer cells. *PLOS Computational Biology*, 21(9):e1013446, 2025. doi: 10.1371/journal.pcbi.1013446.
- [19] David G. Kendall. On the generalized “birth-and-death” process. *Annals of Mathematical Statistics*, 19(1):1–15, 1948. doi: 10.1214/aoms/1177730285.
- [20] Marek Kimmel and David E. Axelrod. *Branching Processes in Biology*. Springer, New York, 2nd edition, 2015. doi: 10.1007/978-1-4939-1559-0.
- [21] Edo Kussell and Stanislas Leibler. Phenotypic diversity, population growth, and information in fluctuating environments. *Science*, 309(5743):2075–2078, 2005. doi: 10.1126/science.1114383.
- [22] Salvador E. Luria and Max Delbrück. Mutations of bacteria from virus sensitivity to virus resistance. *Genetics*, 28(6):491–511, 1943. doi: 10.1093/genetics/28.6.491.
- [23] James R. Norris. *Markov Chains*. Cambridge University Press, Cambridge, 1997. doi: 10.1017/CBO9780511810633.
- [24] David Schnoerr, Guido Sanguinetti, and Ramon Grima. Approximation and inference methods for stochastic biochemical kinetics—a tutorial review. *Journal of Physics A: Mathematical and Theoretical*, 50(9):093001, 2017. doi: 10.1088/1751-8121/aa54d9.
- [25] Sydney M. Shaffer, Margaret C. Dunagin, Stefan R. Torborg, Eduardo A. Torre, Benjamin Emert, Clemens Krepler, Marilda Beqiri, Katrin Sproesser, Patricia A. Brafford, Min Xiao, Elliott Eggan, Ioannis N. Anastopoulos, Cesar A. Vargas-Garcia, Abhyudai Singh, Katherine L. Nathanson, Meenhard Herlyn, and Arjun Raj. Rare cell variability and drug-induced reprogramming as a mode of cancer drug resistance. *Nature*, 546(7658):431–435, 2017. doi: 10.1038/nature22794.

- [26] Sreenath V. Sharma, Diana Y. Lee, Bihua Li, Margaret P. Quinlan, Fumiyuki Takahashi, Shyamala Maheswaran, Ultan McDermott, Nancy Azizian, Lee Zou, Michael A. Fischbach, Kwok-Kin Wong, Kathleyn Brandstetter, Ben Wittner, Sridhar Ramaswamy, Marie Classon, and Jeff Settleman. A chromatin-mediated reversible drug-tolerant state in cancer cell subpopulations. *Cell*, 141(1):69–80, 2010. doi: 10.1016/j.cell.2010.02.027.
- [27] Abhyudai Singh and João P. Hespanha. Approximate moment dynamics for chemically reacting systems. *IEEE Transactions on Automatic Control*, 56(2):414–418, 2011. doi: 10.1109/TAC.2010.2088631.
- [28] The mathlib Community. The Lean mathematical library. In *Proceedings of the 9th ACM SIGPLAN International Conference on Certified Programs and Proofs (CPP 2020)*, pages 367–381. ACM, 2020. doi: 10.1145/3372885.3373824.
- [29] Aad W. van der Vaart. *Asymptotic Statistics*. Cambridge University Press, Cambridge, 1998. doi: 10.1017/CBO9780511802256.
- [30] G. Udny Yule. A mathematical theory of evolution, based on the conclusions of Dr. J. C. Willis, F.R.S. *Philosophical Transactions of the Royal Society of London. Series B*, 213: 21–87, 1925. doi: 10.1098/rstb.1925.0002.