Conservative molecular inheritance and population risk: extinction and variance ordering, approximation limits, and regrowth

Abstract

When a dividing cell distributes a conserved set of molecular marks between its daughters, the two sister states are dependent even though their later lineages evolve independently. We study a finite reader–writer chromatin motif with N modifiable sites, explicit cell division and a state-dependent death hazard, and compare the exact complementary (conservative) allocation of old marks with the common simplification that draws the two daughters independently from the same marginal law. The two models have identical type-resolved mean dynamics. We prove that, for every N, every choice of nonnegative rate constants, constant division rate and any death hazard that is monotone in the natural order on mark states, complementary allocation gives the smaller extinction probability, the smaller population generating function at every time, and the smaller population variance; under irreducibility the extinction inequality is strict at every founder state whenever survival is possible. An exact identity writes the extinction difference as the conditional covariance of daughter continuation risks propagated by a nonnegative response operator, and positive-vector bounds make it quantitative. The raw pair laws are far apart in total variation, so any accuracy of the independence approximation is observable-specific. Exact rational certificates and a validated Taylor integration give, for a two-site source, a difference of more than 0.098 in the probability of ever reaching 300 cells and of more than 0.084 in reaching 300 cells by time 300. For a 16-site source with a sigmoidal hazard, a near-balanced founder has an extinction gap above 0.0128 uniformly over the slope interval $[7, 9]$, while the all-active founder has a gap below 3.6×10^{-4}. For a fixed smooth readout we prove that survival from balanced founders tends to zero as $N \to \infty$ at an explicit polynomial rate, so the certified finite-size contrast is a finite-size effect; a discontinuous readout behaves differently and is left open, as is state-dependent division, for which we exhibit an exact obstruction to the proof. We also derive what paired-clone and barcode observations can and cannot identify. Probability arguments are conventional; selected finite algebra is checked in Lean, and all numerical claims are exact rational comparisons. The examples are mechanistic and are not clinically calibrated.

1 Introduction

A mean growth law cannot say whether a small population dies out or reaches a consequential size. That question depends on the joint law of the offspring, and in particular on how a mother's state is divided between her daughters. Non-genetic cellular memory is often carried by a finite stock of molecular material: modified nucleosomes are recycled to the two daughter strands at the replication fork in roughly equal shares [1, 31, 39], and random partitioning of molecules at division is a major source of phenotypic heterogeneity [19, 20]. Chromatin-mediated drug-tolerant states [34, 33, 30] make the link to population survival concrete: a lineage persists only while enough of its cells remain in a favourable molecular state.

Conservation has a simple probabilistic consequence. If one daughter inherits many old activating marks, her sister inherits few. The two sister *states* are therefore negatively dependent, even in a model in which their future subtrees are conditionally independent. A modeller who knows only the law of a single daughter may be tempted to draw both daughters independently from that marginal. This *independent comparator* preserves every first moment of the population, but it generally violates conservation of the mother's material across the pair. We ask:

> *Does ignoring the sister dependence change population risk in a definite direction? How large is the change, and which preparations, readouts and clocks govern it?*

Results. We work with a finite reader–writer motif with recruited erasure [11, 7, 6] on N sites, a constant division rate b, and a death hazard d that decreases with activating and increases with repressive marks. Section 2 constructs the chronological branching population and its embedded offspring maps.

1. *Direction* (Theorem 5). For every N and every founder state, complementary allocation has the smaller extinction probability and the smaller scalar population generating function at every time. The same holds for finitely many molecular-independent Erlang division phases. The proof combines a monotone coupling of the molecular chain, an order-preserving coupling of allocations, and Harris' inequality [17, 14] after reversing the repressive allocation bits.

2. *Variance* (Corollary 6). The population size has the same mean and a smaller variance under complementary allocation, at every time; the second factorial moment obeys a linear equation forced by the covariance of daughter continuation *means*.

3. *Strictness* (Proposition 7). With positive basal rates and a nonconstant hazard there is a dichotomy: either both models die out almost surely, or the extinction inequality is strict at every state.

4. *Magnitude* (Theorem 9, Proposition 10). The extinction gap equals a nonnegative response operator applied to the conditional covariance of daughter continuation *risks*. Weighted contraction bounds enclose it. The pair laws themselves are at total-variation distance $1 - \binom{2a}{a}\binom{2r}{r}4^{-(a+r)}$, which tends to one, so closeness of the two models can only be observable-specific.

5. *Certified instances* (Section 5). Exact rational risk boxes give $0.1144144 < q_I - q_J < 0.1144331$ for an all-active two-site founder, and a 16-site sigmoidal-hazard certificate that is uniform over the slope interval $[7, 9]$.

6. *Large N* (Theorem 12). For a fixed Lipschitz readout that is subcritical on the symmetry line, survival from founders within $O(N^{-1/2})$ of that line is at most $O(N^{-\gamma}(\log N)^{3/2})$ with an explicit $\gamma > 0$, in both models. The proof is a first-moment (many-to-one) argument and uses no population coupling.

7. *Regrowth* (Theorems 14–15). An exact identity converts extinction boxes into bounds on reaching large population thresholds, and a validated integration of the microscopic generating function gives a finite-deadline statement.

8. *Observation* (Section 8). Paired sister-clone extinction has nonpositive conditional covariance; we give a sensitivity interval under parental mixing, misclassification and missingness, a finite-sample radius, and the exact effect of independent cell capture on barcode-type clone counts.

Relation to prior work. General multitype branching processes already allow an arbitrary joint law of offspring types [2, 23]; conditional independence of subtrees does not require independent sister types. Crump and Mode studied correlated sister lifetimes [9], Olofsson developed branching with local dependencies and showed that growth is governed by the marginals [29], and Duffy and Subramanian showed for lymphocyte models that related-cell correlation leaves the mean unchanged while changing variability [12]; correlated division times are observed experimentally [32]. Sibling dependence through a shared random environment [38] is a different mechanism. Hence "correlation changes risk but not the mean" is not by itself new. Closest in spirit is the division-within-division literature initiated by Kimmel [22], where a cell's content is shared between daughters [3, 4, 25]; there the partitioning kernel influences the fate of the content and of the cell population. Our contribution is different: a *comparison* theorem, for an explicit chemical source, between the conservative joint law and its independent-marginal replacement, with a sign, an exact magnitude identity, certified finite instances and an asymptotic limit. On the chromatin side, transfer-matrix treatments of partition and recovery [26] and replication-driven switching [27] are established; we do not claim a new switching asymptotic, and we keep molecular reactions, death and division active on one clock rather than assuming recovery between divisions. Models that connect inherited phenotype to treatment response prescribe the inheritance rule [8]; here it is derived from material conservation.

Evidence classes. Stochastic constructions, couplings and limit arguments are conventional proofs given in full. Every printed numerical inequality is an exact rational comparison produced by the accompanying scripts. A subset of the finite algebra is additionally checked in Lean 4 with mathlib [10, 37]; Section 9 states precisely which. Floating-point curves are labelled as diagnostics.

2 Molecular source and chronological population

2.1 Reader–writer source

Fix $N \geq 1$ and let $S_N = \{(a, r) \in \mathbb{N}^2 : a + r \leq N\}$, with $u = N - a - r$. A state records the numbers of activating (A), repressive (R) and unmodified (U) sites. The molecular generator Q has the four nearest-neighbour rates

$$q_{(a,r),(a+1,r)} = u\,(w_A + k_A a/N), \qquad q_{(a,r),(a,r+1)} = u\,(w_R + k_R r/N), \tag{1}$$

$$q_{(a,r),(a-1,r)} = a\,(e_A + h_A r/N), \qquad q_{(a,r),(a,r-1)} = r\,(e_R + h_R a/N), \tag{2}$$

with nonnegative constants: basal writing w, read–write feedback k, intrinsic erasure e and erasure recruited by the opposite mark h. Vanishing counts enforce the boundary of the triangle; diagonal entries of Q are minus the outgoing rates; $|S_N| = (N+1)(N+2)/2$. The factors $1/N$ fix how effective interactions scale with the number of sites; a different contact scaling is a different model. The names "activating" and "repressive" designate a modelled functional order and assert nothing universal about particular chromatin marks. The constants e_A, e_R are *intrinsic* erasure rates: dilution by replication is represented explicitly by division below, so an effective erasure rate from an undivided model that already averages replication loss [7] must not be imported, or that loss would be counted twice.

2.2 Complementary and independent daughters

At division each old mark independently goes to daughter one with probability 1/2, and new unmodified sites restore each daughter to N sites. For a mother $x = (a, r)$ the daughter pair is

$$(Y, Z) = ((I, J), (a - I, r - J)), \qquad I \sim \mathrm{Bin}(a, \tfrac{1}{2}),\ J \sim \mathrm{Bin}(r, \tfrac{1}{2}) \text{ independent.} \tag{3}$$

Write $D_1 f(x) = \mathbb{E}_x f(Y) = \mathbb{E}_x f(Z)$ and $D_J(f,g)(x) = \mathbb{E}_x[f(Y)g(Z)]$. The *independent comparator* draws two independent copies of Y, so that $D_I(f,g) = (D_1 f)(D_1 g)$. It has the same marginals, but it does not conserve the mother's marks across the pair: it is an approximation to be evaluated, not an alternative physical redistribution. The vector $(I, a-I)$ is a two-cell multinomial and is negatively associated in the sense of [21]; what we need below is a sign for functions monotone in a *mixed* order, which requires the argument of Lemma 4.

2.3 Population process

A cell of type x evolves by Q, divides at constant rate $b > 0$ and dies at rate $d(x) \geq 0$. At division it is replaced by two daughters with joint law $D \in \{D_J, D_I\}$; the subtrees of distinct cells are independent given their birth types. This prescribes a continuous-time Markov chain on finite configurations of typed cells.

Lemma 1 (Construction)**.** *For either daughter law and any finite founder configuration:*

(i) The process is nonexplosive, its size $K(t)$ is stochastically dominated by a Yule process of rate b, and $\sup_{s\leq t} K(s)$ has finite moments of all orders.

(ii) With

$$\mathcal{H} = \operatorname{diag}(b+d) - Q, \qquad R = \mathcal{H}^{-1} = \int_0^\infty e^{(Q-\operatorname{diag}(b+d))t}\,dt \ \geq 0, \qquad T = bR, \tag{4}$$

the extinction probabilities $q = (q_x)_{x\in S_N}$ from one founder are the least fixed point in $[0,1]^{S_N}$ of the offspring map

$$F_J(v) = Rd + T\,D_J(v,v) \qquad \text{or} \qquad F_I(v) = Rd + T\,(D_1 v)^2, \tag{5}$$

and $q = \lim_n F^n(0)$. Squares of vectors are coordinatewise. Any $\bar{v} \in [0,1]^{S_N}$ with $F(\bar{v}) \leq \bar{v}$, equivalently $\mathcal{H}\bar{v} \geq d + bD(\bar{v},\bar{v})$, satisfies $q \leq \bar{v}$.

(iii) For $0 \leq z \leq 1$ the generating function $u_x(t;z) = \mathbb{E}_x z^{K(t)}$ $(0^0 = 1)$ is the unique solution of

$$\dot{u} = Qu + d\,(\mathbf{1} - u) + b\,\{D(u,u) - u\}, \qquad u(0;z) = z\mathbf{1}, \tag{6}$$

which leaves $[0,1]^{S_N}$ invariant; $u_x(t;0)$ is the probability of extinction by time t.

(iv) $m_x(t) = \mathbb{E}_x K(t)$ satisfies $\dot{m} = Am$, $m(0) = \mathbf{1}$, with the common *mean generator*

$$A = Q + b\,(2D_1 - I) - \operatorname{diag}(d). \tag{7}$$

The same holds for every type-resolved first moment, so the two models have identical mean dynamics.

Proof. (i) Each cell divides at rate b whatever its type and deaths only remove cells, so a coupling with common division clocks bounds K by a Yule process, whose value at time t is geometric. On $\{K \leq n\}$ all rates are bounded; localisation at the hitting time of size n and the Yule bound give nonexplosion and the moment bound.

(ii) The killed semigroup $e^{(Q-\operatorname{diag}(b+d))t}$ is nonnegative with row sums at most e^{-bt}, which gives (4) and $R \geq 0$. Starting from x, $R_{xy}d_y$ and $R_{xy}b$ are the probabilities that the first terminal event of the founder is a death, respectively a division, occurring in molecular state y. Conditioning on it and using independence of the two subtrees given (Y,Z) shows that $F^n(0)_x$ is the probability that the family tree of the founder has fewer than n generations. Since $\mathcal{H}\mathbf{1} = b\mathbf{1} + d$ we have $F(\mathbf{1}) = \mathbf{1}$: death mass is retained. The monotone limit q^* is the least fixed point of the polynomial map F and lies below every supersolution. Every cell has a terminal event almost surely because $b > 0$, so a finite tree dies out in finite time; conversely an infinite, locally finite binary tree contains an infinite line of descent, which by nonexplosion is not completed in finite time. Thus q^* is the probability of eventual extinction and not merely an algebraic root. The equivalence of the two supersolution conditions follows on multiplying by $R \geq 0$.

(iii) Condition on the first event in $[0, \delta]$; by (i) two or more events have probability $O(\delta^2)$ uniformly in the founder type, and after a division the branching property gives $u_Y(t; z)u_Z(t; z)$ averaged over the *joint* law of (Y, Z). The vector field is polynomial, hence locally Lipschitz, and points inward on the faces of the cube. Zero is absorbing, so $u_x(t; 0) = \mathbb{P}_x(K(t) = 0)$.

(iv) Differentiate (6) at $z = 1$, which is justified by (i), and use $D(f, \mathbf{1}) = D(\mathbf{1}, f) = D_1 f$ for both laws. For a nonnegative weight φ on types the same computation applied to $\mathbb{E}_x \sum_{\text{cells}} \varphi$ gives $e^{tA}\varphi$. □

3 Direction of the inheritance effect

Order the states by

$$(a, r) \preceq (a', r') \iff a \leq a' \text{ and } r \geq r'.$$

An upper state has more activating and less repressive material. A function on S_N is called decreasing (increasing) if it is nonincreasing (nondecreasing) in $\preceq$.

Lemma 2 (Monotone molecular coupling)**.** *For any nonnegative constants in* (1)–(2) *and any* $x \preceq x'$ *there is a Markovian coupling* (X_t, X'_t) *of two copies of the chain* Q *started at* (x, x') *with* $X_t \preceq X'_t$ *for all* t.

Proof. Couple each of the four channels at the minimum of its two rates, letting the excess rate fire in one chain only. A joint jump leaves both coordinate differences unchanged. A solitary jump changes one coordinate difference by one, so it can break the order only if that coordinate is currently equal in the two chains. Let $x = (a, r) \preceq x' = (a', r')$. If $a = a'$ then $u \leq u'$ because $r \geq r'$, so the activating-write rate $u(w_A + k_A a/N)$ of the lower chain does not exceed that of the upper chain, and the activating-erasure rate $a(e_A + h_A r'/N)$ of the upper chain does not exceed that of the lower chain. If $r = r'$ then $u' \leq u$, so the repressive-write rate of the upper chain does not exceed that of the lower one, and the repressive-erasure rate $r(e_R + h_R a/N)$ of the lower chain does not exceed that of the upper one. In each of the four cases the excess jump therefore occurs in the chain for which it preserves the order. At $a + r = N$ the write rates vanish, and at $a = 0$ or $r = 0$ the corresponding erasure rate vanishes, so the coupling respects the triangle. □

Lemma 3 (Ordered allocation and population embedding)**.** *Assume* b *is constant and* d *is decreasing. For* $x \preceq x'$ *and either daughter law, the populations started from one founder of type* x *and one of type* x' *can be coupled so that at every time there is an injection from the living cells of the first into the living cells of the second under which types are ordered by* $\preceq$*. Consequently, in each model,* $x \mapsto u_x(t; z)$ *and* $x \mapsto q_x$ *are decreasing,* $x \mapsto m_x(t)$ *is increasing, and* F_J*,* F_I *map the cone*

$$\mathcal{C} = \{f : S_N \to [0, 1] \text{ decreasing}\}$$

into itself. Both maps are also monotone for the coordinatewise order.

Proof. Run a matched pair of cells with the coupling of Lemma 2, a common division clock of rate b, a common death clock at the smaller hazard $d(X'_t)$ and an additional death clock of rate $d(X_t) - d(X'_t) \geq 0$ for the lower cell alone. If the lower cell dies alone, the upper cell and its descendants continue unmatched. At a common division with mother states $(a, r) \preceq (a', r')$ use shared fair bits for the a common activating marks and the r' common repressive marks, and fresh independent bits for the $a' - a$ additional activating marks of the upper mother and the $r - r'$ additional repressive marks of the lower mother. If (I, J') counts the shared marks sent to daughter one and I'', J'' the additional ones, the daughters are

$$(I,\, J' + J'') \preceq (I + I'',\, J') \qquad \text{and} \qquad (a - I,\, r - J' - J'') \preceq (a' - I - I'',\, r' - J'),$$

the second relation because $I'' \leq a' - a$ and $J'' \leq r - r'$. Both daughter pairs are thus ordered and each mother sees her own complementary law. For independent daughters apply the same marginal coupling separately and independently to daughter one and to daughter two. Recursion over matched pairs, with nonexplosion from Lemma 1, gives the embedding; hence $K^x(t) \leq K^{x'}(t)$ pathwise, which yields the three monotonicity statements. For the cone, stop the matched founders at the first terminal event and pay 1 for a death and $f(Y)f(Z)$ for a division: a solitary death of the lower cell pays $1 \geq$ any payoff of the upper cell, a common death pays 1 to both, and a common division pays $f(Y)f(Z) \geq f(Y')f(Z')$ because f is decreasing and nonnegative. Coordinatewise monotonicity holds because R, T and the coefficients of D are nonnegative. □

Lemma 4 (Sign of the sister covariance)**.** *If $f : S_N \to \mathbb{R}$ is monotone in $\preceq$ (in either direction), then for every mother state x*

$$\operatorname{Cov}_x\big(f(Y), f(Z)\big) \leq 0, \qquad \textit{that is,} \qquad D_J(f, f) \leq (D_1 f)^2. \tag{8}$$

Proof. Let $\alpha_1, \ldots, \alpha_a$ and $\beta_1, \ldots, \beta_r$ be the independent fair bits indicating that an activating, respectively repressive, mark goes to daughter one, and put $\tilde\beta_j = 1 - \beta_j$. Then $Y = (\sum \alpha_i,\ r - \sum \tilde\beta_j)$ and $Z = (a - \sum \alpha_i,\ \sum \tilde\beta_j)$, so in the order $\preceq$ the state Y is increasing and Z is decreasing in each of the independent bits $(\alpha, \tilde\beta)$. Hence $f(Y)$ and $f(Z)$ are monotone in opposite directions in every bit. Harris' inequality [17, 14] states that two functions of independent variables that are increasing in each variable have nonnegative covariance; apply it to $-f(Y)$ and $f(Z)$ or to $f(Y)$ and $-f(Z)$. For completeness: condition on the last bit ξ. For functions G, G' that are increasing in each bit, the law of total covariance gives $\operatorname{Cov}(G, G') = \mathbb{E}\operatorname{Cov}(G, G' \mid \xi) + \frac{1}{4}(g_1 - g_0)(g'_1 - g'_0)$, where g_i, g'_i are the conditional means given $\xi = i$; the first term is nonnegative by induction on the number of bits and the second because both factors are nonnegative. □

The cone is essential: for an all-active mother with two marks, $f(0) = f(2) = 1$, $f(1) = 0$ gives covariance $+1/4$.

Theorem 5 (Source comparison)**.** *Let Q be given by (1)–(2) with arbitrary nonnegative constants, let $b > 0$ be constant and let $d \geq 0$ be decreasing in $\preceq$. Then for every N, founder state x, $t \geq 0$ and $z \in [0, 1]$,*

$$q_J(x) \leq q_I(x), \qquad u_{J,x}(t; z) \leq u_{I,x}(t; z).$$

In particular $\mathbb{P}_x(K_J(t) = 0) \leq \mathbb{P}_x(K_I(t) = 0)$ for all t. The same conclusions hold when the exponential division clock is replaced by m successive exponential phases whose rates do not depend on the molecular state, with molecular reactions and mortality active in every phase, division at the end of the last phase, and daughters born in the first phase.

Proof. Extinction. Let $f_n = F_J^n(0)$ and $g_n = F_I^n(0)$. By Lemma 3, $f_n \in \mathcal{C}$. If $f_n \le g_n$ then, by (8) and (5) and then by monotonicity of F_I,

$$f_{n+1} = F_J(f_n) \le F_I(f_n) \le F_I(g_n) = g_{n+1}.$$

Let $n \to \infty$ and use Lemma 1(ii).

Generating functions. Write (6) as $\dot u = G_D(u)$. By Lemma 3, $u_J(t; z) \in \mathcal{C}$ for all t, so (8) gives $\dot u_J = G_J(u_J) \le G_I(u_J)$: the function u_J is a subsolution of the independent system with the same initial value. On the cube the off-diagonal partial derivatives of G_I are $Q_{xy} + 2b(D_1 v)_x D_1(x, y) \ge 0$, so G_I is quasimonotone and the Kamke–Müller comparison theorem [35, Ch. 3] yields $u_J \le u_I$.

Erlang clocks. Append the phase $j \in \{1, \ldots, m\}$ to the type and compare types with equal phase only. Phase clocks do not depend on (a, r) and are coupled identically, death is coupled as before in every phase, and both daughters restart in phase one, so Lemmas 2 and 3 hold on the finite augmented space with the cone of functions decreasing in (a, r) at each phase. Lemma 4 is applied to $f(\cdot, 1)$, and the two steps above are unchanged. □

This is an ordering of extinction probabilities and of scalar generating functions. It is not stochastic dominance of population sizes, which for nondegenerate cases would contradict the equality of means, and it does not order every first-passage probability.

3.1 Variance ordering

Corollary 6 (Equal means, smaller variance)**.** *Under the hypotheses of Theorem 5 the second factorial moment $v_{D,x}(t) = \mathbb{E}_x[K(t)(K(t) - 1)]$ is the solution of*

$$\dot v_D = A v_D + 2b\, D(m, m), \qquad v_D(0) = 0, \tag{9}$$

where $m(t) = e^{tA}\mathbf{1}$ is the common mean. Consequently

$$v_I(t) - v_J(t) = -2b \int_0^t e^{(t-s)A} \big(\mathrm{Cov}_x(m_Y(s), m_Z(s))\big)_x \, ds \;\ge 0, \qquad \mathrm{Var}_x\, K_J(t) \le \mathrm{Var}_x\, K_I(t) \tag{10}$$

for every x and t. The inequality persists for any common initial mixture of founder types and for sums over finitely many independent founders.

Proof. By Lemma 1(i) the generating function is twice differentiable at $z = 1$. Differentiating (6) twice at $z = 1$, where $u = \mathbf{1}$, $\partial_z u = m$, and using $\partial_z^2 D(u, u) = 2D_1 \partial_z^2 u + 2D(m, m)$, gives (9). By Lemma 3 $m(s)$ is increasing, so Lemma 4 gives $(D_1 m)^2 - D_J(m, m) \ge 0$; the matrix A has nonnegative off-diagonal entries, so $e^{(t-s)A} \ge 0$ and variation of constants gives (10). The means agree, so the variances are ordered. For a mixture the between-founder variance of the conditional mean is common to the two models, and for independent founders variances add. □

Alternatively, $u_I(t; z) - u_J(t; z) \ge 0$ on $[0, 1]$ vanishes together with its first derivative at $z = 1$, which forces a nonnegative second derivative there. The mechanism is visible in (10): the covariance of daughter continuation *means* forces the variance, whereas Section 4 shows that the covariance of daughter continuation *risks* forces the extinction gap. Smaller variance at equal mean does not in general imply smaller extinction probability for branching processes; here both follow from the same association argument, and neither orders higher moments or first-passage events. Figure 1C shows the variance ratio for the two worked sources of Section 5: at $t = 60$ it is $91.9/127.6 \approx 0.72$ for the two-site source and $15510/16276 \approx 0.95$ for the 16-site source (diagnostic values).

3.2 Strictness

The inequalities of Theorem 5 are not strict in general: if both models die out almost surely then $q_J = q_I = \mathbf{1}$. Under mild conditions this is the only exception.

Proposition 7 (Dichotomy)**.** *In addition to the hypotheses of Theorem 5, assume $w_A, w_R, e_A, e_R > 0$ and that d is not constant. Let $M = 2TD_1$ be the common mean offspring matrix of the embedded generation process and $\rho(M)$ its spectral radius. If $\rho(M) \le 1$ then $q_J = q_I = \mathbf{1}$. If $\rho(M) > 1$ then*

$$q_J(x) < q_I(x) < 1 \qquad \text{for every } x \in S_N.$$

Proof. Positive basal rates make Q irreducible on S_N, so R and T are entrywise positive, and $D_1(y,y) = 2^{-(a+r)} > 0$ makes M entrywise positive. The generation process is a multitype Galton–Watson process, with joint offspring types under D_J and independent ones under D_I, which is positively regular and nonsingular with mean matrix M in both models. By the classical criterion [2, Ch. V.3], $q = \mathbf{1}$ if $\rho(M) \le 1$ and $q < \mathbf{1}$ in every coordinate if $\rho(M) > 1$.

Assume $\rho(M) > 1$. If q_J were a constant $c < 1$, then $\mathcal{H}q_J = d + bD_J(q_J, q_J)$ would read $(b+d_x)c = d_x + bc^2$, that is $(1-c)(d_x - bc) = 0$, for every x, contradicting that d is not constant. So the decreasing function q_J is not constant, and since the path $(0,N), (0,N-1), \dots, (0,0), (1,0), \dots, (N,0)$ is increasing in $\preceq$, the function q_J is nonconstant on $\{(i,0)\}_{i \le N}$ or on $\{(0,j)\}_{j \le N}$. In the first case take the mother $(N,0)$. The last step of the proof of Lemma 4 applied to $G = -q_J(Y)$, $G' = q_J(Z)$ gives

$$\mathrm{Cov}_{(N,0)}(q_J(Y), q_J(Z)) \le -\tfrac{1}{4}\Big(\mathbb{E}\big[q_J(I',0) - q_J(I'+1,0)\big]\Big)^2 < 0, \qquad I' \sim \mathrm{Bin}(N-1, \tfrac{1}{2}),$$

where the two conditional-mean differences coincide by the symmetry $I' \mapsto N-1-I'$, and the expectation is strictly positive because I' charges every value in $\{0, \dots, N-1\}$ and $q_J(\cdot, 0)$ is decreasing and nonconstant. The second case is identical with the mother $(0,N)$. By (11) below the forcing $\delta = -T(\mathrm{Cov}_x(q_J(Y), q_J(Z)))_x$ is then strictly positive at *every* state because $T > 0$, and Theorem 9 gives $q_I - q_J \ge \delta > 0$. ☐

If d is constant the type is irrelevant and $q_J = q_I = \min(1, d/b)\mathbf{1}$.

3.3 State-dependent division: an obstruction and an open question

For a bounded state-dependent division rate $b(x)$ the construction of Section 2 goes through with $T = R\,\mathrm{diag}(b)$ and $A = Q + \mathrm{diag}(b)(2D_1 - I) - \mathrm{diag}(d)$, and the two models still share their mean. The direction theorem, however, rests on cone preservation, and monotonicity of b does not restore it: a faster division clock in a well-marked cell spreads its marks over two less-marked cells sooner.

Proposition 8 (Exact cone obstruction)**.** *Let $N = 4$, $Q = 0$, d constant, and for some $c \in (0,1)$ let $f(a,r) = 1$ for $a \le 2$ and $f(a,r) = c$ for $a \ge 3$, a decreasing function. Then*

$$D_J(f,f)(3,0) - c = \tfrac{3}{4}(1-c), \qquad D_J(f,f)(4,0) - c = \tfrac{3}{8}(1-c),$$

so if b is increasing with $b(4,0) > 2\,b(3,0)$, the vector field of (6) at f is strictly larger at $(4,0)$ than at $(3,0)$ although $f(3,0) = f(4,0)$ and $(3,0) \preceq (4,0)$: it points out of $\mathcal{C}$.

Proof. For the mother $(3,0)$ the product $f(Y)f(Z)$ equals c exactly when one daughter receives all three marks, which has probability $1/4$, and equals 1 otherwise. For $(4,0)$ it equals c unless the split is $(2,2)$, which has probability $3/8$. The molecular and death terms of the vector field coincide at the two states. ☐

This is a failure of the proof route at a test vector; it is not an example with $q_J > q_I$. A random search followed by hill climbing over roughly 8×10^3 instances with $N \le 5$, b increasing and d decreasing, including $Q = 0$, found no reversal of the extinction order beyond floating point noise (diagnostic). Whether Theorem 5 survives state-dependent division under a natural condition, perhaps only along the actual generating-function flow, is open. The Erlang extension of Theorem 5 accommodates a variable cell-cycle duration, but not a trade-off between proliferation and molecular state.

4 Magnitude: covariance and population response

For any vector v, conditioning on the complete pre-division mother state gives

$$F_J(v) - F_I(v) = T\big(\mathrm{Cov}_x(v(Y), v(Z))\big)_{x \in S_N}. \tag{11}$$

The resolvent T weights the molecular states actually visited before division, and the covariance weights daughters by their *continuation risk*, not by raw mark abundance.

Theorem 9 (Exact response and enclosures)**.** *Let $g = q_I - q_J$,*

$$\delta = F_I(q_J) - F_J(q_J) = -T\big(\mathrm{Cov}_x(q_J(Y), q_J(Z))\big)_x, \qquad B = T\,\mathrm{diag}\big(D_1(q_I + q_J)\big)D_1.$$

Then $g = \delta + Bg$, and under Theorem 5 $\delta \ge 0$, so $g \ge \sum_{k=0}^{n} B^k\delta$ for every n; if $\rho(B) < 1$,

$$g = (I - B)^{-1}\delta. \tag{12}$$

If $q_I \le \bar{q}$ and a vector $\nu > 0$ satisfies $DF_I(\bar{q})\nu \le \kappa\nu$ with $\kappa < 1$, where $DF_I(v) = 2T\,\mathrm{diag}(D_1 v)D_1$, then

$$\|g\|_\nu \le \frac{\|\delta\|_\nu}{1-\kappa}, \qquad \|v\|_\nu = \max_x |v_x|/\nu_x. \tag{13}$$

Given boxes $0 \le \ell_J \le q_J \le v_J \le \mathbf{1}$ and $0 \le \ell_I \le q_I \le v_I \le \mathbf{1}$, put

$$\delta_- = \max\{0,\, T[(D_1\ell_J)^2 - D_J(v_J, v_J)]\}, \qquad \delta_+ = T[(D_1 v_J)^2 - D_J(\ell_J, \ell_J)],$$
$$B_- = T\,\mathrm{diag}(D_1(\ell_I + \ell_J))D_1, \qquad B_+ = T\,\mathrm{diag}(D_1(v_I + v_J))D_1.$$

If $B_+\nu \le \kappa_+\nu$ for some $\nu > 0$, $\kappa_+ < 1$, then

$$(I - B_-)^{-1}\delta_- \;\le\; g \;\le\; (I - B_+)^{-1}\delta_+. \tag{14}$$

Proof. Subtract $q_J = F_J(q_J)$ from $q_I = F_I(q_I)$ and insert $\pm F_I(q_J)$: $g = \delta + T[(D_1 q_I)^2 - (D_1 q_J)^2] = \delta + Bg$ by the difference of two squares. Iterating with $B \ge 0$, $\delta \ge 0$, $g \ge 0$ gives the partial-sum bound. Entrywise $0 \le B \le DF_I(q_I) \le DF_I(\bar{q})$, so the induced ν-norm of B is at most κ, the Neumann series converges to a nonnegative inverse of norm at most $(1-\kappa)^{-1}$, and (12)–(13) follow. The polynomial coefficients are nonnegative, so $\delta_- \le \delta \le \delta_+$ and $B_- \le B \le B_+$ entrywise; termwise comparison of Neumann series gives (14). $\square$

The factor $(1-\kappa)^{-1}$ is an explicit dependence on the distance from criticality. It does not mean that the gap diverges there: as criticality is approached the forcing δ decreases, and in the subcritical regime it vanishes. We prove this margin-dependent bound and make no claim of a uniform near-critical expansion. The premise concerns a Jacobian bound on a valid *enclosing box*; a Jacobian evaluated at an approximate numerical root would not suffice.

Proposition 10 (Limits of an independence approximation)**.** *For a mother* (a,r)*,*

$$\|D_J - D_I\|_{\mathrm{TV}} = 1 - \binom{2a}{a}\binom{2r}{r}4^{-(a+r)}. \tag{15}$$

If changing one allocation bit changes $f(Y)$ *by at most* L/N*, then*

$$\left|\mathrm{Cov}_x(f(Y), f(Z))\right| \le \frac{(a+r)L^2}{4N^2} \le \frac{L^2}{4N}. \tag{16}$$

If this holds for $f = q_J$ *with* $L = L_N$*, and* (13) *holds, then* $\|q_I - q_J\|_\nu \le \|T\mathbf{1}\|_\nu\, L_N^2/[4N(1-\kappa)]$*.*

Proof. With $p_{ij} = \mathbb{P}(I = i, J = j)$, the joint law puts mass p_{ij} on the pair $((i,j),(a-i,r-j))$ and nothing off this conservation graph, while the product law puts mass $p_{ij}p_{a-i,r-j} = p_{ij}^2$ there. Hence the distance is $1 - \sum p_{ij}^2$, and $\sum_i \binom{a}{i}^2 = \binom{2a}{a}$ gives (15). For (16), the Efron–Stein inequality [13] for a function of $a+r$ independent fair bits with bounded differences L/N gives $\mathrm{Var}\, f(Y) \le (a+r)L^2/(4N^2)$; the same bound holds for $f(Z)$ because Z has the same marginal law, and Cauchy–Schwarz concludes. The last assertion follows from $\delta \le T\mathbf{1}\cdot L_N^2/(4N)$ and (13). □

For an all-active mother the distance (15) is $1 - \binom{2N}{N}4^{-N} \sim 1 - (\pi N)^{-1/2} \to 1$: fractional concentration does not make the pair laws close, and no argument through small total variation can justify the independent comparator. What can be small is the covariance of a *regular* observable. A bounded regularity constant L_N, bounded $\|T\mathbf{1}\|_\nu$ and a stability margin bounded away from zero would give an $O(1/N)$ error. These are premises, not established uniform properties of this source: for the smooth readout of Section 5 the diagnostic constant $N \max|q_J(a+1,r) - q_J(a,r)|$ grows from 1.2 at $N = 2$ to 9.0 at $N = 32$, because the extinction probability varies rapidly near the symmetry line.

5 Certified source instances

Throughout this section the constants are symmetric, $w = 0.01$, $k = 1$, $h = 0.5$, $e = 0.01$, $b = 0.1$ (Table 1). The *step* readout has $d = 0.01$ when $a > r$ and $d = 0.3$ otherwise. The *smooth* readout is

$$d_{N,s}(a,r) = 0.01 + \frac{0.29}{1 + \exp(s(a-r)/N)}, \qquad s > 0. \tag{17}$$

Both are decreasing in $\preceq$, so Theorem 5 applies to every N and every $s > 0$. Smoothing changes intermediate hazards as well as removing the discontinuity, and neither map is fitted to data.

5.1 Two sites: exact risk response

Order the six states as UU, UR, RR, AU, AR, AA, that is $(0,0), (0,1), (0,2), (1,0), (1,1), (2,0)$, and use the step readout. Three certificates play distinct roles.

Coarse supersolutions. The rational vectors

$$\bar q_I = (.948115,\ .983103,\ .987779,\ .378814,\ .775706,\ .335634),$$
$$\bar q_J = (.94,\ .982,\ .987,\ .27,\ .735,\ .235)$$

satisfy $\mathcal{H}\bar q \ge d + bD(\bar q, \bar q)$ for their respective laws; the smallest independent residual is $435828959/(4\times 10^{13})$. By Lemma 1(ii), $q_I \le \bar q_I < 0.988\,\mathbf{1}$ and $q_J \le \bar q_J$. Eight iterations of F_I from zero, each rounded down to a multiple of 10^{-6}, give $q_I(AA) \ge 0.313291$, hence already $q_I(AA) - q_J(AA) > 0.078$.

Quantity	Value	Provenance and role
N	2 or 16 sites	Mathematical witnesses; not a measured locus size.
w, e	0.01 per time	Symmetric basal writing and *intrinsic* erasure; exploratory.
k, h	1, 0.5 per time	Feedback at unit mark fraction and recruited erasure; contact factor $1/N$.
b	0.1 per time	Exponential division clock; mean cycle 10 in the absence of death.
d	0.01 to 0.3 per time	Assumed functional mortality; not a fitted drug response.
s	$[7.999, 8.001]$; $[7, 9]$	Dimensionless slope; robustness intervals of the certificates.
m	$1, 2, 4$ phases	Same-mean Erlang clocks with phase rate $m/10$ (diagnostic).
$H,\ L$	300; 300, 2000 cells	Deadline and thresholds of Section 7.

Table 1: Parameters of the worked examples. Time is a declared model unit; the ratios $k/b = 10$, $d_{\max}/b = 3$, $d_{\min}/b = 0.1$ are what matter, and rescaling time cannot calibrate them. No entry is a measured range.

Contraction weight. On the whole box $[0, \bar{q}_I]$,

$$\nu = (2.148,\ 2.011,\ 1.997,\ 4.880,\ 2.882,\ 4.895) \quad \text{satisfies} \quad DF_I(\bar{q}_I)\,\nu \le 0.796\,\nu$$

by exact evaluation, so (13) holds with $\kappa = 0.796$.

Tight response. Two hundred and fifty downward-rounded iterations (denominator 10^{12}) give lower boxes; adding a small multiple of the positive vector $(I - DF(\ell))^{-1}\mathbf{1}$ gives vectors whose supersolution residuals are checked exactly. The resulting boxes have width below 5×10^{-6}, for instance $0.2207216 < q_J(AA) < 0.2207253$ and $0.3351447 < q_I(AA) < 0.3351497$, and (14) yields

$$0.1144144 < q_I(AA) - q_J(AA) < 0.1144331. \tag{18}$$

The lower response bound is strictly positive at all six states, from 0.00246 at RR to 0.12347 at AU, in agreement with Proposition 7. Neighbouring-state increments of q_J are bounded from the same boxes, so the regularity premise of Proposition 10 is checkable at fixed N without being extrapolated.

5.2 Sixteen sites: readout slope and preparation

Proposition 11 (Smooth readout, $N = 16$)**.** *For $N = 16$, the constants of Table 1, the exponential clock and the hazard* (17)*:*

(a) for every $s \in [7.999, 8.001]$,

$$0.48558095 < q_J(9,7) < 0.48561304, \qquad 0.49988773 < q_I(9,7) < 0.49991904,$$
$$q_I(9,7) - q_J(9,7) > 0.01427, \qquad 0 \le q_I(16,0) - q_J(16,0) < 0.000122,$$
$$\max_x q_I(x) < 0.997;$$

(b) for every $s \in [7, 9]$,

$$q_I(9,7) - q_J(9,7) > 0.0128, \qquad 0 \le q_I(16,0) - q_J(16,0) < 0.00036, \qquad \max_x q_I(x) < 0.997.$$

Proof. There are 153 states. Fix a slope cell $[s_0, s_1]$. At a fixed state the hazard (17) is monotone in s, decreasing if $a > r$ and increasing if $a < r$, so on the cell it lies between the smaller and the

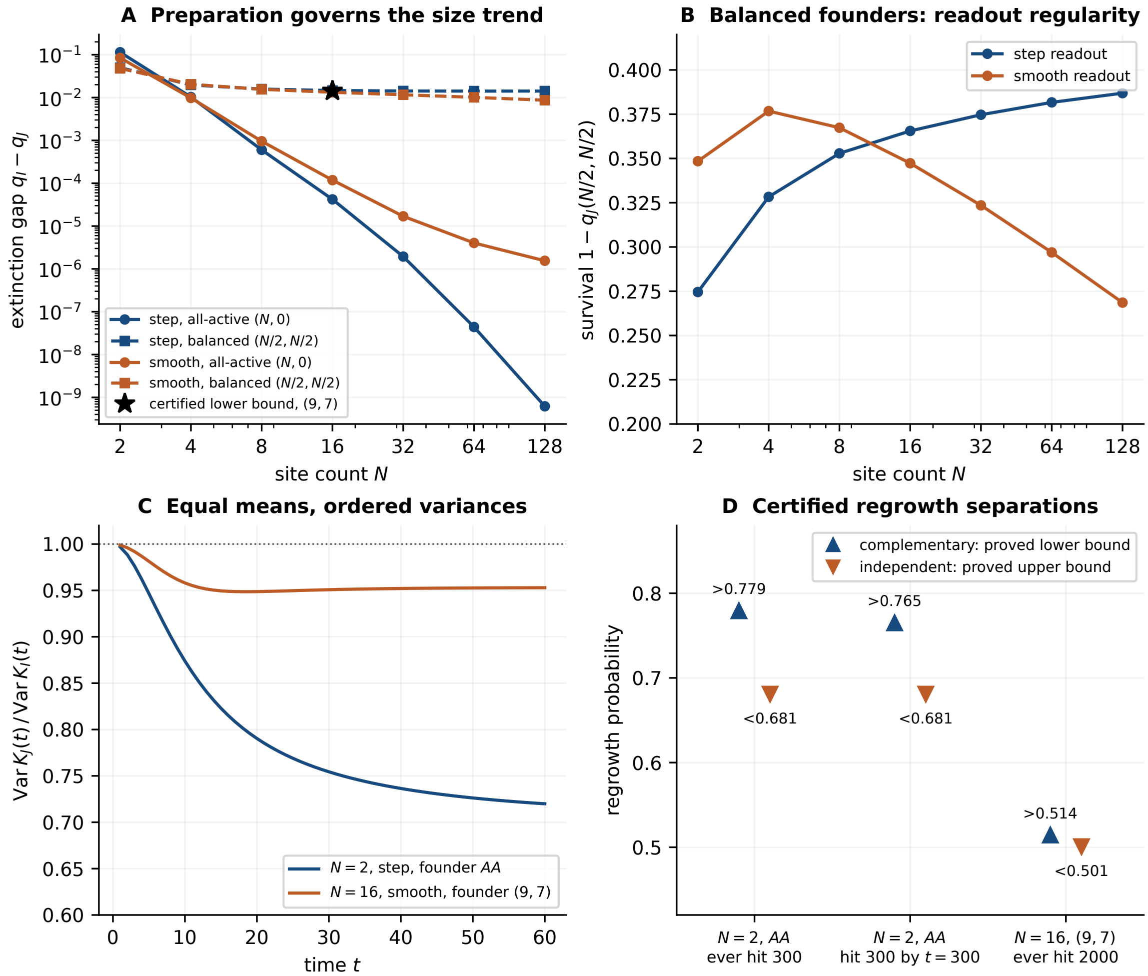

Figure 1: A: extinction gap against site count for all-active and balanced founders under both readouts (floating-point diagnostics); the star is the certified lower bound of Proposition 11(a) for the founder $(9, 7)$. B: survival of a balanced founder under complementary allocation; under the smooth readout it decays, as Theorem 12 requires asymptotically, while under the step readout it does not (diagnostic; the limit for the step readout is open). C: ratio of population variances, which Corollary 6 bounds by one (diagnostic values). D: proved bounds of Theorem 15, not estimates; the independent upper bound for 300 cells is an eventual bound and therefore also bounds the deadline probability.

larger of its two endpoint values. These endpoint values are enclosed by bounding e^y, $0 \le y \le 9$, below by its degree-120 Taylor polynomial and above by that polynomial plus a geometric bound on the tail, inverting for negative arguments, and rounding outward at denominator 10^{14}. This gives rational vectors $d_- \le d_{N,s} \le d_+$ for all s in the cell. Note that d_- and d_+ are statewise envelopes and neither is the hazard at an endpoint slope: the hazards at two different slopes are not ordered, and extinction need not be monotone in s.

A pointwise coupling of death clocks shows that in either model q is nondecreasing in the hazard vector. For d_+ we exhibit rational vectors below one and verify every supersolution inequality

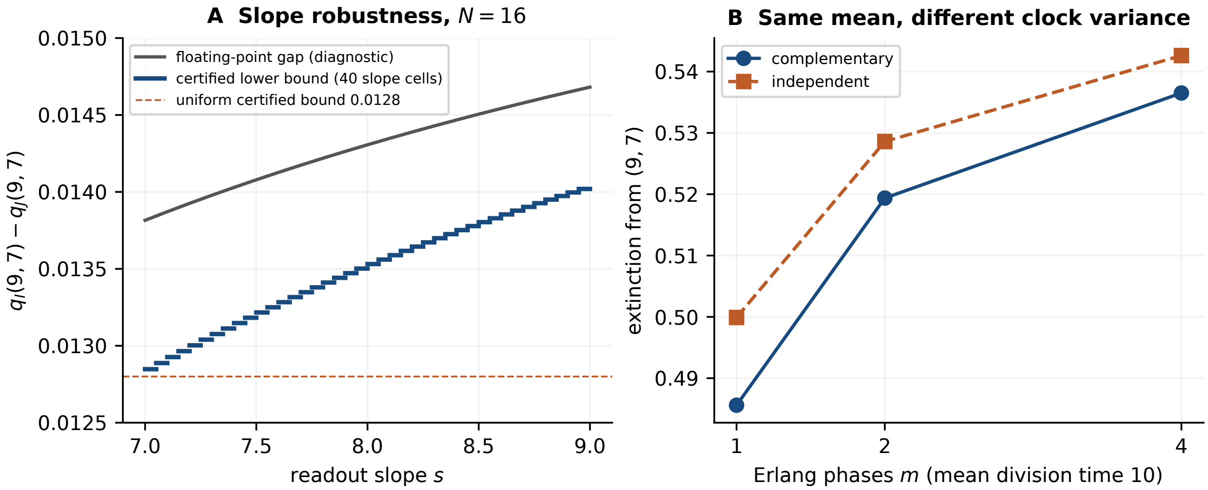

Figure 2: A: certified lower bounds on $q_I(9,7) - q_J(9,7)$ on each of the 40 slope cells of Proposition 11(b), with the floating-point gap for comparison. B: extinction from $(9,7)$ at $N = 16$, $s = 8$ for Erlang division clocks with the same mean (diagnostic); Theorem 5 proves the sign for every m.

$\mathcal{H}\bar{q} \geq d_+ + bD(\bar{q}, \bar{q})$ exactly. For d_- we iterate from zero the positive map

$$G(z)_x = \frac{d_x + \sum_{y \neq x} Q_{xy} z_y + b\, D(z,z)_x}{d_x + b + \sum_{y \neq x} Q_{xy}}, \tag{19}$$

which has the same fixed points as (5), is monotone, and satisfies $G(z) \leq z'$ whenever $z \leq z'$ and z' is a fixed point; rounding all nonnegative coefficients and iterates downward at denominator 10^{14} keeps every iterate below the least fixed point by induction. For (a) a single cell is used; 466 joint and 474 independent iterations give the displayed bounds, and the largest box width over all 153 coordinates and both models is below 3.58×10^{-5}. For (b) the interval is divided into 40 cells of width 0.05; the certified lower bound on the gap at $(9,7)$ increases from 0.012847 on $[7, 7.05]$ to 0.014019 on $[8.95, 9]$ (Figure 2A). All comparisons are exact rational comparisons, and no solver tolerance is used. □

Enlarging the slope interval costs accuracy in a transparent way: a single cell $[7,9]$ gives an envelope so wide that the resulting sufficient bound is negative, although the true gap stays between 0.0138 and 0.0147 (diagnostic). A failed sufficient inequality is not evidence about the dynamics; partitioning the interval repairs it.

The proposition is a source-specific obstruction to promoting an interior estimate to a guarantee that is uniform over preparations: at the same N an all-active error below 1.3×10^{-4} coexists with a near-balanced error above 1.4×10^{-2}. Figure 1A shows the corresponding floating-point trends for $N = 2, \ldots, 128$. From the all-active founder the gap falls to 6×10^{-10} (step) and 1.5×10^{-6} (smooth) at $N = 128$. From the balanced founder $(N/2, N/2)$ it stays between 0.008 and 0.016 for $8 \leq N \leq 128$. A third preparation, obtained by holding an all-active cell for one time unit with division and death suspended, has law $e^{Q^\mathsf{T}} \delta_{(N,0)}$ and gaps within one percent of the all-active family for $N \leq 32$; this is a controlled pre-division hold, not a treatment model. The certified preparation of Proposition 11 is $(9,7)$, the state of largest gap at $N = 16$ under the smooth readout; it should not be confused with $(10,6)$ or $(8,8)$.

5.3 Division-time variance

For m phases of rate $m/10$ let $V = (m/10)\big((m/10)I + \operatorname{diag}(d) - Q\big)^{-1}$; successive phase conditioning gives the offspring map $\mathbf{1} - V^m\mathbf{1} + V^m D(v, v)$. At $N = 16$, $s = 8$ and founder $(9, 7)$, diagnostic values of (q_J, q_I) are $(0.48560, 0.49990)$, $(0.51938, 0.52860)$ and $(0.53651, 0.54262)$ for $m = 1, 2, 4$, with gaps 0.01431, 0.00922, 0.00611 and Malthusian growth rates 0.0849, 0.0709, 0.0644 (Figure 2B). Equal mean division time preserves neither molecular recovery nor risk. Theorem 5 proves the sign for all these clocks, but the quantitative certificate of Proposition 11 is for $m = 1$.

6 Large site counts: balanced founders under a smooth readout

Figure 1A might suggest that the balanced-founder gap tends to a positive limit. For a fixed smooth readout it does not. The reason is that the symmetry line $a = r$ is invariant for the mean-field dynamics, a founder near it stays near it for a time of order $\log N$, and on the line the smooth hazard of our example makes the population subcritical.

Assume symmetric constants $w_A = w_R = w$, $k_A = k_R = k$, $e_A = e_R = e$, $h_A = h_R = h$, and let the hazard be a function of the contrast,

$$d_N(a, r) = \varphi(c(a, r)), \qquad c(a, r) = \frac{a - r}{N} \in [-1, 1],$$

where $\varphi \geq 0$ is Lipschitz with constant ℓ and satisfies $\theta := \varphi(0) - b > 0$. Put

$$\psi = \varphi(0) - \inf \varphi, \qquad \Lambda = \max\{0,\, k - e - \tfrac{3}{4}b\}, \qquad \sigma^2 = \max\{2w + k,\ e + \tfrac{h}{2}\} + \tfrac{b}{2}.$$

Theorem 12 (Balanced founders die out as $N \to \infty$)**.** *Under these assumptions, for either daughter law, every N, every founder $x \in S_N$ and every $t \geq 0$,*

$$1 - q_D(x) \leq \mathbb{E}_x K(t) \leq e^{-\theta t}\Big[1 + \ell\, t\, e^{(\psi+\Lambda)t}\big(c(x)^2 + \sigma^2 t/N\big)^{1/2}\Big]. \tag{20}$$

Consequently, if $|c(x_N)| \leq C_0 N^{-1/2}$ and $\psi + \Lambda > 0$, then with $\gamma = \theta/[2(\psi + \Lambda)]$ and $t_N = \log N/[2(\psi + \Lambda)]$,

$$0 \leq q_I(x_N) - q_J(x_N) \leq 1 - q_J(x_N) \leq N^{-\gamma}\Big[1 + \ell\, t_N\big(C_0^2 + \sigma^2 t_N\big)^{1/2}\Big] \xrightarrow[N\to\infty]{} 0, \tag{21}$$

where the first inequality requires φ to be nonincreasing (Theorem 5) and the others do not.

Proof. Survival forever implies $K(t) \geq 1$, so $1 - q_D(x) \leq \mathbb{P}_x(K(t) \geq 1) \leq \mathbb{E}_x K(t) = (e^{tA}\mathbf{1})_x$, which by Lemma 1(iv) is the same in the two models. Write $A = \mathcal{L} + \operatorname{diag}(b - d)$, where

$$\mathcal{L} = Q + 2b\,(D_1 - I)$$

is the generator of a conservative Markov chain X on S_N: the molecular chain together with a halving $x \mapsto Y$ at rate $2b$. By the Feynman–Kac formula for finite chains (the many-to-one formula [4, 16]),

$$\mathbb{E}_x K(t) = \mathbb{E}_x \exp\Big(\int_0^t (b - d)(X_s)\, ds\Big) = e^{-\theta t} n_x(t), \qquad n(t) = e^{t(\mathcal{L} + \operatorname{diag}\chi)}\mathbf{1}, \quad \chi = \varphi(0) - \varphi \circ c.$$

Here $\chi \leq \ell|c|$ and $\chi \leq \psi$. Since $e^{t\mathcal{L}}\mathbf{1} = \mathbf{1}$, the Duhamel formula gives

$$n_x(t) = 1 + \int_0^t \mathbb{E}_x\Big[\chi(X_s)\exp\Big(\int_0^s \chi(X_\tau)\, d\tau\Big)\Big]ds \leq 1 + \ell\int_0^t e^{\psi s}\,\mathbb{E}_x|c(X_s)|\, ds. \tag{22}$$

It remains to bound the second moment of the contrast along X. Under Q the contrast jumps by $\pm 1/N$. With symmetric constants the difference of the drifts of a and r is $u\,k(a-r)/N - e(a-r)$, the recruited-erasure terms $h\,ar/N$ cancelling, so the drift of c is $(k\,u/N - e)\,c$. The total jump rate is

$$u\big(2w + k\tfrac{a+r}{N}\big) + e(a+r) + 2h\tfrac{ar}{N} \le u(2w+k) + (a+r)\big(e + \tfrac{h}{2}\big) \le N \max\{2w+k,\, e+\tfrac{h}{2}\},$$

using $ar \le (a+r)N/4$. At a halving, $I - J$ has mean $(a-r)/2$ and variance $(a+r)/4$, so $\mathbb{E}_x c(Y)^2 = c^2/4 + (a+r)/(4N^2)$. Therefore

$$\mathcal{L}c^2 = 2\big(k\tfrac{u}{N} - e\big)c^2 + \frac{\text{jump rate}}{N^2} + 2b\Big(\frac{a+r}{4N^2} - \frac{3}{4}c^2\Big) \le 2\Lambda\, c^2 + \frac{\sigma^2}{N}.$$

By Dynkin's formula and Gronwall's inequality, $\mathbb{E}_x c(X_s)^2 \le e^{2\Lambda s}\big(c(x)^2 + \sigma^2 s/N\big)$. Insert $\mathbb{E}|c| \le (\mathbb{E}c^2)^{1/2}$ into (22) and bound the integrand by its value at $s = t$ to obtain (20). For (21) take $t = t_N$, so that $e^{(\psi+\Lambda)t_N} = N^{1/2}$ and $e^{-\theta t_N} = N^{-\gamma}$. □

For the smooth readout (17) with $s = 8$ and the constants of Table 1 one has $\varphi(0) = 0.155$, $\theta = 0.055$, $\ell = 0.29s/4 = 0.58$, $\psi < 0.145$, $\Lambda = 0.915$, $\sigma^2 = 1.07$ and $\gamma = 0.055/2.12 \approx 0.0259$. The theorem covers the balanced founders $(N/2, N/2)$, the square-root family $(\lfloor (N+\sqrt{N})/2\rfloor,\, N - \lfloor (N+\sqrt{N})/2\rfloor)$, and any fixed compact slope interval with the constants taken at the largest slope.

Remark 13 (What the limit does and does not say). (i) The bound (21) is asymptotic and the rate is slow; with the constants above it is below one only for astronomically large N. This is in the nature of the problem and not only of the proof: a founder near the line must survive a subcritical phase whose length grows like $\log N$ before its contrast is macroscopic, which suggests a small negative power of N. The diagnostics agree: under the smooth readout the survival probability of $(N/2, N/2)$ falls from 0.377 at $N = 4$ to 0.268 at $N = 128$ (Figure 1B) and the balanced gap falls from 0.0156 at $N = 8$ to 0.0087 at $N = 128$. A gap of order 10^{-2} at moderate N and a zero limit are therefore compatible, and the certificates of Proposition 11 are finite-size statements of independent interest.

(ii) The argument uses only the common mean, so it cannot distinguish the two models; it shows that both survival probabilities vanish.

(iii) Continuity of φ at zero is essential. For the step readout arbitrarily small contrasts of either sign change the hazard by 0.29, the bound $\chi \le \ell|c|$ fails, and in the diagnostics the balanced survival probability increases with N while the gap stays near 0.014. The limit for a discontinuous readout, or for a slope growing with N, requires a fluctuation analysis that we do not attempt. Regularity of the biological readout thus decides the asymptotic question.

(iv) If $\varphi(0) < b$ the population is supercritical on the symmetry line and the theorem does not apply.

7 Substantial regrowth

Let τ_0 be the extinction time and τ_L the first time the population reaches $L \ge 2$ cells, from one founder, and let $p_L(x) = \mathbb{P}_x(\tau_L < \tau_0)$.

Theorem 14 (Threshold identity)**.** *Assume $d(x) \ge d_* > 0$ on S_N. Then $\tau_0 \wedge \tau_L < \infty$ almost surely and*

$$1 - q_x = \mathbb{E}_x\Big[\Big(1 - \prod_{j=1}^{L} q_{X_j(\tau_L)}\Big)\mathbf{1}_{\{\tau_L < \tau_0\}}\Big], \tag{23}$$

where $X_1(\tau_L), \dots, X_L(\tau_L)$ *are the types present at* τ_L. *If* $q_{\max} = \max_x q_x < 1$,

$$1 - q_x \;\le\; p_L(x) \;\le\; \frac{1-q_x}{1-q_{\max}^L}. \tag{24}$$

For $z \in (0,1)$ *and a deadline* $H > 0$,

$$\mathbb{P}_x(\tau_L \le H) \;\ge\; 1 - u_x(H;0) - z^{-(L-1)}\big[u_x(H;z) - u_x(H;0)\big]. \tag{25}$$

Proof. Before $\tau_0 \wedge \tau_L$ the process lives on the finite set of configurations with fewer than L cells. From each of them, fewer than L successive deaths within one time unit lead to extinction with a probability bounded below uniformly, since $d \ge d_*$ and all other rates are bounded. The strong Markov property gives a geometric tail for $\tau_0 \wedge \tau_L$. Divisions increase the size by exactly one, so the population at τ_L has exactly L cells, and given their types its extinction probability is the product of their single-founder extinction probabilities. Extinction without hitting L has probability $1 - p_L$, so $q_x = 1 - p_L(x) + \mathbb{E}_x[\prod_j q_{X_j(\tau_L)}; \tau_L < \tau_0]$, which is (23). Bounding the bracket between $1 - q_{\max}^L$ and 1 gives (24). Finally $z^k \ge z^{L-1}$ for $1 \le k < L$ gives $\mathbb{P}_x(1 \le K(H) < L) \le z^{-(L-1)}[u_x(H;z) - u_x(H;0)]$; subtract this from $1 - \mathbb{P}_x(K(H) = 0)$ and use $\{K(H) \ge L\} \subseteq \{\tau_L \le H\}$. □

The inclusion at the end is not an equality: (25) uses the event $K(H) \ge L$ as a lower bound for first passage. After collecting terms the coefficient of $u_x(H;0)$ is $z^{-(L-1)} - 1 > 0$, so a certificate needs a *lower* enclosure of $u_x(H;0)$ and an *upper* enclosure of $u_x(H;z)$.

Theorem 15 (Certified risk separations)**.** (a) *For the two-site step-readout source and one AA founder,*

$$p_{300}^J > 0.779, \qquad p_{300}^I < 0.681, \qquad \mathbb{P}_J(\tau_{300} \le 300) > 0.765,$$

so the probability of ever reaching 300 *cells differs by more than* 0.098 *and the probability of reaching* 300 *cells by time* 300 *by more than* 0.084.

(b) *For the* 16*-site smooth source and one* $(9,7)$ *founder, uniformly in* $s \in [7.999, 8.001]$,

$$p_{2000}^J > 0.51438, \qquad p_{2000}^I < 0.50138, \qquad p_{2000}^J - p_{2000}^I > 0.013,$$

and uniformly in $s \in [7,9]$, $p_{2000}^J - p_{2000}^I > 0.0122$.

Proof. (a) Insert the tight boxes of Section 5, $q_J(AA) < 0.2207253$, $q_I(AA) > 0.3351447$ and $\max_x q_I(x) < 0.9875805$, into (24): $p_{300}^J \ge 1 - q_J(AA) > 0.779$ and $p_{300}^I \le 0.6648553/(1 - 0.9875805^{300}) < 0.681$. For the deadline take $z = 299/300$. Appendix A proves $u_{J,AA}(300;0) > 0.2207177$ and $u_{J,AA}(300;299/300) < 0.2259640$, and (25) gives

$$\mathbb{P}_J(\tau_{300} \le 300) \ge 1 - 0.2207177 - (300/299)^{299}\,(0.2259640 - 0.2207177) > 0.765.$$

The eventual bound p_{300}^I also bounds $\mathbb{P}_I(\tau_{300} \le 300)$. With the coarse constants $q_J(AA) \le 0.235$, $q_I(AA) \ge 0.313291$, $q_{I,\max} < 0.988$, $u_J(300;0) > 0.22$ and $u_J(300;z) < 0.227$ the same computation gives the weaker values 0.765, 0.706 and 0.760; these are the versions whose arithmetic is checked in Lean.

(b) Use Proposition 11: in case (a) of that proposition, $q_J(9,7) < 0.48562$, $q_I(9,7) > 0.49988$, $q_{I,\max} < 0.997$ and $0.50012/(1 - 0.997^{2000}) < 0.50138$. For $[7,9]$ the same computation is carried out exactly on each of the 40 slope cells with the boxes of that cell, and the smallest resulting gap exceeds 0.0122. □

These are unconditional probabilities for one living founder of the declared type at time zero. There is no preceding treatment phase and no selection on later survival; the 16-site statement is eventual and no finite deadline is asserted for it. Thresholds of hundreds or thousands of cells establish a substantial clonal-count event within a dilute branching model; they do not calibrate a relapse and do not validate independent growth at such densities.

8 What observations can identify

8.1 Paired sister clones

Enrol sister pairs prospectively at birth from a specified mother preparation, culture the two daughters separately, and record whether each clone is extinct at follow-up time H. Given the daughter states, the two indicators are independent with probabilities $f_H(Y), f_H(Z)$, where $f_H(x) = u_x(H;0)$. Hence, for a fixed complete mother state x,

$$\mathrm{Cov}\big(\text{two extinction indicators} \mid X = x\big) = \mathrm{Cov}_x\big(f_H(Y), f_H(Z)\big) \leq 0$$

under complementary allocation, by Lemmas 3 and 4, and it is zero under the independent comparator. This measures continuation fate and not a thresholded molecular reporter.

Pooling over heterogeneous mothers can reverse the sign. For an imperfectly measured parental stratum W, total covariance gives

$$C_W = \mathbb{E}[c_X \mid W] + \mathrm{Var}(\mu_X \mid W), \qquad \mu_X = D_1 f_H(X), \quad c_X = \mathrm{Cov}_X(f_H(Y), f_H(Z)).$$

For example, for newborn "active" indicators $\mathbf{1}_{\{a>r\}}$ in the two-site model, an AA mother has covariance $1/2 - (3/4)^2 = -1/16$ and an RR mother has covariance zero, but an equal mixture of AA and RR mothers has pooled covariance $1/4 - (3/8)^2 = +7/64$. If μ_X has range at most Δ within the stratum, then $\mathrm{Var}(\mu_X \mid W) \leq \Delta^2/4$. If each binary fate is misclassified with probability at most ζ, the joint probability and the product of marginals each change by at most 2ζ. If a fraction at most ϵ of *prospectively enrolled* pairs is missing, each of the three probabilities changes by at most ϵ. The average within-mother covariance therefore lies in

$$\big[\, C_{\mathrm{obs}} - \Delta^2/4 - 4\zeta - 3\epsilon, \quad C_{\mathrm{obs}} + 4\zeta + 3\epsilon \,\big]. \tag{26}$$

This is a deterministic sensitivity interval. A sampling radius is separate: for n independent enrolled pairs, Hoeffding's inequality [18] applied to the two marginal frequencies and the joint frequency, with a union bound, puts the empirical covariance within $3\sqrt{\log(6/\alpha)/(2n)}$ of its expectation with probability at least $1-\alpha$. The sampling unit is the independent pair and not the descendant cell. Without justified bounds on Δ and on missingness a pooled covariance does not identify partition as its cause; unrecorded enrolment cannot be repaired afterwards, and sampling extant cells rather than divisions introduces a growth bias that this analysis does not cover.

8.2 Clone counts under independent capture

High-throughput lineage barcoding [15, 24, 5, 30] reports clone sizes after incomplete sampling and usually does not identify the two immediate sister branches. Suppose that a clone has $K(t)$ living cells and each is captured independently with a known probability p, giving a count $C \mid K \sim \mathrm{Bin}(K, p)$.

Corollary 16 (Thinned clone counts). *For either model,* $\mathbb{E}_x z^C = u_x(t; 1-p+pz)$, $\mathbb{E}_x C = p\, m_x(t)$ *and* $\mathrm{Var}_x\, C = p^2\, \mathrm{Var}_x\, K(t) + p(1-p) m_x(t)$. *Under the hypotheses of Theorem 5, for a common founder law and a common* p,

$$\mathbb{P}(C_J = 0) \leq \mathbb{P}(C_I = 0), \qquad \mathbb{E}C_J = \mathbb{E}C_I, \qquad \mathrm{Var}\, C_J \leq \mathrm{Var}\, C_I.$$

Proof. $\mathbb{E}[z^C \mid K] = (1-p+pz)^K$; then apply Theorem 5 at $z' = 1-p \in [0,1]$ and Corollary 6. $\square$

Nondetection is not extinction: $\mathbb{P}(C=0) = u(t; 1-p) \geq u(t; 0)$. Independent thinning is an explicit assumption. Sequencing to a fixed total depth induces dependence between clones, and state-dependent capture, barcode collisions or silencing and the passage from reads to cells need their own observation model. Clone-size dispersion is affected by many mechanisms, so Corollary 16 is a prediction for matched founder preparations and mean growth; it does not make barcode dispersion a substitute for the paired-clone covariance, which is the statistic tied directly to the mechanism and which requires the mother preparation and the two daughter branches to be identified, for example by imaging, indexed isolation or sufficiently resolved lineage recording.

9 Scope, formalization and reproducibility

Chain of results. Molecular order and conserved allocation fix the sign (Theorem 5), also for the variance (Corollary 6) and strictly so under irreducibility (Proposition 7). Continuation-risk covariance and a stable population response fix the magnitude (Theorem 9). Exact boxes give finite-size contrasts between preparations (Proposition 11) and consequences at large population thresholds (Theorem 15), while Theorem 12 shows that, for a smooth readout, the balanced contrast is a finite-size effect. The magnitude is not universal: it depends strongly on preparation, on the regularity of the readout and on the division clock.

What is not claimed. There is no unconditional advantage of conservation: the theorem needs the stated order, a hazard monotone in it and a division clock independent of the molecular state, and Section 3.3 shows where the proof breaks otherwise. Nonmonotone hazards and fitness trade-offs are outside the theorem. Changing a symmetric erasure rate removes marks of both kinds, so no monotonicity in that parameter follows. The source supplies no empirical locus size, drug hazard or competition law; hazards, site counts and clocks are exploratory, and no clinical calibration is implied.

Formalization boundary. A Lean 4/mathlib development (`PhenotypeMemory`, fourteen modules, compiled with warnings as errors and without `sorry`; its only axioms are `propext`, `Classical.choice` and `Quot.sound`) checks: normalization of the partition weights for all (a, r); the two-site source formulas; the supersolution inequalities for $\bar{q}_I$ and, uniformly for $e \in [0.0099, 0.0101]$, for $\bar{q}_J$; the bound $\bar{q}_I < 0.988$; the eight downward iterations and $q_I(AA) > 0.31$; the weighted Jacobian inequality with $\kappa = 0.796$; reusable algebra for the state order, the one-bit conditional-covariance step of Lemma 4, the response identity and the threshold inequalities; and the coarse arithmetic $0.686709/(1 - 0.988^{300}) < 0.706$ and $1 - 0.22 - (300/299)^{299}(0.227 - 0.22) > 0.760$. The root declaration is

`PhenotypeMemory.continuation_finite_certificate`.

Everything else is conventional: the chronological construction, the couplings, the induction over allocation bits, the comparison theorem for the generating functions, Corollary 6, Propositions 7 and 8, Theorem 12, the strong Markov argument, the soundness of the 153-state rational checker and of the validated integrator, and the tight two-site boxes. Exact computation is not presented as an end-to-end formal probability theorem.

Reproducibility. The source archive contains the scripts that produce every number: `slope_certificate.py` (Proposition 11; about two seconds per slope cell), `deadline_validated.py` (Appendix A; about eighty seconds), `check_paper.py` (an independent exact replay of all two-site constants, the threshold arithmetic, the constants of Theorem 12 and the small exact examples), `diagnostics.py` and `make_figures.py` (floating-point figures). Rational boxes are stored as exact fractions, and the printed decimals are checked against them.

A Validated microscopic generating functions

Write the six-state joint vector field (6) as $\Phi(v) = c_0 + \mathsf{L}v + \mathsf{P}(v, v)$ with rational coefficients, where $c_0 = d$, $\mathsf{L} = Q - \mathrm{diag}(d) - bI$ and $\mathsf{P} = bD_J$. Here $\|Q\|_\infty = 1.06$, $d \le 0.3$, $b = 0.1$, so in the complex ball $\|v\|_\infty \le 2$,

$$\|\Phi(v)\|_\infty \le 1.06 \cdot 2 + 0.3\,(1 + 2) + 0.1\,(4 + 2) = 3.62 < 4.$$

Local analyticity. Let $v_0 \in [0, 1]^6$. The ball of radius one about v_0 lies in $\{\|v\|_\infty \le 2\}$. By the Cauchy–Picard theorem for analytic vector fields the solution through v_0 extends analytically to complex times as long as it stays in that ball, and along any ray it cannot leave the ball before time $1/3.62 > 1/4$. Thus it is analytic and bounded by 2 on $|t| \le 1/4$, Cauchy's estimate bounds its jth Taylor coefficient by $2 \cdot 4^j$, and with step $\Delta = 1/50$ the truncation after degree 20 has error at most

$$\eta = \frac{2\,(2/25)^{21}}{1 - 2/25} < 2.01 \times 10^{-23}.$$

Arithmetic. The Taylor coefficients obey the exact recurrence $(j + 1)c_{j+1} = \mathsf{L}c_j + \mathbf{1}_{\{j=0\}}c_0 + \sum_{i=0}^{j} \mathsf{P}(c_i, c_{j-i})$. Every operation is enclosed by integer intervals at denominator 10^{40}, the polynomial is evaluated by an interval Horner scheme, and the new centre is the rounded midpoint projected coordinatewise to $[0, 1]$. Projection does not increase the distance to the exact solution, which lies in the cube. The largest distance from the new centre to an endpoint of the enclosure, over all 15,000 steps and both solves, is $\varepsilon_{\mathrm{rd}} = 10^{-40}$ (one unit of the last place); this number is recorded by the code.

Propagation. On the real cube the Jacobian of Φ has nonnegative off-diagonal entries and row sums $-d_x - b + b\sum_y \partial_y D_J(v, v)_x \le -d_x + b \le 0.09$, so its logarithmic ∞-norm is at most 0.09 [36], and two solutions that remain in the cube, which is forward invariant, separate by at most a factor $e^{0.09t}$. The usual telescoping argument for one-step methods [28] bounds the global error at $t = 300$ by

$$e^{27}\{15000\,(\eta + \varepsilon_{\mathrm{rd}}) + 10^{-40}\} < 1.8046 \times 10^{-7},$$

where $e^{27} < 6 \times 10^{11}$ is proved by a rational degree-200 Taylor sum with a geometric tail. Almost all of the budget is the analytic truncation term; the arithmetic term contributes less than 10^{-23}.

initial value z	lower bound for $u_{J,AA}(300; z)$	upper bound for $u_{J,AA}(300; z)$
0	0.2207177851	0.2207181461
299/300	0.2259636331	0.2259639941

The integrator's coefficients are checked independently against the rational source and partition formulas. No floating-point solver tolerance enters an enclosure. The enclosure for $z = 0$ is consistent with $u_{J,AA}(300; 0) \le q_J(AA) < 0.2207253$.

B Certificate data and proof dependencies

The files `data/slope_certificate_7.999_8.001_1.json` and `data/slope_certificate_7_9_40.json` contain, per slope cell, the rational risk boxes at $(9, 7)$, the upper bounds on the all-active gap and on max q_I, iteration counts and smallest supersolution residuals; the single-cell file also contains all 153 hazard brackets and both full 153-coordinate boxes. `data/deadline_validated.json` holds the enclosures of Appendix A as fractions, `data/paper_checks.json` the replayed two-site constants, and `data/diagnostics.json` the floating-point diagnostics behind the figures.

Conclusion	Ingredients
Direction, all N	Monotone molecular coupling; ordered allocation; Harris inequality on reversed bits; monotone iteration; Kamke comparison.
Variance ordering	Second factorial moment equation; increasing mean; same association inequality; nonnegative semigroup.
Strictness	Irreducibility; classical extinction criterion; one-bit covariance bound; response identity.
Magnitude	Resolvent; covariance forcing; difference of squares; weighted contraction; rational boxes.
Smooth $N = 16$ contrast	Rational exponential bounds; statewise slope envelopes; hazard coupling; lower iteration; supersolutions.
Large-N limit	Common mean; Feynman–Kac; Duhamel; second moment of the contrast; Gronwall.
Regrowth	Almost-sure stopping; strong Markov; extinction boxes; validated generating functions for the deadline.
Observation	Conditional independence; total covariance; range–variance bound; Hoeffding; binomial thinning.

References

[1] C. Alabert and A. Groth. Chromatin replication and epigenome maintenance. *Nat. Rev. Mol. Cell Biol.*, 13(3):153–167, 2012.

[2] K. B. Athreya and P. E. Ney. *Branching Processes*, volume 196 of *Grundlehren der mathematischen Wissenschaften*. Springer, Berlin, 1972.

[3] V. Bansaye. Proliferating parasites in dividing cells: Kimmel's branching model revisited. *Ann. Appl. Probab.*, 18(3):967–996, 2008.

[4] V. Bansaye, J.-F. Delmas, L. Marsalle, and V. C. Tran. Limit theorems for Markov processes indexed by continuous time Galton–Watson trees. *Ann. Appl. Probab.*, 21(6):2263–2314, 2011.

[5] B. A. Biddy, W. Kong, K. Kamimoto, C. Guo, S. E. Waye, T. Sun, and S. A. Morris. Single-cell mapping of lineage and identity in direct reprogramming. *Nature*, 564:219–224, 2018.

[6] S. Bruno, F. A. Campos, Y. Fu, D. Del Vecchio, and R. J. Williams. Analysis of singularly perturbed stochastic chemical reaction networks motivated by applications to epigenetic cell memory. *SIAM J. Appl. Dyn. Syst.*, 23(4):2695–2731, 2024.

[7] S. Bruno, R. J. Williams, and D. Del Vecchio. Epigenetic cell memory: the gene's inner chromatin modification circuit. *PLoS Comput. Biol.*, 18(4):e1009961, 2022.

[8] T. Cassidy, D. Nichol, M. Robertson-Tessi, M. Craig, and A. R. A. Anderson. The role of memory in non-genetic inheritance and its impact on cancer treatment resistance. *PLoS Comput. Biol.*, 17(8):e1009348, 2021.

[9] K. S. Crump and C. J. Mode. An age-dependent branching process with correlations among sister cells. *J. Appl. Probab.*, 6(1):205–210, 1969.

[10] L. de Moura and S. Ullrich. The Lean 4 theorem prover and programming language. In *Automated Deduction – CADE 28*, volume 12699 of *Lecture Notes in Comput. Sci.*, pages 625–635. Springer, Cham, 2021.

[11] I. B. Dodd, M. A. Micheelsen, K. Sneppen, and G. Thon. Theoretical analysis of epigenetic cell memory by nucleosome modification. *Cell*, 129(4):813–822, 2007.

[12] K. R. Duffy and V. G. Subramanian. On the impact of correlation between collaterally consanguineous cells on lymphocyte population dynamics. *J. Math. Biol.*, 59(2):255–285, 2009.

[13] B. Efron and C. Stein. The jackknife estimate of variance. *Ann. Statist.*, 9(3):586–596, 1981.

[14] J. D. Esary, F. Proschan, and D. W. Walkup. Association of random variables, with applications. *Ann. Math. Statist.*, 38(5):1466–1474, 1967.

[15] C. Gutierrez, A. M. Al'Khafaji, E. Brenner, K. E. Johnson, S. H. Gohil, Z. Lin, et al. Multifunctional barcoding with ClonMapper enables high-resolution study of clonal dynamics during tumor evolution and treatment. *Nat. Cancer*, 2(7):758–772, 2021.

[16] R. Hardy and S. C. Harris. A spine approach to branching diffusions with applications to L^p-convergence of martingales. In *Séminaire de Probabilités XLII*, volume 1979 of *Lecture Notes in Math.*, pages 281–330. Springer, Berlin, 2009.

[17] T. E. Harris. A lower bound for the critical probability in a certain percolation process. *Proc. Cambridge Philos. Soc.*, 56(1):13–20, 1960.

[18] W. Hoeffding. Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.*, 58:13–30, 1963.

[19] D. Huh and J. Paulsson. Non-genetic heterogeneity from stochastic partitioning at cell division. *Nat. Genet.*, 43(2):95–100, 2011.

[20] D. Huh and J. Paulsson. Random partitioning of molecules at cell division. *Proc. Natl. Acad. Sci. USA*, 108(36):15004–15009, 2011.

[21] K. Joag-Dev and F. Proschan. Negative association of random variables, with applications. *Ann. Statist.*, 11(1):286–295, 1983.

[22] M. Kimmel. Quasistationarity in a branching model of division-within-division. In K. B. Athreya and P. Jagers, editors, *Classical and Modern Branching Processes*, volume 84 of *IMA Vol. Math. Appl.*, pages 157–164. Springer, New York, 1997.

[23] M. Kimmel and D. E. Axelrod. *Branching Processes in Biology*, volume 19 of *Interdisciplinary Applied Mathematics*. Springer, New York, 2 edition, 2015.

[24] W. Kong, B. A. Biddy, K. Kamimoto, J. M. Amrute, E. G. Butka, and S. A. Morris. CellTagging: combinatorial indexing to simultaneously map lineage and identity at single-cell resolution. *Nat. Protoc.*, 15(3):750–772, 2020.

[25] A. Marguet and C. Smadi. Parasite infection in a cell population: role of the partitioning kernel. *J. Math. Biol.*, 89:Paper No. 14, 2024.

[26] M. A. Micheelsen, N. Mitarai, K. Sneppen, and I. B. Dodd. Theory for the stability and regulation of epigenetic landscapes. *Phys. Biol.*, 7(2):026010, 2010.

[27] A. Movilla Miangolarra and M. Howard. Theory of epigenetic switching due to stochastic histone mark loss during DNA replication. *Phys. Biol.*, 22(1):016005, 2025.

[28] N. S. Nedialkov, K. R. Jackson, and G. F. Corliss. Validated solutions of initial value problems for ordinary differential equations. *Appl. Math. Comput.*, 105(1):21–68, 1999.

[29] P. Olofsson. Branching processes with local dependencies. *Ann. Appl. Probab.*, 6(1):238–268, 1996.

[30] Y. Oren, M. Tsabar, M. S. Cuoco, L. Amir-Zilberstein, H. F. Cabanos, J.-C. Hütter, et al. Cycling cancer persister cells arise from lineages with distinct programs. *Nature*, 596:576–582, 2021.

[31] N. Petryk, M. Dalby, A. Wenger, C. B. Strømme, A. Strandsby, R. Andersson, and A. Groth. MCM2 promotes symmetric inheritance of modified histones during DNA replication. *Science*, 361(6409):1389–1392, 2018.

[32] O. Sandler, S. P. Mizrahi, N. Weiss, O. Agam, I. Simon, and N. Q. Balaban. Lineage correlations of single cell division time as a probe of cell-cycle dynamics. *Nature*, 519:468–471, 2015.

[33] S. M. Shaffer, M. C. Dunagin, S. R. Torborg, E. A. Torre, B. Emert, C. Krepler, et al. Rare cell variability and drug-induced reprogramming as a mode of cancer drug resistance. *Nature*, 546:431–435, 2017.

[34] S. V. Sharma, D. Y. Lee, B. Li, M. P. Quinlan, F. Takahashi, S. Maheswaran, et al. A chromatin-mediated reversible drug-tolerant state in cancer cell subpopulations. *Cell*, 141(1):69–80, 2010.

[35] H. L. Smith. *Monotone Dynamical Systems: An Introduction to the Theory of Competitive and Cooperative Systems*, volume 41 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 1995.

[36] G. Söderlind. The logarithmic norm. History and modern theory. *BIT Numer. Math.*, 46(3):631–652, 2006.

[37] 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, New York, 2020. ACM.

[38] V. A. Vatutin and E. E. Dyakonova. Branching processes in random environment with sibling dependence. *J. Math. Sci. (N.Y.)*, 246(4):569–579, 2020.

[39] C. Yu, H. Gan, A. Serra-Cardona, L. Zhang, S. Gan, S. Sharma, E. Johansson, A. Chabes, R.-M. Xu, and Z. Zhang. A mechanism for preventing asymmetric histone segregation onto replicating DNA strands. *Science*, 361(6409):1386–1389, 2018.