Reliable finite-time output from stochastic autocatalytic reactors:

retained rewards, uncertain background catalysis, and quantitative startup

20 September 2026

Abstract

A reflexively autocatalytic and food-generated (RAF) set certifies that a reaction network could regenerate its catalysts from food. It supplies no deadline, stock, export quota or reliability for a finite reactor started from food, and the surrounding chemistry that may destroy a selected autocatalyst is usually unknown. We prove a finite-time output certificate for marked jump processes that uses this disturbance. A one-sided logarithmic reward retains every adverse fluctuation of a selected count and only its designated self-catalytic birth; a drift inequality charges all other catalytic activity to nonfood occupation, the compensator of collected output; and a continuous-time Bernstein bound controls the reward without paying for favourable births. In a fed, diluted, reversible binary-polymer reactor with falling-factorial propensities, for every maximum length $n \geq 4$, count scale $V \geq \max\{2 \times 10^{22}, 10^{8}n\}$, arbitrary nonfood catalytic background and arbitrary food catalysis of strength at most 10^{-5}, a self-catalytic witness started from food meets two export quotas, three stock observations, a mass corridor and a food budget on one trajectory with probability above $1 - 10^{-10}$. Every time window receives a quota on one event. For fixed n the failure bound tends to zero, and averaging over a capped-Zipf source gives success probability $\Omega(2^{-n})$ along the critical sequence. The internal chemistry is detailed-balanced for every admitted assignment. An earlier food-silent theorem at $V \geq \max\{10^{24}, 10^{13}n\}$ with reliability 0.9989 is machine-checked in Lean 4, as are the new algebraic and scalar components; the new stochastic theorems are proved conventionally. The sufficient scale remains fifteen orders of magnitude above a necessary first-birth scale, and we isolate the establishment estimate that controls the gap. The model is schematic and the output is aggregate material, not a functional product.

1 Introduction

Collectively autocatalytic sets were proposed by Kauffman [13, 23] as a route by which a random chemistry could sustain itself without a single self-replicating molecule. Their combinatorial formalization, the reflexively autocatalytic and food-generated set or RAF [18, 31, 32], is a closure statement: a set of reactions can produce all of its reactants from a food set, and each of its reactions has a catalyst so produced. In the binary polymer model RAFs appear once the mean number of reactions catalysed per molecule grows linearly with the maximum word length [22, 26], and heterogeneous, power-law catalysis changes both their appearance and their size [20, 21].

A closure statement is not an operating guarantee. It supplies no initial state, no kinetic law, no deadline, no collection rule and no resource account. Kinetic studies of autocatalytic sets are as old as the model [13]. Hordijk and Steel simulated molecular flow in RAF networks [19]; Filisetti et al.

and Vasas et al. studied stochastic emergence and compartmentalized dynamics [14, 37]; Giri and Jain separated the catalytic strength needed to maintain a large-molecule state from that needed to reach it from food [17]; Ravoni studied how composition shapes stochastic dynamics [30]; the stoichiometric theory of autocatalysis classifies motifs that can amplify their own species [6, 35]; seed-dependent activation shows why food-only initiation is a substantive restriction [28]; and discreteness of molecule numbers is known to change the behaviour of small autocatalytic systems qualitatively [34]. A companion paper [3] fixed one random model in which an operating question can be posed exactly: the capped-Zipf binary polymer model in a fed, diluted, stochastic mass-action reactor started from food, with a two-window operating requirement on one uninterrupted trajectory. It proved a positive lower bound on success, a matching converse and the rarity order $\Theta(2^{-n})$ along the critical Zipf sequence, at the sufficient count scale $10^{60}(n+1)^2$.

The question. Can one certify collected output from a food-only preparation, with an explicit deadline and an explicit probability, *uniformly* over a large and potentially disruptive catalytic background? Nonfood molecules can catalyse reactions that destroy a selected autocatalyst, and in a random or poorly characterized chemistry one cannot list them. A guarantee that holds only when the background is absent, or only when no unlisted reaction ever fires, is of little use at macroscopic counts.

The idea. The disturbance has a measurable cost. A nonfood catalyst that accelerates the destruction of the selected species is itself nonfood material, and nonfood material is washed out at unit rate and collected. The total nonfood monomer concentration m is simultaneously the quantity that bounds every harmful catalytic speed and the compensator of the collected output. A drift inequality of the form

$$(\text{drift of a potential}) \;\geq\; A - B\,m$$

therefore says: either the selected autocatalyst grows, or the background is already producing output. Since the potential cannot grow by more than a bounded amount over a window, occupation $\int m$ is forced, and output follows. This avoids requiring the whole network to stay near a reference differential equation; quantitative fluid approximation of Markov chains is a well-developed alternative [9, 25], but the observable here serves a different purpose, certifying a service quantity directly. The second ingredient is one-sidedness. The potential is a logarithm of the selected count; only adverse jumps can threaten a lower bound, so the reward retains every loss and one designated birth, and sets the logarithmic increment of every other favourable birth to zero. The chemistry is unchanged. Only the observable is thinned, and its variance no longer pays for reactions that help.

Contributions.

(i) *A reward-to-output certificate* (Section 4). For a marked jump process, a reward dominated by the increments of a potential, with drift at least $A - Bg$, predictable quadratic rate at most q and jumps at most b, forces output with compensator $\int g$ above $(Ad - D - z)/B - a$ on a window of length d, outside two Bernstein-type events (Theorem 4.2). A two-sided version covers every sub-window of a horizon on one event (Corollary 4.3). The exponential inequality is classical [11, 15, 36]; we give the short continuous-time marked proof with the stopping convention that the application needs. [conventional]

(ii) *A robust finite-volume theorem for every n* (Theorem 3.1). For $n \geq 4$, integer $V \geq \max\{2 \times 10^{22}, 10^8 n\}$, arbitrary nonfood catalytic background and arbitrary food catalytic incidences with coefficients at most 10^{-5}, the two-window requirement of [3] holds with probability greater than $1 - 10^{-10}$; with food coefficients up to 10^{-4} the bound is $1 - 10^{-6}$. Table 1 lists further points of the scale–robustness–confidence trade-off. Every added reaction is allowed to fire; food catalysis costs a finite drift term 1287η, not a coupling penalty that grows with V. [conventional; algebraic and scalar components Lean]

(iii) *All windows on one event* (Theorem 3.2). For $V \geq \max\{10^{24}, 10^9 n\}$, outside an event of probability below 10^{-20}, every window $(s,t] \subseteq (1,199]$ exports more than $\lceil (2.182(t-s) - 80)/624 - \frac{1}{20} \rceil V$. Consequently export over $(1,199]$ exceeds $0.51V$, recovery exceeds $1/4700$ of all supplied monomer, and every window longer than 79.6 residence times exports more than $V/10$ (Corollary 3.3). [conventional]

(iv) *Limits and source averages* (Corollaries 3.4 and 3.5). For fixed n and food coefficients up to 5×10^{-4} the certified failure tends to zero as $V \to \infty$. Averaging over the capped-Zipf source gives $\liminf_n 2^n \mathbb{P}(F_{n,V_n}) \geq (1-10^{-10})(6/\pi^2)^6\, 9/(2\pi^2)$ along the critical sequence for every $V_n \geq \max\{2 \times 10^{22}, 10^8 n\}$, and 1600 independent redraws of chemistry suffice for a 95% screening guarantee at $n = 4$. [conventional; normalized source limit Lean]

(v) *A compiled food-silent theorem* (Theorem 3.6). For $V \geq \max\{10^{24}, 10^{13} n\}$ and empty food rows, the requirement holds with probability at least 9989/10000, every window receives the quota $(\frac{19}{10}d - 82)/624 - \frac{1}{20}$ on one event, and successful trajectories have net internal synthesis above $V/5$. This statement, for the literal continuous-time jump law and the literal source, is machine-checked. [Lean]

(vi) *Internal detailed balance* (Proposition 3.7). Because each reversible pair shares one dimensionless coefficient, every admitted assignment of coefficients and catalysts makes the internal chemistry detailed-balanced, deterministically and at the level of the count process with falling-factorial repeated roles. The uncertain background changes kinetic access, not the internal equilibrium family; the reactor is driven only by feed and dilution. [conventional]

(vii) *What sets the scale* (Section 12). The sufficient scale is set by a basal count floor that does not use selected reproduction. We show exactly where the certificate stops ($V \approx 1.5 \times 10^{22}$), state the joint establishment–output implication that would lower it, and record a structural reason, the pairing of every catalysed loss with a catalysed gain, why the uniform loss envelope is pessimistic.

Three kinds of evidence. Statements marked [Lean] have a named declaration in a Lean 4 [10] and Mathlib [33] development with a strict compiler receipt (Appendix A). Statements marked [conventional] are proved in full in the text. *Exact arithmetic*, rational enclosures with Taylor lower bounds for exponentials, is replayed by a script distributed with the source; *numerical diagnostics* are labelled as such and are never premises. The robust mission theorem is conventional: its physical generator bounds and its probability assembly are in Sections 5–8, while the completed-square drift inequality, the weighted-square variance inequality, the potential range, the rounding of the count floor, and the scalar margins and error budget of the headline operating point are compiled.

What is not claimed. The chemistry is schematic: binary words, ordered split identities, one basal scale, paired reversible coefficients, uniform washout and a well-mixed vessel. The count scale is sufficient, not a threshold, and remains conservative by many orders of magnitude. The witness is a single self-catalytic ligation embedded in an arbitrary background, not a collectively autocatalytic cycle. Output is aggregate nonfood monomer effluent, not a purified or functional product. The robustness theorem is pointwise in a modified kinetic family; it does not remove the food-silence factor from the random-source average, and it does not make the probability of drawing a suitable chemistry tend to one. No matching upper bound or selection theorem is transferred to the new scale. No thermodynamic account of the maintained feed is given.

Relation to earlier drafts. This paper merges and supersedes two unpublished drafts of 17 and 20 September 2026. The first contained the compiled food-silent theorem and the retained-reward mechanism at a fixed cutoff; the second introduced the volume-proportional cutoff, the optimized Bernstein certificate and structural weak-food robustness at $n = 4$ and $V \geq 5 \times 10^{22}$. The present version adds the exact quadratic minimum in the drift, sharper logarithmic and potential-range constants, the all-n statement, the robust all-window theorem and the detailed-balance proposition. Static branching certificates for RAF availability that appeared in the first draft concern a different probability space and are reported with the critical-window work [1, 2].

Organization. Section 2 fixes the model, the uncertainty class and the mission. Section 3 states the results. Section 4 proves the general certificate. Sections 5–7 establish the physical bounds, the guard and count floor, and the retained reward; Section 8 assembles the theorems. Section 9 treats source averaging and the scope of robustness, Section 10 the internal equilibrium structure, Section 11 dimensions and readout, and Section 12 the remaining scale. Appendix A maps statements to evidence.

2 Model, uncertainty class and mission

2.1 Catalogue and kinetics

Fix $n \geq 4$. Let $\mathcal{X}_n$ be the nonempty binary words of length at most n, and let $\mathcal{R}_n$ be the ordered split identities $r = (v, w, vw)$ with $|v|, |w| \geq 1$ and $|vw| \leq n$. Then

$$|\mathcal{X}_n| = 2^{n+1} - 2, \qquad R_n = |\mathcal{R}_n| = \sum_{\ell=2}^{n} (\ell - 1) 2^{\ell} = (n-2)2^{n+1} + 4. \tag{2.1}$$

Each identity has a ligation direction $v + w \to vw$ and a cleavage direction $vw \to v + w$. The food set is $\mathcal{F} = \{0, 1, 00, 01, 10, 11\}$; all other words are *nonfood*.

Each identity r has a basal coefficient b_r, and each pair (r, c) of an identity and a catalyst word c may carry a catalytic coefficient k_{rc}. One coefficient is used for both directions of the pair. A channel with coefficient κ and input multiplicities $\nu = (\nu_x)_{x \in \mathcal{X}_n}$, *including catalyst roles*, has propensity

$$a_\nu(N) = \kappa V^{1-|\nu|} \prod_x (N_x)_{\nu_x}, \qquad |\nu| = \sum_x \nu_x, \qquad (N)_j = N(N-1)\cdots(N-j+1). \tag{2.2}$$

Thus the catalysed channel of (r, c) has inputs $v + w + c$ and outputs $vw + c$ in the ligation direction, and inputs $vw + c$ and outputs $v + w + c$ in the cleavage direction; coincident substrate and catalyst roles are combined before the falling factorial is taken. Catalysts cancel from the net stoichiometric change, not from the propensity. This is the ordered-encounter convention without factorial divisors; other conventions absorb such factors into κ [4, 16]. Here V is a positive integer *count scale*.

Each food species arrives at rate V and every molecule washes out at rate one, so time is measured in residence times. Initially every food count is V and every nonfood count is zero; the initial molecule number is $6V$ and the initial monomer mass is $10V$. Write $x_z = N_z/V$ and

$$L = \sum_z |z| x_z, \qquad m = \sum_{|z|>2} |z| x_z, \qquad S = \sum_z x_z, \qquad u = x_{00}, \quad w = x_{11}, \quad K = N_{0011}. \tag{2.3}$$

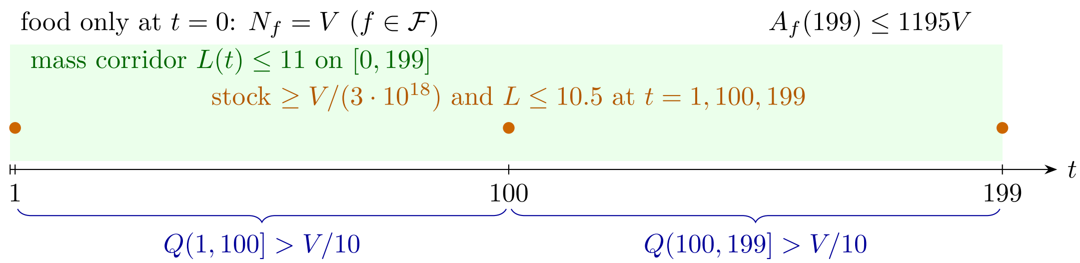

Figure 1: The single-trajectory requirement $F_{n,V}$. Stock is observed at three times; the corridor and the supply allowance cover the whole horizon, including startup from the food-only state.

L is the total and m the nonfood monomer concentration; K is a count, not a concentration. Internal reactions conserve monomer mass, and feed and washout give the generator identity $\mathcal{L}L = 10 - L$. On every bounded interval the process is nonexplosive and visits finitely many states: mass is bounded by the initial mass plus the monomer weight of finitely many Poisson arrivals [4, 27]. At fixed V and n there are finitely many states and channels with $L \le 11$.

2.2 The uncertainty class

Put $\varepsilon = 1/500\,000\,000$. For $\eta \ge 0$ let $\mathcal{U}(\eta)$ be the class of kinetic assignments with

$$\varepsilon \le b_r \le 4\varepsilon \quad (r \in \mathcal{R}_n), \qquad k_{rc} \in \{0\} \cup [4, 16] \quad (c \notin \mathcal{F}), \qquad 0 \le k_{rc} \le \eta \quad (c \in \mathcal{F}), \tag{2.4}$$

subject to the single requirement $k_{r_*,0011} \in [4, 16]$ for $r_* = (00, 11, 0011)$. This *selected incidence* is the pair

$$00 + 11 + 0011 \rightleftharpoons 0011 + 0011. \tag{2.5}$$

Every other nonfood word may catalyse any subset of $\mathcal{R}_n$, including identities that destroy 0011, and every food word may catalyse any subset with strength at most η. All theorems are pointwise in the assignment: no independence, sparsity or degree bound on the background is assumed. The class $\mathcal{U}(0)$ is the *food-silent* class of [3].

2.3 The mission

Let $Q(s, t]$ be the monomer equivalents removed by nonfood washouts during $(s, t]$, $A_f(t)$ the number of food arrivals by time t, and $B_f(t) \le 2A_f(t)$ their monomer equivalents.

Definition 2.1 (Two-window requirement [3])**.** The event $F_{n,V}$ requires, on one trajectory:

(M1) at each of the times 1, 100 and 199, $L \le \frac{21}{2}$ and some species of length four has at least $V/(3 \times 10^{18})$ copies;

(M2) $Q(1, 100] > V/10$ and $Q(100, 199] > V/10$;

(M3) $L(t) \le 11$ for all $0 \le t \le 199$, and $A_f(199) \le 1195V$.

There is no reset, resampling of chemistry or fresh initial law at time 100. The stock condition is existential over length-four species; the proofs follow the single species 0011. The output counts all nonfood species; it is not a selected-product yield. Figure 1 displays the requirement.

Lemma 2.2 (Material identity [3])**.** *Let $J(t)$ be the signed nonfood monomer mass produced by internal reactions up to time t. On every trajectory from the food-only state,*

$$m(t)V + Q(0,t] = J(t), \tag{2.6}$$

and the total monomer supplied by time t is $10V + B_f(t) \le 10V + 2A_f(t)$. [Lean]

Proof. A food arrival does not change nonfood mass; a nonfood washout lowers it by exactly what Q records; an internal reaction changes it by exactly what J records. Summing over the finitely many jumps on $[0,t]$ and using $m(0) = 0$ gives (2.6). $\square$

2.4 The random source

For source averages we use the capped-Zipf source of [3]. For a real exponent $\alpha > 1$, independently for each word $z \in \mathcal{X}_n$ draw $J_z \ge 1$ with $\mathbb{P}(J_z = k) = k^{-\alpha}/\zeta(\alpha)$, set the degree $D_z = \min(J_z, R_n) - 1$, and let the row $C_z \subseteq \mathcal{R}_n$ of identities catalysed by z be a uniformly random D_z-subset. The whole Zipf tail sits at the capped degree; rows of different words are independent and incidences within a row are strongly dependent. Every incidence receives a coefficient in $[4, 16]$ and every identity a basal coefficient in $[\varepsilon, 4\varepsilon]$ by an arbitrary mark law. Put

$$e_\alpha = \mathbb{P}(D = 0) = \frac{1}{\zeta(\alpha)}, \qquad p_{\alpha,n} = \mathbb{P}(r \in C_z) = \frac{\mathbb{E}D}{R_n}, \tag{2.7}$$

and let $\mathcal{A}_n$ be the source event that all six food rows are empty and $r_* \in C_{0011}$. Row independence and exchangeability within a row give $\mathbb{P}(\mathcal{A}_n) = e_\alpha^6\, p_{\alpha,n}$. On $\mathcal{A}_n$ the assignment lies in $\mathcal{U}(0)$. No other entry of C_{0011} and no other nonfood row is restricted.

3 Main results

3.1 The robust mission theorem

Define the count-floor constants

$$\ell = \frac{1}{1358} - \frac{1}{50000}, \qquad h = \left\lceil \frac{V}{3\times 10^{18}} \right\rceil, \qquad m_0 = \left\lfloor \frac{V}{2\times 10^{18}} \right\rfloor, \qquad N_V = \left\lfloor \frac{V}{10^{20}} \right\rfloor, \tag{3.1}$$

the reward constants, for $h > 2$,

$$A = \frac{11}{5} - 468\varepsilon - 1287\eta - \frac{1500}{h-2} - \frac{768000}{V}, \tag{3.2}$$

$$q = \frac{101}{100}\Big(\frac{121}{h} + \frac{2952\,h}{(h-2)^2}\Big) + \frac{2\,482\,176\,000}{V}, \tag{3.3}$$

$$b = \frac{2}{h-2} + \frac{32}{V}, \tag{3.4}$$

and the Bernstein function

$$\mathcal{B}(z, v, \beta) = \exp\Big\{-\frac{z^2}{2(v + \beta z/3)}\Big\}. \tag{3.5}$$

Count condition (all $n \geq 4$)	$\eta \leq$	z	exponent	quota per window >	failure <
$V \geq \max\{2 \times 10^{22}, 10^8 n\}$	10^{-5}	48	24.97	$0.101\,V$	10^{-10}
$V \geq \max\{2 \times 10^{22}, 10^8 n\}$	10^{-4}	37	14.84	$0.1002\,V$	10^{-6}
$V \geq \max\{5 \times 10^{22}, 10^8 n\}$	10^{-5}	40	43.38	$0.135\,V$	10^{-18}
$V \geq \max\{5 \times 10^{22}, 10^8 n\}$	2×10^{-4}	37	37.11	$0.101\,V$	10^{-15}
$V \geq \max\{10^{23}, 10^8 n\}$	10^{-5}	40	86.77	$0.142\,V$	10^{-37}
$V \geq \max\{10^{24}, 10^8 n\}$	5×10^{-4}	8	34.71	$0.100\,V$	10^{-14}

Table 1: Operating points of Theorem 3.1. "Exponent" is $z^2/[2(99q + bz/3)]$ at the smallest admitted V, rounded down. Quota floors and failure ceilings are certified in exact rational arithmetic with Taylor lower bounds for exponentials. The rows are different points of one trade-off; no row gives all advantages simultaneously.

For a reward allowance $z > 0$ set

$$q_*(V, \eta, z) = \frac{99A - 52 - z}{624} - \frac{1}{20}, \tag{3.6}$$

$$E_G(V, n) = 2\,\mathcal{B}\Big(\tfrac{1}{4}, \tfrac{200(18+11n)}{V}, \tfrac{n}{V}\Big) + 4\,\mathcal{B}\Big(10^{-5}, \tfrac{200\cdot 96000}{V}, \tfrac{2}{V}\Big), \tag{3.7}$$

$$E_K(V) = 3 \times 10^8\,(N_V + 1)\,2^{-N_V}, \tag{3.8}$$

$$E(V, n, \eta, z) = E_G + E_K + 2\,\mathcal{B}(z, 99q, b) + 2\,\mathcal{B}\Big(\tfrac{1}{20}, \tfrac{99\cdot 11n}{V}, \tfrac{n}{V}\Big) + e^{-V/2500}. \tag{3.9}$$

The five groups of terms are the mass and food guard, the count floor, the retained reward on two windows, the export counter on two windows, and the food supply.

Theorem 3.1 (Robust mission)**.** *Let $n \geq 4$, $0 \leq \eta \leq 3/100$, and let $V \geq 10^{22}$ be an integer. Let $z > 0$ satisfy $q_*(V, \eta, z) \geq 1/10$. Then for every assignment in $\mathcal{U}(\eta)$,*

$$\mathbb{P}(F_{n,V}) \;\geq\; 1 - E(V, n, \eta, z). \tag{3.10}$$

More precisely, outside an event of probability at most $E(V, n, \eta, z)$: $K(t) \geq h$ for every $1 \leq t \leq 199$; $L(t) \leq \frac{21}{2}$ and $u(t), w(t) \geq \ell$ for every $0 \leq t \leq 199$; $Q(1, 100]$ and $Q(100, 199]$ both exceed $q_* V$; $A_f(199) \leq 1195V$; and $J(199) > 2q_* V$. The bound is nonincreasing in V for fixed n/V and fixed z, η. In particular:

- *(a) for $V \geq \max\{2 \times 10^{22}, 10^8 n\}$ and $\eta \leq 10^{-5}$, $\mathbb{P}(F_{n,V}) > 1 - 10^{-10}$, and each window exports more than $0.101V$;*
- *(b) for $V \geq \max\{2 \times 10^{22}, 10^8 n\}$ and $\eta \leq 10^{-4}$, $\mathbb{P}(F_{n,V}) > 1 - 10^{-6}$;*
- *(c) the further operating points of Table 1 hold.*

[conventional; scalar margins of (a), (b) Lean]

3.2 All windows on one event

For the whole horizon $(1, 199]$, with export tolerance $1/40$, define

$$E_{\text{all}}(V, n, \eta, z) = E_G + E_K + 2\,\mathcal{B}(z, 198q, b) + 2\,\mathcal{B}\Big(\tfrac{1}{40}, \tfrac{198\cdot 11n}{V}, \tfrac{n}{V}\Big) + e^{-V/2500}. \tag{3.11}$$

Theorem 3.2 (All windows)**.** *Let $n \geq 4$, $0 \leq \eta \leq 3/100$, $V \geq 10^{22}$ an integer, $z > 0$, and any assignment in $\mathcal{U}(\eta)$. Outside an event of probability at most $E_{\rm all}(V, n, \eta, z)$, the guard and floor conclusions of Theorem 3.1 hold and, simultaneously for all $1 \leq s < t \leq 199$,*

$$\int_s^t m(v)\,\mathrm{d}v \;\geq\; \frac{A(t-s) - 52 - 2z}{624}, \qquad \frac{Q(s,t]}{V} \;>\; \frac{A(t-s) - 52 - 2z}{624} - \frac{1}{20}. \tag{3.12}$$

In particular, for $\eta \leq 10^{-5}$:

- *(a) if $V \geq \max\{10^{24}, 10^9 n\}$ then, outside an event of probability below 10^{-20}, every window satisfies $Q(s,t] > \big[(2.182\,(t-s) - 80)/624 - \frac{1}{20}\big]V$;*
- *(b) if $V \geq \max\{10^{23}, 10^9 n\}$ then, outside an event of probability below 10^{-7}, every window satisfies $Q(s,t] > \big[(2.142\,(t-s) - 104)/624 - \frac{1}{20}\big]V$.*

[conventional]

The event quantifies over a continuum of windows, but it contains a measurable event defined by the suprema of two càdlàg martingales, so no measurability question arises.

Corollary 3.3 (Output, recovery and deadlines)**.** *On the event of Theorem 3.2*(a)*:*

- *(a) $Q(1, 100]$ and $Q(100, 199]$ each exceed $0.16\,V$;*
- *(b) $Q(1, 199] > 0.51\,V$ and $J(199) > 0.51\,V$;*
- *(c) $Q(0, 199]/(10V + B_f(199)) > 1/4700$;*
- *(d) every window $(s,t] \subseteq (1, 199]$ with $t - s \geq 79.6$ exports more than $V/10$, so two disjoint such windows fit in the horizon; three disjoint windows of the guaranteed length do not.*

On the event of Theorem 3.2(b) *the corresponding values are $0.12\,V$, $0.46\,V$ and 92.3 residence times.* [conventional]

The recovery ratio is a guaranteed floor in monomer equivalents, including the initial inventory. It is not a measured or typical conversion efficiency, and nothing here bounds typical recovery from above.

Corollary 3.4 (Vanishing conditional failure)**.** *Fix $n \geq 4$ and $0 \leq \eta \leq 5 \times 10^{-4}$. Then*

$$\inf_{\mathcal{U}(\eta)} \mathbb{P}(F_{n,V}) \longrightarrow 1 \qquad \text{as integer } V \to \infty.$$

[conventional]

The conclusion is for fixed n: the mass and export budgets depend on n/V. It is a statement about conditional reliability within the witness family, not about the probability of drawing a suitable chemistry.

Corollary 3.5 (Source average and rarity)**.** *For the capped-Zipf source, every $\alpha > 1$, every $n \geq 4$ and every integer $V \geq \max\{2 \times 10^{22}, 10^8 n\}$,*

$$\mathbb{P}(F_{n,V}) \;\geq\; (1 - 10^{-10})\, e_\alpha^6\, p_{\alpha,n}. \tag{3.13}$$

Along $\alpha_n = 2 - 2/n$ and any such sequence V_n,

$$\liminf_{n\to\infty} 2^n\, \mathbb{P}(F_{n,V_n}) \;\geq\; (1 - 10^{-10})\Big(\frac{6}{\pi^2}\Big)^6 \frac{9}{2\pi^2} \approx 2.30 \times 10^{-2}. \tag{3.14}$$

At $\alpha = 2$, $n = 4$, the witness mass is $e_2^6 p_{2,4} \approx 0.0018736515$, and the probability of at least one success in 1600 independent redraws of chemistry, coefficients and trajectory exceeds 95%. [conventional; normalized source limit Lean]

	Companion [3]	Theorem 3.6	Theorems 3.1, 3.2
Background	arbitrary nonfood; food silent	arbitrary nonfood; food silent	arbitrary nonfood; arbitrary food incidences $\leq \eta$
Sufficient scale	$10^{60}(n+1)^2$	$\max\{10^{24}, 10^{13}n\}$	$\max\{2 \times 10^{22}, 10^8 n\}$
Mission failure	positive success bound	$\leq 1.1 \times 10^{-3}$	$< 10^{-10}$ ($\eta \leq 10^{-5}$)
$V \to \infty$, fixed n	–	bound does not improve	failure bound $\to 0$
All-window slope, cost	–	1.9, 82	2.182, 80 at $V \geq 10^{24}$
Observable	product concentration, symmetric noise	retained reward, fixed cutoff	retained reward, cutoff $\propto V$, Bernstein tilt
Evidence	Lean	Lean	conventional; components Lean

Table 2: Results for the same mission, preparation and nonfood background class. The necessary first-birth scale in food-silent environments is 1.43×10^7 [3].

3.3 The compiled food-silent theorem

Put $V_*(n) = \max\{10^{24}, 10^{13}n\}$ and $q_{17}(d) = \left(\frac{19}{10}d - 82\right)/624 - \frac{1}{20}$.

Theorem 3.6 (Compiled startup theorem)**.** *Let $n \geq 4$, $\alpha > 1$ and let $V \geq V_*(n)$ be an integer.*

(a) For every assignment in $\mathcal{U}(0)$ arising from rows in $\mathcal{A}_n$, $\mathbb{P}(F_{n,V}) \geq 9989/10000$.

(b) The fully averaged success probability satisfies $\mathbb{P}(F_{n,V}) \geq \frac{9989}{10000}\, e_\alpha^6\, p_{\alpha,n}$.

(c) For every source and mark configuration, almost surely on $F_{n,V}$, $J(199) > V/5$.

(d) Let $E_{n,V}$ be the event that $F_{n,V}$ holds and $Q(s,t] > q_{17}(t-s)V$ for every $1 \leq s \leq t \leq 199$. The outer probability of its complement is less than $11/10000$, and almost surely on $E_{n,V}$, $J(199) > 6V/25$.

[Lean: `StartupMarked.startup_quantitative_resolution`, `StartupMarked.startup_enhanced_failure`, `RandomViability.physical_enhanced_accounts`]

Theorem 3.6 is the statement whose every step, including the jump law, the source weights, the stopping arguments and the exponential supermartingale, is checked by the compiler. It uses a fixed logarithmic cutoff $K_0 = 333\,334$, the drift constant 19/10 and a second-order exponential bound, which is why its reliability saturates. Theorems 3.1 and 3.2 improve the range, the robustness and the confidence by conventional proofs. Table 2 and Figure 2 compare them.

3.4 Internal equilibrium structure

Proposition 3.7 (Internal detailed balance)**.** *Fix any assignment in $\mathcal{U}(\eta)$, $\eta \geq 0$, and any $\theta_0, \theta_1 > 0$. Put $x_z^* = \theta_0^{\#_0(z)}\theta_1^{\#_1(z)}$, where $\#_i(z)$ counts the letters i in z.*

(a) In the deterministic mass-action system of the internal reactions, every basal and every catalysed reversible pair has equal forward and reverse flux at x^*.

(b) In the count process without feed and washout, on every closed communicating class the normalized restriction of

$$\pi(N) \;\propto\; \prod_{z \in \mathcal{X}_n} \frac{(V x_z^*)^{N_z}}{N_z!} \tag{3.15}$$

satisfies detailed balance for every reversible pair, including pairs with repeated substrate or catalyst roles. Each class is finite, so π is normalizable and stationary.

Neither statement extends to the fed and diluted reactor. [conventional]

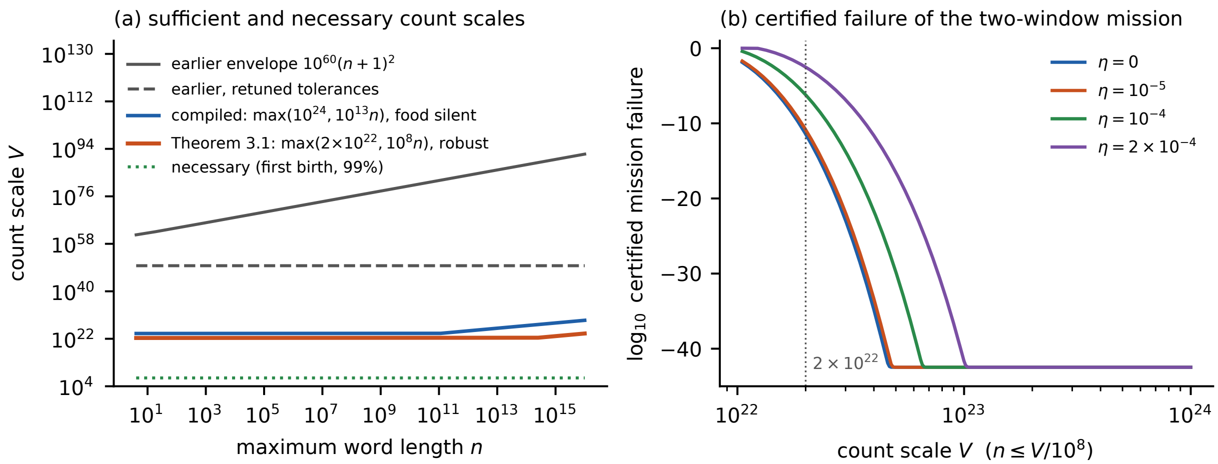

Figure 2: (a) Sufficient count scales for the same two-window requirement, together with the necessary first-birth scale for 99% reliability in food-silent environments. The curves have different logical roles and are not estimates of a transition. (b) Certified failure of the mission from (3.9) with the largest admissible z, for $n \le V/10^8$. The plateau near $10^{-42.5}$ is the conservative ceiling $9e^{-100}$ used for the clock terms, not a property of the reactor.

4 A reward-to-output certificate

Consider a nonexplosive marked jump process X on a countable state space. Mark j has intensity $a_j(X_{t-})$ and a state increment determined by the mark; different marks may have the same increment. Fix a deterministic window $(s,t]$ of length d and a set $\mathcal{G}$ of states, the *active region*. A *reward* $r_j(x)$ is assigned to the actual marked jump out of the pre-jump state x; it need not be the increment of a state function.

Localization. Rewards are disabled if $X_s \notin \mathcal{G}$. Otherwise they are retained up to and including the first jump that leaves $\mathcal{G}$, and then turned off. The indicator that a jump is rewarded depends only on the strict past and the pre-jump state, so it is predictable. We assume that on $\mathcal{G}$ the total rate is bounded and $|r_j| \le b$. The compensated reward is

$$Z_v = \sum_{s<\sigma\le v} r_{J_\sigma}(X_{\sigma-}) - \int_s^v \sum_j a_j(X_{\sigma-})\, r_j(X_{\sigma-})\, \mathrm{d}\sigma, \tag{4.1}$$

with the stopping indicator understood in both terms. It is a bounded-jump martingale on $[s,t]$.

Lemma 4.1 (Continuous-time Bernstein bound)**.** *If $|r_j| \le b$ and $\sum_j a_j r_j^2 \le q$ on $\mathcal{G}$, then for every $z > 0$,*

$$\mathbb{P}\Big(\inf_{s\le v\le t} Z_v \le -z\Big) \;\le\; \mathcal{B}(z, dq, b) = \exp\Big\{-\frac{z^2}{2(dq + bz/3)}\Big\}, \tag{4.2}$$

and the same bound holds for $\mathbb{P}(\sup_{s\le v\le t} Z_v \ge z)$.

Proof. For $|y| < 3$, the series of the exponential and $k! \geq 2 \cdot 3^{k-2}$ for $k \geq 2$ give

$$e^y - 1 - y \;\leq\; \frac{y^2}{2} \sum_{k\geq 2} \Big(\frac{|y|}{3}\Big)^{k-2} = \frac{y^2}{2(1-|y|/3)}. \tag{4.3}$$

Fix $0 < \theta < 3/b$. By the exponential formula for marked point processes [7], or directly by applying the bounded generator of the pair (X, Z) to $(x, \zeta) \mapsto e^{-\theta\zeta}$,

$$M_v = \exp\Big\{-\theta Z_v - \int_s^v \sum_j a_j (e^{-\theta r_j} - 1 + \theta r_j)\, \mathrm{d}\sigma\Big\}$$

is a nonnegative martingale with $M_s = 1$. By (4.3) with $y = -\theta r_j$, the integral is at most $\theta^2 dq/[2(1-\theta b/3)]$ for $v \leq t$. On $\{\inf_v Z_v \leq -z\}$ the martingale therefore reaches $\exp\{\theta z - \theta^2 dq/[2(1-\theta b/3)]\}$, and Doob's maximal inequality gives

$$\mathbb{P}\Big(\inf_{s\leq v\leq t} Z_v \leq -z\Big) \leq \exp\Big\{-\theta z + \frac{\theta^2 dq}{2(1-\theta b/3)}\Big\}.$$

Choose $\theta = z/(dq + bz/3)$, for which $\theta b < 3$ and $1 - \theta b/3 = dq/(dq + bz/3)$. The exponent becomes $-z^2/(dq + bz/3) + z^2/[2(dq + bz/3)]$, which is (4.2). If $dq = 0$ the martingale Z is constant and the claim is trivial. Replacing r by $-r$ gives the upper tail. $\square$

This is the bounded-jump Bernstein–Freedman mechanism [15]; continuous-time versions appear in [11, 36]. Nothing in it is new. What matters below is that the bound is unconditional for the stopped process, so the application never conditions on the future survival of a guard.

Theorem 4.2 (Service certificate)**.** *Let F be a function on states and $g \geq 0$ a rate function. Suppose that on $\mathcal{G}$, with $B > 0$,*

$$\Delta_j F \;\geq\; r_j, \qquad \sum_j a_j r_j \;\geq\; A - Bg, \qquad \sum_j a_j r_j^2 \;\leq\; q, \qquad |r_j| \leq b, \tag{4.4}$$

where $\Delta_j F$ is the increment of F under mark j. Let H be an event on which $X_v \in \mathcal{G}$ for all $s \leq v \leq t$ and $F(X_t) - F(X_s) \leq D$. Let $Y(s,t]$ be a normalized output whose compensator is $\int_s^t g(X_v)\,\mathrm{d}v$ and whose compensated error, stopped at the exit from $\mathcal{G}$, has quadratic rate at most q_Y and jumps at most b_Y. Then for all $z, a > 0$,

$$\mathbb{P}\Big(H \cap \Big\{Y(s,t] \leq \frac{Ad - D - z}{B} - a\Big\}\Big) \;\leq\; \mathcal{B}(z, dq, b) + \mathcal{B}(a, dq_Y, b_Y). \tag{4.5}$$

For finitely many prescribed windows one adds the bounds and $\mathbb{P}(H^c)$; neither independence nor a restart is used.

Proof. On H the localization never acts, so the first premise telescopes over the jumps in $(s,t]$ and the second bounds the compensator in (4.1):

$$D \;\geq\; F(X_t) - F(X_s) \;\geq\; \sum_{s<\sigma\leq t} r_{J_\sigma} \;=\; \int_s^t \sum_j a_j r_j \,\mathrm{d}v + Z_t \;\geq\; Ad - B\int_s^t g(X_v)\,\mathrm{d}v + Z_t.$$

Outside $\{Z_t \leq -z\}$ this forces $\int_s^t g \geq (Ad - D - z)/B$. Outside the event that the compensated output error is at most $-a$, $Y(s,t] > \int_s^t g - a$. Lemma 4.1 bounds both exceptional events for the stopped processes, without conditioning on H. $\square$

Corollary 4.3 (Every sub-window)**.** *Under the hypotheses of Theorem 4.2 on a horizon* $(s_0, t_0]$ *of length* d_0*, outside an event of probability at most* $\mathbb{P}(H^c) + 2\mathcal{B}(z, d_0 q, b) + 2\mathcal{B}(a, d_0 q_Y, b_Y)$*, every window* $s_0 \le s < t \le t_0$ *satisfies*

$$\int_s^t g(X_v)\,\mathrm{d}v \ \ge\ \frac{A(t-s) - D - 2z}{B}, \qquad Y(s,t] \ >\ \frac{A(t-s) - D - 2z}{B} - 2a,$$

provided $F(X_t) - F(X_s) \le D$ *for all such* $s < t$ *on* H*.*

Proof. Define Z from time s_0. On $\{\sup_v |Z_v| < z\}$ every increment satisfies $Z_t - Z_s > -2z$, and likewise for the output error with $2a$. The telescoping argument applies verbatim to $(s, t]$. Lemma 4.1 bounds the four one-sided suprema. □

For k windows of length d and a reward-error allowance ρ, write $\Lambda = \log(k/\rho)$. Solving $k\,\mathcal{B}(z, dq, b) = \rho$ gives the inverse form

$$z = \frac{b\Lambda}{3} + \sqrt{\Big(\frac{b\Lambda}{3}\Big)^2 + 2dq\Lambda}, \qquad y_{\text{cert}} = \frac{Ad - D - z}{B} - a, \tag{4.6}$$

a deadline–quota–confidence certificate. The allowance ρ covers only the reward tails; output, guard and floor errors are additional.

5 Physical bounds that survive background and food catalysis

Throughout Sections 5–8 the assignment is any element of $\mathcal{U}(\eta)$ with $0 \le \eta \le 3/100$. On the mass corridor $\{L \le 11\}$ define

$$C = 4\varepsilon + \frac{16}{3}m + 11\eta. \tag{5.1}$$

For each identity r the *speed* $b_r + \sum_c k_{rc}x_c$, the basal coefficient plus concentration-weighted catalytic coefficients, is at most C: every nonfood catalyst has length at least three, so the nonfood number concentration is at most $m/3$, and the food number concentration is at most $S \le L \le 11$. Falling factorials are bounded by powers, so they only lower propensities below the power-product envelope. Since $m \le L \le 11$,

$$C \le 4\varepsilon + \frac{176}{3} + \frac{33}{100} < 59. \tag{5.2}$$

Lemma 5.1 (Flux and noise envelopes)**.** *At every state with* $L \le 11$*:*

- *(E1) the copy-weighted loss rate of* K *satisfies* $D_K := \sum_{\Delta K<0} |\Delta K|\, a_j \le (1 + 25C)K \le 1476K$*, and every jump lowers* K *by at most two copies;*
- *(E2) the selected forward channel has rate* $\lambda = k_{r_*,0011}\, uwK$*, so* $4uwK \le \lambda \le 16uwK \le 121K$*;*
- *(E3)* $\mathcal{L}u \ge 1 - u - 23Cu \ge 1 - 1358u$*, and likewise for* w*;*
- *(E4) the total rate of all channels is at most* $24000V$*; every jump changes* u *and* w *by at most* $2/V$ *each, so each of* u, w *has quadratic rate at most* $96000/V$*;*
- *(E5) total mass* L *has quadratic rate at most* $(18 + 11n)/V$ *and jumps at most* n/V*; the normalized output* Q/V *has compensator* $\int m$*, quadratic rate at most* $11n/V$ *and jumps at most* n/V*.*

Constant	Origin	Role
11	chosen mass corridor	bounds S, m and partner concentrations
23, 25	two ligation positions against partners of total concentration ≤ 11, plus one (dimer) or three (tetramer) cleavage positions	food depletion; selected-copy loss
59	ceiling of C, (5.2)	uniform envelopes
1358, 1476	$1 + 23 \cdot 59$, $1 + 25 \cdot 59$	food restoring rate; copy-loss ceiling
121	$16uw \le 121$ when $2u + 2w \le 11$	selected birth rate per copy
117	25 from copy loss plus 46 for each of two food deficits	catalytic drift penalty
624, 1287, 468	$117 \cdot \frac{16}{3}$, $117 \cdot 11$, $117 \cdot 4$	cost of nonfood activity, of food catalysis and of basal reactions
16/5	exact minimum of $4(1-\alpha)(1-\beta) + 8(\alpha^2 + \beta^2)$	gross drift; gives $\frac{11}{5}$ in (3.2)
1500	$\frac{121}{8} + 1476$ rounded up	second-order logarithmic correction
96000	four times the rate ceiling 24000	coordinate quadratic rate
768000	$4 \cdot 2 \cdot 96000$	second-order food-penalty term
2 482 176 000	$101 \cdot 128 \cdot 2 \cdot 96000$	food part of the reward variance
1558/9	$\frac{19}{162} \cdot 1476$	loss term of the power tilt
52	$\log(\frac{33}{4} \cdot 10^{18}) + 8 < 52$	range of the potential

Table 3: Provenance of the constants. None is fitted; each is a rounded consequence of the corridor $L \le 11$, the coefficient box (2.4) and the mission's relative stock threshold.

Proof. (E1) and (E3). A word g is consumed by ligation as a left substrate, $g + y \to gy$, and as a right substrate, $y + g \to yg$. Summing the envelope $Cx_g x_y V$ over partners y gives at most $2C\,S\,N_g$ copies per unit time. The identity (g, g, gg) removes two copies per firing and occurs once in each of the two sums, so its copy-weighted loss is covered. A word is consumed by cleavage at its $|g| - 1$ split positions, with copy-weighted rate at most $(|g| - 1)CN_g$. When a substrate coincides with the catalyst the propensity carries one more falling factor, which is below the envelope, and the catalyst copy is returned, so the net loss is unchanged. Washout adds N_g. With $S \le 11$ the copy-weighted loss of g is at most $[1 + C(22 + |g| - 1)]N_g$, which is $(1 + 25C)K$ for $g = 0011$ and $(1 + 23C)N_g$ for a dimer. By (5.2), $1 + 25C \le 1476$ and $1 + 23C \le 1358$. No internal reaction removes more than two copies of one word. For (E3), feed contributes $+1$ to $\mathcal{L}u$ and production is discarded.

(E2). The three inputs 00, 11 and 0011 are distinct words, so the propensity (2.2) is exactly $kV^{-2}N_{00}N_{11}K = k\,uwK$. Since $2u + 2w \le L \le 11$, $uw \le (11/4)^2$ and $16uw \le 121$.

(E4). Every speed is at most $4\varepsilon + 16S \le 177$. Summing the ligation power products over all ordered splits gives at most $S^2 \le 121$, and summing cleavage inputs with their split multiplicity gives $\sum_z (|z| - 1)x_z \le L \le 11$. The internal rate is thus at most $177 \cdot 132\,V$; feed contributes $6V$ and washout $SV \le 11V$. The total is at most $23381V < 24000V$. A jump changes one count by at most two, so the quadratic rate of u is at most $(2/V)^2 \cdot 24000V$.

(E5). Internal reactions do not change mass. The six feed clocks contribute $V \sum_f |f|^2/V^2 = 18/V$, and washout contributes $\sum_z |z|^2 x_z/V \le nL/V \le 11n/V$. Restricting the washout sum to nonfood words gives the output statements; its compensator is $\sum_{|z|>2} |z|N_z/V = m$. □

These bounds charge every catalyst, food or nonfood, and every repeated role. They replace the empty-food-row hypothesis of the earlier work by the single term 11η in (5.1). Table 3 records where each constant comes from.

6 Guard and count floor

The *guard* is $\mathfrak{g} = \{L \le 11,\ u \ge \ell,\ w \ge \ell\}$ with ℓ from (3.1).

Lemma 6.1 (Mass and food guard)**.** *Outside an event of probability at most $E_G(V,n)$, (3.7), every state occupied during $[0,199]$ satisfies $L \le \frac{21}{2}$ and $u, w \ge \ell$.*

Proof. Let τ be the first exit time from $\{L \le 11\}$, and stop the compensated processes Z_L, Z_u, Z_w of L, u, w at $\tau \wedge 200$, including the exit jump. By Lemma 5.1(E4)–(E5) and Lemma 4.1, outside an event of probability at most E_G,

$$\sup_{v\le 200} |Z_L| < \tfrac{1}{4}, \qquad \sup_{v\le 200} |Z_u| < 10^{-5}, \qquad \sup_{v\le 200} |Z_w| < 10^{-5}.$$

Mass. Since $\mathcal{L}L = 10 - L$ and $L(0) = 10$, the deviation $e = L - 10$ solves $\mathrm{d}e = -e\,\mathrm{d}t + \mathrm{d}Z_L$ up to and including τ, so

$$e(t) = \int_0^t e^{-(t-v)}\,\mathrm{d}Z_L(v) = Z_L(t) - \int_0^t e^{-(t-v)} Z_L(v)\,\mathrm{d}v, \qquad |e(t)| \le (2 - e^{-t}) \sup_{v\le t} |Z_L| < \tfrac{1}{2}. \quad (6.1)$$

The identity holds at $t = \tau$ as well, because the stopped martingale includes the exit jump; hence $L(\tau) \le \frac{21}{2}$, which contradicts $L(\tau) > 11$. So $\tau > 200$ and $L \le \frac{21}{2}$ throughout.

Food. On $[0,\tau)$ write $\mathcal{L}u = 1 - 1358u + \sigma$ with $\sigma \ge 0$ by (E3). Then $y = u - 1/1358$ satisfies $\mathrm{d}y = -1358\,y\,\mathrm{d}t + \sigma\,\mathrm{d}t + \mathrm{d}Z_u$ with $y(0) > 0$, and

$$y(t) = e^{-1358t}y(0) + \int_0^t e^{-1358(t-v)}\sigma\,\mathrm{d}v + Z_u(t) - 1358\int_0^t e^{-1358(t-v)} Z_u(v)\,\mathrm{d}v \ \ge\ -2\sup_{v\le t}|Z_u|.$$

Integration by parts is pathwise for the finite-variation process Z_u, so the bound holds across jumps. Hence $u \ge 1/1358 - 2\cdot 10^{-5} = \ell$, and likewise for w. The union bound over two mass tails and four food tails gives E_G. □

Lemma 6.2 (Killed count tail and two-state crossing bound)**.** *Let $V \ge 10^{22}$. The probability that there is a time $t \in [1,199]$ with $K(t) < h$ such that the guard $\mathfrak{g}$ held at every state occupied before t is at most*

$$10\{1 + 1476\,(h+1)\cdot 198\}\Big(\frac{9}{10}\Big)^{m_0-h-1} \ \le\ 3\times 10^8\,(N_V+1)\,2^{-N_V} = E_K(V). \qquad (6.2)$$

Proof. Power tilt. Kill the process at the first jump into a state violating $\mathfrak{g}$. Before killing, only the finitely many states with $L \le 11$ are visited. Let $f(K) = (9/10)^K$. Upward jumps of K lower f; among them we retain only the basal firing of r_*, whose inputs are distinct and whose rate is $b_{r_*}N_{00}N_{11}/V \ge \varepsilon\ell^2 V$ under the guard, contributing $-\varepsilon\ell^2 V f/10$. A loss of $j \in \{1,2\}$ copies multiplies f by $(10/9)^j$, and $[(10/9)^j - 1]/j \le 19/162$. By (E1),

$$\mathcal{L}f \ \le\ \Big(-\frac{\varepsilon\ell^2 V}{10} + \frac{1558}{9}K\Big)f. \qquad (6.3)$$

Put $c = V/10^{17}$ and $B_0 = 1558/9$. Exact rational arithmetic gives $c + B_0 m_0 \le \varepsilon\ell^2 V/10$, with margin $6.08\times 10^{-18}V$, and $B_0 m_0/c < 9$. If $K \le m_0$, (6.3) gives $\mathcal{L}f \le -cf$. If $K > m_0 \ge 10$ then $K(9/10)^K \le m_0(9/10)^{m_0}$, because $k(9/10)^k$ decreases for $k \ge 9$. In both cases

$$\mathcal{L}f \ \le\ -cf + B_0 m_0 (9/10)^{m_0}. \qquad (6.4)$$

Killed tail. The generator of the killed chain applied to f is at most $\mathcal{L}f$, because killing replaces nonnegative values of f by zero. Kolmogorov's equation for the finite killed chain and (6.4) give, for $\phi(t) = \mathbb{E}[f(K_t); \text{not killed by } t]$,

$$\phi(t) \le e^{-ct}\phi(0) + (1 - e^{-ct})\frac{B_0 m_0}{c}\Big(\frac{9}{10}\Big)^{m_0} \le e^{-ct} + 9\Big(\frac{9}{10}\Big)^{m_0}.$$

Since $c \ge 20m_0$ and $e^{-20} < 9/10$, for $t \ge 1$ we have $\phi(t) \le 10(9/10)^{m_0}$, and Markov's inequality gives, for every integer $j \le m_0$,

$$\mathbb{P}\{\text{not killed by } t,\ K(t) \le j\} \le 10\,(9/10)^{m_0 - j} \qquad (t \ge 1). \tag{6.5}$$

Crossings. A first entry into $\{K < h\}$ during $[1, 199]$ is either present at time one, or occurs by a jump after time one from a state with $K \in \{h, h+1\}$; the second boundary state is necessary because a jump can remove two copies. By (E1) the rate of downward jumps from such a state is at most $1476(h + 1)$. The expected number of such crossings by the unkilled process, including a crossing by the killing jump itself, is the integral over $(1, 199]$ of this rate against the probability in (6.5) with $j = h + 1$. Adding the time-one term $10(9/10)^{m_0-h+1}$ gives the first expression in (6.2). The argument does not assume that the guard survives; Lemma 6.1 supplies that.

Rounding. For $V \ge 10^{22}$: $m_0 \ge 10$, $N_V \ge 100$, $10N_V \le m_0 - h - 1$ and $h + 1 \le 100(N_V + 1)$, directly from the floor and ceiling inequalities. [Lean: `StartupCount.count_rounding_bounds_small_volume`] Since $(9/10)^{10} < 1/2$ and $10(1 + 1476 \cdot 198 \cdot 100) < 3 \times 10^8$, the second inequality of (6.2) follows. $\square$

Let H be the intersection of the good events of Lemmas 6.1 and 6.2. Then $\mathbb{P}(H^c) \le E_G + E_K$, and on H: $L \le \frac{21}{2}$ and $u, w \ge \ell$ on $[0, 199]$, and $K \ge h$ on $[1, 199]$. At $V = 2 \times 10^{22}$ the parameters are $h = 6667$, $m_0 = 10000$ and $N_V = 200$.

7 The retained reward

Let $\delta(x) = (1 - x)_+$ and define the potential

$$F = \log K + \psi(u, w), \qquad \psi(u, w) = -4\big[\delta(u)^2 + \delta(w)^2\big]. \tag{7.1}$$

At a pre-jump state with $K > 2$, the reward of mark j is

$$r_j = \lambda_j + \Delta_j\psi, \qquad \lambda_j = \begin{cases} \log(1 + 1/K), & j \text{ is the selected forward mark of (2.5)}, \\ \log(1 - a/K), & \Delta_j K = -a < 0, \\ 0, & \text{otherwise.} \end{cases} \tag{7.2}$$

The logarithmic part keeps the selected self-catalytic birth and every loss of K. It is set to zero on every other birth of 0011: the basal ligation, ligations catalysed by other words, and cleavages of longer words. The food part $\Delta_j\psi$ is kept in full on *every* mark, including marks whose logarithmic part is discarded; it may be adverse. Every omitted logarithmic increment is nonnegative, which gives the first premise of Theorem 4.2:

$$\Delta_j F \ \ge\ r_j \qquad \text{at every state with } K > 2. \tag{7.3}$$

[Lean: `retained_path_sum_le`] No channel is deleted or thinned from the reactor. Only the observable changes.

Lemma 7.1 (Drift, variance, jump and range)**.** *On $\mathcal{G} = \{L \le 11,\ K \ge h\}$ with $h > 2$, the premises (4.4) hold with $g = m$, $B = 624$, and A, q, b given by (3.2)–(3.4). If two states have $L \le 11$ and $K \ge h$, their potentials differ by at most $D = 52$.*

Proof. Logarithmic drift. For $K \ge 3$, $\log(1 + 1/K) \ge 1/K - 1/(2K^2)$, so by (E2) the selected birth contributes at least $(\lambda/K)(1 - \frac{1}{2K}) \ge 4uw - 2uw/K \ge 4uw - \frac{121}{8}/K$. For $0 \le x < 1$, $-\log(1 - x) = \sum_{k \ge 1} x^k/k \le x + x^2/[2(1 - x)]$. With $x = a/K$ and $a \in \{1, 2\}$ this gives

$$\log\Big(1 - \frac{a}{K}\Big) \ \ge\ -\frac{a}{K} - \frac{a^2}{2K(K-a)} \ \ge\ -\frac{a}{K} - \frac{a}{K(K-2)},$$

since $a^2/2 \le a$ and $K - a \ge K - 2$. Summing against the loss rates and using (E1),

$$\sum_j a_j \lambda_j \ \ge\ 4uw - (1 + 25C) - \frac{121/8}{K} - \frac{1476}{K-2} \ \ge\ 4uw - 1 - 25C - \frac{1500}{K-2}. \tag{7.4}$$

Food drift. The function $\delta(\cdot)^2$ has derivative -2δ, which is 2-Lipschitz, so $\delta(y)^2 \le \delta(x)^2 - 2\delta(x)(y - x) + (y - x)^2$ for all x, y, including across $x = 1$. Summing over marks and using (E4),

$$\sum_j a_j \Delta_j \psi \ \ge\ 8\delta(u)\mathcal{L}u + 8\delta(w)\mathcal{L}w - 4 \cdot 2 \cdot \frac{96000}{V}.$$

By (E3) and $\delta(u)(1 - u) = \delta(u)^2$, $u\,\delta(u) \le \frac{1}{4}$: $8\delta(u)\mathcal{L}u \ge 8\delta(u)^2 - 46C$.

Completed square. Put $\alpha = \delta(u)$, $\beta = \delta(w)$. Since $u, w \ge 0$ we have $\alpha, \beta \in [0, 1]$, $u \ge 1 - \alpha \ge 0$ and $w \ge 1 - \beta \ge 0$, so $uw \ge (1 - \alpha)(1 - \beta)$ and

$$4uw + 8(\alpha^2 + \beta^2) \ \ge\ 4(1 - \alpha)(1 - \beta) + 8(\alpha^2 + \beta^2) = \frac{16}{5} + 5\Big(\alpha + \beta - \frac{2}{5}\Big)^2 + 3(\alpha - \beta)^2 \ \ge\ \frac{16}{5}. \tag{7.5}$$

The minimum is attained at $\alpha = \beta = 1/5$. [Lean: `StartupCount.food_product_deficit_sharp`] Adding the three parts, with $117C = 468\varepsilon + 624m + 1287\eta$ and $K \ge h$,

$$\sum_j a_j r_j \ \ge\ \frac{16}{5} - 1 - 117C - \frac{1500}{K-2} - \frac{768000}{V} \ \ge\ A - 624\,m.$$

[Lean, given (7.4): `StartupCount.count_log_occupation_sharp`]

Variance and jumps. On the selected birth $|\lambda_j| \le 1/K$, and on a loss $|\lambda_j| \le a/(K-a) \le a/(K-2)$. With $a^2 \le 2a$ and (E1)–(E2),

$$\sum_j a_j \lambda_j^2 \ \le\ \frac{121}{K} + \frac{2952\,K}{(K-2)^2},$$

and both terms decrease in $K > 2$. The function ψ is 8-Lipschitz in each argument on the nonnegative axes, so $(\Delta\psi)^2 \le 128[(\Delta u)^2 + (\Delta w)^2]$ and $\sum_j a_j(\Delta_j \psi)^2 \le 128 \cdot 2 \cdot 96000/V$. The two parts of the reward may be correlated; the weighted square inequality $(x + y)^2 \le \frac{101}{100}x^2 + 101y^2$ uses no independence and gives (3.3). [Lean: `StartupCount.weighted_square`] The same Lipschitz bound and $|\Delta u|, |\Delta w| \le 2/V$ give (3.4).

Range. We have $-8 \le \psi \le 0$ and $4K \le LV \le 11V$. For two states with $K \ge h \ge V/(3 \times 10^{18})$ the potential difference is at most

$$\log\frac{11V}{4h} + 8 \ \le\ \log\big(\tfrac{33}{4} \times 10^{18}\big) + 8 \ <\ 52,$$

because $e^{44} > 8.25 \times 10^{18}$. [Lean: `StartupCount.potential_range_exp`] □

For each window, rewards are disabled if the state at its start is outside $\mathcal{G}$ and are stopped just after the first exit. At an exit jump the pre-jump count is at least h and the post-jump count at least $h-2>0$, so every reward is well defined and bounded by b. The start of the second window is the actual random state at time 100.

8 Assembly

Proof of Theorem 3.1. Apply Theorem 4.2 on the windows $(1,100]$ and $(100,199]$ with $\mathcal{G}=\{L\leq 11, K\geq h\}$, $d=99$, $Y=Q/V$, $g=m$, $a=1/20$ and the event H of Section 6. Lemma 7.1 supplies the reward premises and $D=52$; (E5) supplies $q_Y=11n/V$ and $b_Y=n/V$. On H the localization is inactive on $[1,199]$. Outside an event of probability at most $E_G+E_K+2\mathcal{B}(z,99q,b)+2\mathcal{B}(\frac{1}{20},\frac{99\cdot 11n}{V},\frac{n}{V})$, both windows therefore satisfy

$$\frac{Q}{V} \;>\; \frac{99A-52-z}{624}-\frac{1}{20}=q_* \;\geq\; \frac{1}{10},$$

which is (M2). The event H gives (M1) and the corridor part of (M3), because $h\geq V/(3\times 10^{18})$. The number of food arrivals by time 199 is Poisson with mean $1194V$; the Chernoff bound at $1195V$ has exponent $V[1195\log(1195/1194)-1]>V/2500$, using $\log(1+x)\geq x-x^2/2$ at $x=1/1194$. This is the last term of (3.9). By Lemma 2.2, $J(199)\geq Q(1,100]+Q(100,199]>2q_*V$.

Monotonicity. As V increases, h and N_V are nondecreasing, so A increases while q and b decrease; $(N+1)2^{-N}$ is nonincreasing; and every clock exponent increases when n/V does not increase. The quota and every term of E therefore improve. The parameter η enters only through -1287η in A and through (5.2).

Instances. For $n\geq 4$ we have $18+11n\leq\frac{31}{2}n$. With $n/V\leq 10^{-8}$ the mass exponent exceeds 1008 and the export exponent exceeds 114; for $V\geq 2\times 10^{22}$ the food exponent exceeds 5×10^4, the feed exponent exceeds 8×10^{18}, and $E_K\leq 3\times 10^8\cdot 201\cdot 2^{-200}<10^{-49}$. At $V=2\times 10^{22}$, $\eta=10^{-5}$ and $z=48$, exact rational arithmetic gives

$$A>1.9620,\qquad q<0.4659,\qquad \frac{48^2}{2(99q+16b)}>24.97,\qquad q_*>0.10103.$$

Replacing each e^{-x} by the reciprocal of a positive Taylor partial sum gives $E<2.87\times 10^{-11}$. This proves (a). For (b), $\eta=10^{-4}$ and $z=37$ give $A>1.8462$, exponent above 14.84, $q_*>0.10028$ and $E<7.27\times 10^{-7}$. The remaining rows of Table 1 are evaluated in the same way by the replay script. [Lean for the margins and the budget of (a), (b): `StartupCount.robust_operating_point_margins`, `StartupCount.robust_operating_point_margins_wide`, `StartupCount.robust_error_budget`] □

Proof of Theorem 3.2. Apply Corollary 4.3 on the horizon $(1,199]$ with $d_0=198$, export tolerance $a=1/40$, and the same H; the range bound $D=52$ of Lemma 7.1 holds for every pair of times in $[1,199]$ on H. This is (3.12). With $n/V\leq 10^{-9}$ the export exponent exceeds 143. For (a), at $V=10^{24}$, $\eta=10^{-5}$, $z=14$: $A>2.182$, the exponent $14^2/[2(198q+14b/3)]$ exceeds 53, and $E_{\mathrm{all}}<1.7\times 10^{-23}$. For (b), at $V=10^{23}$ and $z=26$: $A>2.142$, the exponent exceeds 18.3 and $E_{\mathrm{all}}<2.2\times 10^{-8}$. Monotonicity in V is as before. □

Proof of Corollary 3.3. Put $q_{\mathrm{a}}(d)=(2.182d-80)/624-\frac{1}{20}$. Then $q_{\mathrm{a}}(99)>0.16$ and $q_{\mathrm{a}}(198)>0.51$, which give (a) and the first part of (b); $J(199)=m(199)V+Q(0,199]\geq Q(1,199]$ by (2.6). For (c), the mission bound $A_f(199)\leq 1195V$ gives $10V+B_f(199)\leq 2400V$, and $0.51/2400>1/4700$. For (d), $q_{\mathrm{a}}(d)>1/10$ as soon as $d>173.6/2.182$, which is below 79.6; two disjoint windows of that length fit in $(1,199]$. Even with the limiting slope $\frac{11}{5}$ and no reward allowance, a quota above 1/10 needs $d > 145.6/2.2 > 66$, so three disjoint guaranteed windows would need more than 198 residence times. The values for (b) of Theorem 3.2 follow from $(2.142d - 104)/624 - \frac{1}{20}$ in the same way. $\square$

Windows shorter than $(52 + 2z + 31.2)/A$ receive no guarantee; the drift must first pay for the range of the potential and the endpoint noise.

Proof of Corollary 3.4. As $V \to \infty$ with n fixed, q and b are $O(1/V)$ and $A \to A_\infty = \frac{11}{5} - 468\varepsilon - 1287\eta$. For $\eta \le 5 \times 10^{-4}$ the margin $\mu = 99A_\infty - 52 - 93.6$ is positive; in fact it is positive up to $\eta = 5.66 \times 10^{-4}$ and negative from 5.67×10^{-4}. Take $z = \mu/2$. For all large V, $q_*(V, \eta, z) > 1/10$, and $\mathcal{B}(z, 99q, b) \to 0$. Every other term of (3.9) tends to zero for fixed n, and (5.2) holds. $\square$

	two-window reward-error allowance ρ			
V	10^{-3}	10^{-6}	10^{-9}	10^{-12}
2×10^{22}	0.1355	0.1193	0.1067	0.0960
5×10^{22}	0.1725	0.1623	0.1543	0.1475
10^{23}	0.1875	0.1803	0.1746	0.1699
10^{24}	0.2069	0.2046	0.2028	0.2013

Table 4: Normalized quota per window from the inverse formula (4.6) with $k = 2$, $d = 99$, $\eta = 10^{-5}$, $a = 1/20$. These are descriptive floating-point evaluations of the analytic certificate, rounded down; guard, floor, export and feed errors are additional to ρ. Figure 3(b) shows the same trade-off.

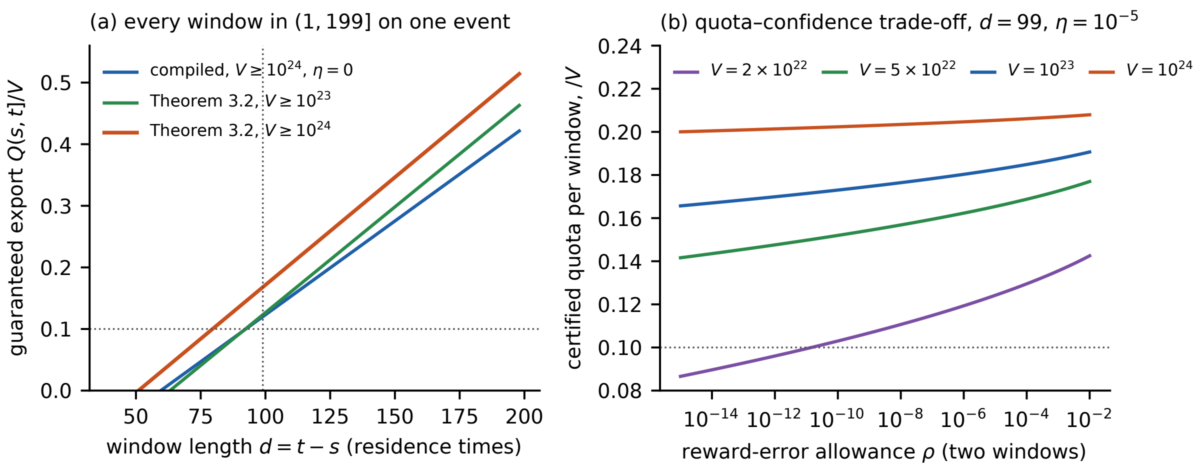

Figure 3: (a) Guaranteed export of every window of length d inside $(1, 199]$ on one event: the compiled food-silent quota q_{17} of Theorem 3.6, and the two instances of Theorem 3.2. (b) Certified quota per 99-residence-time window against the reward-error allowance, from (4.6).

Where the improvement over the compiled theorem comes from. Three changes are structural and three are constants. The cutoff of the logarithm is $h \propto V$ instead of the fixed $K_0 = 333\,334$, so q and b decay like $1/V$ and the failure bound is no longer floored; the Bernstein tilt $\theta = z/(dq + bz/3)$ replaces the fixed tilt $\theta = 1$ with the bound $e^y \le 1 + y + y^2$; and the generator envelopes are re-derived with food catalysis present, instead of coupling to a food-silent process.

The constants are the exact minimum $16/5$ in (7.5) in place of 3, the second-order logarithmic term $1500/(K-2)$ in place of $3000/(K-2)$, and the weights $(\frac{101}{100}, 101)$ and range 52 in place of $(\frac{17}{16}, 17)$ and 54. With the constants of the 20 September draft the certificate at $V = 2 \times 10^{22}$ has exponent below $\frac{1}{2}$; with the present constants it exceeds 24.9.

9 Source averaging and the scope of robustness

Proof of Corollary 3.5. On $\mathcal{A}_n$ the assignment lies in $\mathcal{U}(0)$, so Theorem 3.1(a) with $\eta = 0$ gives conditional success probability above $1 - 10^{-10}$ for every such environment and every mark law. Integrating over $\mathcal{A}_n$, whose mass is $e_\alpha^6 p_{\alpha,n}$, gives (3.13). The conditional confidence multiplies the witness mass; it is not an additive unconditioned error. Along $\alpha_n = 2 - 2/n$: $e_{\alpha_n} \to 6/\pi^2$; $\mathbb{E}D/n \to 9/\pi^2$, because $\sum_{k \le R} k^{1-\alpha_n} = \frac{n}{2}(R^{2/n} - 1) + O(1)$ with $R_n^{2/n} \to 4$ and the capped tail contributes $O(1)$; and $R_n/(n2^n) \to 2$. Hence $2^n p_{\alpha_n,n} \to 9/(2\pi^2)$, which gives (3.14). The normalized source limit is formal, `startup_witness_normalized_limit`; its combination with the conventional bound (3.13) is conventional.

At $n = 4$ and $\alpha = 2$: $R_4 = 68$, $e_2 = 6/\pi^2$ and

$$\mathbb{E}D = 67 - e_2 \sum_{k=1}^{67} \frac{68-k}{k^2} \approx 2.5239961939, \qquad e_2^6 \, \frac{\mathbb{E}D}{68} \approx 0.0018736515.$$

Rational enclosures of π from Machin's formula certify a witness mass above 187365×10^{-8} and $(1 - 187365 \times 10^{-8}(1 - 10^{-10}))^{1600} < 1/20$. $\square$

The screening statement assumes that each trial independently redraws chemistry, coefficients and trajectory. Repeated runs in one sampled chemistry are conditionally independent given that chemistry but dependent after averaging, and the calculation does not apply to them. These are lower bounds for a screening design, not predicted success frequencies.

What is not transferred. The companion converse $\mathbb{P}(F_{n,V}) \le 224\, p_{\alpha,n} + \bar{r}$ [3] has a residual containing $Ce^{-cV_n/n}$, which is negligible relative to 2^{-n} only when $cV_n/n - n\log 2 \to \infty$. A scale linear in n does not guarantee this, so neither the converse, nor singleton selection, nor the order $\Theta(2^{-n})$ is claimed at the new scale. Failure of this transfer is not evidence that the converse fails; it is simply not proved. A lower operating certificate also says nothing about other productive mechanisms.

Robustness is pointwise, not averaged. For $\eta > 0$, Theorem 3.1 concerns a modified kinetic family: any witness environment may be extended by arbitrary food incidences of strength at most η. The capped-Zipf source assigns coefficients in $[4, 16]$ to all of its incidences, food rows included, and $4 \gg \eta$. The theorem therefore does not remove the factor e_α^6 from (3.13). An averaged theorem without that factor would need a source with weak food incidences and a computation of its witness mass; none is claimed.

Comparison with a coupling argument. A natural route to robustness couples the modified process to the food-silent one through the random time-change representation [4, 12] and asks that no added channel fires. On the corridor the added channels have total rate at most $1452\,\eta V$, so an event $F \subseteq \{L \le 11 \text{ on } [0,T]\}$ satisfies $\mathbb{P}^\eta(F) \ge e^{-1452\eta VT}\mathbb{P}(F)$. At $V = 10^{24}$ and $T = 199$ a loss below 10^{-3} needs $\eta \le 3.5 \times 10^{-33}$: the admissible perturbation shrinks like $1/V$, whereas a macroscopic reactor fires enormously many weak reactions. The present argument lets all of them fire and pays the finite drift cost 1287η in (3.2), together with 11η in every envelope. This structural difference, not a better coupling constant, is the robustness statement. At fixed V and d the certified quota decreases by $1287\, d\, \Delta\eta/624$ under an increase $\Delta\eta$.

10 Internal equilibrium structure

Proof of Proposition 3.7. (a) For every split, $x^*_v x^*_w = x^*_{vw}$, because letter counts add under concatenation. A basal pair with coefficient b_r has forward flux $b_r x^*_v x^*_w$ and reverse flux $b_r x^*_{vw}$; a pair catalysed by c has the same two fluxes multiplied by $k_{rc} x^*_c$. They agree, whatever the coefficients.

(b) Let $\nu \rightleftharpoons \nu'$ be a reversible pair of input multiplicity vectors with common coefficient κ; a catalyst appears in both ν and ν'. Using $(N_z)_{\nu_z}/N_z! = 1/(N_z - \nu_z)!$ for $N_z \geq \nu_z$,

$$\pi(N)\, a_\nu(N) \;\propto\; \kappa V\, (x^*)^\nu \prod_z \frac{(V x^*_z)^{N_z - \nu_z}}{(N_z - \nu_z)!}, \qquad (x^*)^\nu = \prod_z (x^*_z)^{\nu_z},$$

and both sides vanish when some $N_z < \nu_z$. The reverse channel fires from $N' = N - \nu + \nu'$ with the same expression in which $(x^*)^\nu$ is replaced by $(x^*)^{\nu'}$ and $N - \nu$ by $N' - \nu' = N - \nu$. By (a), $(x^*)^\nu = (x^*)^{\nu'}$. Hence $\pi(N) a_\nu(N) = \pi(N') a_{\nu'}(N')$ for every pair, with falling factorials and repeated roles included. Internal reactions conserve the numbers of letters 0 and 1, so every closed class is finite, π is normalizable there, and pairwise balance implies stationarity. □

The direct argument is consistent with the product-form theory for complex-balanced stochastic networks [5] and with classical reversibility [24]; it does not require the network to have deficiency zero.

Interpretation. It is sometimes said of schematic polymer models that no thermodynamic constraint is imposed. For this model that is too strong. Sharing one dimensionless coefficient between the two directions of each pair fixes the equilibrium constant of every ligation, in units of the reference concentration, to one, and makes the two-parameter family x^* an equilibrium of every admitted chemistry. Arbitrary catalytic backgrounds change kinetic access and rates while preserving this internal equilibrium family; the uncertainty class is in this sense thermodynamically consistent, and also restricted, since it excludes letter-dependent bond energies. The open reactor is driven by feed and removal alone. A thermodynamic realization of the maintained reservoirs, their chemical potentials, finite capacities, work and dissipation is absent here [29], and the ideal internal equilibrium structure is not an empirical validation of any polymer chemistry.

11 Dimensions, a designed example and readout

Dimensions. For a physical volume Ω in litres, a reference concentration c_* in mol/L and a residence time τ,

$$V = N_A\, \Omega\, c_*, \qquad t_{\text{phys}} = \tau t, \qquad k_{\text{phys}} = \frac{\kappa}{\tau\, c_*^{|\nu|-1}}, \tag{11.1}$$

for a channel with $|\nu|$ input roles, catalyst included, where $N_A = 6.02214076 \times 10^{23}\,\text{mol}^{-1}$ exactly [8]. The encounter convention of (2.2) must match the convention used for k_{phys}. Each food is supplied at c_*/τ per unit volume and time, and dilution is $1/\tau$. Basal ligation has two input roles and cleavage one; catalysis adds one role in each direction. Equal dimensionless coefficients in the two directions therefore do not mean equal dimensional rate constants.

Element	Specification (residence-time and reference-concentration units)
Every split $v + w \rightleftharpoons vw$	basal coefficient in $[2, 8] \times 10^{-9}$
Required $00 + 11 + 0011 \rightleftharpoons 0011 + 0011$	catalytic coefficient in $[4, 16]$
Other nonfood catalytic incidences	any subsets; each present coefficient in $[4, 16]$
Food catalytic incidences	any subsets; coefficients in $[0, 10^{-5}]$
Feed and washout	six feed clocks of rate V; washout rate one per molecule
Preparation	V copies of each food; no nonfood
Count scale and collection	$V = 2 \times 10^{22}$; two windows of 99 residence times

Table 5: A designed uncertainty class at $n = 4$: 30 words and 68 reversible split identities. The certificate tolerates every unknown incidence pattern within the class. This is a reproducible model specification, not a claim that binary labels identify a characterized chemical system.

A designed example. For the class of Table 5 with the illustrative choices $c_* = 1\,\mathrm{mM}$ and $\tau = 1\,\mathrm{h}$, the count scale 2×10^{22} corresponds to $33.2\,\mathrm{L}$ and the mission lasts $199\,\mathrm{h}$. The selected forward catalytic coefficient becomes $[1111, 4444]\,\mathrm{M}^{-2}\mathrm{s}^{-1}$ and the reverse $[1.11, 4.44]\,\mathrm{M}^{-1}\mathrm{s}^{-1}$; basal ligation is $[0.56, 2.2] \times 10^{-9}\,\mathrm{M}^{-1}\mathrm{s}^{-1}$. Each window collects more than $0.101V/N_A \approx 3.35\,\mathrm{mmol}$ of monomer equivalents. At $V = 5 \times 10^{22}$, that is $83.0\,\mathrm{L}$, each window collects more than $0.135V/N_A \approx 11.2\,\mathrm{mmol}$, and at $V = 10^{24}$, that is $1661\,\mathrm{L}$, the whole-horizon guarantee $0.51V$ is about $0.85\,\mathrm{mol}$. Meanwhile the guaranteed selected stock is a relative concentration of only $3.3 \times 10^{-19} c_*$: the requirement controls quantities on very different scales, and the export must not be read as the concentration of the selected tetramer. Raising c_* to shrink Ω changes the physical rate constants required by a fixed dimensionless model through (11.1) and may invalidate dilute mass action; it is not a free engineering improvement. The count guarantee is not a small-volume design claim.

Readout. An ideal assay assigning response $|z|$ to every washed-out nonfood molecule reads exactly Q. If the nonnegative response weights satisfy $w_z \geq \gamma|z|$ for all nonfood z and some $\gamma > 0$, and the total negative measurement error is at most eV, then a certified quota q_*V yields a measured signal above $(\gamma q_* - e)V$; for example $\gamma = 0.95$, $e = 0.002$ and $q_* = 0.135$ give $0.12625V$. This is a deterministic adapter requiring those calibration bounds. A selective assay that ignores some exported nonfood species has $\gamma = 0$, and the adapter is then uninformative: aggregate output is not functional yield. For an actual installation, measured reaction orders, coefficient intervals, mixing validity and assay bounds remain missing inputs.

12 What sets the remaining scale

Anatomy of 2×10^{22}. The mission demands only $V/(3 \times 10^{18})$ copies of a tetramer. The certificate uses that stock as the cutoff h of its logarithm. Its two costs are the second-order drift term $1500/(h-2)$ and the variance $q \approx 3100/h$ per residence time. Requiring $99A - 52 - 93.6 = z$ with $z^2/(2 \cdot 99q) \gtrsim 25$ gives $h \approx 6.7 \times 10^3$, that is $V \approx 2 \times 10^{22}$. The replay script records where the certificate stops: at $V = 1.5 \times 10^{22}$ it still proves the mission with failure below 2×10^{-6} for $\eta = 0$; at $V = 10^{22}$ the best exponent is 4.1; at $V \leq 5 \times 10^{21}$ the drift margin is negative; and below $V = 6 \times 10^{18}$ the cutoff is at most two and the logarithmic reward is undefined. These are failures of a sufficient certificate, not counterexamples to reactor success.

The weak link. The count floor of Lemma 6.2 balances basal immigration, at food concentrations allowed to fall to $\ell \approx 7.2 \times 10^{-4}$, against the uniform loss ceiling $1476K$. It never uses selected reproduction, so it certifies only the tiny relative stock $V/(2 \times 10^{18})$. A cutoff larger than h would improve everything: with a floor H in place of h, the correction becomes $1500/(H-2)$, the variance $\approx 3100/H$ and the range $\log(11V/(4H)) + 8$. Declaring such a floor does not help. It creates the unproved implication

$$\mathbb{P}\{\tau_H \leq s,\ K_v \geq H \text{ for } s \leq v \leq 199,\ \text{guard on } [0, 199]\} \ \geq\ 1 - \delta_H, \tag{12.1}$$

for an establishment time τ_H, uniformly over the background. A floating-point planning calculation at $V = 10^{20}$ gives conditional quotas near 0.15 for $H = 10^4$ and 0.18 for $H = 2 \times 10^4$ at reward error 10^{-6}, whereas the basal cut there is only $m_0 = 50$. It shows what (12.1) would buy; it is not a theorem about the reactor started from food.

Why the loss envelope is pessimistic. By the shared coefficients behind Proposition 3.7, every catalysed cleavage of 0011 is paired with a catalysed ligation. For the split $(00, 11)$ and a catalyst c the two rates are $k\, x_c K$ and $k\, x_c V\, uw$, so the net catalysed flux through this split is $k\, x_c V (uw - x_{0011})$, which is *into* 0011 whenever $x_{0011} < uw$; under the guard and at the tiny stocks in question this always holds. Envelope (E1) charges the loss, and the reward discards the paired gain as an unused favourable birth. A background catalyst of the split $(00, 11)$ is therefore a net producer of 0011 during establishment, not an adversary. The genuinely adverse channels are the two trimer splits $0|011$ and $001|1$ when the trimers are kept depleted, and, for $n \geq 5$, ligations $0011 + y \to 0011y$ whose products are depleted. We record this as a heuristic, not a result: an establishment estimate that exploits it needs a multi-species potential that tracks those reactants.

The exact open implication. A proof of (12.1) must exploit selected reproduction during establishment while charging high background activity to collected output, including on paths that never establish: either m is small, in which case C is small, food is near one and the selected channel has net per-copy growth near $4uw - 1 \approx 3$, or m is large and output is already being produced. At small counts the logarithmic reward has variance of order $3000/K$ per residence time, so a branching comparison, not a Bernstein bound for $\log K$, is needed below $K \approx 10^3$. We state this as the next target and do not use simulation to declare it proved.

Necessary versus sufficient. In every food-silent environment the mission probability is at most $1 - \exp[-(484/3)\varepsilon V]$ [3], so 99% reliability needs $V \geq 1.43 \times 10^7$, about 24 pL at 1 μM. The sufficient scale 2×10^{22} covers far more than first birth: stock persistence, food and mass control, and two export windows. The ratio 1.4×10^{15} is unresolved sharpness. The first-birth bound is not a converse for $\mathcal{U}(\eta)$ with $\eta > 0$, where food-catalysed reactions create additional first births. Three scopes should be kept apart: a *uniform* theorem covers every admitted background, as here; a *designed-network* theorem may use a specified smaller reaction set and measured harmful-flux bounds, which can be far better than the catalogue envelope; and a *typical-source* theorem must control the probability of adverse random backgrounds before averaging. Progress in one scope does not prove another.

Diagnostics. Small computations shaped the proof and are not premises. A literal-generator panel at $n = 4$ with every food row full at coefficient 10^{-5} and nonfood rows of degree 67, at floor, active-trimer, depleted-food and large-count states, together with a repeated-role check at $n = 8$, satisfied every envelope of Lemma 5.1 and gave reward variances between 3×10^{-4} and 2.2×10^{-3}

against the uniform bound in force at that time, 0.196; the conservatism is real but does not prove a smaller uniform bound. At $n = 8$ and $K = 333\,334$ with all foods at V, the retained variance was 1.5×10^{-5} whereas favourable non-selected births contributed $\sum a(\Delta \log K)^2 \approx 1.8 \times 10^4$ to a symmetric budget: discarding them is what makes the certificate possible. In a chain on $\{0, \ldots, 30\}$ with two-copy losses, omitting the second boundary state undercounts crossings by three orders of magnitude.

13 Conclusion

The service certificate links a local inequality for a marked generator to a finite-time collected output. Its efficiency comes from retaining adverse count fluctuations while letting unused favourable births help the telescoping potential, and from charging harmful catalytic activity to the occupation that generates effluent. In the polymer realization this gives, for every n, a mission theorem at $V \geq \max\{2 \times 10^{22}, 10^8 n\}$ with failure below 10^{-10}, structural tolerance of weak food catalysis, quotas for every window on one event, failure tending to zero for fixed n, and a source-averaged rarity bound at a scale linear in n. The compiled theorem at 10^{24} remains the statement checked end to end by machine; the new stochastic assembly is conventional, with its algebraic and scalar components compiled.

The method transfers as a proof contract: (1) specify the chemistry or its source, the preparation and the coefficient uncertainty; (2) specify stock, output, deadlines, feed and losses as one trajectory event; (3) separate the probability of acquiring the chemistry from the probability of completing the task given it; (4) choose an observable whose fluctuations are exactly the ways the task can fail; (5) find the occupation that both bounds the harmful activity and compensates the output; (6) verify floor, drift, variance, jump and range for the actual generator; (7) report conditional and averaged guarantees with their units and unvalidated assumptions.

The next mathematical advance is a joint establishment–output estimate such as (12.1); the next empirical advance is a calibrated reaction and assay model. Further open problems are a converse and selection at scales linear in n; collective witnesses, where a mutually catalytic pair or cycle replaces the self-catalytic ligation and the floor and reward need multi-species potentials; a formal mission root for the modified family; and an energetic account of the maintained feed. Neither a sharp startup threshold nor a functional-product guarantee follows from the present certificate.

Data and code availability. The repository location will be supplied with the author information. The source package contains this manuscript, its bibliography, the figure script, the exact-arithmetic replay script and its output, the new Lean module with its strict verification receipt, and the verification manifest of the earlier formal development (Appendix A).

A Evidence map and reproducibility

The formal development is a single Lean 4 project pinned to `leanprover/lean4:v4.30.0` with a pinned Mathlib manifest. Verification uses a wrapper that treats every warning as an error and records a strict receipt with the source hash. Declaration probes report only the standard axioms `propext`, `Classical.choice` and `Quot.sound`; no admitted proof or additional axiom occurs. In Theorem 3.6 the physical process is the literal continuous-time jump law with the propensities (2.2), all channels present, and the source is the capped-Zipf configuration weight with a uniform finite mark law. Table 6 maps statements to evidence. Declarations live under `proofs/StartupMarked/`, proofs/StartupCount/, proofs/RepeatedFunction/ and proofs/RandomViability/; the module added for this paper is proofs/StartupCount/RobustSharpening.lean.

Table 6: Statements of this paper and their evidence.

Statement	Evidence	Declarations or remark
Theorem 3.6(a)–(c)	Lean	StartupMarked.startup_quantitative_resolution (root; uses 9989/10000, the exact witness mass and the signed synthesis account)
Theorem 3.6(d)	Lean	StartupMarked.startup_enhanced_failure, StartupMarked.source_window_export_quota, RandomViability.physical_enhanced_accounts
Lemma 2.2	Lean	physical_mission_accounts, actual_net_synthesis_identity
Lemma 4.1, Theorem 4.2, Corollary 4.3	conventional	proof in Section 4; the fixed-tilt analogue is Lean (gamma_reward_interval_tail)
Lemma 5.1	conventional	food-silent analogues are Lean (unbounded_total_rate_mass_eleven, physical_crossing_rate_le); the 11η extension is proved in the text
Lemma 6.1	conventional	food-silent analogue Lean (count_guard_from_noise, physical_mass_localization, physical_food_floor_from_noise)
Lemma 6.2	conventional + Lean	tilt, killed tail and crossing: food-silent analogues Lean (physical_power_affine_generator, source_count_tail, source_count_crossing_bound); rounding for $V \geq 10^{22}$: StartupCount.count_rounding_bounds_small_volume
Inequality (7.3)	Lean	retained_reward_le_potential_increment, retained_path_sum_le
Lemma 7.1	conventional + Lean	StartupCount.food_product_deficit_sharp, StartupCount.count_log_occupation_sharp, StartupCount.weighted_square, StartupCount.potential_range_exp; logarithm series bound and generator sums conventional
Theorem 3.1	conventional + Lean	assembly in Section 8; margins and budget of (a), (b): StartupCount.robust_operating_point_margins, StartupCount.robust_operating_point_margins_wide, StartupCount.robust_error_budget; Table 1: exact arithmetic
Theorem 3.2, Corollaries 3.3, 3.4	conventional	exact arithmetic for all printed values

Statement	Evidence	Declarations or remark
Corollary 3.5	conventional + Lean	`startup_witness_normalized_limit`, `startup_normalized_source_lower` (at the compiled scale); Machin enclosure: exact arithmetic
Proposition 3.7	conventional	proof in Section 10
Table 4, Figures 2, 3	descriptive	floating-point evaluation of explicit formulas
Section 12 diagnostics	numerical	never used as premises

Reproduction. The script `check_paper.py` (standard library only, a few seconds) replays every printed number. In exact rational arithmetic it checks the provenance of the constants in Table 3, the completed-square identity, the tilt margin and rounding of the floor, the potential range, the feed exponent, every row of Table 1 with monotonicity spot checks at larger and non-round volumes, both instances of Theorem 3.2 and the values in Corollary 3.3, the limits in Corollary 3.4, the source enclosure and the screening bound; it also replays the numbers of the 20 September draft and evaluates the descriptive tables and unit conversions. The script `make_figures.py` draws both figures. Formal verification of the new module uses the repository wrapper with an explicit declaration, for example `StartupCount.robust_error_budget` in `proofs/StartupCount/RobustSharpening.lean`; its receipt and the receipts of `StartupMarked.startup_quantitative_resolution` and `StartupCount.count_rounding_bounds_small_volume` are distributed with the package. No dependency update is performed.

References

[1] Author name to be inserted. Critical-window emergence of autocatalytic sets under reaction-channel quotienting. Companion manuscript, 2026; arXiv preprint in preparation, 2026.

[2] Author name to be inserted. A nontrivial critical window for RAF emergence in the binary polymer model. Companion manuscript, 2026; arXiv preprint in preparation, 2026.

[3] Author name to be inserted. Two consecutive productive windows in random polymer reactors: startup, resource bounds, and selection by successful operation. Companion manuscript, 2026; arXiv preprint in preparation, 2026.

[4] D. F. Anderson and T. G. Kurtz. Continuous time Markov chain models for chemical reaction networks. In H. Koeppl, G. Setti, M. di Bernardo, and D. Densmore, editors, *Design and Analysis of Biomolecular Circuits*, pages 3–42. Springer, New York, 2011. doi: 10.1007/978-1-4419-6766-4_1.

[5] D. F. Anderson, G. Craciun, and T. G. Kurtz. Product-form stationary distributions for deficiency zero chemical reaction networks. *Bulletin of Mathematical Biology*, 72(8):1947–1970, 2010. doi: 10.1007/s11538-010-9517-4.

[6] A. Blokhuis, D. Lacoste, and P. Nghe. Universal motifs and the diversity of autocatalytic systems. *Proceedings of the National Academy of Sciences*, 117(41):25230–25236, 2020. doi: 10.1073/pnas.2013527117.

[7] P. Brémaud. *Point Processes and Queues: Martingale Dynamics*. Springer Series in Statistics. Springer, New York, 1981. doi: 10.1007/978-1-4684-9477-8.

[8] Bureau International des Poids et Mesures. *The International System of Units (SI)*. BIPM, Sèvres, 9th edition, 2019. Definition of the mole, https://www.bipm.org/en/publications/si-brochure.

[9] R. W. R. Darling and J. R. Norris. Differential equation approximations for Markov chains. *Probability Surveys*, 5:37–79, 2008. doi: 10.1214/07-PS121.

[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 Computer Science*, pages 625–635. Springer, 2021. doi: 10.1007/978-3-030-79876-5_37.

[11] K. Dzhaparidze and J. H. van Zanten. On Bernstein-type inequalities for martingales. *Stochastic Processes and their Applications*, 93(1):109–117, 2001. doi: 10.1016/S0304-4149(00)00086-7.

[12] S. N. Ethier and T. G. Kurtz. *Markov Processes: Characterization and Convergence*. Wiley, New York, 1986. doi: 10.1002/9780470316658.

[13] J. D. Farmer, S. A. Kauffman, and N. H. Packard. Autocatalytic replication of polymers. *Physica D: Nonlinear Phenomena*, 22(1–3):50–67, 1986. doi: 10.1016/0167-2789(86)90233-2.

[14] A. Filisetti, A. Graudenzi, R. Serra, M. Villani, D. De Lucrezia, R. M. Füchslin, S. A. Kauffman, N. Packard, and I. Poli. A stochastic model of the emergence of autocatalytic cycles. *Journal of Systems Chemistry*, 2:2, 2011. doi: 10.1186/1759-2208-2-2.

[15] D. A. Freedman. On tail probabilities for martingales. *The Annals of Probability*, 3(1):100–118, 1975. doi: 10.1214/aop/1176996452.

[16] D. T. Gillespie. Exact stochastic simulation of coupled chemical reactions. *The Journal of Physical Chemistry*, 81(25):2340–2361, 1977. doi: 10.1021/j100540a008.

[17] V. Giri and S. Jain. The origin of large molecules in primordial autocatalytic reaction networks. *PLoS ONE*, 7(1):e29546, 2012. doi: 10.1371/journal.pone.0029546.

[18] W. Hordijk and M. Steel. Detecting autocatalytic, self-sustaining sets in chemical reaction systems. *Journal of Theoretical Biology*, 227(4):451–461, 2004. doi: 10.1016/j.jtbi.2003.11.020.

[19] W. Hordijk and M. Steel. Autocatalytic sets extended: dynamics, inhibition, and a generalization. *Journal of Systems Chemistry*, 3:5, 2012. doi: 10.1186/1759-2208-3-5.

[20] W. Hordijk and M. Steel. Autocatalytic sets in polymer networks with variable catalysis distributions. *Journal of Mathematical Chemistry*, 54(10):1997–2021, 2016. doi: 10.1007/s10910-016-0666-z.

[21] W. Hordijk and M. Steel. Chasing the tail: the emergence of autocatalytic networks. *Biosystems*, 152:1–10, 2017. doi: 10.1016/j.biosystems.2016.12.002.

[22] W. Hordijk, S. A. Kauffman, and M. Steel. Required levels of catalysis for emergence of autocatalytic sets in models of chemical reaction systems. *International Journal of Molecular Sciences*, 12(5):3085–3101, 2011. doi: 10.3390/ijms12053085.

[23] S. A. Kauffman. Autocatalytic sets of proteins. *Journal of Theoretical Biology*, 119(1):1–24, 1986. doi: 10.1016/S0022-5193(86)80047-9.

[24] F. P. Kelly. *Reversibility and Stochastic Networks*. Wiley, Chichester, 1979.

[25] T. G. Kurtz. The relationship between stochastic and deterministic models for chemical reactions. *The Journal of Chemical Physics*, 57(7):2976–2978, 1972. doi: 10.1063/1.1678692.

[26] E. Mossel and M. Steel. Random biochemical networks: the probability of self-sustaining autocatalysis. *Journal of Theoretical Biology*, 233(3):327–336, 2005. doi: 10.1016/j.jtbi.2004.10.011.

[27] J. R. Norris. *Markov Chains*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, 1997. doi: 10.1017/CBO9780511810633.

[28] Z. Peng, J. Linderoth, and D. A. Baum. The hierarchical organization of autocatalytic reaction networks and its relevance to the origin of life. *PLoS Computational Biology*, 18(9):e1010498, 2022. doi: 10.1371/journal.pcbi.1010498.

[29] R. Rao and M. Esposito. Nonequilibrium thermodynamics of chemical reaction networks: wisdom from stochastic thermodynamics. *Physical Review X*, 6(4):041064, 2016. doi: 10.1103/PhysRevX.6.041064.

[30] A. Ravoni. Impact of composition on the dynamics of autocatalytic sets. *Biosystems*, 198: 104250, 2020. doi: 10.1016/j.biosystems.2020.104250.

[31] M. Steel. The emergence of a self-catalysing structure in abstract origin-of-life studies. *Applied Mathematics Letters*, 13(3):91–95, 2000. doi: 10.1016/S0893-9659(99)00191-3.

[32] M. Steel, W. Hordijk, and J. C. Xavier. Autocatalytic networks in biology: structural theory and algorithms. *Journal of the Royal Society Interface*, 16(151):20180808, 2019. doi: 10.1098/rsif.2018.0808.

[33] 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.

[34] Y. Togashi and K. Kaneko. Transitions induced by the discreteness of molecules in a small autocatalytic system. *Physical Review Letters*, 86(11):2459–2462, 2001. doi: 10.1103/PhysRevLett.86.2459.

[35] J. Unterberger and P. Nghe. Stoechiometric and dynamical autocatalysis for diluted chemical reaction networks. *Journal of Mathematical Biology*, 85:26, 2022. doi: 10.1007/s00285-022-01798-0.

[36] S. van de Geer. Exponential inequalities for martingales, with application to maximum likelihood estimation for counting processes. *The Annals of Statistics*, 23(5):1779–1801, 1995. doi: 10.1214/aos/1176324323.

[37] V. Vasas, C. Fernando, M. Santos, S. Kauffman, and E. Szathmáry. Evolution before genes. *Biology Direct*, 7:1, 2012. doi: 10.1186/1745-6150-7-1.