

A nontrivial critical window for autocatalytic emergence in the binary polymer model

16 October 2026

Abstract

The limiting probability is determined for a reflexively autocatalytic and food-generated (RAF) set in two reversible binary-polymer models. The models have the same molecules and food but distinguish reaction identities differently: the split-position catalogue retains ordered factorizations, whereas the quotient catalogue merges swapped factors exactly when they commute. With independent molecule–reaction catalysis, one mark shared by both directions, and mean catalysis f_n per molecule satisfying $f_n/n \rightarrow \lambda > 0$, the RAF probability converges in either model to

$$\Theta_\nu(\lambda) = S_\nu(1 - e^{-\lambda}), \quad \nu \in \{\text{sp}, \text{qt}\},$$

where S_ν is the survival probability of infinite ordinary reversible food closure in the corresponding independent open-reaction field. Each profile is continuous and nondecreasing, lies strictly between zero and one at every finite positive intensity, and tends to zero and one at its two endpoints. Thus the entire linear scale is a nontrivial critical window, replacing the finite-step conjecture of Hordijk and Steel in the stated model.

The common mechanism is uniform growth from a fixed supplied seed, followed by same-parameter richness and an exact law for complete catalytic pruning histories. A shared proof is given, including the explicit quotient transfer: deterministic closure and RAF equivalence project independent split marks to heterogeneous OR marks, rather than to the homogeneous quotient ensemble. The profiles satisfy $\Theta_{\text{sp}}(\lambda/2) \leq \Theta_{\text{qt}}(\lambda) \leq \Theta_{\text{sp}}(\lambda)$; productive food gateways give the conventional bounds $\Theta_{\text{sp}}(\lambda) \leq 1 - e^{-32\lambda}$ and $\Theta_{\text{qt}}(\lambda) \leq 1 - e^{-30\lambda}$. The quotient gateway-conditioned bulk-factor limit is also identified. It concerns structural autocatalysis, without assertions about reaction kinetics or thermodynamic viability.

Keywords: RAF theory; binary polymer model; reaction quotient; reversible closure; critical window; product probability; formal verification.

1 The literature question and the two reaction ensembles

Random polymer models offer a precise setting for asking how much catalysis permits a chemistry to support itself. The structural object is a reflexively autocatalytic and food-generated reaction set, or RAF: its molecules can be generated from food using its reactions, and each reaction has a catalyst in the resulting closure. This separates availability and collective catalytic support from the time evolution of a chemical reactor. The framework originates in work of Kauffman and Steel and the RAF detection algorithm of Hordijk and Steel [4, 8, 2].

Mossel and Steel [7] identified linear mean catalysis as the relevant scale. Hordijk and Steel [3, Section 7, Eq. (21)] subsequently proposed a finite constant γ separating limiting RAF probability zero and one within that scale. For the exact-linear sequence $f_n = \lambda n$, their proposed law would give zero below γ and one above it. Their experiments use the six monomers and dimers as food and usually count a ligation and its reverse cleavage together. The theorem

here fixes these choices explicitly and shows that every positive finite λ instead has a limiting probability strictly between zero and one.

Reaction identity is essential to that statement. A split description is an ordered pair (u, v) of nonempty words, representing $u + v \rightleftharpoons uv$. A product word together with a split position retains all these descriptions. The public generator accompanying Varanasi and Korenaga [12, 11] suppresses duplicate displayed reactions. In the convention studied here, (u, v) and (v, u) merge only if $uv = vu$. For example, $A + AA \rightleftharpoons AAA$ and its swapped description merge, but $A + B \rightleftharpoons AB$ and $B + A \rightleftharpoons BA$ have different products and remain distinct.

The split-position and quotient laws share a supplied-seed growth estimate and a common probability argument. Both are developed in this self-contained account, with ordinary RAF existence defined separately from the reaction catalogue and catalytic sampling law. The gateway-conditioned limit follows from the exact gateway counts and the RAF limit (Section 8). The two reaction ensembles remain distinct throughout.

The deterministic quotient is exact, but the probability laws differ. Giving two descriptions independent catalytic opportunities and then merging them produces an OR probability $2p - p^2$. Merging first and assigning one Bernoulli(p) coordinate produces p . The relatively small number of deleted descriptions cannot justify identifying those ensembles. Section 3 proves the correct comparisons; Sections 4–7 establish the common limit without such an identification.

2 Model, normalization, and principal conclusions

Let $\Sigma = \{A, B\}$ and

$$X_n = \bigcup_{j=1}^n \Sigma^j, \quad M_n = |X_n| = 2^{n+1} - 2, \quad F = X_2.$$

All finite models below have $n \geq 2$; arbitrary initial indices can be totalized without changing any limit. The food set has six molecule types. Availability is set-valued: a word type can be used again after it becomes available. No molecular counts are attached to closure.

Let R_n consist of all ordered pairs (u, v) with $|u| + |v| \leq n$. Define J_n by identifying pairs with the same product and the same unordered pair of factors. A fibre of the quotient map $\pi : R_n \rightarrow J_n$ has size one or two. Write

$$\mathcal{R}_n^{\text{sp}} = R_n, \quad \mathcal{R}_n^{\text{qt}} = J_n, \quad N_n^\nu = |\mathcal{R}_n^\nu|.$$

Each identity denotes a reversible reaction. For a reaction family A and supplied molecules B , $\mathcal{C}_n^\nu(A; B)$ is the least subset of X_n containing B such that an open reaction $u + v \rightleftharpoons uv$ permits both rules

$$u, v \text{ available} \implies uv \text{ available}, \quad uv \text{ available} \implies u, v \text{ available}.$$

In particular, $u = v$ requires availability of that type, not two counted copies.

Definition 2.1 (Catalysis and RAF). Independently for every $(x, r) \in X_n \times \mathcal{R}_n^\nu$, set $\xi_{x,r} = 1$ with probability p . One coordinate serves both orientations. A nonempty family A is a RAF when every reaction in A has one complete side in $\mathcal{C}_n^\nu(A; F)$ and a catalyst x in that closure with $\xi_{x,r} = 1$. Once one complete side is in closure, both sides are; thus this is equivalently the usual all-endpoints food-generation condition. Let $P_n^\nu(p)$ be the probability that some RAF exists.

Food molecules can catalyze reactions, and a RAF whose entire closure is food is allowed. Catalysts need not appear before the first use of their reactions in a food derivation. This is ordinary structural RAF semantics.

The repository implementation stores ligation and cleavage as adjacent directed records, but its uniform sampler draws once for each pair and assigns the same catalyst to both. Its normalization divides the number of directed records by two [11]. These observations identify the reaction catalogue and catalytic sampling law. The RAF event in this paper is independently defined by ordinary food closure, including all food catalysts; no identification with the output of a repository detection routine is assumed. The channel convention above makes the shared-coordinate choice explicit. Independent marks on the two orientations would define a different model. The generator's special food choice at $n = 2$ is not used here; the fixed six-word food convention applies throughout, and that single initial index has no effect on the limits.

The mean number of reaction identities catalyzed by one molecule is $f_n = pN_n^\nu$. For an arbitrary real sequence f_n , define $p_n^\nu = \min(1, \max(0, f_n/N_n^\nu))$. We prove eventual removal of this clamp rather than imposing a sign condition on an irrelevant initial segment.

For $\nu \in \{\text{sp}, \text{qt}\}$, independently open each reaction identity of $\mathcal{R}_\infty^\nu = \bigcup_n \mathcal{R}_n^\nu$ with probability a . Let $\mathcal{C}_\infty^\nu$ be ordinary reversible closure of F in this field, and put

$$S_\nu(a) = \mathbb{P}_a^\nu\{\mathcal{C}_\infty^\nu \text{ has unbounded word length}\}, \quad \Theta_\nu(\lambda) = S_\nu(1 - e^{-\lambda}). \quad (1)$$

Membership has a finite derivation, so it is measurable in the countable product field. The quotient respects every finite cap because it preserves the product word.

Theorem 2.2 (The two critical-window laws). *For either ν and every real sequence f_n with $f_n/n \rightarrow \lambda \in (0, \infty)$, the parameter is eventually exactly $p_n^\nu = f_n/N_n^\nu$ and*

$$P_n^\nu(p_n^\nu) \rightarrow \Theta_\nu(\lambda). \quad (2)$$

Both profiles are continuous and nondecreasing on $(0, \infty)$, positive at every positive intensity, tend to zero as $\lambda \downarrow 0$, and tend to one as $\lambda \rightarrow \infty$. More precisely,

$$0 < \Theta_{\text{sp}}(\lambda) \leq 1 - e^{-32\lambda} < 1, \quad 0 < \Theta_{\text{qt}}(\lambda) \leq 1 - e^{-30\lambda} < 1, \quad (3)$$

$$\Theta_{\text{sp}}(\lambda/2) \leq \Theta_{\text{qt}}(\lambda) \leq \Theta_{\text{sp}}(\lambda). \quad (4)$$

Let G_n^ν be the event that some food gateway is catalyzed by at least one molecule of X_n , and set $g_{\text{sp}} = 36$, $g_{\text{qt}} = 34$. Then

$$\mathbb{P}(G_n^\nu) \rightarrow 1 - e^{-g_\nu \lambda}, \quad \mathbb{P}(\text{RAF}_n^\nu \mid G_n^\nu) \rightarrow \frac{\Theta_\nu(\lambda)}{1 - e^{-g_\nu \lambda}} < 1. \quad (5)$$

The constants 32 and 30 count gateways that can leave food; the total counts 36 and 34 also include four food-only channels. Section 8 proves the productive-gateway bounds and the strict conditional bounds. The same probability-ratio argument gives the conditioned formula in both models.

Corollary 2.3 (The conjectured step and the outer regimes). *Neither model has a finite positive step threshold within $f_n \asymp n$. For valid Bernoulli parameters $0 \leq f_n \leq N_n^\nu$, the RAF probability tends to zero if $f_n/n \rightarrow 0$ and to one if $f_n/n \rightarrow \infty$. The same statements hold under the clamped convention for arbitrary real initial values.*

No convergence rate in n is asserted. The limiting survival event already includes the food boundary and every finite trap; the conditional expression in (5) is a probability ratio, not a decomposition into independent ignition and bulk events.

3 What quotienting preserves, and what it changes

Lemma 3.1 (Deterministic projection and the OR law). *For every set of split reactions A and supplied set B , their split closure equals the quotient closure of $\pi(A)$. If catalytic marks are projected by*

$$\bar{\xi}_{x,j} = \bigvee_{r \in \pi^{-1}(j)} \xi_{x,r},$$

a split RAF exists if and only if a quotient RAF exists for $\bar{\xi}$. Independent Bernoulli(p) split coordinates project to independent coordinates with probabilities $1 - (1 - p)^{|\pi^{-1}(j)|}$.

Proof. Project or lift each finite reaction application; all representatives have identical sides. Induction on derivation length proves the closure assertion. Project a split RAF to obtain a quotient RAF, keeping its catalytic witnesses. Conversely, lift each member of a quotient RAF by choosing a representative with a catalytic witness in its closure. The lifted family has the same closure, so all its chosen catalysts remain available and it is a RAF. Different molecule-channel fibres are disjoint collections of independent coordinates, which proves independence of their ORs and the displayed probabilities. $\square$

Example 3.2 (A four-molecule discriminator). Take molecule types $\{A, AA, AAA, AAAA\}$, food $\{A, AA\}$, and split reactions (A, AA) , (AA, A) , (A, AAA) . There are two quotient channels. At $p = 1/2$, the split and OR-projected RAF probabilities are 4077/4096, whereas homogeneous quotient catalysis gives 239/256. These values can be checked without simulation. Write $q = 1 - p$. The first channel, if catalyzed by one of its three closure molecules, already forms a singleton RAF. Otherwise it requires catalysis by $AAAA$, which is available only when both channels are present; the second must then have a catalyst in the four-molecule closure. With fibre sizes $d = 1$ or 2 for the first channel, the failure probability is

$$q^{3d}[q^d + (1 - q^d)q^4].$$

For $q = 1/2$ it is 17/256 when $d = 1$ and 19/4096 when $d = 2$, proving the values. For $0 < p < 1$ the $d = 2$ failure probability is strictly smaller. This finite fixture distinguishes ensembles; it does not prove strict inequality between their infinite profiles.

Lemma 3.3 (Product-measure comparisons). *If $a \in [0, 1]$ and $b = 1 - \sqrt{1 - a}$, then the OR projection of split openness b is dominated by homogeneous quotient openness a . The weaker choice $b = a/2$ also works. Homogeneous quotient openness a is dominated, for closure events, by split openness a . Consequently*

$$S_{\text{sp}}(1 - \sqrt{1 - a}) \leq S_{\text{qt}}(a) \leq S_{\text{sp}}(a). \quad (6)$$

Proof. On a double fibre the OR parameter at b is $1 - (1 - b)^2 = a$; on a singleton it is $b \leq a$. Couple the projected and homogeneous quotient fields by common independent uniform variables on channels. For $b = a/2$, the largest OR parameter is $a - a^2/4 \leq a$. For the upper comparison, read each quotient mark from its canonical split representative at openness a . These representative marks are independent. Every open quotient derivation then lifts to an open split derivation. These couplings hold in the countable fields, or first at finite caps and then by passage to the decreasing escape events. $\square$

Substituting $a = 1 - e^{-\lambda}$ gives (4). Neither comparison asserts equality of the two profiles or orders their differently normalized conditional ratios.

Lemma 3.4 (Counts and source normalization). *For $n \geq 2$,*

$$|R_n| = (n - 2)2^{n+1} + 4, \quad D_n := |R_n| - |J_n| \leq nM_{\lfloor (n-1)/2 \rfloor}.$$

Thus $nM_n/N_n^\nu \rightarrow 1$ for both catalogues. If $f_n/n \rightarrow \lambda > 0$, then eventually $0 < f_n/N_n^\nu < 1$, and $p_n^\nu \rightarrow 0$, $M_n p_n^\nu \rightarrow \lambda$.

Proof. The split count is $\sum_{\ell=2}^n (\ell-1)2^\ell$, giving the formula. Give a binary word u of length r its length- r binary code $c(u)$. If $uv = vu$ with $|v| = s$, then

$$(2^r - 1)c(v) = (2^s - 1)c(u).$$

For fixed u, s this determines v . Equal-length commuting words are identical. Each deleted noncanonical description is therefore injectively encoded by its shorter factor u and longer-factor length s . The shorter length is at most $\lfloor (n-1)/2 \rfloor$ and $s \leq n$, proving the bound. Divide by nM_n ; the loss tends to zero exponentially, while $|R_n|/(nM_n) \rightarrow 1$. Inverting proves normalization. Finally $f_n/N_n^\nu = (f_n/n)(n/N_n^\nu) \rightarrow 0$ with eventual positivity, and multiplication by M_n proves the intensity limit. $\square$

The sharper exact count follows from the commuting-word theorem [6, 5]:

$$D_n = \sum_{d \leq n/3} P_d \left\lfloor \frac{(\lfloor n/d \rfloor - 1)^2}{4} \right\rfloor, \quad (7)$$

where P_d counts primitive binary words of length d . Indeed commuting unequal factors are unequal powers w^a, w^b of a unique primitive word; for $a < b$ and $a + b \leq K$, their number is $\lfloor (K-1)^2/4 \rfloor$. The elementary bound above suffices for every limit proof and avoids using this stronger classification as a hidden premise.

4 The shared growth theorem

For a cap N and supplied cutoff L , write $\mathcal{C}_{N,L}^\nu$ for closure of $X_{\min(L,N)}$ in the independent open field. The key estimate is uniform in the ambient cap; generating every fixed word alone would not imply it.

Theorem 4.1 (Uniform growth from a supplied seed). *For every $a_0 \in (0, 1]$ and integer $m \geq 2$ there is L such that, simultaneously for every $a \in [a_0, 1]$ and $N \geq 1$,*

$$\mathbb{P}_a^\nu\{m|\mathcal{C}_{N,L}^\nu| < (m-1)M_N\} \leq 1/m. \quad (8)$$

For split reactions, put

$$A = 2 \cdot 7^{64} \cdot 81, \quad h = A(1 - a_0^2)^k, \quad \delta = \frac{h}{1-h}, \quad L = 10k. \quad (9)$$

Any $k \geq 1$ with $h < 1$ and $\delta \leq 1/[m(m+1)]$ suffices. For quotient channels the same prescription at $b = a_0/2$ suffices in place of a_0 .

The theorem is the indispensable contribution shared by the two original arguments. We prove its probabilistic part here and the finite geometric facts in Appendix A, so the combined paper has no unproved supplied-seed assumption.

Contract all words of length at most L to a root in the split word graph; words above L have basic edges to their first-bit and last-bit deletions. For each edge and each $1 \leq j \leq w = 5k$, delete j letters at its chosen end. The two endpoints can reach this common descendant through at most two literal reactions, whose removed factors have length at most $w \leq L$. When both coordinates are open the endpoints are in the same reversible closure component. We call the basic edge effective if at least one such detour is open. All these paths remain inside cap N .

The geometric facts needed are: (i) each minimal nonempty cut, or bond, is connected in a diamond adjacency of degree at most nine; (ii) for any actual molecule, at most A^r bonds of size r have it on a smaller side, counting every contracted food molecule separately; (iii) each literal reaction occurs in at most three detour supports, so wr detours contain at least $wr/5 = kr$ disjoint supports; (iv) distinct effective components are separated by a bond of non-effective edges. Appendix A proves all four.

Proof of Theorem 4.1. At split openness a_0 , failure of every detour of a size- r bond has probability at most $(1 - a_0^2)^{kr}$: select the disjoint literal supports before invoking independence. A molecule is bad if some such failed bond has it on a molecularly smaller side. The envelope and a union bound give a marginal probability at most

$$\sum_{r \geq 1} A^r (1 - a_0^2)^{kr} = \delta.$$

Any two nonbad molecules belong to the same effective component. Otherwise a failed separating bond has at least one on its smaller side. Fix a food monomer as root representative, and let B count bad molecules. If the root is not bad, all nonbad molecules are in its effective component, hence in $\mathcal{C}_{N,L}^{\text{sp}}$. Since $\mathbb{E}B \leq M_N \delta$, a union bound and Markov's inequality give, for every integer $d \geq 1$,

$$\mathbb{P}_{a_0}^{\text{sp}}\{|\mathcal{C}_{N,L}^{\text{sp}}| + d \leq M_N\} \leq \delta + \frac{M_N \delta}{d}. \quad (10)$$

Take $d = \lfloor M_N/m \rfloor + 1$. Failure in (8) implies deficit at least d , and $M_N/d \leq m$, proving the estimate. Larger openness only reduces failure. If $N \leq L$ the supplied set is all X_N and failure is impossible.

For quotient channels use the split estimate at $b = a_0/2$. Lemma 3.1 identifies projected split closure with closure under its OR marks. Lemma 3.3 dominates those marks by homogeneous quotient openness a_0 , and then by every $a \geq a_0$. The mass failure event is decreasing, so its quotient probability is at most the split failure probability. This completes the transfer; no quotient contour count is assumed. $\square$

Remark 4.2 (Quantitative meaning and conservatism). The explicit sufficient condition is

$$A(1 - b^2)^k \leq \frac{1}{m(m+1) + 1}.$$

For $m = 10$ and quotient lower openness $a_0 = 0.1$, the inherited conservative choice $b = 0.05$, $k = 53667$, $L = 536670$ suffices. At $a_0 = 0.5$, $b = 0.25$, $k = 2082$, $L = 20820$ suffices. These are theoretical seed cutoffs, not useful finite-size error bars. The loss comes from a uniform contour envelope, disjoint-support selection, a union bound, and Markov's inequality. Lemma 3.3 permits replacing $a_0/2$ by $1 - \sqrt{1 - a_0}$ in this conventional transfer, reducing this particular coupling loss; it does not remove the large geometric constant. We do not infer a practical RAF probability from these sufficient bounds.

Proposition 4.3 (Positive survival at every positive openness). *For either catalogue and every $a > 0$, $S_\nu(a) > 0$.*

Proof. For the split field take L from Theorem 4.1 with $m = 2$. Force every reaction with product length at most L open, an event of probability $a^{|R_L|} > 0$. It generates X_L from food. Supplied-seed closure does not depend on these short coordinates, whose endpoints are already supplied, so the event of half molecular mass at any $N \geq L$ has probability at least $a^{|R_L|}/2$ from food. For each fixed K , take N so large that half mass cannot fit in X_K . The probability of escape beyond K has the same lower bound. Decrease these escape events to survival to obtain $S_{\text{sp}}(a) \geq a^{|R_L|}/2$. Quotient positivity follows from $S_{\text{qt}}(a) \geq S_{\text{sp}}(a/2) > 0$. $\square$

5 Survival, richness, and continuity in the same field

The argument in this section applies to either catalogue. All probabilities are at the original openness a ; there is no independent sprinkling field.

Lemma 5.1 (Finite witnesses and escape). *Every finite collection of generated words has a common finite derivation cap. If closure of X_L reaches beyond $K \geq L$, it already does so within cap $2K$. Thus for closure from F , the event E_K^ν of reaching a word longer than K is a finite-coordinate event on $\mathcal{R}_{2K}^\nu$. Its probability $e_K^\nu(a)$ is a polynomial, and $e_K^\nu(a) \downarrow S_\nu(a)$.*

Proof. Take the largest length occurring in finitely many finite derivations. The first derivation step leaving X_K cannot be cleavage, and ligates two words of lengths at most K , so its product has length at most $2K$. Every preceding word is in X_K . This proves finite determination. Escape events decrease to unbounded closure; continuity of probability from above gives the last assertion. $\square$

Lemma 5.2 (Same-parameter richness). *At every $a > 0$, almost surely either closure from food is finite or it contains every nonempty binary word. In particular survival implies acquisition of every fixed finite supplied seed, almost surely.*

Proof. Unbounded closure has arbitrarily long own-cap records: take a finite derivation reaching above a chosen length, and a longest word appearing in it. That word is generated within its own length cap. Fix a target $v = v_1 \cdots v_\ell$. From a record w , successively append the food letters $v_1, \dots, v_\ell$, then cleave wv into w and v . At most $\ell + 1$ reaction coordinates suffice. Each product is longer than w , also after quotient canonicalization, so these coordinates are outside the prefix certifying the record. Their simultaneous opening has probability at least $a^{\ell+1}$.

Let D be the event of survival without v . For a cylinder event C , take a product-length boundary B containing its coordinates. Partition the event of some own-cap record above B by the least record length and the complete finite prefix assignment at that length, choosing a record word on each nonempty atom. These are disjoint cylinder atoms and cover D . On each atom the prescribed repair is independent of that prefix and of C ; on D it must fail. Summing over atoms therefore yields

$$\mathbb{P}_a^\nu(D \cap C) \leq (1 - a^{\ell+1})\mathbb{P}_a^\nu(C).$$

Here the atoms cover the event of a first record, rather than asserting that survival itself has a finite cylinder partition. Approximate the measurable event D in symmetric-difference probability by cylinders C . Letting their error tend to zero gives $\mathbb{P}_a^\nu(D) \leq (1 - a^{\ell+1})\mathbb{P}_a^\nu(D)$, hence $\mathbb{P}_a^\nu(D) = 0$. A countable union over target words proves the dichotomy. $\square$

Let H_L^ν mean that every word in X_L is generated from food, and $H_{N,L}^\nu$ require this within cap N . Write $h_{N,L}^\nu(a) = \mathbb{P}_a^\nu(H_{N,L}^\nu)$. Finite witnesses give $H_{N,L}^\nu \uparrow H_L^\nu$; each $h_{N,L}^\nu$ is a polynomial. Richness gives $S_\nu(a) \leq \mathbb{P}_a^\nu(H_L^\nu)$.

Lemma 5.3 (Uniform approximation by seed acquisition). *For every $a_0 > 0$ and $m \geq 2$, some L satisfies*

$$S_\nu(a) \leq \mathbb{P}_a^\nu(H_L^\nu) \leq S_\nu(a) + 1/m \quad (a \in [a_0, 1]).$$

Proof. Choose $L \geq 2$ from Theorem 4.1 and take $N \geq L$. In the same open field, put

$$B_{N,j}^\nu = \{m|\mathcal{C}_{N,j}^\nu| \geq (m-1)M_N\}, \quad j \in \{2, L\}.$$

On $H_{N,L}^\nu$, every word of X_L already lies in $\mathcal{C}_{N,2}^\nu$. The latter set is closed under all the open reactions in cap N , so minimality of supplied-seed closure gives $\mathcal{C}_{N,L}^\nu \subseteq \mathcal{C}_{N,2}^\nu$. Consequently

$$H_{N,L}^\nu \cap B_{N,L}^\nu \subseteq B_{N,2}^\nu, \quad H_{N,L}^\nu \subseteq B_{N,2}^\nu \cup (B_{N,L}^\nu)^c.$$

Taking probabilities and applying Theorem 4.1 to the last complement gives

$$h_{N,L}^\nu(a) \leq \mathbb{P}_a^\nu\{m|\mathcal{C}_{N,2}^\nu| \geq (m-1)M_N\} + 1/m. \quad (11)$$

No independence between seed acquisition and supplied-seed growth is needed for this union bound. For fixed K , the mass event implies E_K^ν once N is large. First let $N \rightarrow \infty$, then $K \rightarrow \infty$, to prove the upper estimate. The lower estimate is richness. $\square$

Proposition 5.4 (Continuity). *For either catalogue, S_ν is nondecreasing on $[0, 1]$ and continuous on $(0, 1]$ relative to $[0, 1]$, with $S_\nu(1) = 1$.*

Proof. Common uniform coordinates prove monotonicity. Let $a_i \rightarrow a > 0$. Since $S_\nu(a_i) \leq e_K^\nu(a_i)$ and e_K^ν is a polynomial, $\limsup_i S_\nu(a_i) \leq e_K^\nu(a)$; let $K \rightarrow \infty$. For the lower bound fix $a_0 < a$, $m \geq 2$, and the L of Lemma 5.3. Eventually $a_i \geq a_0$, and for each fixed N ,

$$S_\nu(a_i) \geq h_{N,L}^\nu(a_i) - 1/m.$$

Pass first to $i \rightarrow \infty$ at fixed N , then $N \rightarrow \infty$ and use richness, and finally $m \rightarrow \infty$. Thus $\liminf_i S_\nu(a_i) \geq S_\nu(a)$. At $a = 1$, all coordinates are open almost surely; successive monomer ligations generate every word. $\square$

Seed acquisition is not a deterministic certificate of survival: it can occur on a finite closure. Lemma 5.3 controls the probability of that discrepancy uniformly above a positive openness. This uniformity closes the possible jump that mere monotonicity would leave in the limit proof.

6 Exact catalytic pruning histories

We prove once the probabilistic identity that removes the endogenous-catalyst obstruction. Fix a finite universe X with M molecules, a finite set of reaction columns, and any deterministic monotone map Φ from reaction families to molecule sets. For independent Bernoulli(p) coordinates let $A_\xi(D)$ be the columns with a catalyst in D . Starting with $D_0 = X$, set

$$A_t = A_\xi(D_t), \quad D_{t+1} = \Phi(A_t).$$

Both sequences decrease. Fix an ordering of X and let I_k be its first k molecules. The prefix process starts at $\hat{D}_0 = X$ and uses $\hat{D}_{t+1} = I_{|\Phi(\hat{A}_t)|}$, $\hat{A}_t = A_\xi(\hat{D}_t)$.

Lemma 6.1 (Complete-history replacement). *For every finite T , $(A_0, \dots, A_T)$ and $(\hat{A}_0, \dots, \hat{A}_T)$ have the same joint law, and their pool-cardinality histories have the same law.*

Proof. Fix a proposed decreasing column history $A_0^*, \dots, A_T^*$. Its prescribed pools $D_0^* = X$, $D_{t+1}^* = \Phi(A_t^*)$ are deterministic and nested. The actual process follows this history exactly when every column gives its prescribed response in every one of those pools, as follows successively in t . Put $s = 1 - p$ and $k_t = |D_t^*|$. A column can only disappear once. If its first absence is at d , its constraint probability is

$$s^{k_0} \quad (d = 0), \quad s^{k_d} - s^{k_{d-1}} \quad (1 \leq d \leq T), \quad 1 - s^{k_T} \quad (d = T + 1). \quad (12)$$

The middle difference requires no catalyst in D_d^* and at least one in $D_{d-1}^* \setminus D_d^*$. Different columns are independent, so multiply (12) over columns. Replacing the pools by prefixes of the same sizes preserves all factors and hence every history probability. Histories that are not decreasing are impossible in either process. Pool cardinalities are deterministic functions of the preceding column history. $\square$

The result compares complete dynamics. Conditioning on a terminal random pool and then declaring its catalytic columns independent would not be justified by this argument.

Apply this with $X = X_n$, $\Phi(A) = \mathcal{C}_n^\nu(A; F)$. At stabilization write A_* for the active family and $D_* = \Phi(A_*)$. Every RAF belongs to every active family, by induction: its closure remains inside the current pool, so its catalysts keep it active. Delete from A_* reactions with endpoints outside D_* . This does not change closure, since every reaction used in a derivation has all endpoints in the final closure. The remaining family is a RAF whenever nonempty, and is nonempty if $|D_*| > 6$. This terminal test allows food-only RAFs as well; the mass condition is sufficient, not necessary.

Lemma 6.2 (A fixed static mass barrier). *For $6 < k \leq M_n$ and $q(p, k) = 1 - (1 - p)^k$,*

$$\mathbb{P}_{q(p,k)}^\nu \{|\mathcal{C}_{n,2}^\nu| \geq k\} \leq P_n^\nu(p). \quad (13)$$

Proof. In the prefix process, the fixed family $A_\xi(I_k)$ has independent open columns of parameter $q(p, k)$. If its closure has at least k molecules, every prefix pool retains at least k molecules: the initial pool contains I_k , and induction preserves containment of $A_\xi(I_k)$ in the active family. The terminal pool therefore has more than six molecules and supplies a RAF. Transfer this terminal-cardinality probability through Lemma 6.1. Stabilization occurs after at most the number of columns in strict decreases (allowing one additional pool index), so a finite history suffices. $\square$

7 Proof of the critical-window laws

We establish a general statement before using source normalization. Fix either catalogue, and let $p_n \in [0, 1]$ satisfy $p_n \rightarrow 0$ and $M_n p_n \rightarrow \lambda > 0$. Set $a = 1 - e^{-\lambda}$. For fixed $m \geq 2$ let $k_{n,m} = (m - 1)\lfloor M_n/m \rfloor$. Since $M_n p_n^2 \rightarrow 0$, expansion of $\log(1 - p_n)$ gives

$$q(p_n, M_n) \rightarrow a, \quad q(p_n, k_{n,m}) \rightarrow 1 - e^{-(1-1/m)\lambda}. \quad (14)$$

For the lower bound, fix $b \in (0, a)$ and any sufficiently large m with $b < 1 - e^{-(1-1/m)\lambda}$. Choose L by Theorem 4.1 at openness b . Reversing the union bound (11), then using richness and finite witnesses, gives

$$\liminf_n \mathbb{P}_b^\nu \{m|\mathcal{C}_{n,2}^\nu| \geq (m - 1)M_n\} \geq S_\nu(b) - 1/m.$$

Eventually $q(p_n, k_{n,m}) \geq b$ and $k_{n,m} > 6$. The mass event implies closure of size at least $k_{n,m}$, so monotonicity and Lemma 6.2 imply $\liminf_n P_n^\nu(p_n) \geq S_\nu(b) - 1/m$. Let $m \rightarrow \infty$ at fixed b , then $b \uparrow a$ and use Proposition 5.4:

$$\liminf_n P_n^\nu(p_n) \geq S_\nu(a). \quad (15)$$

For the upper bound, the family catalyzed from all X_n has independent column openness $q(p_n, M_n)$. Every RAF is contained in it, and monotonicity of closure places each RAF's own food closure inside this ambient closure. Fix $K \geq 2$ and $n \geq 2K$. Either the ambient closure escapes X_K , or all these RAF closures are confined to X_K . In the latter case every RAF catalyst is in X_K , because its witness must belong to its own food closure by Definition 2.1. Moreover, every nonempty food-generated family contains a food gateway: otherwise no reaction could be enabled at food, so closure could not generate the endpoints required by that family. Thus a RAF confined to X_K requires a gateway-catalyst coordinate with catalyst in the deterministic set X_K . There are g_ν gateways (counted below), and therefore

$$P_n^\nu(p_n) \leq e_K^\nu(q(p_n, M_n)) + g_\nu M_K p_n. \quad (16)$$

No conditioning preserves independence here; the second term is an unconditional union bound over fixed coordinates. At fixed K , polynomial continuity and (14) give $\limsup_n P_n^\nu(p_n) \leq e_K^\nu(a)$. Now $K \rightarrow \infty$ gives the opposite inequality to (15). Thus the limit is $S_\nu(a)$.

Lemma 3.4 applies this general statement to every real $f_n/n \rightarrow \lambda > 0$ with eventual unclamping. Positivity, monotonicity, and continuity follow from Propositions 4.3 and 5.4. Lemma 3.3 gives the profile comparisons. Section 8 supplies the caps, endpoints, and conditioned formula, completing Theorem 2.2.

8 Food gateways, sharper caps, and scaling regimes

A gateway has at least one complete side in food. Any nonempty food-generated family contains a gateway: if none were enabled at food, no closure step could occur, contradicting food generation of its nonempty family. This argument includes a family entirely within food.

For $n \geq 4$, the split gateways are the $6^2 = 36$ ordered pairs of food words. Four have dimer products already in food, sixteen have trimer products, and sixteen have length-four products. In the quotient, the trimer descriptions for $\mathbf{A} + \mathbf{AA}$ and $\mathbf{AA} + \mathbf{A}$ merge, as do their $\mathbf{B}$ counterparts. Equal-length commuting factors are equal, so no other distinct descriptions merge among food pairs. Thus the quotient counts are $4 + 14 + 16 = 34$. The numbers of productive gateways, with nonfood product, are

$$t_{\text{sp}} = 32, \quad t_{\text{qt}} = 30, \quad g_\nu = t_\nu + 4.$$

Proposition 8.1 (Finite and limiting gateway obstructions). *For $n \geq 4$,*

$$\mathbb{P}(G_n^\nu) = 1 - (1 - p)^{g_\nu M_n}, \quad P_n^\nu(p) \leq \mathbb{P}(G_n^\nu), \quad (17)$$

$$P_n^\nu(p) \leq 1 - (1 - p)^{t_\nu M_n + 24}, \quad S_\nu(a) \leq 1 - (1 - a)^{t_\nu}. \quad (18)$$

The additional 24 coordinates in the finite obstruction retain RAFs whose entire closure is food.

Proof. Gateway columns are disjoint, with M_n independent possible catalysts each, proving (17). In the static field unbounded closure needs a first nonfood molecule. That first step is a ligation of two food words through a productive gateway. If all t_ν such channels are closed, closure stays at food, proving the second inequality in (18).

For the finite refinement, split by whether a RAF closure exceeds food. If it does, the RAF contains a productive gateway with a catalyst somewhere in X_n . If it does not, all its reaction endpoints are food. There are only four food-only channels (the monomer-to-dimer reactions), and each must have its catalyst among the six food types. Hence a RAF requires at least one of $t_\nu M_n + 4 \cdot 6$ distinct coordinates. Their absence has probability $(1 - p)^{t_\nu M_n + 24}$. The food-catalyzed singleton $\mathbf{A} + \mathbf{A} \rightleftharpoons \mathbf{AA}$ shows why those extra coordinates cannot be omitted. $\square$

In the critical window, passing (18) to the limit, or substituting $a = 1 - e^{-\lambda}$ into its static version, gives (3). Since $\text{RAF}_n^\nu \subseteq G_n^\nu$,

$$\mathbb{P}(\text{RAF}_n^\nu \mid G_n^\nu) = \frac{P_n^\nu(p_n)}{1 - (1 - p_n)^{g_\nu M_n}} \rightarrow \frac{\Theta_\nu(\lambda)}{1 - e^{-g_\nu \lambda}}.$$

The positive denominator and $t_\nu < g_\nu$ give strictness and the useful conventional estimates

$$\frac{\Theta_\nu(\lambda)}{1 - e^{-g_\nu \lambda}} \leq \frac{1 - e^{-t_\nu \lambda}}{1 - e^{-g_\nu \lambda}} < 1, \quad \limsup_{\lambda \downarrow 0} \frac{\Theta_\nu(\lambda)}{1 - e^{-g_\nu \lambda}} \leq \frac{t_\nu}{g_\nu}. \quad (19)$$

The ratios 32/36 and 30/34 are upper bounds, not small-intensity asymptotic equalities. There is no intensity-independent gap from one implied by these estimates.

At $a = 0$, closure is food. The productive cap gives $S_\nu(a) \rightarrow 0$ as $a \downarrow 0$, so S_ν is continuous there as well. At $a = 1$, Proposition 5.4 gives value one and left continuity. This proves the profile endpoints. If a finite $\gamma > 0$ were a step threshold, $\lambda = \gamma/2$ would have positive limit instead of zero, and every finite $\lambda > \gamma$ would have limit below one. Finally RAF probability increases in each catalytic coordinate. For $f_n/n \rightarrow 0$, compare to every fixed cn from above and let $c \downarrow 0$. For $f_n/n \rightarrow \infty$, compare from below and let $c \rightarrow \infty$. The clamp preserves these comparisons, proving Corollary 2.3 without assuming $p_n \rightarrow 0$ in the superlinear regime.

Corollary 8.2 (Independent heterogeneous catalysis). *Keep either catalogue and independent coordinates with probabilities $p_{n,x,r}$ between ℓ_n and u_n . If $M_n \ell_n \rightarrow \lambda_- > 0$ and $M_n u_n \rightarrow \lambda_+ < \infty$, then*

$$\Theta_\nu(\lambda_-) \leq \liminf_n \mathbb{P}(\text{RAF}_n^\nu) \leq \limsup_n \mathbb{P}(\text{RAF}_n^\nu) \leq \Theta_\nu(\lambda_+).$$

Proof. Couple all coordinates by the same independent uniforms to sandwich the heterogeneous RAF event between the two homogeneous events. Both envelope sequences tend to zero and satisfy the mass limits used in Section 7; apply that general limit statement. $\square$

This is an envelope result under independence, not a theorem for arbitrary correlations between catalytic assignments.

9 Interpretation and finite-model examples

The map $\lambda \mapsto 1 - e^{-\lambda}$ aggregates catalytic opportunities from the molecular universe. Survival in the infinite reversible field then determines whether the food closure remains finite. The history identity is the bridge back to a RAF with catalysts retained in its own closure. An ambient-catalyzed reaction set alone would supply only an upper bound; the mass barrier supplies the missing self-support.

The exact count normalization matters at accessible sizes. For $f_n = \lambda n$, define the effective full-pool intensity

$$r_n^\nu = \frac{nM_n}{N_n^\nu}, \quad q_n^\nu(\lambda) = 1 - \left(1 - \frac{\lambda n}{N_n^\nu}\right)^{M_n},$$

whenever the catalytic parameter lies in $[0, 1]$. The mean number of catalysts available to one channel from the full pool is λr_n^ν , rather than exactly λ . The leading finite-size factor is already visible in the split count $N_n^{\text{sp}} = (n-2)2^{n+1} + 4$, which gives $r_n^{\text{sp}} \simeq n/(n-2)$; quotient deletions add a smaller correction. Table 1 gives exact-formula evaluations from the reaction counts above. At $n = 12$ and $\lambda = 1$, $r_n^{\text{qt}} \simeq 1.20106$ and full-pool channel openness is about 0.69915, rather than the limiting $1 - e^{-1} \simeq 0.63212$. Thus a nominal finite-network intensity need not have its asymptotic input openness. These are channel-openness calculations, not sampled RAF probabilities or an error estimate for $P_n^\nu - \Theta_\nu$. They explain a parameter shift that must be separated from the finite closure and catalytic-pruning effects before interpreting an apparent transition.

n	M_n	$ R_n $	$ J_n $	$nM_n/ J_n $	$1 - (1 - n/ J_n)^{M_n}$
4	30	68	64	1.87500	0.855743
8	510	3076	3048	1.33858	0.738244
12	8190	81924	81828	1.20106	0.699150
16	131070	1835012	1834798	1.14297	0.681131
20	2097150	37748740	37748322	1.11112	0.670811
24	33554430	738197508	738196506	1.09091	0.664089

Table 1: Finite normalization at intensity one. The last column is full-pool openness, not RAF probability. Values are exact-formula evaluations, rounded for display.

The quotient example also gives a practical modeling lesson: chemical equivalence must be specified before probabilities are assigned, and independent opportunities must be distinguished from duplicate observations of the same event. If two database records describe the same catalytic observation, their marks can coincide; the independent-input OR formula then does not apply. Products, compartments, stoichiometry, and reaction direction may carry distinctions absent from this particular word model.

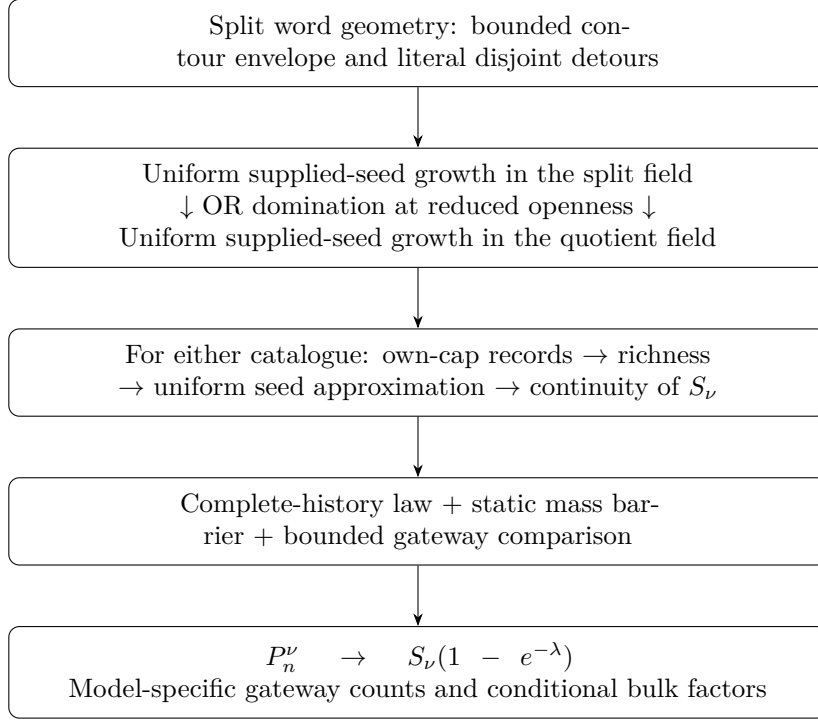


Figure 1: Dependency structure of the consolidated proof. The two product measures remain distinct throughout.

Cooperative catalytic RNA networks motivate structural autocatalysis [10], but the present theorem supplies no measured rate, concentration, energy budget, or probability of sustained reactor operation. The intensity λ is dimensionless. Food is fixed as six word types; catalysis is independent; closure is reversible and set-valued. Inhibition, depletion, kinetic persistence, and thermodynamic consistency require separate models.

10 Remaining questions

The weak comparisons leave pointwise strict inequality $\Theta_{\text{qt}}(\lambda) < \Theta_{\text{sp}}(\lambda)$ as an unresolved question. Finite fixture differences do not imply strictness of infinite survival probabilities. The gateway-conditioned ratios across conventions are not compared here.

Practical rates in n , sharper geometric cutoffs, correlated catalysis, effective evaluation of the profiles, and finer low-intensity asymptotics remain questions beyond the present results. A finite seed cutoff from the proof is not by itself a usable numerical evaluator: an effective computation requires a proved error contract. Neither the normalization table nor the limit theorem alone determines the shape of a finite RAF probability curve.

A Finite word geometry underlying the seed estimate

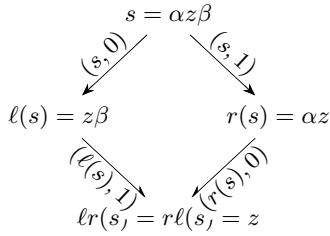
This appendix supplies the deterministic geometric argument for Theorem 4.1. The quotient theorem uses the resulting split-field estimate through the comparison proved in Section 3, rather than a new quotient enumeration.

Fix $N \geq L \geq 1$. The root o represents all words of length at most L ; all longer words up to N are distinct vertices. Let ℓ and r delete the first and last bit respectively, followed by contraction; both fix o and commute. A basic edge has a nonroot product s and deletion index 0 for ℓ , or 1 for r ; its endpoints are s and that parent. Parallel indexed edges are retained. The four slots

$$(s, 0), \quad (s, 1), \quad (r(s), 0), \quad (\ell(s), 1) \quad (20)$$

form a diamond because $\ell r(s) = r\ell(s)$. A slot at the root is empty. A cut is the set of edges on which a Boolean vertex label changes; it has even parity around each diamond, with multiplicities for repeated slots.

(a) A deletion diamond



(b) A literal left-edge detour

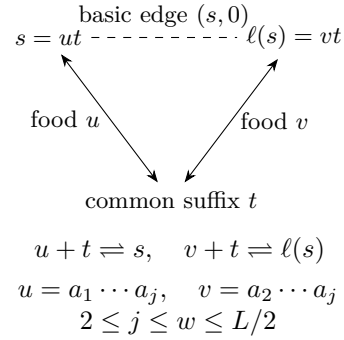


Figure 2: Deletion diamonds and food-assisted detours, shown before contraction. In (a), arrows indicate deletion maps; the indexed basic edges are undirected. In (b), the two solid arrows represent literal reversible reactions with supplied factors $u, v \in X_L$. Words of length at most L are identified with the root, whose slots are empty.

Figure 2(a) displays the four slots of (20), with letters $\alpha, \beta \in \Sigma$ and $|z| > L$, so the four displayed vertices are uncontracted. For the left-edge detour in (b), write $s = a_1 \cdots a_j t$. Removing the supplied prefix $u = a_1 \cdots a_j$ from s and the supplied prefix $v = a_2 \cdots a_j$ from $\ell(s)$ reaches the same nonempty descendant t . These are the literal splits (u, t) and (v, t) , with food-factor lengths j and $j - 1$. At $j = 1$, only $a_1 + \ell(s) \rightleftharpoons s$ is used; no empty factor is introduced. If the parent is contracted food, the descendant is food too and the lower reaction is unnecessary. The right-edge construction uses short suffixes in the same way.

Lemma A.1 (Integration and connected bonds). *Every basic-edge indicator with even diamond parity is a cut. Every bond is connected under adjacency of edges sharing a diamond. This adjacency has degree at most nine, and at most 81^{b-1} connected b-edge sets contain a specified edge.*

Proof. Label the root zero and label each word by the parity along its repeated-right-deletion path to the root. Right-edge differences have the required indicator. Induction on word height and diamond parity give the same for a left edge; commutation identifies the two paths below its diamond. Thus the indicator integrates to a cut.

Take a diamond-adjacency component of a bond's edges. On each diamond it retains either all cut edges or none, so still has even parity and is itself a cut. Minimality makes it the entire bond. An edge belongs to its own upper diamond and at most two lower diamonds,

corresponding to the two possible one-bit preimages at the relevant end. Each diamond adds at most three neighbors, giving degree nine. Choose a deterministic depth-first traversal of a spanning tree of any connected b -set from the specified edge. Its $2(b-1)$ adjacency steps encode the set and give at most $9^{2(b-1)}$ possibilities. $\square$

End-edit distance permits insertion or deletion of one letter at either end, including passage through the empty word. Each word has at most six neighbors, so its radius- ρ ball has at most 7^ρ elements (pad shorter walks with idle steps). Diamond-adjacent product words differ by at most four edits. Products of all edges in a connected b -edge bond are therefore within $4(b-1)$ edits of any one. An interior substring protected by $4b$ letters at both ends persists in every boundary product.

Lemma A.2 (Molecular anchoring). *Let a bond have $b \geq 1$ edges and let s_e be the product of any one of them. Every actual molecule on either smaller side of the bond has end-edit distance at most $64b$ from s_e . Side sizes count every food word separately.*

Proof. If $|s_e| \leq 8b$, all boundary products have length at most $M = |s_e| + 4b \leq 12b$. If $N \leq M$, every molecule has length at most M . If $N > M$, all words in the band $[M, N]$ have the same cut label: a one-bit shift of a length- M word inserts one bit then deletes one, using products above M ; successive shifts connect all length- M words, and truncation connects higher lengths. This band has more than half of the actual molecules. Hence every molecule on a smaller side has length at most M , and deletion through the empty word gives distance at most $M + |s_e| \leq 20b$.

If $|s_e| > 8b$, put $\rho = 4b$ and remove ρ letters from each end to leave a nonempty core z . Every boundary product contains z and has length at most $|z| + 3\rho$. Every word avoiding z stays on the food side during right deletion, since no boundary product can occur on that path. Let $M = |z| + 3\rho$. If $N \leq M$, all words already have length at most M . If $N > M$, the high band is constant as above and contains a constant word opposite to the first letter of z . That word avoids z , so the entire band has the food label. In either case every nonfood-side word contains z and has length at most M .

If the nonfood side is smaller, remove exterior letters of that word and of s_e to reach z and return: distance at most $3\rho + 2\rho = 20b$. If the food side is smaller, the nonfood side has at least $M_N/2$ molecules. The number of words containing z of length at most $|z| + k$ is bounded by

$$\sum_{j=0}^k (j+1)2^j < 2^{2(k+1)}.$$

Choose an occurrence of z , its left-padding length, and the exterior letters; multiple occurrences only overcount. With $k = 3\rho$ this bounds the nonfood side, forcing $N \leq 2(3\rho + 1)$ from $M_N = 2^{N+1} - 2$. Any molecule can then reach s_e via the empty word in at most $2N \leq 12\rho + 4 \leq 64b$ edits. The argument counts food molecules by their actual number, not as a weight-one root. $\square$

Lemma A.3 (Envelope and separation). *For every actual molecule s , at most A^b size- b bonds have s on a smaller side, with $A = 2 \cdot 7^{64} \cdot 81$. Distinct effective components are separated by a bond consisting only of non-effective edges.*

Proof. Anchoring gives at most 7^{64b} possible product words for an anchor edge and two directions per product. Lemma A.1 gives at most 81^{b-1} connected edge sets through each anchor. Thus the count is at most $2 \cdot 7^{64b} \cdot 81^{b-1} \leq A^b$. For separation, contract effective components in the connected basic graph and choose a minimal cut separating the two specified components. A minimal separating cut is a bond (both shores are connected); lift its two connected shores back through the connected effective components. No lifted edge is effective. $\square$

Lemma A.4 (Literal detours and bounded overlap). *For $w \leq L/2$, each of the w detours assigned to a basic edge uses at most two literal split reactions and connects its endpoints reversibly using supplied short factors. Each literal reaction occurs in at most three supports, and among the wb detours of a bond one can select at least $wb/5$ pairwise disjoint supports.*

Proof. Delete $j \leq w$ bits at the edge’s end to a common descendant. The upper reaction removes j letters from the product, and, for $j > 1$, the lower removes $j - 1$ from its one-bit parent. Removed factors belong to X_L . If the parent is contracted food, the descendant is food too and the lower reaction is unnecessary. Nonempty descendants exist because products have length greater than $L \geq 2w$.

A literal reaction is specified by its product, deletion end, and removed length. It is upper in at most one support; as a lower reaction it has at most two one-bit preimage owners. The two ends cannot describe the same relevant split: this would require both factor lengths at most w and a nonfood product of length at most $2w \leq L$. Omitted food reactions introduce no new incidences. Thus each support meets at most four others, and greedy selection retains at least one of every five. Opening its one or two independent coordinates has probability at least a^2 , proving the detour estimate used in (10). $\square$

End-edit walks are a counting device, whereas literal detours are chemical derivations. Only the latter enter the product-measure calculation. This distinction and the molecular food weight are necessary for the uniform estimate.

B Formalization context

Lean 4 and its Mathlib library [1, 9] support the split and quotient formal developments. The selected split theorem is registered as PALOMAR-2026-10-01-000017, version 1 at source commit `f90984f2b107aad62c6efcceb0d8cb9da0fe62a4`; the selected quotient theorem is registered as PALOMAR-2026-10-02-000012, version 1 at source commit `12b505dcc1ccc5ee2d06b32692af201d452a2a7c`. Both use Lean v4.35.0-rc2 and Mathlib revision 065356127b1dc0016f66b7283ce0ce2c4055aa55.

These historical checks cover `HordijkSteelThreshold.palomar_critical_window_variable_intensity` and `RAFReactionQuotient.palomar_quotient_critical_window`, respectively, rather than every result of this article. Their selected total-gateway ceilings remain 36 and 34; the sharper productive ceilings 32 and 30 and the other extensions identified below have conventional proofs here. Broader declarations in the source developments and historical research receipts have their own scope. Lean sources and dependency pins are available through the linked public records and are separate from this TeX deposit. No new Lean verification or numerical replay was performed for this edition. Table 2 records the distinction.

Part of the argument	Formalization boundary
Reaction catalogues and projection	Earlier formal sources include counts, deterministic closure projection, the independent OR law, and RAF projection.
Supplied-seed growth and survival	Earlier developments cover split growth, its transfer to quotient channels, survival identification, positivity, monotonicity, and continuity.
Catalytic pruning and total gateways	Earlier sources cover the history law and mass barrier, total-gateway ceilings 36 and 34, the quotient conditioned formula, and the lower profile comparison.
Productive gateways and comparisons	The 32 and 30 ceilings, finite +24 correction, strict conditional bounds, split conditioned formula, and upper profile comparison are conventional arguments here.
Endpoints and heterogeneous catalysis	Split outer regimes have earlier formal support; quotient endpoints and outer regimes, and the independent heterogeneous envelope, are conventional consequences here.

Table 2: Formal support and the scope of the conventional proofs.

Provenance. The text is primarily AI-generated. AI systems assisted the mathematical development, computations and Lean formalization under human direction; human understanding covers some parts of the work.

References

- [1] Leonardo de Moura and Sebastian 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.
- [2] Wim Hordijk and Mike Steel. Detecting autocatalytic, self-sustaining sets in chemical reaction systems. *Journal of Theoretical Biology*, 227(4):451–461, 2004.
- [3] Wim Hordijk and Mike Steel. Autocatalytic sets in polymer networks with variable catalysis distributions. *Journal of Mathematical Chemistry*, 54(10):1997–2021, 2016.
- [4] Stuart A. Kauffman. Autocatalytic sets of proteins. *Journal of Theoretical Biology*, 119(1):1–24, 1986.
- [5] M. Lothaire. *Combinatorics on Words*. Cambridge Mathematical Library. Cambridge University Press, Cambridge, 1997.
- [6] Roger C. Lyndon and Marcel-Paul Schützenberger. The equation $a^m = b^n c^p$ in a free group. *Michigan Mathematical Journal*, 9(4):289–298, 1962.
- [7] Elchanan Mossel and Mike Steel. Random biochemical networks: the probability of self-sustaining autocatalysis. *Journal of Theoretical Biology*, 233(3):327–336, 2005.
- [8] Mike Steel. The emergence of a self-catalysing structure in abstract origin-of-life studies. *Applied Mathematics Letters*, 13(3):91–95, 2000.
- [9] 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.
- [10] Niles Vaidya, Michael L. Manapat, Irene A. Chen, Ramon Xulvi-Brunet, Eric J. Hayden, and Niles Lehman. Spontaneous network formation among cooperative RNA replicators. *Nature*, 491(7422):72–77, 2012.

- [11] Varun Varanasi. Emergence-of-Autocatalysis-in-Prebiotic-Reaction-Networks: public source and data repository. <https://github.com/VJ-Varanasi/Emergence-of-Autocatalysis-in-Prebiotic-Reaction-Networks/tree/314f1e664cd7a8ccc97eed950f28d33311d0b5ef>, 2026. Reaction catalogue and shared-direction catalytic sampler; immutable commit 314f1e664cd7a8ccc97eed950f28d33311d0b5ef, inspected 15 October 2026. RAF detection is a separate implementation question.
- [12] Varun Varanasi and Jun Korenaga. Emergence of autocatalysis in prebiotic reaction networks. *Physical Review E*, 113:044304, 2026.