

Exact Overlap Partitions and Gateway Scaling in Kauffman RAF Networks

Abstract

We study the probability that a reflexively autocatalytic and food-generated (RAF) set exists in Kauffman’s binary-polymer model with random catalysis, in the exact reaction convention of the public code accompanying the recent finite-core theory of RAF emergence of Varanasi and Korenaga. That theory assigns activation probabilities to minimal catalytic cores and, for tractability, treats the cores as non-overlapping. We remove the approximation. For any finite catalogue of candidate cores we prove an exact two-level inclusion–exclusion formula for the probability that at least one core is activated, valid both for independent Bernoulli catalysis and for a uniformly random set of exactly Q catalysis assignments, and we prove that for the exhaustive catalogue of food-generated reaction supports the formula computes precisely the probability of the RAF event. The sufficient statistic is not a matrix of pairwise core overlaps: it is, for each reaction channel, the family of cardinalities of all unions of the eligible-catalyst sets of the selected cores. A four-molecule example shows that pairwise data cannot determine the answer. The independent-core formula of the finite-core theory is recovered exactly under the sufficient hypothesis of coordinate-disjoint cores. We then show that the fixed food set imposes a finite gateway. For every maximal polymer length $n \geq 4$ the repository convention has exactly 34 reaction channels usable from food, and every RAF contains a catalysed one. Consequently, if f_n is the expected number of channels catalysed per molecule, then $\Pr(\text{RAF}_n) \leq 272f_n/n$ for $n \geq 3$; $\Pr(\text{RAF}_n) \rightarrow 0$ whenever $f_n = o(n)$, so the sharp finite-size transition observed at approximately constant f is not the shadow of an asymptotic transition at constant f ; and for every finite $\lambda > 0$ the RAF probability at $f_n = \lambda n$ does not converge to 1. A closed-form count of the repository’s reaction channels, proved via the Lyndon–Schützenberger theorem on commuting words, gives the sharp gateway limit $1 - e^{-34\lambda}$ on the linear scale. With the exception of that closed form and its corollary, every theorem in the paper has been formalised and checked in Lean 4 with warnings promoted to errors, against the literal RAF predicate of the source model.

Keywords. autocatalytic sets; RAF theory; Kauffman binary polymer model; random catalysis; inclusion–exclusion; overlapping cores; phase transition; formal verification.

MSC 2020. 92E20, 60C05, 05A19, 05A05, 68V20.

1 Introduction

A reflexively autocatalytic and food-generated set (RAF) is a nonempty set of reactions in a catalytic reaction system whose reactants can all be built from a designated food set and each of whose reactions is catalysed by a molecule that is either food or is produced by the set. The notion goes back to Kauffman’s autocatalytic sets of proteins [12, 13], was formalised by Steel [18] and by Hordijk and Steel [7], and has become a well-developed structural theory [11, 19, 6, 20] with a central role in discussions of the origin of metabolism [4, 9, 10]. The model system throughout this literature is Kauffman’s binary polymer model: molecules are words over a two-letter alphabet up to some maximal length n , the food set consists of the short words, reactions are ligations and their reverse cleavages, and catalysis is assigned at

random. The basic quantitative question is how much catalysis is needed for a RAF to exist. Mossel and Steel [16] proved that in the model with independent uniform catalysis the required expected number of reactions catalysed per molecule grows only linearly in n , and Hordijk, Kauffman and Steel [5] showed numerically that a constant expected number between one and two reactions per molecule suffices for the polymer lengths accessible to simulation.

Varanasi and Korenaga [23] recently proposed a finite-core theory of this transition. They assign probabilities to minimal catalytic structures (*cores*), describe RAF emergence in a finite-size regime as the activation of at least one core, and combine the one-core probabilities through a restricted partition function. To make the combination tractable they treat the cores as *non-overlapping*, and they identify overlap and core-size information as the main limitation to be addressed. The public repository accompanying their paper [22] contains the generator of the reaction network, the catalysis sampler, the RAF detector, the simulated RAF frequencies, and a set of single-catalyst witnesses.

This paper answers two questions raised by that programme. First, what is the *exact* law for the probability that at least one of a family of overlapping cores is activated, and what finite statistic of the overlap geometry does it depend on? Second, what does the exact law, together with the literal reaction convention of the repository, imply for the asymptotic interpretation of the observed finite-size transition?

Main results

Write X_n for the set of nonempty binary words of length at most n , J_n for the set of reversible reaction channels in the repository convention (Section 2), and F for the food set of words of length at most 2. A catalytic configuration is a subset of the coordinate set $X_n \times J_n$: the coordinate (x, j) present means that molecule x catalyses both directions of channel j . Under the Bernoulli ensemble each coordinate is present independently with probability p ; under the fixed- Q ensemble a uniformly random set of exactly Q coordinates is present. Following the repository, the Bernoulli parameter is expressed through the expected number f of channels catalysed by a fixed molecule, $p = f/|J_n|$.

- (A) **Exact finite overlap law** (Theorems 4.2 and 5.1). For a finite catalogue $\mathcal{K}$ of cores, each core k demanding on each channel j that the fibre $A_j \subseteq X$ of molecules catalysing j meet an eligible set $W_{k,j}$, the probability that at least one core is activated equals

$$\sum_{\emptyset \neq \mathcal{A} \subseteq \mathcal{K}} (-1)^{|\mathcal{A}|+1} \prod_{j \in J} H_{j,\mathcal{A}}(1-p), \quad H_{j,\mathcal{A}}(q) = \sum_{\mathcal{T} \subseteq \mathcal{R}_{j,\mathcal{A}}} (-1)^{|\mathcal{T}|} q^{|\bigcup \mathcal{T}|},$$

where $\mathcal{R}_{j,\mathcal{A}}$ is the set of distinct nonempty eligible sets requested on channel j by the cores in $\mathcal{A}$. An analogous generating-polynomial identity gives the exact fixed- Q probability. The exact sufficient statistic is, channel by channel, the family of union cardinalities $|\bigcup \mathcal{T}|$.

- (B) **Semantic identification** (Theorem 3.3). For the exhaustive catalogue of all nonempty food-generated reaction supports of the repository system, with $W_{S,j}$ equal to the set of molecules appearing on the reactions of S or in the food, the union of the activation events is *exactly* the set of configurations for which some support is a RAF in the sense of the closure-based RAF predicate. The formulas in (A) are therefore laws for the RAF event itself, not for a proxy.
- (C) **Pairwise data are insufficient** (Theorem 6.1). Two families of three two-element subsets of a four-element set have identical member sizes and identical pairwise intersection sizes, but their channel hit polynomials differ by $q^3 - q^4$. No universally exact overlap correction can be a function of member sizes and pairwise intersections alone.

(D) **Recovery of the non-overlap formula** (Theorem 7.2). If cores occupy pairwise disjoint blocks of coordinates, all joint activation probabilities factor and the exact law collapses to the independent-core product $1 - \prod_k (1 - [1 - (1 - p)^{w_k}]^m)$ used in the finite-core theory. The approximation therefore lies entirely in the replacement of the shared coordinate geometry by dedicated blocks, not in the one-core marginals.

(E) **Exact finite gateway** (Theorems 8.1, 8.2 and 8.3). For every $n \geq 4$ the repository convention has exactly 34 channels with a food-only side. Every RAF contains a catalysed such channel. Hence, with $K_n = 34|X_n| = 34(2^{n+1} - 2)$,

$$0 \leq \Pr_p(\text{RAF}_n) \leq G_n(p) := 1 - (1 - p)^{K_n}.$$

(F) **Asymptotic consequences** (Theorems 9.2 and 9.4). With $p_n = f_n/|J_n|$ and $0 \leq f_n \leq |J_n|$, $\Pr(\text{RAF}_n) \leq 272 f_n/n$ for $n \geq 3$, so $\Pr(\text{RAF}_n) \rightarrow 0$ whenever $f_n = o(n)$; in particular for every constant f . For every finite $\lambda > 0$, the sequence $\Pr(\text{RAF}_n)$ at $f_n = \lambda n$ does not converge to 1.

(G) **Sharp gateway constant** (Proposition 2.1 and Corollary 9.6). The repository has $|J_n| = (n - 2)2^{n+1} + 4 - C(n)$ channels, where $C(n)$ counts unordered pairs of distinct commuting words and is $O(n^2 2^{n/3})$. Consequently $G_n(\lambda n/|J_n|) \rightarrow 1 - e^{-34\lambda}$. These two statements are proved by hand and are the only results of the paper not checked in Lean.

Result (F) corrects the asymptotic reading of the finite-core picture. The simulated RAF frequencies of [22] rise sharply near $f \approx 1.3$ for $n = 6, \dots, 10$, and one might extrapolate to a constant- f phase transition as $n \rightarrow \infty$. Theorem 9.2 shows that no such phase exists in the fixed-food model: at any constant f the probability that even one of the 34 entry channels is catalysed tends to zero. The finite-size curve is nevertheless not explained by the gateway, which is essentially saturated throughout the observed transition window (Section 10); the gateway controls the scaling, not the shape. The first scale at which the gateway is non-degenerate is $f_n = \lambda n$, in agreement with the linear growth established for the split-position model by Mossel and Steel [16], and there Theorem 9.4 shows that no finite λ makes RAF existence asymptotically certain. For the split-position catalogue, a companion manuscript [1] determines the limit on this scale as $\Theta(\lambda) = S(1 - e^{-\lambda})$, with S a survival probability of an infinite reversible reaction field and $0 < \Theta(\lambda) \leq 1 - e^{-36\lambda} < 1$. In the repository quotient the analogous bulk factor is isolated exactly (Proposition 9.5) and its limit is left open.

Relation to previous work

The exact formula in (A) is finite inclusion–exclusion, and each layer is classical; the contribution is the identification of the correct *objects* to which inclusion–exclusion must be applied so that the result is the RAF event in the source convention, and the proof that nothing coarser suffices. In particular the reaction identity used by the repository differs from the textbook split-position catalogue: ligations $u + v \rightarrow uv$ and $v + u \rightarrow vu$ are identified when $uv = vu$. The difference is visible already at $n = 3$, where the repository has 18 channels and the split-position catalogue has 20, and it changes the gateway count from 36 to 34. The finite-gateway mechanism itself was identified, for the split-position catalogue, in a companion manuscript on the critical window of the uniform binary-polymer model [1]; here it is established in the literal repository quotient and coupled to the exact overlap law. Bonferroni-type truncations of inclusion–exclusion [3] are the natural framework for controlled approximations of the exact law, and Theorem 6.1 states precisely why the second-order truncation is an approximation and not an identity. Variable catalysis distributions [8] and partitioned networks [17] are outside the homogeneous Bernoulli and fixed- Q ensembles treated here; Section 12 records what would change.

Formal verification

Every theorem, lemma and corollary below, except Proposition 2.1 and Corollary 9.6, is accompanied by the name of a Lean 4 declaration [2, 21] that states and proves it against the model of Section 2. The development compiles with warnings promoted to errors and contains no `sorry`. The finite incidence facts of Theorem 6.1 and the cardinalities $|J_3| = 18$ and $|\text{Seed}_2| = 34$ are established by kernel evaluation. Section 11 lists the modules and their hashes. We have taken care that the terminal theorem quantifies over the actual RAF event of the repository system rather than over an abstract probability sequence, and that no containment or independence hypothesis is supplied by hand.

Organisation

Section 2 defines the model. Section 3 constructs the catalogue and proves the semantic identification. Sections 4 and 5 prove the exact overlap law, Section 6 the insufficiency of pairwise data, Section 7 the recovery of the independent-core formula. Section 8 proves the exact 34-channel gateway and the probability cap, Section 9 the asymptotic consequences and the sharp constant. Section 10 reports the numerical diagnostics, Section 11 the verification receipts, and Section 12 discusses scope, non-claims and open problems.

2 The repository model

2.1 Molecules, food, and the commuting-factor quotient

Fix $n \geq 1$. A molecule is a nonempty binary word of length at most n ; write X_n for the set of molecules, so $|X_n| = 2^{n+1} - 2$. In the formalisation a molecule is a pair (length $\ell \in \{1, \dots, n\}$, code $c \in \{0, \dots, 2^\ell - 1\}$), and concatenation is the mixed-radix map $(u, v) \mapsto c_u 2^{|v|} + c_v$ on codes with lengths added. The food set with horizon t is $F_t = \{x \in X_n : |x| \leq t\}$; the repository uses $t = 2$ (with the exception $t = 1$ when $n = 2$, which plays no role below), so $F = F_2$ consists of the six words A, B, AA, AB, BA, BB.

A ligation channel is determined by an ordered pair (u, v) of molecules with $|u| + |v| \leq n$; it stands for the reversible reaction $u + v \rightleftharpoons uv$. The textbook *split-position* catalogue keeps every such ordered pair; equivalently every (product word, split position). The repository generator instead compares displayed reactions and suppresses duplicates. When $uv \neq vu$ both (u, v) and (v, u) are kept, since they display different products. When $uv = vu$ the two displayed reactions coincide, and exactly one is kept, namely the one whose first factor precedes the other in the generator’s molecule order (shorter words first, then increasing code). We write $u \preceq v$ for this total preorder and record the resulting reaction set as the formal subtype

$$J_n = \{(u, v) \in X_n \times X_n : |u| + |v| \leq n \text{ and } (uv \neq vu \text{ or } u \preceq v)\} \quad (\text{RepositoryChannel } n).$$

Here $\preceq$ is reflexive, so a self-ligation (u, u) is kept once, and it is antisymmetric on distinct molecules, so a commuting pair $u \neq v$, $uv = vu$ is kept exactly once. Kernel evaluation gives $|J_3| = 18$ (`card_repository_channels_three`), whereas the split-position catalogue has $\sum_{\ell=2}^3 (\ell - 1) 2^\ell = 20$ channels. Two ordered pairs are lost at $n = 3$: $(A, AA) \sim (AA, A)$ and $(B, BB) \sim (BB, B)$.

The exact size of J_n is governed by the classical theorem that two words commute if and only if they are powers of a common word [15, 14]. Let P_d denote the number of primitive binary words of length d , so that $P_d = 2^d - \sum_{e|d, e < d} P_e$.

Proposition 2.1 (Exact channel count; not formalised). *For every $n \geq 1$,*

$$|J_n| = (n - 2)2^{n+1} + 4 - C(n), \quad C(n) = \sum_{d=1}^{\lfloor n/3 \rfloor} P_d \left\lfloor \frac{(\lfloor n/d \rfloor - 1)^2}{4} \right\rfloor.$$

In particular $C(n) \leq n^2 2^{n/3+1}$ and $|J_n| = (n-2)2^{n+1} (1 + o(1))$.

Proof. The number of ordered pairs (u, v) of nonempty words with $|u| + |v| \leq n$ is $\sum_{\ell=2}^n (\ell-1)2^\ell = (n-2)2^{n+1} + 4$. The quotient removes exactly one element for each unordered pair $\{u, v\}$ of distinct words with $uv = vu$ and $|u| + |v| \leq n$. By the Lyndon–Schützenberger theorem, $uv = vu$ with $u \neq v$ if and only if $u = w^a$ and $v = w^b$ for a unique primitive word w and positive integers $a \neq b$. Given $|w| = d$, the constraint is $a + b \leq m := \lfloor n/d \rfloor$, and the number of pairs $a < b$ of positive integers with $a + b \leq m$ is $\lfloor (m-1)^2/4 \rfloor$ (there are $\lfloor (s-1)/2 \rfloor$ such pairs with $a + b = s$, and $\sum_{s=3}^m \lfloor (s-1)/2 \rfloor = \lfloor (m-1)^2/4 \rfloor$). Since $a + b \geq 3$ forces $d \leq n/3$, summing over primitive roots gives $C(n)$. For the bound, $P_d \leq 2^d$ and $\lfloor (m-1)^2/4 \rfloor \leq n^2/(4d^2) \leq n^2$, so $C(n) \leq n^2 \sum_{d \leq n/3} 2^d \leq n^2 2^{n/3+1}$. $\square$

The formula reproduces the literal generator for $n = 2, \dots, 8$ (Table 1). The formal development uses only the weaker bound

$$|J_n| \geq \left\lfloor \frac{n-1}{2} \right\rfloor 2^n \quad (n \geq 3) \quad (1)$$

(`repositoryChannel_card_lower`), obtained from the injective family of channels (u, v) with $|u| < |v|$ and $|u| + |v| = n$, which are always canonical representatives.

2.2 The reversible catalytic reaction system

A *reversible catalytic reaction system* on molecule type M and channel type R consists of finite sets $\text{lhs}(r), \text{rhs}(r) \subseteq M$ for each channel r and a food set $F \subseteq M$. The repository system $\mathcal{Q}_{n,t}$ (`repositoryCRS`) has $M = X_n$, $R = J_n$, $\text{lhs}(u, v) = \{u, v\}$, $\text{rhs}(u, v) = \{uv\}$ and food F_t . A catalysis relation is a predicate $C(x, r)$, “molecule x catalyses channel r ”; catalysis is attached to the channel and therefore shared by the ligation and cleavage directions, as in the repository sampler.

For a set $S \subseteq R$ of channels the *reversible closure* of the food is the increasing sequence $\text{cl}_S^0 = F$,

$$\text{cl}_S^{k+1} = \text{cl}_S^k \cup \bigcup_{r \in S} \left([\text{lhs}(r) \subseteq \text{cl}_S^k] \text{rhs}(r) \cup [\text{rhs}(r) \subseteq \text{cl}_S^k] \text{lhs}(r) \right),$$

where $[\cdot]W$ denotes W if the bracketed condition holds and $\emptyset$ otherwise. The set S is *food-generated* if for every $r \in S$ there is k with $\text{lhs}(r) \cup \text{rhs}(r) \subseteq \text{cl}_S^k$; it is *reflexively autocatalytic* for C if every $r \in S$ has a catalyst x with $C(x, r)$ and $x \in \text{cl}_S^k$ for some k ; and

$$\text{IsRevRAF}(\mathcal{Q}, C, S) :\iff S \neq \emptyset \wedge S \text{ food-generated} \wedge S \text{ reflexively autocatalytic for } C.$$

This is the closure-based RAF definition of [7, 5] for a reversible reaction set, and it agrees with the repository’s iterative pruning detector. The RAF event is $\text{RAF}(C) :\iff \exists S, \text{IsRevRAF}(\mathcal{Q}, C, S)$.

A channel r is a *seed* (`RevSeedReaction`) if $\text{lhs}(r) \subseteq F$ or $\text{rhs}(r) \subseteq F$: at least one side of the reaction can fire from food alone. Write $\text{Seed}_t(n) = \{r \in J_n : r \text{ seed for } \mathcal{Q}_{n,t}\}$.

2.3 Catalysis ensembles

A catalytic configuration is described by its channel fibres $A = (A_r)_{r \in J_n}$ with $A_r \subseteq X_n$ the set of molecules catalysing r ; the associated relation is $C_A(x, r) :\iff x \in A_r$. The *Bernoulli ensemble* with parameter $p \in [0, 1]$ gives configuration A the weight

$$w_p(A) = \prod_{r \in J_n} p^{|A_r|} (1-p)^{|X_n| - |A_r|},$$

so that coordinates are independent. Following the repository we parametrise p by the expected number of channels catalysed by one molecule, $p_n(f) = f/|J_n|$ (`repositoryCatalysisP`). The *fixed- Q ensemble* is the uniform distribution on configurations with $\sum_r |A_r| = Q$; probabilities are ratios of an integer count and $\binom{|X_n||J_n|}{Q}$, both cast to $\mathbb{R}$ before division. The two ensembles are treated separately; no equivalence of ensembles is used.

3 Cores, catalogues, and the semantic bridge

The generic part of the theory is stated for an arbitrary finite molecule type M , a finite set J of channels, and a finite index set $\mathcal{K}$ of *cores*. A core k is specified by its *requirement map* $j \mapsto W_{k,j} \subseteq M$; the convention $W_{k,j} = \emptyset$ means that k does not use channel j . The activation event of k is

$$E_k = \{A : A_j \cap W_{k,j} \neq \emptyset \text{ for every } j \text{ with } W_{k,j} \neq \emptyset\},$$

and the catalogue event is $E_{\mathcal{K}} = \bigcup_{k \in \mathcal{K}} E_k$. In the source specialisation the catalogue is derived, not assumed.

Definition 3.1 (Exhaustive repository catalogue). For the system $\mathcal{Q}_{n,t}$ let $\mathcal{K}_{n,t}$ be the set of all nonempty food-generated $S \subseteq J_n$ (`repositoryCoreSupports`). For $S \in \mathcal{K}_{n,t}$ put

$$\text{core}(S) = F_t \cup \bigcup_{r \in S} (\text{lhs}(r) \cup \text{rhs}(r)), \quad W_{S,j} = \begin{cases} \text{core}(S) & j \in S, \\ \emptyset & j \notin S, \end{cases}$$

(`repositorySupportRequirements`).

Lemma 3.2 (Reachability). *If S is food-generated in a reversible system $\mathcal{Q}$, then $x \in \text{core}(S)$ if and only if $x \in \text{cl}_S^k$ for some k .*

Lean: `mem_revCoreMolecules_iff_reachable`.

Proof. Every stage of the closure lies in $\text{core}(S)$, by induction on k : the step adds only sides of channels in S . Conversely food lies in cl_S^0 , and a molecule on a side of $r \in S$ lies in cl_S^k for the k provided by food generation of r . $\square$

Theorem 3.3 (Semantic identification). *For every n, t and every configuration A ,*

$$A \in \bigcup_{S \in \mathcal{K}_{n,t}} E_S \iff \exists S \subseteq J_n : \text{IsRevRAF}(\mathcal{Q}_{n,t}, C_A, S).$$

Lean: `mem_repository_catalogue_iff_exists_isRevRAF, repository_catalogue_eq_raf_configurations`.

Proof. Fix $S \in \mathcal{K}_{n,t}$. By construction $A \in E_S$ if and only if for every $r \in S$ some $x \in A_r$ lies in $\text{core}(S)$; here we use that $\text{core}(S) \neq \emptyset$, since it contains the product of any $r \in S$, so no requirement is dropped by the empty-set convention. By Lemma 3.2 this is reflexive autocatalysis of S for C_A , and S is nonempty and food-generated because $S \in \mathcal{K}_{n,t}$; hence $A \in E_S$ if and only if $\text{IsRevRAF}(\mathcal{Q}_{n,t}, C_A, S)$ (`mem_singleton_repository_support_iff_isRevRAF`). Conversely, if $\text{IsRevRAF}(\mathcal{Q}_{n,t}, C_A, S)$ then S is nonempty and food-generated, so $S \in \mathcal{K}_{n,t}$, and the same equivalence gives $A \in E_S$. $\square$

The theorem is the bridge between algebra and semantics. The formulas of the next two sections are proved for $E_{\mathcal{K}}$ with an arbitrary catalogue; Theorem 3.3 says that for $\mathcal{K} = \mathcal{K}_{n,t}$ they are laws for the RAF event itself. The catalogue is exhaustive and finite; it is a proof device, not an algorithm.

4 The exact local overlap law

Fix a selected family $\mathcal{A} \subseteq \mathcal{K}$ of cores and a channel j . The distinct nonempty requirements placed on j by the cores of $\mathcal{A}$ form the finite set

$$\mathcal{R}_{j,\mathcal{A}} = \{W_{k,j} : k \in \mathcal{A}, W_{k,j} \neq \emptyset\} \quad (\text{selectedRequirementFamily}).$$

Duplicates are removed because $\mathcal{R}_{j,\mathcal{A}}$ is a set; unused channels give $\mathcal{R}_{j,\mathcal{A}} = \emptyset$. For a finite family $\mathcal{T}$ of subsets of M write $\bigcup \mathcal{T}$ for its union.

Definition 4.1 (Local hit polynomials).

$$H_{j,\mathcal{A}}(q) = \sum_{\mathcal{T} \subseteq \mathcal{R}_{j,\mathcal{A}}} (-1)^{|\mathcal{T}|} q^{|\bigcup \mathcal{T}|}, \quad G_{j,\mathcal{A}}(z) = \sum_{\mathcal{T} \subseteq \mathcal{R}_{j,\mathcal{A}}} (-1)^{|\mathcal{T}|} (1+z)^{|M| - |\bigcup \mathcal{T}|}.$$

Lean: `localHitPolynomial`, `localFixedQHitPolynomial`.

Theorem 4.2 (Exact local law). *Let $\text{Hit}_j(\mathcal{A}) = \{B \subseteq M : B \cap W \neq \emptyset \text{ for all } W \in \mathcal{R}_{j,\mathcal{A}}\}$ be the set of fibres on channel j that satisfy every requirement of $\mathcal{A}$.*

- (i) $\sum_{B \in \text{Hit}_j(\mathcal{A})} p^{|B|} (1-p)^{|M|-|B|} = H_{j,\mathcal{A}}(1-p);$
- (ii) $\sum_{\substack{B \in \text{Hit}_j(\mathcal{A}) \\ \text{in } G_{j,\mathcal{A}}}} z^{|B|} = G_{j,\mathcal{A}}(z),$ so the number of hitting fibres of size q is the coefficient of z^q in $G_{j,\mathcal{A}}$.

Lean: `localHitProbability_eq_sum_fibreBernoulliWeight`, `localFixedQHitPolynomial_counts_configurations`.

Proof. $\text{Hit}_j(\mathcal{A})$ is the complement of the union over $W \in \mathcal{R}_{j,\mathcal{A}}$ of the miss events $\{B : B \cap W = \emptyset\}$. Inclusion–exclusion over $\mathcal{T} \subseteq \mathcal{R}_{j,\mathcal{A}}$ expresses its indicator as $\sum_{\mathcal{T}} (-1)^{|\mathcal{T}|} \mathbf{1}[B \cap \bigcup \mathcal{T} = \emptyset]$. The sets B disjoint from $U = \bigcup \mathcal{T}$ are the subsets of $M \setminus U$; their Bernoulli mass is $(1-p)^{|U|}$ and their size-generating polynomial is $(1+z)^{|M|-|U|}$. $\square$

Because distinct channels use disjoint blocks of coordinates, the joint activation of every core in $\mathcal{A}$ multiplies across channels.

Corollary 4.3 (Joint activation). $\Pr_p\left(\bigcap_{k \in \mathcal{A}} E_k\right) = \prod_{j \in J} H_{j,\mathcal{A}}(1-p)$, and the number of configurations of total size Q activating every $k \in \mathcal{A}$ is the coefficient of z^Q in $\prod_{j \in J} G_{j,\mathcal{A}}(z)$.

Lean: `jointBernoulliHitProbability_eq_sum_fibre_weights`, `jointFixedQHitPolynomial_counts_fibre_configurations`.

Remark 4.4 (The exact sufficient statistic). For homogeneous Bernoulli or uniform fixed- Q catalysis, Theorem 4.2 shows that the joint law of any selected family is a function of the channel-indexed collection of union cardinalities $\{|\bigcup \mathcal{T}| : \mathcal{T} \subseteq \mathcal{R}_{j,\mathcal{A}}\}_{j \in J}$ (`jointChannelHitProbability_congr_trace`). Member sizes and pairwise intersections are the $|\mathcal{T}| \leq 2$ projections of this collection. Section 6 shows that the projection loses information.

Two elementary cases will be used repeatedly: $H_{j,\mathcal{A}} = 1$ when $\mathcal{R}_{j,\mathcal{A}} = \emptyset$, and $H_{j,\mathcal{A}}(q) = 1 - q^{|W|}$ when $\mathcal{R}_{j,\mathcal{A}} = \{W\}$ (`localHitProbability_empty`, `localHitProbability_singleton`).

5 The exact global partition

Theorem 5.1 (Exact finite overlap law). *For every finite M , J , catalogue $\mathcal{K}$ and requirement map W :*

- (i) (Bernoulli) $\Pr_p(E_{\mathcal{K}}) = \sum_{\emptyset \neq \mathcal{A} \subseteq \mathcal{K}} (-1)^{|\mathcal{A}|+1} \prod_{j \in J} H_{j,\mathcal{A}}(1-p);$
- (ii) (fixed Q) with $Z_{\mathcal{K}}(z) = \sum_{\emptyset \neq \mathcal{A} \subseteq \mathcal{K}} (-1)^{|\mathcal{A}|+1} \prod_{j \in J} G_{j,\mathcal{A}}(z)$, the number of configurations of total size Q in $E_{\mathcal{K}}$ is $[z^Q]Z_{\mathcal{K}}(z)$, and

$$\Pr_Q(E_{\mathcal{K}}) = \frac{[z^Q]Z_{\mathcal{K}}(z)}{\binom{|M||J|}{Q}}.$$

Lean: `rafBernoulliProbability_eq_sum_catalogue_weights`, `rafFixedQCountPolynomial_counts_catalogue_configurations`, `rafFixedQProbability_exact`.

Proof. Inclusion–exclusion over the finite union $\bigcup_{k \in \mathcal{K}} E_k$ gives $\mathbf{1}_{E_{\mathcal{K}}} = \sum_{\emptyset \neq \mathcal{A}} (-1)^{|\mathcal{A}|+1} \mathbf{1}_{\bigcap_{k \in \mathcal{A}} E_k}$.

Summing against the Bernoulli weight, respectively against $z^{\sum_j |A_j|}$, and inserting Corollary 4.3 gives (i) and (ii). In Lean the outer step is `Finset.inclusion_exclusion_sum_inf_comp` applied to the actual configuration sets, after the inner identities have been proved for those sets. $\square$

Corollary 5.2 (Repository RAF law). *For every n, t, Q and $p = p_n(f)$, the expressions of Theorem 5.1 with $\mathcal{K} = \mathcal{K}_{n,t}$ equal, respectively, the Bernoulli mass of $\{A : \text{RAF}(C_A)\}$ and the ratio of the number of its members of total size Q to $\binom{|X_n||J_n|}{Q}$.*

Lean: `hasExactRepositoryRAFLaw`.

Two mechanisms of dependence are resolved separately. *Inner* overlap, where several cores request overlapping eligible sets on the same channel, is handled by the union exponents inside $H_{j,\mathcal{A}}$. *Outer* overlap, where one configuration activates several cores, is handled by the alternating sum over $\mathcal{A}$. Retaining the empty family with sign $(-1)^{|\mathcal{A}|}$ gives the exact no-RAF probability and hence an independent check of normalisation (`noRAFBernoulliProbability`).

Remark 5.3 (Complexity). The formula is exponential in $|\mathcal{K}|$ and in $|\mathcal{R}_{j,\mathcal{A}}|$. This is not a defect of the theorem: the probability of a union of dependent events is generically exponential to compute exactly. The value of Theorem 5.1 is to fix the exact target against which any truncation must be measured, and to identify (Remark 4.4) the state variable that a truncation must approximate.

6 Pairwise overlap data are insufficient

The natural first correction to a non-overlap theory would retain, for each channel, the sizes of the eligible sets and the sizes of their pairwise intersections. The following fixture shows that no exact law can be built from these data alone.

Theorem 6.1 (Pairwise insufficiency). *On the four-element set $\{0, 1, 2, 3\}$ let*

$$\mathcal{F}_I = \{\{0, 1\}, \{0, 2\}, \{0, 3\}\}, \quad \mathcal{F}_{II} = \{\{0, 1\}, \{0, 2\}, \{1, 2\}\}.$$

Every member of either family has size 2, and every pair of distinct members meets in exactly one element. Nevertheless

$$H_{\mathcal{F}_I}(q) = 1 - 3q^2 + 3q^3 - q^4, \quad H_{\mathcal{F}_{II}}(q) = 1 - 3q^2 + 2q^3, \quad H_{\mathcal{F}_I}(q) - H_{\mathcal{F}_{II}}(q) = q^3 - q^4.$$

Lean: `firstTripleFamily_uniform_pair_data`, `secondTripleFamily_uniform_pair_data`, `pairwise_data_miss_triple_term`.

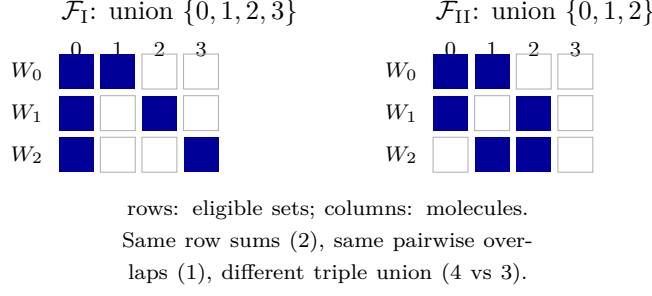


Figure 1: Incidence matrices of the two three-set families of Theorem 6.1. All statistics of order at most two agree; the third-order union term $q^{|W_0 \cup W_1 \cup W_2|}$ differs.

Proof. The incidence facts are finite and are checked by evaluation. The three-requirement instance of the inclusion–exclusion formula (`localHitProbability_triple`) reads $H(q) = 1 - \sum_i q^{|W_i|} + \sum_{i < j} q^{|W_i \cup W_j|} - q^{|W_1 \cup W_2 \cup W_3|}$. All pair unions have size 3 in both families; the triple union has size 4 for $\mathcal{F}_I$ and 3 for $\mathcal{F}_{II}$. $\square$

Corollary 6.2. *Any formula for the activation probability of a family of cores that depends only on the eligible-set sizes and pairwise intersection sizes on each channel assigns the same value to a channel carrying $\mathcal{F}_I$ and to one carrying $\mathcal{F}_{II}$, and is therefore wrong for at least one of them whenever $0 < q < 1$.*

The corollary does not say that pair-corrected formulas are useless: they are second-order Bonferroni-type truncations of Theorem 5.1 [3], and their error is governed by the third and higher union terms. It says that “pair-corrected” is not synonymous with “exact”.

7 Recovery of the non-overlap formula

The finite-core theory of [23] combines one-core probabilities π_k as if the cores were independent, $\Pr(\text{some core}) = 1 - \prod_k (1 - \pi_k)$. We give a sufficient hypothesis under which this is an identity: the cores occupy pairwise disjoint blocks of catalysis coordinates. Theorem 7.2 does not assert a converse.

Definition 7.1 (Dedicated coordinates). Given eligible sets $W_k \subseteq M$ ($k \in \mathcal{K}$) and a finite set Φ of fibre labels, the dedicated channel set is $\mathcal{K} \times \Phi$ and the requirement of core k on channel (k', φ) is W_k if $k' = k$ and $\emptyset$ otherwise (`dedicatedRequirements`).

Theorem 7.2 (Non-overlap specialisation). *With dedicated coordinates, all $W_k \neq \emptyset$, and $m = |\Phi|$:*

- (i) *for every $\mathcal{A} \subseteq \mathcal{K}$, $\Pr_p(\bigcap_{k \in \mathcal{A}} E_k) = \prod_{k \in \mathcal{A}} \Pr_p(E_k)$;*
- (ii) $\Pr_p(E_k) = [1 - (1 - p)^{|W_k|}]^m$;
- (iii) $\Pr_p(E_{\mathcal{K}}) = 1 - \prod_{k \in \mathcal{K}} (1 - [1 - (1 - p)^{|W_k|}]^m)$.

Lean: `dedicated_intersections_factorize`, `jointBernoulliHitProbability_dedicated`, `rafBernoulliProbability_dedicated_nonoverlap`.

Proof. On the dedicated channel (k', φ) the requirement family of $\mathcal{A}$ is $\{W_{k'}\}$ if $k' \in \mathcal{A}$ and empty otherwise, so by the two elementary cases of Section 4 the joint probability is $\prod_{k \in \mathcal{A}} (1 - (1 - p)^{|W_k|})^m$, which is (i) and (ii). Substituting a factorising family into Theorem 5.1(i) and using the identity $\sum_{\emptyset \neq \mathcal{A}} (-1)^{|\mathcal{A}|+1} \prod_{k \in \mathcal{A}} \pi_k = 1 - \prod_k (1 - \pi_k)$ (`independentCoreUnionExpansion`) gives (iii). $\square$

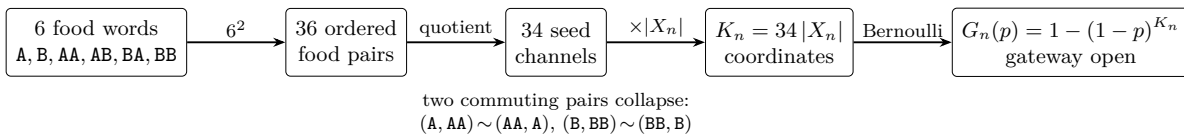


Figure 2: The repository gateway at food horizon 2 and $n \geq 4$. Exactly two of the 36 ordered food pairs are identified by the commuting-factor quotient, leaving 34 seed channels; each carries $|X_n|$ independent catalysis coordinates.

Theorem 7.2 locates the approximation of the finite-core theory precisely. The one-core marginals $[1 - (1 - p)^w]^m$ are correct, and are exactly the singleton case of Theorem 4.2. What the non-overlap assumption changes is the geometry of intersections: shared molecules on shared channels are replaced by private copies. The general formula, Theorem 5.1, keeps the marginals and replaces the product of marginals by the union trace of Remark 4.4. Comparing two models with the same marginals and different traces isolates the effect of network geometry from the effect of the catalysis level.

8 The finite food gateway

The results so far are exact at every finite n . The asymptotic content of the paper comes from a structural feature of the fixed-food model: while $|X_n|$ and $|J_n|$ grow exponentially, the number of channels through which a RAF can be entered from food does not grow at all.

8.1 Exactly 34 seed channels

Products of two food words have length at most 4. Once $n \geq 4$ every ordered pair of food words is an admissible channel, and the seed channels are exactly the (quotiented) food pairs: a channel with $\text{rhs} \subseteq F$ has a product of length at most 2, hence factors of length 1, and is a food pair as well. Among the 36 ordered food pairs the commuting distinct pairs are $\{A, AA\}$ and $\{B, BB\}$, each collapsed once.

Theorem 8.1 (Exact seed count). *For every $n \geq 4$, $|\text{Seed}_2(n)| = 34$.*

Lean: `repositorySeedChannels_card_eq_34`.

Proof. Let Γ be the intrinsic gateway type: ordered pairs (u, v) of words of length at most 2 with $uv \neq vu$ or $u \preceq v$ (`RepositoryGateway 2`). Kernel evaluation gives $|\Gamma| = 34$ (`card_repository_gateway_binary_t2`). *Upper bound*, valid for all n : a seed channel has both factors in F (if $\text{lhs} \subseteq F$ this is immediate; if $\text{rhs} \subseteq F$ the product has length at most 2, so each factor has length 1). Truncating the two factors to words of length at most 2 preserves length and code, hence preserves the canonicity condition, and gives an injection $\text{Seed}_2(n) \rightarrow \Gamma$ (`repositorySeedChannels_card_le_34`). *Lower bound*, for $n \geq 4$: embedding a pair in Γ into $X_n \times X_n$ preserves lengths and codes; the product fits because $|u| + |v| \leq 4 \leq n$; the canonicity condition transfers verbatim; and $\text{lhs} = \{u, v\} \subseteq F$, so the image is a seed channel. This map is injective (`repositoryGatewayToSeedChannel_injective`). $\square$

The theorem is a stabilisation statement across all $n \geq 4$, not a computation at one n . It also explains the difference between the two conventions: in the split-position catalogue the same argument yields 36 seed reactions, the count used in [1].

8.2 Every RAF crosses the gateway

Theorem 8.2 (Seed necessity). *Let $\mathcal{Q}$ be any reversible system, C a catalysis relation and S a set of channels with $\text{IsRevRAF}(\mathcal{Q}, C, S)$. Then there exist a seed channel r and a molecule x*

with $C(x, r)$; moreover $r \in S$ and x is reachable in the closure of S .

Lean: `rev_raf_implies_seedOpen, repository_raf_has_catalyzed_gateway`.

Proof. If no channel of S were a seed, no side of any channel of S would be contained in $\text{cl}_S^0 = F$, so $\text{cl}_S^1 = F$, and by induction $\text{cl}_S^k = F$ for all k (`revClosureAt_eq_food_of_no_seed`). Take any $r \in S$; food generation gives k with $\text{lhs}(r) \subseteq \text{cl}_S^k = F$, so r is a seed after all. Reflexive autocatalysis supplies its catalyst. $\square$

8.3 The probability cap

Let Γ and $K_n = |X_n| \cdot |\Gamma| = 34(2^{n+1} - 2)$ be as above (`repositoryGatewayCoordinateCount`), and define the gateway-open probability

$$G_n(p) = 1 - (1 - p)^{K_n} \quad (\text{repositoryGatewayOpenProbability}).$$

For $n \geq 4$ this is exactly the Bernoulli probability that some seed channel is catalysed by some molecule; for smaller n it is an upper bound for that probability.

Theorem 8.3 (Gateway cap). *For every $n \geq 1$ and $p \in [0, 1]$, with $K = K_{n,2}$,*

$$0 \leq \Pr_p(\text{RAF}_n) \leq \Pr_p(\text{some seed channel catalysed}) \leq G_n(p) \leq K_n p.$$

Lean: `jointFibreBernoulliWeight_nonneg, repositoryRAFBernoulliProbability_le_seedOpen, repositoryRAFBernoulliProbability_le_gateway, repositoryGatewayOpenProbability_le`.

Proof. Model the seed-open event as a catalogue in the sense of Section 3: one core per seed channel r , with requirement $W_{r,j} = X_n$ if $j = r$ and $\emptyset$ otherwise (`repositorySeedRequirements`). Its cores are coordinate-disjoint, so Theorem 7.2 gives the exact probability $1 - (1 - p)^{|X_n| |\text{Seed}_2(n)|}$ (`repositorySeedOpenProbability_exact`). By Theorems 3.3 and 8.2, every configuration in the RAF event lies in the seed-open event (`repositoryRAFFibreConfigurations_subset_seedOpenConfigurations`), and Bernoulli weights are nonnegative, which gives the first two inequalities. The third is $|\text{Seed}_2(n)| \leq 34$ from Theorem 8.1, and the last is the union bound $1 - (1 - p)^K \leq Kp$. $\square$

We stress that the containment in Theorem 8.3 is proved for the actual RAF event, using the exhaustive catalogue and Theorem 3.3. An earlier version of the development stated the asymptotic results for an arbitrary probability sequence dominated by G_n ; this was replaced because such a statement leaves the essential step, that the RAF event is inside the seed-open event, to the reader.

9 Asymptotic consequences

Throughout this section the food horizon is $t = 2$, $p_n = f_n/|J_n|$, and $\Pr(\text{RAF}_n)$ denotes the Bernoulli probability of the actual RAF event (`repositoryRAFBernoulliProbability`).

9.1 Sublinear catalysis

Proposition 9.1 (Quantitative cap). *For $n \geq 3$ and $0 \leq f_n \leq |J_n|$, $G_n(p_n) \leq 272 f_n/n$.*

Lean: `repositoryGatewayOpenProbability_le_sublinearScale`.

Proof. By Theorem 8.3, $G_n(p_n) \leq K_n p_n = 34 f_n |X_n|/|J_n|$. Now $|X_n| \leq 2 \cdot 2^n$ and, by (1), $|J_n| \geq \lfloor (n-1)/2 \rfloor 2^n \geq (n/4) 2^n$ for $n \geq 3$. Hence $|X_n|/|J_n| \leq 8/n$ and $G_n(p_n) \leq 272 f_n/n$. $\square$

The constant $272 = 34 \cdot 8$ is deliberately coarse. Proposition 2.1 would replace 8 by $1 + o(1)$, but the formal development uses only (1), which is all that is needed for the scale separation.

Theorem 9.2 (Sublinear catalysis forces RAF probability to vanish). *Let $f : \mathbb{N} \rightarrow \mathbb{R}$ satisfy $0 \leq f_n \leq |J_n|$ for all n and $f_n/n \rightarrow 0$. Then $\Pr(\text{RAF}_n) \rightarrow 0$. In particular this holds for every constant f .*

Lean: `sublinear_catalysis_forces_actual_repository_raf_vanishing`.

Proof. For $n \geq 3$, Theorem 8.3 and Proposition 9.1 give $0 \leq \Pr(\text{RAF}_n) \leq 272f_n/n \rightarrow 0$. $\square$

9.2 Linear catalysis

Every fixed f is sublinear, so the first scale not excluded by Theorem 9.2 is $f_n = \lambda n$. On that scale the expected number of open gateway coordinates stays bounded.

Lemma 9.3. *For $\lambda \geq 0$ and $n \geq 3$, $p_n K_n \leq 272\lambda$ when $f_n = \lambda n$; and $K_n \rightarrow \infty$.*

Lean: `repository_linear_rate_mul_gatewayCoordinateCount_le`,
`repositoryGatewayCoordinateCount_tendsto`.

Proof. $p_n K_n = 34\lambda n |X_n|/|J_n| \leq 34\lambda n \cdot 8/n$ by the estimates in the proof of Proposition 9.1; and $K_n \geq 2^n$. $\square$

Theorem 9.4 (Finite linear catalysis cannot force RAF probability to one). *For every real $\lambda > 0$, the sequence $\Pr(\text{RAF}_n)$ at $f_n = \lambda n$ does not converge to 1.*

Lean: `linear_catalysis_actual_repository_raf_not_tendsto_one`.

Proof. Put $C = 272\lambda$. By Lemma 9.3, eventually $C \leq K_n$ and $p_n \leq C/K_n$. Since $(1 - C/K_n)^{K_n} \rightarrow e^{-C}$, there is $\delta > 0$ with $(1 - p_n)^{K_n} \geq (1 - C/K_n)^{K_n} \geq \delta$ for all large n (`bernoulli_absent_eventually_pos`). By Theorem 8.3, $\Pr(\text{RAF}_n) \leq 1 - (1 - p_n)^{K_n} \leq 1 - \delta$ eventually, which is incompatible with convergence to 1. $\square$

The theorem says that the boundary keeps a non-vanishing influence at every finite linear rate: with positive probability bounded away from zero, none of the 34 entry channels is catalysed and no RAF can exist. It does not determine $\lim \Pr(\text{RAF}_n)$, nor assert that the limit exists.

9.3 The exact boundary–bulk factorisation

Proposition 9.5 (Gateway-conditioned bulk factor). *For every n and p with $G_n(p) \neq 0$ define $B_n(p) = \Pr_p(\text{RAF}_n)/G_n(p)$. Then $\Pr_p(\text{RAF}_n) = G_n(p) B_n(p)$.*

Lean: `repositoryRAFBernoulliProbability_gateway_factorization`.

The identity is algebraic and asserts no independence. By Theorem 8.3, $0 \leq B_n(p) \leq 1$ for $p \in (0, 1]$, and for $n \geq 4$, where $G_n(p)$ is exactly the seed-open probability, $B_n(p)$ is the conditional probability of a RAF given that some seed channel is catalysed. Its purpose is to name the object that remains to be understood: all unresolved large- n behaviour on the linear scale lives in $B_n(\lambda n/|J_n|)$.

9.4 The sharp gateway constant

Corollary 9.6 (Sharp gateway limit; not formalised). *For every $\lambda \geq 0$, $G_n(\lambda n/|J_n|) \rightarrow 1 - e^{-34\lambda}$.*

Proof. By Proposition 2.1, $|J_n| = (n-2)2^{n+1}(1+o(1))$, so $p_n K_n = 34\lambda n(2^{n+1}-2)/|J_n| \rightarrow 34\lambda$, while $p_n \rightarrow 0$. Hence $K_n \log(1-p_n) = -p_n K_n(1+O(p_n)) \rightarrow -34\lambda$ and $(1-p_n)^{K_n} \rightarrow e^{-34\lambda}$. $\square$

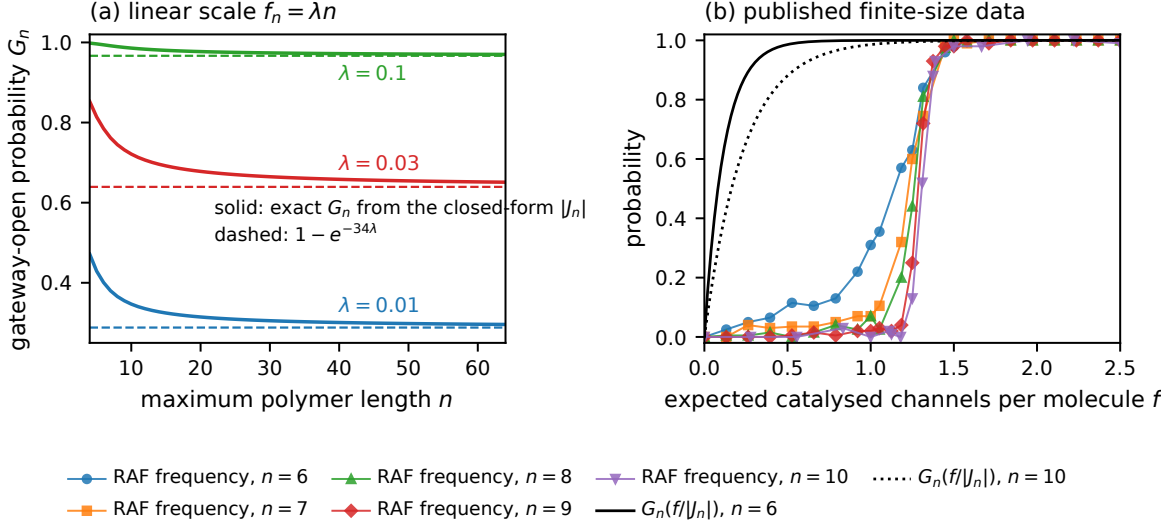


Figure 3: (a) Exact gateway-open probability $G_n(\lambda n/|J_n|)$ for $\lambda \in \{0.01, 0.03, 0.1\}$ with $|J_n|$ from Proposition 2.1, against the limits $1 - e^{-34\lambda}$ of Corollary 9.6. (b) The 111 RAF frequencies published in [22] for $n = 6, \dots, 10$ as a function of f , with the exact gateway curves $G_n(f/|J_n|)$ for $n = 6$ and $n = 10$. In the finite transition window the gateway is almost surely open; the finite curve is not shaped by it.

Figure 3(a) shows the exact finite curves $G_n(\lambda n/|J_n|)$ computed from the closed form of Proposition 2.1; at $\lambda = 0.03$ the value moves from 0.7848 at $n = 6$ to 0.6511 at $n = 64$ against the limit $1 - e^{-1.02} \approx 0.6394$. The convergence is slow, of order $1/n$, because $n|X_n|/|J_n| = (n/(n-2))(1 + o(1))$. Together with Proposition 9.5 the corollary gives the exact asymptotic decomposition on the linear scale,

$$\Pr(\text{RAF}_n) = (1 - e^{-34\lambda} + o(1)) B_n(\lambda n/|J_n|),$$

and reduces the emergence problem at that scale to the bulk factor. The corollary is not needed for any of Theorems 9.2 and 9.4, which rest only on (1).

10 Numerical diagnostics

Numerical work was used to discriminate conventions and to motivate statements, never as evidence for a theorem. Three small exact computations were decisive.

Reaction census. A semantic re-implementation of the repository generator (which records displayed reactant and product lists and suppresses duplicates) was compared with Proposition 2.1 for $n = 2, \dots, 8$ (Table 1). The counts agree, every directed record is followed by its exact reverse, and the seed count stabilises at 34 from $n = 4$. The mismatch with the split-position catalogue at $n = 3$ (18 versus 20) was the observation that forced the reconstruction of the quotient before any probability theorem was attempted.

Single-catalyst witnesses. The repository file `single_solutions.json` lists 96 single-catalyst RAF witnesses at $n = 6$. Each is a pair of adjacent directed indices, that is one reversible channel, together with one catalyst molecule lying in the closure of that channel. The 96 witnesses are supported on exactly 34 distinct channels, the seed channels of Theorem 8.1. This illustrates the distinction, built into Section 3, between a reaction-support core and a

Table 1: Exact census of the repository generator (binary alphabet, food horizon 2). “Split” is the split-position count $(n - 2)2^{n+1} + 4$; “ $C(n)$ ” is the commuting-pair correction of Proposition 2.1.

n	$ X_n $	directed reactions	$ J_n $	split	$C(n)$	seed channels
2	6	8	4	4	0	4
3	14	36	18	20	2	18
4	30	128	64	68	4	34
5	62	376	188	196	8	34
6	126	1004	502	516	14	34
7	254	2528	1264	1284	20	34
8	510	6096	3048	3076	28	34

catalyst-labelled witness of it: treating the 96 objects as 96 cores in the outer inclusion–exclusion would count the same event many times.

Gateway versus the finite transition. Figure 3(b) overlays the 111 published RAF frequencies with the exact gateway curve. At the observed transition the gateway is already almost surely open: at $n = 6$ and $f \approx 1.316$ the RAF frequency is 0.84 while $G_6 = 0.99999$; at $n = 10$ and $f = 1.31$ the 25-run estimate is 0.52 while $G_{10} = 0.9962$. The gateway therefore does not shape the finite-size curve. Its role is asymptotic: at fixed f the same curve cannot persist, because $G_n(f/|J_n|) = O(f/n)$.

Performance. No enumeration of catalysis configurations was performed anywhere on the proof path. The pairwise-insufficiency fixture is evaluated exhaustively inside Lean on a four-element set; the channel and gateway counts at $n = 3$ and horizon 2 are kernel evaluations on types of size at most 14^2 .

11 Formal verification

The development is organised in four layers under `proofs/OverlapCorrectedRAF/`: **Source** (the repository quotient, its reversible system, the semantic catalogue, the gateway dock, and the exact seed count), **Core** (finite core events and one-core probabilities), **Overlap** (local hit polynomials, semantic counting, the finite partition, the pairwise counterexample, and the non-overlap specialisation), and **Asymptotic** (gateway probability, sublinear and linear catalysis, and the bulk factorisation). The terminal theorem, `OverlapCorrectedRAF.correctedEmergenceResolution`, takes as hypotheses natural numbers n, t, Q , a real f , a catalysis relation C and a channel set S with `IsRevRAF( $\mathcal{Q}_{n,t}, C, S$ )`, and a sequence $f : \mathbb{N} \rightarrow \mathbb{R}$ with $0 \leq f_m \leq |J_m|$ for all m and $f_m/m \rightarrow 0$. It concludes the conjunction of `HasExactRepositoryRAFLaw  $n t Q$` , of the existence of a seed channel r and a molecule x with $C(x, r)$, and of $\Pr(\text{RAF}_m) \rightarrow 0$, which packages Corollary 5.2, Theorem 8.2 and Theorem 9.2. The post-proof modules add Theorems 8.1 and 9.4 and Proposition 9.5. All modules compile under Lean 4.30.0 against a frozen Mathlib with `-DwarningAsError=true`, exit code zero, and no `sorry`. Table 2 lists the SHA-256 hashes (first 16 hexadecimal digits) of the sources and of the compiled artefacts. The reproduction command from the repository root is

```
py scripts/verify_proof.py
proofs/OverlapCorrectedRAF/CorrectedEmergenceResolution.lean
```

Two remarks on what the formalisation forced. First, the exact formulas of Sections 4 and 5 were initially *definitions*; they became theorems only when each polynomial was proved equal to

Table 2: Verification receipts. Hashes are SHA-256, truncated to 16 hexadecimal digits; the full values are recorded in the repository handoff.

Module	source	artefact
Overlap/FinitePartition.lean	a3c94dced502d88c	21113c5713af0117
Overlap/SemanticCounting.lean	d091c8ceca281272	0105b89a98eb287a
Overlap/TriplePairwiseFailure.lean	9622c13de8d246fb	40fd1d100a89eb15
Overlap/Nonoverlap.lean	3bd660e9ccd9d688	42b28a1d9f013ecb
Source/RepositorySemanticCounting.lean	4929e5fe8f5d8318	83d27e083680a59e
Source/ActualGatewayDock.lean	7c22a10632f2201a	827f33ad5ccff87a
Source/ExactSeedChannels.lean	1a4abdf9ea110b2a	bf93795fce3b7271
Asymptotic/SublinearCatalysis.lean	ec241b57f7766fee	dd94a0946dfaf8ca
Asymptotic/LinearCatalysis.lean	23d37db67fb5ebff	66e0cd9d11de5195
Asymptotic/GatewayConditionedBulk.lean	aac959e41cda9510	d07f7b82eed0f38a
CorrectedEmergenceResolution.lean	7aa9018c0c15aadb	27ec5bbe0e46db17

the generating polynomial of an explicit finite set of configurations (`fibreHitConfigurations`, `jointFibreHitConfigurations`, `catalogueFibreHitConfigurations`) and those sets were identified with the RAF event. Second, the fixed- Q probability has codomain $\mathbb{R}$, with the integer numerator and binomial denominator cast before division; a natural-number quotient would not define a probability. Both points are invisible in informal presentations and both were caught by the type checker.

12 Discussion

What is and is not claimed

The paper proves an exact finite law, identifies its sufficient statistic, shows that pairwise data are not sufficient, recovers the independent-core formula under coordinate disjointness, establishes the exact 34-channel gateway, and derives the sublinear vanishing and finite-linear obstruction theorems for the actual RAF event of the repository model. It also proves by hand, outside Lean, the closed-form channel count and the sharp gateway limit.

It does not claim a positive RAF limit at $f_n = \lambda n$, a transition width, a Poisson law for the number of cores, or a decay rate for connected overlap clusters. It does not claim that the gateway explains the shape of the finite transition; Section 10 shows that it does not. It does not claim computational efficiency for Theorem 5.1. It does not identify the repository quotient with the split-position model, nor the Bernoulli with the fixed- Q ensemble, nor the 96 catalyst witnesses with 96 cores. The homogeneous ensembles are essential to Remark 4.4: for molecule-dependent or correlated catalysis distributions [8] the local law would need the full set-valued trace $\{\cup \mathcal{T}\}$ rather than its cardinalities, and the fixed- Q ensemble would no longer be determined by Q alone.

Interpretation

Three structural facts, rather than one, were hidden in the phrase “overlapping cores”. The first is that dependence between cores is a third- and higher-order phenomenon on each channel, so the correct state variable is a union trace. The second is that the reaction identity of the source is a quotient of the textbook catalogue, so an exact theorem must be proved in the quotient. The third is that a fixed food set is a boundary of bounded size in a system whose bulk grows exponentially; the exponentially many catalyst molecules do not compensate, because the per-coordinate probability shrinks at exactly the compensating rate $f/|J_n|$. The finite-size

transition near $f \approx 1.3$ is real, and the finite-core theory describes it; what the gateway forbids is the extrapolation of a constant- f transition to $n \rightarrow \infty$.

Open problems

- (1) *The bulk factor.* Determine the behaviour of $B_n(\lambda n/|J_n|)$ as $n \rightarrow \infty$. For the split-position catalogue the companion manuscript [1] proves that the RAF probability itself converges to $\Theta(\lambda) = S(1 - e^{-\lambda})$, a nondecreasing continuous function with $0 < \Theta < 1$ on $(0, \infty)$, so that no finite λ separates limit 0 from limit 1. Transferring that analysis to the repository quotient, where Corollary 9.6 fixes the gateway factor at $1 - e^{-34\lambda}$, would determine $\lim B_n$; it is open whether the two conventions have the same bulk limit.
- (2) *Compression of the trace.* Quotient the union trace to a coverage polynomial whose coefficients group equal union cardinalities, and determine when this smaller statistic is exact for homogeneous catalysis.
- (3) *Support cores versus witnesses.* Formalise the resummation from catalyst-labelled minimal witnesses to reaction-support cores that explains the $96 \rightarrow 34$ collapse in general.
- (4) *Cluster factorisation.* Cores with disjoint coordinate sets are independent under Bernoulli catalysis; factor the no-RAF probability over connected components of the overlap graph and use it as the basis for cluster expansions and bounded-width dynamic programmes.
- (5) *Universality of finite gateways.* Abstract the argument of Section 8 to random catalytic systems with a uniformly bounded mandatory boundary set, per-coordinate probability $f_n/|J_n|$ and molecule-to-channel ratio of order $1/n$, and formalise Proposition 2.1 and Corollary 9.6.

Data and code availability

The Lean sources, the exact census scripts, the audit of the single-catalyst fixture, the 111 published RAF frequencies used in Figure 3, and the script that generates that figure are available in the public repository <https://github.com/mmislán/Erdos> under `proofs/OverlapCorrectedRAF/`, `problem_workspaces/RAF_overlap_corrected_Kauffman_emergence/` and `key_results/RAFs/`.

References

- [1] [Author name to be inserted]. A nontrivial critical window for RAF emergence in the binary polymer model. Companion manuscript, 2026. Treats the split-position reaction catalogue, in which the binary-food gateway has 36 reaction identities, and proves that the RAF probability converges to $S(1 - e^{-\lambda})$ when $f_n/n \rightarrow \lambda$, 2026.
- [2] Leonardo de Moura and Sebastian Ullrich. The Lean 4 theorem prover and programming language. In *Automated Deduction – CCADE 28*, volume 12699 of *Lecture Notes in Computer Science*, pages 625–635. Springer, 2021.
- [3] Janos Galambos and Italo Simonelli. *Bonferroni-type Inequalities with Applications*. Probability and its Applications. Springer, New York, 1996.
- [4] Wim Hordijk, Jotun Hein, and Mike Steel. Autocatalytic sets and the origin of life. *Entropy*, 12(7):1733–1742, 2010.

- [5] Wim Hordijk, Stuart A. Kauffman, and Mike 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.
- [6] Wim Hordijk, Joshua I. Smith, and Mike Steel. Algorithms for detecting and analysing autocatalytic sets. *Algorithms for Molecular Biology*, 10:15, 2015.
- [7] Wim Hordijk and Mike Steel. Detecting autocatalytic, self-sustaining sets in chemical reaction systems. *Journal of Theoretical Biology*, 227(4):451–461, 2004.
- [8] Wim Hordijk and Mike Steel. Autocatalytic sets in polymer networks with variable catalysis distributions. *Journal of Mathematical Chemistry*, 54(10):1997–2021, 2016.
- [9] Wim Hordijk and Mike Steel. Chasing the tail: The emergence of autocatalytic networks. *Biosystems*, 152:1–10, 2017.
- [10] Wim Hordijk and Mike Steel. Autocatalytic networks at the basis of life’s origin and organization. *Life*, 8(4):62, 2018.
- [11] Wim Hordijk, Mike Steel, and Stuart Kauffman. The structure of autocatalytic sets: Evolvability, enablement, and emergence. *Acta Biotheoretica*, 60(4):379–392, 2012.
- [12] Stuart A. Kauffman. Autocatalytic sets of proteins. *Journal of Theoretical Biology*, 119(1):1–24, 1986.
- [13] Stuart A. Kauffman. *The Origins of Order: Self-Organization and Selection in Evolution*. Oxford University Press, New York, 1993.
- [14] M. Lothaire. *Combinatorics on Words*. Cambridge Mathematical Library. Cambridge University Press, Cambridge, 1997.
- [15] 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.
- [16] Elchanan Mossel and Mike Steel. Random biochemical networks: the probability of self-sustaining autocatalysis. *Journal of Theoretical Biology*, 233(3):327–336, 2005.
- [17] Joshua I. Smith, Mike Steel, and Wim Hordijk. Autocatalytic sets in a partitioned biochemical network. *Journal of Systems Chemistry*, 5:2, 2014.
- [18] Mike Steel. The emergence of a self-catalysing structure in abstract origin-of-life models. *Applied Mathematics Letters*, 13(3):91–95, 2000.
- [19] Mike Steel, Wim Hordijk, and Joshua Smith. Minimal autocatalytic networks. *Journal of Theoretical Biology*, 332:96–107, 2013.
- [20] Mike Steel, Wim Hordijk, and Joana C. Xavier. Autocatalytic networks in biology: structural theory and algorithms. *Journal of the Royal Society Interface*, 16(151):20180808, 2019.
- [21] 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.

- [22] Vijay 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>, 2026. Files `raf_results.csv` (blob 8ab510aa) and `single_solutions.json` (blob 1f446029) as retrieved 1 September 2026.
- [23] Vijay Varanasi and Jun Korenaga. Emergence of autocatalysis in prebiotic reaction networks. *Physical Review E*, 113:044304, 2026.