Elementary cores, abundant reactions, and the union-closed boundary of autocatalytic reaction systems

Author name to be inserted

Affiliation to be inserted

13 September 2026

Abstract

A reflexively autocatalytic and food-generated set (RAF) of a catalytic reaction system is a collection of reactions that can be built up from a food set and in which every reaction is catalysed from within. The RAFs of a finite system, with the empty set adjoined, form a union-closed family, and Steel asked whether some reaction always lies in at least half of its members: a RAF form of Frankl's union-closed sets conjecture. We prove this for every system that contains a *viable elementary core*, a nonempty RAF all of whose reactants are food, with no restriction on the remaining reactions or on catalytic feedback into the core. The proof conditions on the reactions selected outside the core, recomputes the core's support relation in that context, and applies the Horn-function counting injection of Lozin and Zamaraev to the resulting support relation under an upward constraint. The occupancy inequality holds separately in every exterior context, survives arbitrary nonnegative weights on contexts, and implies that every catalytic cycle of food-ready reactions meets a globally abundant reaction, that vertex-disjoint cycles yield distinct abundant reactions, and that systems with at most two non-food-ready reactions satisfy the half-frequency bound. Sharpness and a changing-witness example are given, together with exact projection diagnostics for an embedded module. We then locate the boundary of the method. RAF interior operators on a fixed reaction set are characterised as accessible union-closed food feasibility intersected with predecessor support; food-generation families are exactly antimatroids. A producer–gate construction with two reactions per coordinate realises every finite union-closed family containing the empty set, preserving frequencies and all availability queries with one food molecule, one catalyst per reaction, and molecular closure stabilising after two rounds. Its elementary RAFs correspond exactly to the singleton members of the encoded family. Hence a half-frequency theorem for systems without an elementary core is equivalent to the unrestricted union-closed conjecture, which remains open. The main theorems are verified in Lean 4 against Mathlib; the correspondence between the article and the formal development is recorded in an appendix.

Keywords. autocatalytic sets; RAF theory; union-closed sets conjecture; interior operators; antimatroids; Horn functions; formal verification.
MSC 2020. 05D05, 92C42, 06A15, 68V20.

1 Introduction

1.1 Autocatalytic sets and the RAF formalism

Collectively autocatalytic sets were proposed by Kauffman [11, 12] as a route by which a mixture of simple molecules could organise into a self-sustaining chemistry: a set of reactions each of whose products is needed, as reactant or catalyst, by other reactions of the set, so that the set as a whole reproduces itself from an ambient food supply. Hordijk and Steel [9] made this precise in the RAF formalism. A *catalytic reaction system* consists of molecule types, reactions with reactant and product sets, a distinguished *food* set, and a catalysis relation. A set of reactions is *food-generated* if its reactants can all be produced, in some order, from food using only the selected reactions, and it is *reflexively autocatalytic* if every selected reaction has a catalyst that is food or produced by the selected reactions. A RAF is a nonempty set that is both. Hordijk and Steel gave a polynomial-time algorithm for the largest RAF inside any given set of reactions, and the formalism has since been developed extensively; we refer to [10, 15, 21] for structural theory, algorithms, and random-network analysis.

The union of two RAFs is again a RAF. Adjoining the empty set, the RAFs of a finite system therefore form a *union-closed family* of subsets of the reaction set, and the map sending an available reaction set to the largest RAF it contains is an *interior operator*: contractive, monotone and idempotent. Steel [20] studied this operator in detail and, in his concluding section, asked

> "whether there is always a reaction that lies in at least half the RAFs (for either an elementary or general CRS)."

Here *elementary* means that every reaction has all its reactants in the food set. Since the RAF family is union-closed, a positive answer for all systems would follow from Frankl's union-closed sets conjecture [2, 7], which asserts that every finite union-closed family other than $\{\varnothing\}$ has an element lying in at least half of its members. That conjecture is open. After a long period in which only logarithmic bounds were known [23], Gilmer [8] proved a constant lower bound of 0.01 by an entropy argument, quickly improved to $(3-\sqrt{5})/2 \approx 0.382$ and beyond [1, 3, 4, 19]; the constant $1/2$ remains out of reach.

The question for RAF systems has two faces. Does the chemical structure of a RAF family make the half-frequency statement provable where it is not provable for arbitrary union-closed families? Or is every union-closed family already a RAF family in disguise, so that the RAF question is exactly as hard as Frankl's? This paper answers both questions as precisely as we are able: the first positively for a large structurally defined class, the second positively for the complement of that class.

1.2 Main results

Throughout, $\mathcal{F}(\mathcal{Q})$ denotes the family of RAFs of a finite catalytic reaction system $\mathcal{Q}$ together with the empty set, $N = |\mathcal{F}(\mathcal{Q})|$, and $f(r)$ is the number of members containing the reaction r. A reaction is *abundant* if $2f(r) \geq N$. Because the empty set contributes to the denominator but to no numerator, abundance is slightly stronger than the half-frequency statement over nonempty RAFs alone; Steel's question is the weaker form, and all our positive results give the stronger one.

A reaction is *food-ready* if all its reactants are food. An *elementary core* of $\mathcal{Q}$ is a nonempty RAF of $\mathcal{Q}$ consisting of food-ready reactions. The core is a subset of a system that may otherwise be entirely unrestricted.

Theorem 1.1 (Elementary-core abundance; Theorem 4.1)**.** *Let U be an elementary core of a finite catalytic reaction system $\mathcal{Q}$. Then some reaction of U is abundant in $\mathcal{F}(\mathcal{Q})$. More precisely, for every set T of reactions outside U, the core parts $S = W \cap U$ of the members $W \in \mathcal{F}(\mathcal{Q})$ with $W \setminus U = T$ have average size at least $|U|/2$.*

The fibrewise form is the real content. Fixing the exterior selection T changes which core reactions are catalysed from outside; the proof recomputes the internal support relation of U in the presence of T and shows that the family of compatible core subsets is always a "supported family under an upward constraint", for which a counting injection gives average size at least half the ground set. The counting injection is the horizontal/vertical cell map of Lozin and Zamaraev [14], who used it to prove Frankl's conjecture for Horn functions with a *dependency* *property*; an upward constraint is a positive Boolean predicate and satisfies that property vacuously, so the lemma we need (Lemma 3.1) is the aggregate form of their theorem. No assumption is made that the exterior preserves the isolated support structure of the core; indeed we exhibit systems where it does not (Section 6).

The fibrewise inequality has several consequences that a bare existence statement would not give (Section 5). The half bound survives any nonnegative reweighting of exterior contexts; every directed cycle in the food-ready support digraph contains a globally abundant reaction, so the strictly rare food-ready reactions induce an acyclic subgraph and k vertex-disjoint cycles force k distinct abundant reactions; and any system with a RAF and at most two non-food-ready reactions has an abundant reaction, by a separate two-element transversal argument. The constant $1/2$ is attained, and the abundant reaction inside the core may depend on the exterior context (Example 5.2).

The second group of results describes what RAF families can be (Sections 7–8).

Theorem 1.2 (Exact-ground realizability; Theorem 7.5)**.** *An interior operator on* 2^V *is the RAF interior operator of a finite catalytic reaction system with reaction set exactly* V *if and only if its fixed family is the intersection of an accessible union-closed family containing* $\varnothing$ *with a predecessor-support family* $\{S :\ S \cap P(v) \neq \varnothing \text{ for all } v \in S\}$*. In particular the families of food-generated sets are exactly the finite antimatroids (Corollary 7.6).*

Theorem 1.3 (Sparse universality; Theorem 8.2 and Proposition 8.3)**.** *Every finite union-closed family* $\mathcal{D} \subseteq 2^V$ *containing* $\varnothing$ *is, up to the bijection* $A \mapsto \{p_v, g_v : v \in A\}$*, the RAF family of a system with* $2|V|$ *reactions, one food molecule, exactly one catalyst per reaction, no self-catalysis, catalytic product edges forming disjoint two-cycles, and molecular closure stabilising after two rounds for every availability set. The bijection preserves frequencies and translates every maxRAF query into an interior query on complete pairs. The elementary RAFs of this system are the encodings of the singleton members of* $\mathcal{D}$*.*

Since a union-closed family with a singleton member trivially has an abundant element, Theorem 1.3 shows that the class of systems *without* an elementary core carries the full difficulty of the union-closed conjecture: a half-frequency theorem for that class is equivalent to Frankl's conjecture (Theorem 8.4). The pair of theorems thus draws a sharp line. On one side, elementary cores force abundance with room to spare; on the other, no amount of structural sparsity in the catalytic graph helps, because the difficulty is carried entirely by reactant requirements.

1.3 Relation to prior work and what is new

The RAF framework, the maxRAF interior operator, and the question itself are due to Hordijk and Steel [9] and Steel [20]. The characterisation of elementary RAFs by directed cycles in a support digraph, and the shortest-cycle method for a smallest elementary RAF, appear in Steel, Hordijk and Xavier [21]; we use them as certificates for abundance in a surrounding system. The counting kernel is Theorem 4 of Lozin and Zamaraev [14], which proves Frankl's conjecture for Boolean functions admitting a Horn DNF with their dependency property: every variable occurring negatively has a companion variable occurring positively in each term in which it is negative. The RAF family of an isolated elementary system is an instance. What is new here is (i) the exterior-conditioned application, in which the Horn bodies depend on the fixed exterior and food generation together with exterior support enter as a monotone constraint, giving abundance for embedded cores under arbitrary feedback; (ii) the weighting, cycle, disjointness and two-exception consequences; (iii) the exact-ground characterisation and the antimatroid identification of food generation; and (iv) the sparse producer–gate universality theorem with its exact elementary boundary and the resulting equivalence with the unrestricted conjecture.

We make no claim on the unrestricted conjecture. A focused comparison with the sources cited is not a certification of priority across the whole literature.

The classical observation that a union-closed family containing a singleton (or a two-element set) satisfies the conjecture is due to Sarvate and Renaud [18] (see also Poonen [16]). Reimer's average set size theorem [17] gives average member size at least $\frac{1}{2}\log_2 N$ for any union-closed family; our occupancy inequalities are of the same average-size type but give the much larger bound $|U|/2$ under the structural hypothesis, and only for the core coordinates.

1.4 Formal verification

The theorems marked with a Lean identifier in Appendix A have been formalised in Lean 4 [5] against Mathlib [22] directly from the literal RAF semantics of Definition 2.1, compiled with warnings as errors, and depend only on the standard axioms `Classical.choice`, `Quot.sound` and `propext`. The paper is written to be read without the formal development; the appendix states exactly which forms were compiled and which steps (a few reindexings and the examples) are ordinary mathematics only.

1.5 Organisation

Section 2 fixes the RAF semantics and basic facts. Section 3 proves the counting lemma. Section 4 proves the exterior-conditioned abundance theorem. Section 5 derives the weighted, cycle, disjointness and two-exception consequences and gives examples. Section 6 gives exact diagnostics for when an embedded module's isolated behaviour is visible globally. Sections 7 and 8 contain the realisation theorems and the equivalence with the unrestricted conjecture. Section 9 discusses certificates and algorithms, and Section 10 states the limits of the results and open problems. Appendix A records the formal verification.

2 Ordinary RAF semantics

Definition 2.1. A finite *catalytic reaction system* (CRS) is a tuple $\mathcal{Q} = (X, R, F, \rho, \pi, C)$ where X is a finite set of molecule types, R a finite set of reactions, $F \subseteq X$ the *food* set, $\rho(r), \pi(r) \subseteq X$ the reactants and products of $r \in R$, and $C \subseteq X \times R$ the catalysis relation; write $C(r) = \{x : (x, r) \in C\}$. For $W \subseteq R$ define the *molecular closure* stages

$$L_0(W) = F, \qquad L_{j+1}(W) = L_j(W) \cup \bigcup\{\pi(r) : \ r \in W, \ \rho(r) \subseteq L_j(W)\},$$

and the *closure* $\mathrm{cl}_W(F) = \bigcup_{j \geq 0} L_j(W)$. W is *food-generated* if $\rho(r) \subseteq \mathrm{cl}_W(F)$ for every $r \in W$. W is a *RAF* if it is nonempty, food-generated, and $C(r) \cap \mathrm{cl}_W(F) \neq \varnothing$ for every $r \in W$.

Catalysis plays no role in the closure recursion: a reaction contributes its products as soon as its reactants are present, and catalytic support is checked afterwards in the resulting closure. This is the semantics of [9, 20]; we impose no stoichiometry, inhibition or kinetics. Since X is finite the closure stabilises after finitely many stages.

Lemma 2.2 (Products identity)**.** *Closure is monotone: $W \subseteq W'$ implies $L_j(W) \subseteq L_j(W')$ for all j. If W is food-generated then*

$$\mathrm{cl}_W(F) = F \cup \pi(W), \qquad \pi(W) := \bigcup_{r \in W} \pi(r). \tag{1}$$

Proof. Monotonicity is immediate by induction on j. Every stage adds only products of reactions in W, so $\mathrm{cl}_W(F) \subseteq F \cup \pi(W)$. Conversely if W is food-generated and $r \in W$, then $\rho(r) \subseteq L_j(W)$ for some finite j, whence $\pi(r) \subseteq L_{j+1}(W)$. $\square$

Thus, *for a food-generated set*, catalytic support is exactly support by food or by a product of a selected reaction. The qualification matters: products of a selected reaction that never becomes enabled are not available.

Definition 2.3. For a finite CRS $\mathcal{Q}$ put

$$\mathcal{F} = \mathcal{F}(\mathcal{Q}) = \{\varnothing\} \cup \{W \subseteq R : \ W \text{ is a RAF}\}, \qquad N = |\mathcal{F}|, \qquad f(r) = |\{W \in \mathcal{F} : r \in W\}|.$$

A reaction is *abundant* if $2f(r) \geq N$ and *strictly rare* otherwise. A reaction is *food-ready* if $\rho(r) \subseteq F$. An *elementary core* is a nonempty RAF all of whose reactions are food-ready. For $A \subseteq R$, $\mathrm{maxRAF}_{\mathcal{Q}}(A) = \bigcup\{W \in \mathcal{F} : W \subseteq A\}$.

Lemma 2.4. *$\mathcal{F}$ is union-closed.* $\mathrm{maxRAF}_{\mathcal{Q}}(A)$ *is the largest RAF contained in A, or empty if there is none, and $A \mapsto \mathrm{maxRAF}_{\mathcal{Q}}(A)$ is an interior operator on 2^R with fixed family $\mathcal{F}$.*

Proof. If W, W' are RAFs then $W \cup W'$ is nonempty; by monotonicity of closure each reaction of $W \cup W'$ keeps its food generation and its catalyst. The union of all RAFs inside A is therefore a RAF or empty, and the remaining properties are immediate. □

Conversely, every interior operator ψ on a finite power set is determined by its fixed family $\mathrm{Fix}(\psi)$, which contains $\varnothing$ and is union-closed, via $\psi(A) = \bigcup\{S \in \mathrm{Fix}(\psi) : S \subseteq A\}$. We use "interior operator" and "union-closed family containing $\varnothing$" interchangeably. For such a family $\mathcal{D} \subseteq 2^V$ we write $i_{\mathcal{D}}(A) = \bigcup\{D \in \mathcal{D} : D \subseteq A\}$.

Remark 2.5 (Two normalisations of the conjecture). Steel's question concerns nonempty RAFs: is there r with $f(r) \geq (N-1)/2$? Our abundance condition $2f(r) \geq N$ is stronger. For the union-closed conjecture the two normalisations are equivalent: applying the standard conjecture to $\mathcal{D} \cup \{\varnothing\}$ gives the empty-inclusive bound, and the empty-inclusive bound for $\mathcal{D} \cup \{\varnothing\}$ implies the standard bound for $\mathcal{D}$ because the denominator can only grow. We work exclusively with the empty-inclusive form because it is the one that sums correctly over exterior fibres (the empty fibre member is a genuine row of the count).

3 Supported families under an upward constraint

A predicate K on 2^U is *upward* if $K(S)$ and $S \subseteq S'$ imply $K(S')$.

Lemma 3.1 (Occupancy of supported families)**.** *Let U be a finite set with $n = |U|$. For each $r \in U$ let $B_r \subseteq U$ be a nonempty* body*, and let K be an upward predicate on 2^U. Put*

$$\mathcal{B} = \{S \subseteq U : \ K(S) \text{ and } S \cap B_r \neq \varnothing \text{ for every } r \in S\}.$$

Then

$$\sum_{S \in \mathcal{B}} (2|S| - n) \ \geq \ 0, \qquad \textit{equivalently} \qquad n\,|\mathcal{B}| \leq 2 \sum_{S \in \mathcal{B}} |S|. \tag{2}$$

In words: if membership requires each selected element to be "supported" by a selected element of its body, and is additionally restricted by any upward condition, then the members have average size at least $n/2$. In Boolean terms, the complements of the members are the false points of the DNF $P \vee \bigvee_r (\bar{r} \wedge \bigwedge_{p \in B_r} p)$, where P is the positive DNF of the failure of K. This DNF is Horn and has the dependency property of Lozin and Zamaraev [14, Theorem 4]: the only term in which r occurs negatively contains any chosen $d(r) \in B_r$ positively, and the terms of P contain no negative literal. Their proof shows that the matrix of true points has at least as many ones as zeros, which is exactly (2) after complementation; they then deduce a good variable, whereas we need the aggregate inequality itself. We include the argument for completeness and because the formal development treats the upward predicate abstractly rather than as a DNF.

Proof. Work in deletion coordinates $D = U \setminus S$, and call a row $D \subseteq U$ *bad* if $U \setminus D \notin \mathcal{B}$. Set $P(D) :\Leftrightarrow \neg K(U \setminus D)$ and, for $r \in U$,

$$H_r(D) :\Leftrightarrow \; r \notin D \text{ and } B_r \subseteq D.$$

Then D is bad iff $P(D)$ or $H_r(D)$ for some r: indeed $S = U \setminus D$ fails $S \cap B_r \neq \varnothing$ for some $r \in S$ exactly when $r \notin D$ and $B_r \subseteq D$. Since K is upward, P is upward in D.

A *cell* is a pair $(D, r) \in 2^U \times U$; it is a *zero cell* if $r \notin D$ and a *one cell* if $r \in D$. Fix a choice $d(r) \in B_r$ for each r. Define a map φ on the zero cells of bad rows:

(H) if $P(D)$ holds, or $H_q(D)$ holds for some $q \neq r$, let $\varphi(D, r) = (D \cup \{r\}, r)$;

(V) otherwise let $\varphi(D, r) = (D, d(r))$.

Call the first kind of move horizontal and the second vertical.

Targets are one cells of bad rows. In case (H), $r \in D \cup \{r\}$, and the row $D \cup \{r\}$ is bad because P persists upward, or because $q \notin D$ and $q \neq r$ give $q \notin D \cup \{r\}$ while $B_q \subseteq D \subseteq D \cup \{r\}$. In case (V), D is bad but $P(D)$ fails, so some $H_q(D)$ holds, and by the case hypothesis $q = r$; hence $B_r \subseteq D$ and $d(r) \in D$, so $(D, d(r))$ is a one cell of the bad row D.

φ is injective. Two horizontal sources with a common target $(D \cup \{r\}, r)$ have the same second coordinate r and the same row after removing r, so they coincide. Two vertical sources $(D, r), (D, r')$ with a common target lie in the same row; verticality means that every covering term of D has head r, and also head r', and since some covering term exists, $r = r'$. For a mixed collision let (A, a) be horizontal and (D, b) vertical with $(A \cup \{a\}, a) = (D, d(b))$; so $D = A \cup \{a\}$, $a = d(b)$ and $a \notin A$. If $P(A)$ held then $P(D)$ would hold by upwardness, contradicting verticality of (D, b). Hence the horizontal move was justified by some $H_q(A)$ with $q \neq a$. If $q = b$ then $B_b \subseteq A$, so $a = d(b) \in A$, a contradiction. If $q \neq b$ then $q \notin A$ and $q \neq a$ give $q \notin D$, and $B_q \subseteq A \subseteq D$, so $H_q(D)$ holds with $q \neq b$, contradicting verticality. Thus φ is injective.

Counting. Let $Z_{\text{bad}}, O_{\text{bad}}$ be the numbers of zero and one cells in bad rows and $Z_{\text{good}}, O_{\text{good}}$ those in good rows. The injection gives $Z_{\text{bad}} \leq O_{\text{bad}}$. The whole cube has $n2^{n-1}$ zero cells and $n2^{n-1}$ one cells, so $Z_{\text{good}} \geq O_{\text{good}}$, that is $\sum_{D \text{ good}}(n - |D|) \geq \sum_{D \text{ good}} |D|$. Writing $S = U \setminus D$, the good rows are exactly the members of $\mathcal{B}$ and $n - 2|D| = 2|S| - n$, which is (2). ☐

Remark 3.2. The lemma is trivially true when $\mathcal{B} = \varnothing$, and it is applied below with U nonempty. Only the finiteness of U, the nonemptiness of the bodies and the upwardness of K are used; the bodies may contain their own head (a self-supported element), and K may be false everywhere or true everywhere.

4 The exterior-conditioned abundance theorem

Theorem 4.1 (Elementary-core abundance, fibre form)**.** *Let U be an elementary core of a finite CRS $\mathcal{Q}$, $n = |U|$. For $T \subseteq R \setminus U$ define the* exterior fibre

$$\mathcal{B}_T = \{S \subseteq U : \; S \cup T \in \mathcal{F}(\mathcal{Q})\}.$$

Then for every $T \subseteq R \setminus U$,

$$\sum_{S \in \mathcal{B}_T} (2|S| - n) \;\geq\; 0. \tag{3}$$

Consequently

$$2 \sum_{W \in \mathcal{F}} |W \cap U| \;\geq\; nN, \qquad \textit{and some } r \in U \textit{ satisfies } 2f(r) \geq N. \tag{4}$$

Nothing is assumed about the reactions outside U: they may have arbitrary reactants, products and catalysts, including catalysts for reactions of U.

Proof. Fix $T \subseteq R \setminus U$. Call a core reaction $r \in U$ *T-automatic* if $C(r) \cap (F \cup \pi(T)) \neq \varnothing$, and define bodies

$$B_r = \begin{cases} \{r\} & \text{if } r \text{ is } T\text{-automatic}, \\ \{p \in U : \ \pi(p) \cap C(r) \neq \varnothing\} & \text{otherwise.} \end{cases}$$

Each B_r is nonempty: if r is not T-automatic then $C(r) \cap F = \varnothing$, while U is a RAF, so by Lemma 2.2 r has a catalyst in $\mathrm{cl}_U(F) = F \cup \pi(U)$, hence in $\pi(p)$ for some $p \in U$. (Exterior support can only enlarge the set of automatic heads; it never destroys an internal supplier.)

For $S \subseteq U$ let $K_T(S)$ assert that $S \cup T$ is food-generated and that every $t \in T$ satisfies $C(t) \cap (F \cup \pi(S \cup T)) \neq \varnothing$. Then K_T is upward on 2^U: if $S \subseteq S'$, every reaction of $S \cup T$ keeps its food generation in $S' \cup T$ by monotonicity of closure, every reaction of $S' \setminus S$ is food-ready, and $\pi(S \cup T) \subseteq \pi(S' \cup T)$.

We claim the exact identity

$$\mathcal{B}_T = \{S \subseteq U : \ K_T(S) \text{ and } S \cap B_r \neq \varnothing \text{ for every } r \in S\}. \tag{5}$$

If $S \cup T = \varnothing$ both sides contain $S = \varnothing$ (all conditions are vacuous), so assume $S \cup T \neq \varnothing$. Suppose $S \cup T \in \mathcal{F}$, i.e. it is a RAF. Then $S \cup T$ is food-generated, and by (1) every $t \in T$ has a catalyst in $F \cup \pi(S \cup T)$, so $K_T(S)$ holds. Let $r \in S$. If r is T-automatic then $r \in S \cap B_r$. Otherwise r has a catalyst in $F \cup \pi(S \cup T)$ but none in $F \cup \pi(T)$, so some $p \in S$ has $\pi(p) \cap C(r) \neq \varnothing$, i.e. $p \in S \cap B_r$. Conversely suppose $K_T(S)$ and $S \cap B_r \neq \varnothing$ for all $r \in S$. Then $S \cup T$ is nonempty and food-generated, so $\mathrm{cl}_{S \cup T}(F) = F \cup \pi(S \cup T)$; each $t \in T$ is catalysed from this closure by K_T; a T-automatic $r \in S$ is catalysed from $F \cup \pi(T) \subseteq \mathrm{cl}_{S \cup T}(F)$; and a non-automatic $r \in S$ has some $p \in S$ with $\pi(p) \cap C(r) \neq \varnothing$ and $\pi(p) \subseteq \mathrm{cl}_{S \cup T}(F)$. So $S \cup T$ is a RAF.

Lemma 3.1 applied to the bodies B_r and the upward predicate K_T now gives (3). Note that T itself need not be food-generated or a RAF; its products are usable as catalysts only jointly with the requirement that $S \cup T$ be food-generated, which is part of K_T.

For (4), every $W \in \mathcal{F}$ is uniquely $W = S \cup T$ with $T = W \setminus U$ and $S = W \cap U \in \mathcal{B}_T$, and conversely $S \cup T \in \mathcal{F}$ for every $T \subseteq R \setminus U$ and $S \in \mathcal{B}_T$. Summing (3) over T gives $\sum_{W \in \mathcal{F}} (2|W \cap U| - n) \geq 0$. Finally $\sum_{W \in \mathcal{F}} |W \cap U| = \sum_{r \in U} f(r)$, so $\sum_{r \in U} f(r) \geq nN/2$, and as $U \neq \varnothing$ some $r \in U$ has $f(r) \geq N/2$. $\square$

Remark 4.2 (What the theorem does not assert)*.* The abundant reaction is not identified by the proof, only located in U; different exterior contexts may be balanced by different core reactions (Example 5.2). The proof does not interchange quantifiers: it derives a single global witness from the aggregate incidence inequality, not from fibrewise witnesses. The hypothesis is a *viable* elementary core: a food-ready set that fails to be a RAF on its own does not qualify, even if it becomes viable with exterior help.

Remark 4.3 (Fibre slack)*.* Define the slack of a context by $\sigma(T) = \sum_{S \in \mathcal{B}_T} (2|S| - n) \geq 0$. The proof shows the exact identity

$$2 \sum_{r \in U} f(r) - nN = \sum_{T \subseteq R \setminus U} \sigma(T),$$

so the global surplus over the half bound decomposes exactly into the contexts in which the occupancy inequality is strict. Example 5.2 has $\sigma \equiv 0$.

Remark 4.4 (Comparison with an isolated argument)*.* For the isolated elementary system $\mathcal{Q}|_U$ (exterior deleted) the fibre $\mathcal{B}_\varnothing$ is its own RAF family, all $K_\varnothing(S)$ hold, and (3) is the Lozin–Zamaraev theorem for the Horn function "every selected head has a selected supplier". The point of Theorem 4.1 is that an exterior with feedback changes both the bodies (heads become automatic) and the admissible rows (through K_T), and the counting lemma is indifferent to both changes.

Reaction	Reactants	Products	Catalyst
a	f	x	y
b	f	y	x
g	f	y	f
h	f	x	f

Table 1: The changing-witness system of Example 5.2. Food is $\{f\}$ and the designated core is $\{a, b\}$.

5 Consequences

5.1 Weighting and conditioning

Corollary 5.1 (Exterior weighting)**.** *Under the hypotheses of Theorem 4.1, let* $h : 2^{R\setminus U} \to \mathbb{R}_{\geq 0}$. *Then*

$$\sum_{W\in\mathcal{F}} h(W\setminus U)\,(2|W\cap U| - n) \;\geq\; 0. \tag{6}$$

If $Z = \sum_{W\in\mathcal{F}} h(W\setminus U) > 0$, *some* $r \in U$ *satisfies* $\sum_{W\in\mathcal{F},\ r\in W} h(W\setminus U) \;\geq\; Z/2$.

Proof. Multiply (3) by $h(T) \geq 0$ and sum over T; reindexing by $W = S \cup T$ gives (6). Then $2\sum_{r\in U}\sum_{W\ni r} h(W\setminus U) = 2\sum_W h(W\setminus U)|W\cap U| \geq nZ$, and averaging over the n elements of U gives the witness. □

Indicator weights $h = \mathbf{1}_{\mathcal{T}}$ condition on any collection $\mathcal{T}$ of exterior contexts, including a single nonempty fibre. Choosing $h(T) = q(T)/|\mathcal{B}_T|$ for a probability distribution q on contexts with nonempty fibres gives the statement for "choose a context according to q, then a compatible core subset uniformly". Within a context all compatible rows must receive equal weight; arbitrary distributions on individual RAFs are outside the theorem, and no kinetic meaning is attached to h.

Example 5.2 (The witness can change with the context)**.** Let $F = \{f\}$ and take the four food-ready reactions of Table 1, with designated core $U = \{a, b\}$. The pair $\{a, b\}$ is an elementary core: a produces x, which catalyses b, and b produces y, which catalyses a. The exterior reactions g and h are food-catalysed and supply y and x respectively. Since every reaction is food-ready, membership in $\mathcal{F}$ is decided by catalysis alone, and the four fibres are

T	$\mathcal{B}_T$	$\lvert\mathcal{B}_T\rvert$	rows $\ni a$	rows $\ni b$
$\varnothing$	$\varnothing$, $\{a, b\}$	2	1	1
$\{g\}$	$\varnothing$, $\{a\}$, $\{a, b\}$	3	2	1
$\{h\}$	$\varnothing$, $\{b\}$, $\{a, b\}$	3	1	2
$\{g, h\}$	$\varnothing$, $\{a\}$, $\{b\}$, $\{a, b\}$	4	2	2

Thus $N = 12$, $f(a) = f(b) = 6$ and $f(g) = f(h) = 7$. In context $\{g\}$ the conditional frequencies of a, b are $2/3, 1/3$; in context $\{h\}$ they are $1/3, 2/3$. No core reaction is abundant in both of these contexts, yet both are globally abundant, with equality. Every fibre has slack $\sigma(T) = 0$, so every exterior weighting also attains equality in (6). The example also shows that projection to the isolated core can fail: $\{a, g\} \in \mathcal{F}$ while $\{a\}$ is not a RAF of the isolated core $\{a, b\}$.

5.2 Cycles, feedback vertex sets, and disjoint cores

Definition 5.3. The *food-ready support digraph* $\Gamma(\mathcal{Q})$ has vertex set the food-ready reactions, an arc $p \to r$ whenever $\pi(p) \cap C(r) \neq \varnothing$, and a loop at r whenever $C(r) \cap F \neq \varnothing$.

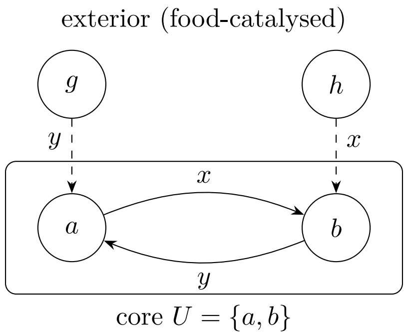

Figure 1: Product–catalysis edges in Example 5.2: an arrow $p \to r$ labelled x means that p produces x and x catalyses r. Selecting g makes a automatic in that context; selecting h makes b automatic.

Lemma 5.4. *The vertex set of any directed cycle of $\Gamma(\mathcal{Q})$ (a loop counts as a cycle of length one) is an elementary core. Conversely every elementary core contains the vertex set of a directed cycle of $\Gamma(\mathcal{Q})$.*

Proof. Let V be the vertex set of a directed cycle. It is nonempty and consists of food-ready reactions, hence food-generated, and $\mathrm{cl}_V(F) = F \cup \pi(V)$. Each $r \in V$ is either food-catalysed (a loop) or has a cycle predecessor $p \in V$ with $\pi(p) \cap C(r) \neq \varnothing$; either way $C(r) \cap \mathrm{cl}_V(F) \neq \varnothing$. Conversely if U is an elementary core then every $r \in U$ has a catalyst in $F \cup \pi(U)$, so r carries a loop or has an in-neighbour in U; following in-neighbours inside the finite set U produces a directed cycle. $\square$

The first half is the elementary-RAF characterisation of [21], where it is also observed that a shortest cycle is a smallest elementary RAF; we return to this in Section 9.

Corollary 5.5 (Cycles meet abundant reactions)**.** *Every directed cycle of $\Gamma(\mathcal{Q})$ contains a globally abundant reaction of $\mathcal{Q}$. Hence the subgraph of $\Gamma(\mathcal{Q})$ induced by the strictly rare food-ready reactions is acyclic, the abundant food-ready reactions form a directed feedback vertex set of $\Gamma(\mathcal{Q})$, and if $\Gamma(\mathcal{Q})$ has k vertex-disjoint directed cycles then $\mathcal{Q}$ has at least k distinct abundant reactions. More generally, pairwise disjoint elementary cores $U_1, \ldots, U_k$ admit distinct abundant reactions $r_i \in U_i$.*

Proof. By Lemma 5.4 and Theorem 4.1 applied in the full system $\mathcal{Q}$, each cycle contains an abundant reaction. A cycle consisting of strictly rare reactions is therefore impossible, which is the acyclicity and feedback-vertex-set statement. For disjoint cores choose a witness in each; disjointness makes the choices distinct. $\square$

In particular the number of abundant reactions is at least the minimum size of a directed feedback vertex set of $\Gamma(\mathcal{Q})$. We do not claim that this minimum, or a maximum cycle packing, is computed by the greedy procedure of Section 9.

5.3 Few non-food-ready reactions

The following set-family lemma is independent of RAF semantics.

Lemma 5.6 (Two-element transversals)**.** *Let $\mathcal{F} \subseteq 2^R$ be union-closed with $\varnothing \in \mathcal{F}$ and at least one nonempty member, $N = |\mathcal{F}|$. If there are $a, b \in R$ (not necessarily distinct) such that every nonempty member of $\mathcal{F}$ contains a or b, then $2f(a) \geq N$ or $2f(b) \geq N$.*

Proof. Let $\mathcal{F}_a, \mathcal{F}_b$ be the members containing a, resp. b; then $\mathcal{F}_a \cup \mathcal{F}_b = \mathcal{F} \setminus \{\varnothing\}$ has $N-1$ elements. If $\mathcal{F}_b = \varnothing$ then $f(a) = N-1 \geq N/2$ since $N \geq 2$; similarly if $\mathcal{F}_a = \varnothing$. Otherwise pick $A \in \mathcal{F}_a$, $B \in \mathcal{F}_b$; by union closure $A \cup B \in \mathcal{F}_a \cap \mathcal{F}_b$, so

$$f(a) + f(b) = |\mathcal{F}_a \cup \mathcal{F}_b| + |\mathcal{F}_a \cap \mathcal{F}_b| \;\geq\; (N-1) + 1 = N,$$

and one of $f(a), f(b)$ is at least $N/2$. □

Corollary 5.7 (At most two non-food-ready reactions)**.** *If a finite CRS has a RAF and at most two of its reactions are not food-ready, then some reaction is abundant.*

Proof. Let H be the set of non-food-ready reactions, $|H| \leq 2$. If some RAF avoids H it is an elementary core and Theorem 4.1 applies. Otherwise every nonempty member of $\mathcal{F}$ meets H; if $H = \{a, b\}$ (or $H = \{a\}$, taking $b = a$) Lemma 5.6 applies. The case $H = \varnothing$ falls under the first branch. □

The overlap term $|\mathcal{F}_a \cap \mathcal{F}_b| \geq 1$ in Lemma 5.6 is exactly what pays for the empty set in the denominator. The same argument does not extend to three exceptional reactions: a three-element transversal of the nonempty members gives only $f(a)+f(b)+f(c) \geq N-1+(\text{overlaps})$, and the average $(N-1)/3$ is well below $N/2$. We show in Section 8 that no argument confined to the non-food-ready reactions can work in general.

5.4 Sharpness, intervention, and a feedback example

Example 5.8 (Sharpness)**.** Let $U = \{r_1, \ldots, r_k\}$ with $r_i :\ f \to c_i$ catalysed only by c_{i-1} (indices modulo k), and no other reactions. Any nonempty RAF containing r_i must contain r_{i-1}, hence all of U; so $\mathcal{F} = \{\varnothing, U\}$, $N = 2$ and $f(r_i) = 1$ for all i. The constant 1/2 in Theorem 4.1 cannot be improved, even for isolated elementary systems. (This is an existence statement about equality, not a classification of equality cases; Example 5.2 attains equality with nontrivial exterior support.)

Remark 5.9 (Intervention stability)*.* Theorem 4.1 is a statement about one system at a time, so it applies to a modified system as soon as the core data survive. Deleting exterior reactions, adding reactions with arbitrary reactants and products, adding catalytic feedback into the core, enlarging the food set or adding catalysts all preserve the hypotheses on U, and each resulting system has an abundant reaction in U. The identity of the witness and the individual frequencies may change, and no monotonicity of any single $f(r)$ is asserted.

Example 5.10 (Genuine feedback into a viable core)**.** Let $F = \{f\}$ and

$$a :\ f \to c_a\ [c_b], \qquad b :\ f \to \{c_b, z\}\ [c_a], \qquad g :\ z \to \{c_b, d\}\ [f],$$

brackets giving the catalysts. The core $\{a, b\}$ is an elementary core; g is not food-ready, consumes the core product z, and feeds the catalyst c_b of a back into the core. The RAF family is $\mathcal{F} = \{\varnothing, \{a, b\}, \{a, b, g\}\}$: g cannot be enabled without b, and a without b has no catalyst other than through g. So $N = 3$, $f(a) = f(b) = 2$, $f(g) = 1$, and a, b are abundant as Theorem 4.1 predicts. This literal example is compiled in the formal development as a consumer of the general theorem (Appendix A).

Example 5.11 (The average over a maximal RAF can fall below one half)**.** Let $F = \{f\}$ and $a :\ f \to z\ [f]$, $g :\ z \to c_g\ [c_l]$, $l :\ f \to c_l\ [c_g]$. Then $\mathcal{F} = \{\varnothing, \{a\}, \{a, g, l\}\}$, $N = 3$, and $(f(a), f(g), f(l)) = (2, 1, 1)$. The maximal RAF $\{a, g, l\}$ has total incidence $4 < \frac{3}{2} \cdot 3$, so the average frequency over a maximal RAF, or over all food-ready reactions (a and l: average 3/2, exactly the bound), is not what the theorem controls. The half-occupancy statement is attached to a viable elementary core (here $\{a\}$), and the aggregate bound (4) concerns incidences in that core only.

6 Diagnostics for an embedded module

Theorem 4.1 needs no relation between the isolated behaviour of a module and the global RAF family. In applications one nevertheless wants to know whether the isolated analysis of a module is faithful. This section gives an exact, polynomial-time test and a quantitative version of the abundance bound that separates faithful from unfaithful rows. Both are formally verified.

Let $E \subseteq R$ be a set of food-ready reactions (a *module*; it need not be a RAF and may contain reactions that are never active). Define the internal predecessors of $r \in E$ by

$$P_E(r) = \{p \in E : \ \pi(p) \cap C(r) \neq \varnothing\}.$$

Say that E is *projection-preserving* if $W \cap E \in \mathcal{F}$ for every $W \in \mathcal{F}$.

Theorem 6.1 (Exact projection criterion)**.** *E is projection-preserving if and only if for every $r \in E$ with $C(r) \cap F = \varnothing$,*

$$r \notin \mathrm{maxRAF}_{\mathcal{Q}}(R \setminus P_E(r)).$$

If the test fails for r, then $W = \mathrm{maxRAF}_{\mathcal{Q}}(R \setminus P_E(r))$ is a RAF containing r whose projection $W \cap E$ is not in $\mathcal{F}$.

Proof. For any $W \subseteq R$ the set $W \cap E$ is food-generated because its reactions are food-ready, so by (1) $W \cap E \in \mathcal{F}$ iff every $r \in W \cap E$ has a catalyst in $F \cup \pi(W \cap E)$, i.e. iff $C(r) \cap F \neq \varnothing$ or $W \cap P_E(r) \neq \varnothing$ for every $r \in W \cap E$.

Suppose the test fails: $r \in E$, $C(r) \cap F = \varnothing$ and $r \in W := \mathrm{maxRAF}_{\mathcal{Q}}(R \setminus P_E(r))$. Then W is a RAF with $W \cap P_E(r) = \varnothing$ and $r \in W \cap E$, so $W \cap E \notin \mathcal{F}$ and E is not projection-preserving; this also proves the witness statement. Conversely suppose every test passes and let $W \in \mathcal{F}$, $r \in W \cap E$ with $C(r) \cap F = \varnothing$. If $W \cap P_E(r) = \varnothing$ then W is a nonempty RAF inside $R \setminus P_E(r)$, so $r \in W \subseteq \mathrm{maxRAF}_{\mathcal{Q}}(R \setminus P_E(r))$, contradicting the test. Hence $W \cap P_E(r) \neq \varnothing$ and $W \cap E \in \mathcal{F}$. ☐

Each test is one maxRAF computation on a restricted availability set, hence polynomial by the RAF algorithm [9]. A sufficient condition, requiring no computation, is that no reaction outside E produces a catalyst of a core reaction that is not food-catalysed. Example 5.10 shows that this sufficient condition can fail while the exact criterion passes: $g \notin E$ produces the catalyst c_b of a, but $\mathrm{maxRAF}(R \setminus P_E(a)) = \mathrm{maxRAF}(\{a, g\}) = \varnothing$ because g needs z from b.

Theorem 6.2 (Defect bound)**.** *Let E be a module of food-ready reactions with $U = \mathrm{maxRAF}_{\mathcal{Q}}(E) \neq \varnothing$, and let $\mathcal{D}_{\mathrm{bad}} = \{W \in \mathcal{F} : \ W \cap E \notin \mathcal{F}\}$. Then*

$$2 \sum_{W \in \mathcal{F}} |W \cap U| \ \geq \ |U|(N - |\mathcal{D}_{\mathrm{bad}}|) + 2 \sum_{W \in \mathcal{D}_{\mathrm{bad}}} |W \cap U|, \tag{7}$$

and consequently some $r \in U$ satisfies $2|U|\, f(r) \geq |U|(N - |\mathcal{D}_{\mathrm{bad}}|) + 2 \sum_{W \in \mathcal{D}_{\mathrm{bad}}} |W \cap U|$.

Proof. Let $\mathcal{G} = \mathcal{F} \setminus \mathcal{D}_{\mathrm{bad}}$ be the good rows. For $W \in \mathcal{G}$, $W \cap E$ is empty or a RAF contained in E, hence $W \cap E \subseteq \mathrm{maxRAF}_{\mathcal{Q}}(E) = U$, so $W \cap E = W \cap U$ and $T := W \setminus U$ is disjoint from E. Fix $T \subseteq R \setminus E$ and consider the good fibre $\mathcal{G}_T = \{S \subseteq U : \ S \cup T \in \mathcal{F}, \ S \in \mathcal{F}\}$; the map $W \mapsto (W \setminus U, W \cap U)$ is a bijection from $\mathcal{G}$ onto $\bigsqcup_{T \cap E = \varnothing} \mathcal{G}_T$.

Since U is an elementary core, define bodies as in the proof of Theorem 4.1 with $T = \varnothing$: $B_r^0 = \{r\}$ if $C(r) \cap F \neq \varnothing$ and $B_r^0 = \{p \in U : \pi(p) \cap C(r) \neq \varnothing\}$ otherwise; these are nonempty. For $S \subseteq U$, $S \in \mathcal{F}$ iff $S \cap B_r^0 \neq \varnothing$ for all $r \in S$. Given $S \in \mathcal{F}$, the condition $S \cup T \in \mathcal{F}$ reduces to $K_T(S)$ of Theorem 4.1, because the support of the reactions of S inside $F \cup \pi(S)$ persists in $F \cup \pi(S \cup T)$. Hence $\mathcal{G}_T = \{S \subseteq U : K_T(S), \ S \cap B_r^0 \neq \varnothing \ \forall r \in S\}$ and Lemma 3.1

gives $\sum_{S\in\mathcal{G}_T}(2|S|-|U|)\geq 0$. Summing over T, $2\sum_{W\in\mathcal{G}}|W\cap U|\geq |U|\,|\mathcal{G}|=|U|(N-|\mathcal{D}_{\mathrm{bad}}|)$. Adding the bad rows gives (7), and $\sum_W |W\cap U|=\sum_{r\in U} f(r)\leq |U|\max_{r\in U} f(r)$ gives the frequency form. □

The bound (7) reproduces (4) when $\mathcal{D}_{\mathrm{bad}}=\varnothing$ and is weaker than it in general (it discards the good half-occupancy in bad rows in exchange for their exact incidence). Its use is descriptive: it tells how much of the core incidence is contributed by rows in which the module's isolated analysis is unfaithful.

Example 6.3 (Projection fails, abundance holds, the defect bound is tight)**.** Let $F=\{f\}$ and $a:\ f\to u\ [f]$, $b:\ f\to z\ [c]$, $g:\ z\to c\ [f]$, with module $E=\{a,b\}$. Isolated, $U=\mathrm{maxRAF}(E)=\{a\}$, because c is not produced inside E. Globally $\mathcal{F}=\{\varnothing,\{a\},\{b,g\},\{a,b,g\}\}$: the pair $\{b,g\}$ is a RAF in which g consumes z from b and returns the catalyst c. Projection fails, $\{b,g\}\cap E=\{b\}\notin\mathcal{F}$, and the deletion query for b (with $P_E(b)=\varnothing$) returns the witness $\mathrm{maxRAF}(R)=\{a,b,g\}$. Here $N=4$, $f(a)=f(b)=f(g)=2$, $\mathcal{D}_{\mathrm{bad}}=\{\{b,g\},\{a,b,g\}\}$, and (7) reads $4\geq 1\cdot 2+2\cdot 1=4$.

Remark 6.4 (Layer identities). For a distinguished reaction g, the rows omitting g and the rows containing g with g erased partition the incidence counts exactly: $N=|\mathcal{F}_{\bar g}|+|\mathcal{F}_g|$, $f(g)=|\mathcal{F}_g|$, and $f(r)=f_{\bar g}(r)+f_g(r)$ for $r\neq g$, where erasure is injective on rows containing g. These identities are useful when testing proposed extension inequalities and are part of the formal development, but the proof of Theorem 4.1 does not need them.

7 Realisation of RAF interior operators on a fixed reaction set

We now ask which interior operators arise as $\mathrm{maxRAF}_{\mathcal{Q}}$. The answer separates the two halves of RAF semantics: food generation, which is an antimatroid, and catalytic support, which is a predecessor condition.

7.1 Blockers

Let $\mathcal{D}\subseteq 2^V$ be union-closed with $\varnothing\in\mathcal{D}$ and interior $i_{\mathcal{D}}$.

Definition 7.1. For $v\in V$, a *blocker* for v is a set $B\subseteq V$ with $v\notin B$ and $v\notin i_{\mathcal{D}}(V\setminus B)$. Write $\mathcal{H}_v$ for the set of blockers of v. The empty set is a blocker for v exactly when v lies in no member of $\mathcal{D}$.

Lemma 7.2 (Blocker test)**.** *For $v\in A\subseteq V$: $v\in i_{\mathcal{D}}(A)$ if and only if $A\cap B\neq\varnothing$ for every $B\in\mathcal{H}_v$.*

Proof. If $A\cap B=\varnothing$ for a blocker B then $A\subseteq V\setminus B$ and monotonicity gives $i_{\mathcal{D}}(A)\subseteq i_{\mathcal{D}}(V\setminus B)\not\ni v$. Conversely if $v\notin i_{\mathcal{D}}(A)$ then $V\setminus A$ is a blocker for v disjoint from A. □

Lemma 7.3 (Minimal blockers suffice)**.** *Every blocker for v contains an inclusion-minimal blocker for v, so the test of Lemma 7.2 may be restricted to minimal blockers. Every blocker is contained in $V\setminus\{v\}$, so $|\mathcal{H}_v|\leq 2^{|V|-1}$ and $\sum_{B\in\mathcal{H}_v}|B|\leq(|V|-1)|\mathcal{H}_v|$.*

Proof. Among the blockers contained in a given blocker choose one of minimum cardinality. The bounds are immediate. □

Minimality does not make the blocker description polynomial. For the threshold family $\mathcal{D}=\{\varnothing\}\cup\{A\subseteq V:|A|\geq k\}$ on n elements, B is a blocker for v iff $v\notin B$ and $|V\setminus B|<k$, so the minimal blockers for v are exactly the $(n-k+1)$-subsets of $V\setminus\{v\}$, of which there are $\binom{n-1}{n-k+1}$.

7.2 Exact-ground characterisation

Call $\mathcal{A} \subseteq 2^V$ an *antimatroid family* if $\varnothing \in \mathcal{A}$, $\mathcal{A}$ is union-closed, and $\mathcal{A}$ is *accessible*: every nonempty $S \in \mathcal{A}$ has an element v with $S \setminus \{v\} \in \mathcal{A}$. We do not require $\bigcup \mathcal{A} = V$, so elements that lie in no member are allowed; up to this convention these are the feasible-set families of antimatroids [6, 13].

Lemma 7.4 (Food generation is accessible)**.** *For any finite CRS, the family of food-generated subsets of R is an antimatroid family.*

Proof. It contains $\varnothing$ and is union-closed by monotonicity of closure. Let S be nonempty and food-generated, and for $r \in S$ let $j(r)$ be the least j with $\rho(r) \subseteq L_j(S)$. Choose $v \in S$ with $j(v)$ maximal. By induction on $j \leq j(v)$ one has $L_j(S \setminus \{v\}) = L_j(S)$: the recursion for L_{j+1} with $j < j(v)$ only uses reactions enabled at $L_j(S)$, and v is not among them. Hence every $r \in S \setminus \{v\}$ has $\rho(r) \subseteq L_{j(r)}(S) = L_{j(r)}(S \setminus \{v\})$, and $S \setminus \{v\}$ is food-generated. □

Theorem 7.5 (Exact-ground realisability)**.** *An interior operator ψ on 2^V, V finite, is the RAF interior operator* $\mathrm{maxRAF}_{\mathcal{Q}}$ *of a finite CRS $\mathcal{Q}$ with reaction set exactly V if and only if there exist an antimatroid family $\mathcal{A} \subseteq 2^V$ and a map $P : V \to 2^V$ such that*

$$\mathrm{Fix}(\psi) = \{S \in \mathcal{A} : \ S \cap P(v) \neq \varnothing \ \textit{for every}\ v \in S\}. \tag{8}$$

Proof. Necessity. Let $\mathcal{A}$ be the food-generated subsets of $R = V$, an antimatroid family by Lemma 7.4, and let $P(v) = \{p \in V : \pi(p) \cap C(v) \neq \varnothing\}$, with v adjoined to $P(v)$ when $C(v) \cap F \neq \varnothing$. By (1), a food-generated S is empty or a RAF iff every $v \in S$ has a catalyst in $F \cup \pi(S)$, iff $S \cap P(v) \neq \varnothing$ for every $v \in S$. This is (8) with $\mathrm{Fix}(\psi) = \mathcal{F}(\mathcal{Q})$.

Sufficiency. Given $\mathcal{A}$ and P, let $\mathcal{H}_v$ be the blockers of v with respect to $\mathcal{A}$. Take molecules: one food molecule f, a private product c_v for each $v \in V$, and a marker $m_{v,B}$ for each $v \in V$ and $B \in \mathcal{H}_v$. Define, for each $v \in V$, the single reaction

$$v : \quad \{f\} \cup \{m_{v,B} : B \in \mathcal{H}_v\} \ \longrightarrow \ \{c_v\} \cup \{m_{u,B} : \ u \in V, \ B \in \mathcal{H}_u, \ v \in B\}, \tag{9}$$

catalysed exactly by the molecules c_p, $p \in P(v)$. Thus the marker $m_{u,B}$ is produced precisely by the reactions indexed by elements of B, and c_p is produced only by p.

We show that the food-generated subsets of this system are exactly $\mathcal{A}$. Let $S \in \mathcal{A}$; we induct on $|S|$, the case $S = \varnothing$ being trivial. Choose $v \in S$ with $S \setminus \{v\} \in \mathcal{A}$; by induction it is food-generated, so $\mathrm{cl}_{S\setminus\{v\}}(F) = F \cup \pi(S \setminus \{v\})$. Every reaction of $S \setminus \{v\}$ has its reactants in $\mathrm{cl}_{S\setminus\{v\}}(F) \subseteq \mathrm{cl}_S(F)$. For v: $f \in F$, and for $B \in \mathcal{H}_v$ we have $v \in S = i_{\mathcal{A}}(S)$, so $S \cap B \neq \varnothing$ by Lemma 7.2; as $v \notin B$ some $u \in S \setminus \{v\}$ lies in B and produces $m_{v,B}$, so $m_{v,B} \in \pi(S \setminus \{v\}) \subseteq \mathrm{cl}_S(F)$. Hence S is food-generated.

Conversely let S be food-generated and nonempty; we induct on $|S|$. By Lemma 7.4 there is $v \in S$ with $S \setminus \{v\}$ food-generated, hence in $\mathcal{A}$ by induction. For $B \in \mathcal{H}_v$ the marker $m_{v,B} \in \rho(v) \subseteq \mathrm{cl}_S(F) = F \cup \pi(S)$ is not food, so it is produced by some $u \in S$ with $u \in B$; thus $S \cap B \neq \varnothing$ for all $B \in \mathcal{H}_v$, and Lemma 7.2 gives $v \in i_{\mathcal{A}}(S)$. Also $S \setminus \{v\} \in \mathcal{A}$ and $S \setminus \{v\} \subseteq S$ give $S \setminus \{v\} \subseteq i_{\mathcal{A}}(S)$. So $i_{\mathcal{A}}(S) = S$ and $S \in \mathcal{A}$. (This direction is what excludes circular marker bootstrapping: a set of reactions that merely produce one another's markers is not food-generated unless some admissible order exists.)

Finally, for a food-generated S the molecule c_p lies in $\mathrm{cl}_S(F) = F \cup \pi(S)$ iff $p \in S$, so S is empty or a RAF iff $S \in \mathcal{A}$ and $S \cap P(v) \neq \varnothing$ for every $v \in S$. The fixed family of $\mathrm{maxRAF}_{\mathcal{Q}}$ is therefore the right-hand side of (8), which equals $\mathrm{Fix}(\psi)$, and interior operators with equal fixed families are equal. □

Corollary 7.6 (Food generation is exactly antimatroidal)**.** *A family $\mathcal{A} \subseteq 2^V$ is the family of food-generated subsets of some finite CRS with reaction set V if and only if $\mathcal{A}$ is an antimatroid family.*

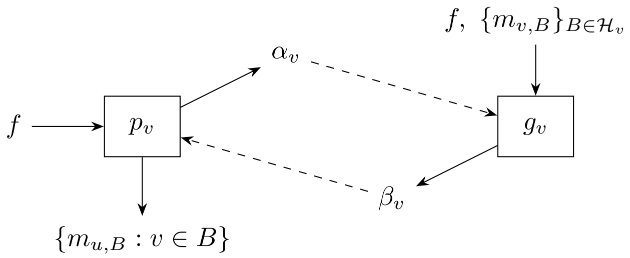

Figure 2: One producer–gate pair of Construction 8.1. Solid arrows are reactant and product relations, dashed arrows catalysis. The producer is food-ready and emits its marker outputs in the first closure round; the gate is enabled exactly when every blocker of v meets the selected coordinates.

Proof. Lemma 7.4 and the first part of the sufficiency proof of Theorem 7.5, which did not use catalysis. □

Theorem 7.5 is a structural factorisation of RAF realisability, not an efficient algorithm for finding a factorisation of a family given explicitly; the marker construction may have exponentially many markers (Lemma 7.3 and the threshold example). It uses $1 + |V| + b$ molecules and $2|V| + b + j + \sum_v |P(v)|$ reactant, product and catalyst incidences, where $b = \sum_v |\mathcal{H}_v|$ and $j = \sum_v \sum_{B\in\mathcal{H}_v} |B|$.

8 Sparse universality and the union-closed boundary

Theorem 7.5 shows that not every interior operator is a RAF interior operator on the same ground set: the fixed family must be an antimatroid family cut down by predecessor support, and a union-closed family with no accessible structure (for instance $\{\varnothing\} \cup \{A : |A| \geq 2\}$) is not of this form. Allowing two reactions per coordinate removes the restriction entirely, and does so with a very sparse catalytic structure.

Construction 8.1 (Producer–gate source)**.** Let $\mathcal{D} \subseteq 2^V$ be union-closed with $\varnothing \in \mathcal{D}$, with blockers $\mathcal{H}_v$ as in Section 7. Take one food molecule f, private molecules α_v, β_v for $v \in V$, and markers $m_{v,B}$ for $B \in \mathcal{H}_v$. For each v take two reactions,

$$p_v : \quad f \longrightarrow \{\alpha_v\} \cup \{m_{u,B} : \ B \in \mathcal{H}_u, \ v \in B\} \qquad \text{catalysed only by } \beta_v, \tag{10}$$

$$g_v : \quad \{f\} \cup \{m_{v,B} : B \in \mathcal{H}_v\} \longrightarrow \{\beta_v\} \qquad \text{catalysed only by } \alpha_v. \tag{11}$$

Call the resulting CRS $\mathcal{Q}_{\mathcal{D}}$. For $A \subseteq V$ put $\mathrm{enc}(A) = \{p_v, g_v : v \in A\}$, and for a set T of reactions of $\mathcal{Q}_{\mathcal{D}}$ let $c(T) = \{v : \ p_v \in T \text{ and } g_v \in T\}$ be its *complete pairs.*

Theorem 8.2 (Sparse two-round universality)**.** *For every finite union-closed $\mathcal{D} \subseteq 2^V$ containing $\varnothing$:*

(i) $\mathcal{F}(\mathcal{Q}_{\mathcal{D}}) = \{\mathrm{enc}(A) : A \in \mathcal{D}\}$, *and* enc *is a bijection from $\mathcal{D}$ onto $\mathcal{F}(\mathcal{Q}_{\mathcal{D}})$ that preserves inclusions and unions; in particular $f(p_v) = f(g_v) = |\{A \in \mathcal{D} : v \in A\}|$ for every v.*

(ii) For every set T of reactions of $\mathcal{Q}_{\mathcal{D}}$, $\mathrm{maxRAF}_{\mathcal{Q}_{\mathcal{D}}}(T) = \mathrm{enc}(i_{\mathcal{D}}(c(T)))$.

(iii) $\mathcal{Q}_{\mathcal{D}}$ has $2|V|$ reactions, one food molecule, exactly one catalyst per reaction, no self-catalysis, and its product–catalysis arcs are exactly the two-cycles $p_v \leftrightarrow g_v$.

(iv) For every set T of reactions, $L_{j+2}(T) = L_2(T)$ for all $j \geq 0$.

Proof. (i) The only source of β_v is g_v and the only source of α_v is p_v; neither is food. In a RAF W, $p_v \in W$ needs β_v in the closure, hence $g_v \in W$, and $g_v \in W$ needs α_v, hence $p_v \in W$. So every RAF is $\mathrm{enc}(A)$ for a unique nonempty A. Now fix $A \subseteq V$ nonempty and consider $\mathrm{enc}(A)$. All producers are food-ready, so $L_1(\mathrm{enc}(A))$ contains f, all α_v ($v \in A$), and exactly the markers $m_{u,B}$ with $B \cap A \neq \varnothing$. Markers have no other source, so g_v ($v \in A$) is ever enabled iff $A \cap B \neq \varnothing$ for all $B \in \mathcal{H}_v$, iff $v \in i_{\mathcal{D}}(A)$ by Lemma 7.2. Hence $\mathrm{enc}(A)$ is food-generated iff $A \subseteq i_{\mathcal{D}}(A)$, iff $A \in \mathcal{D}$. When $A \in \mathcal{D}$, all β_v ($v \in A$) appear in L_2, and every reaction of $\mathrm{enc}(A)$ has its catalyst in the closure, so $\mathrm{enc}(A)$ is a RAF; when $A \notin \mathcal{D}$ it is not even food-generated. Together with $\mathrm{enc}(\varnothing) = \varnothing \in \mathcal{F}$ this proves (i); frequency preservation follows from the bijection, since $p_v, g_v \in \mathrm{enc}(A)$ iff $v \in A$.

(ii) The RAFs contained in T are the sets $\mathrm{enc}(A)$ with $A \in \mathcal{D}$ and $\mathrm{enc}(A) \subseteq T$, i.e. $A \subseteq c(T)$. Their union is enc of the union of the members of $\mathcal{D}$ inside $c(T)$, which is $\mathrm{enc}(i_{\mathcal{D}}(c(T)))$ because enc commutes with unions. A producer whose gate is unavailable (or vice versa) contributes nothing, as it must: an incomplete pair never lies in a RAF.

(iii) is read off from (10)–(11): the only product of g_v is β_v, which catalyses only p_v; among the products of p_v only α_v is a catalyst, and it catalyses only g_v; markers catalyse nothing.

(iv) Let T be any set of reactions. Every producer in T is enabled at stage 0, so $L_1(T)$ contains all markers and all α's that will ever be produced. A gate's reactants are f and markers, so every gate in T that is ever enabled is enabled at $L_1(T)$ and its output β lies in $L_2(T)$. Gate outputs are never reactants. Hence $L_3(T) = L_2(T)$, and by induction $L_{j+2}(T) = L_2(T)$. Nothing here requires T to be paired or a RAF. □

Proposition 8.3 (Elementary RAFs of the universal source)**.** *Let E_{food} be the set of food-ready reactions of $\mathcal{Q}_{\mathcal{D}}$ and $J = \{v \in V : \{v\} \in \mathcal{D}\}$. A gate g_v is food-ready iff $\{v\} \in \mathcal{D}$, and*

$$\mathrm{maxRAF}_{\mathcal{Q}_{\mathcal{D}}}(E_{\mathrm{food}}) = \mathrm{enc}(J). \tag{12}$$

In particular $\mathcal{Q}_{\mathcal{D}}$ has an elementary RAF if and only if $\mathcal{D}$ has a singleton member, and its elementary RAFs are exactly the encodings of the members of $\mathcal{D}$ contained in J.

Proof. Producers are always food-ready; g_v is food-ready iff it has no marker reactants, iff $\mathcal{H}_v = \varnothing$. If $\{v\} \in \mathcal{D}$ then for every $B \not\ni v$ the member $\{v\} \subseteq V \setminus B$ witnesses $v \in i_{\mathcal{D}}(V \setminus B)$, so $\mathcal{H}_v = \varnothing$. If $\mathcal{H}_v = \varnothing$ then Lemma 7.2 at $A = \{v\}$ gives $v \in i_{\mathcal{D}}(\{v\})$, so $\{v\} \in \mathcal{D}$. Thus $c(E_{\mathrm{food}}) = J$. Union closure gives $J \in \mathcal{D}$ (including $J = \varnothing$), so Theorem 8.2(ii) yields (12). An elementary RAF is a RAF inside E_{food}, i.e. $\mathrm{enc}(A)$ with $A \in \mathcal{D}$, $A \neq \varnothing$, $A \subseteq J$. □

8.1 The boundary

If $\{v\} \in \mathcal{D}$ then v is abundant in $\mathcal{D}$: the map $A \mapsto A \cup \{v\}$ injects the members omitting v into the members containing v [18]. On the universal source $\mathcal{Q}_{\mathcal{D}}$, therefore, Theorem 4.1 says exactly this and nothing more: an elementary core exists iff a singleton member exists, and then the core is $\mathrm{enc}(J)$, one of whose coordinates is abundant. The interest of Theorem 4.1 lies in systems that are not of the form $\mathcal{Q}_{\mathcal{D}}$, where elementary cores are common and carry no singleton information about the family.

Conversely, the class of CRSs *without* an elementary core is as hard as the general problem.

Theorem 8.4 (Equivalence with the union-closed conjecture)**.** *The following are equivalent.*

(a) Every finite union-closed family $\mathcal{D}$ with $\varnothing \in \mathcal{D}$ and a nonempty member has an element v with $2|\{A \in \mathcal{D} : v \in A\}| \geq |\mathcal{D}|$ (Frankl's conjecture).

(b) Every finite CRS possessing a RAF has an abundant reaction.

(c) Every finite CRS possessing a RAF, with one food molecule, one catalyst per reaction, no self-catalysis, product–catalysis arcs forming disjoint two-cycles, two-round closure stabilisation, and no elementary RAF, has an abundant reaction.

Proof. (a)⇒(b): $\mathcal{F}(\mathcal{Q})$ is union-closed with $\varnothing \in \mathcal{F}(\mathcal{Q})$ (Lemma 2.4). (b)⇒(c) is trivial. (c)⇒(a): let $\mathcal{D}$ be as in (a). If $\mathcal{D}$ has a singleton member, that element is abundant by the injection above. Otherwise $\mathcal{Q}_{\mathcal{D}}$ has all the properties listed in (c) by Theorem 8.2(iii),(iv) and Proposition 8.3, and it has a RAF because $\mathcal{D}$ has a nonempty member. By (c) some reaction p_v or g_v is abundant in $\mathcal{F}(\mathcal{Q}_{\mathcal{D}})$, and by Theorem 8.2(i) $|\mathcal{F}(\mathcal{Q}_{\mathcal{D}})| = |\mathcal{D}|$ and $f(p_v) = f(g_v) = |\{A \in \mathcal{D} : v \in A\}|$, so v is abundant in $\mathcal{D}$. □

Theorem 8.4 constrains further work. Unique catalysts, absence of self-catalysis, a catalytic graph consisting of private two-cycles, and shallow molecular closure do not, separately or together, remove the difficulty of the unrestricted conjecture: the universal source has all of them and carries the whole difficulty in its reactant requirements. Any additional hypothesis proposed to yield abundance for coreless systems should first be tested against Construction 8.1; if the universal source satisfies it, a theorem under that hypothesis is a theorem about all union-closed families. The two-round stabilisation in Theorem 8.2(iv) also shows that closure depth places no bound on the number of non-food-ready reactions, so Corollary 5.7 does not indicate a sharp threshold at three exceptional reactions.

Remark 8.5 (Size of the encoding). $\mathcal{Q}_{\mathcal{D}}$ has a linear number of reactions but $1 + 2|V| + b$ molecule labels and $6|V| + b + j$ reactant, product and catalyst incidences, where $b = \sum_v |\mathcal{H}_v|$ and $j = \sum_v \sum_{B \in \mathcal{H}_v} |B|$; both can be exponential in $|V|$, and by Lemma 7.3 restricting to minimal blockers does not change this in the worst case. Linear reaction count must not be mistaken for a polynomial-size description. The formal statement (Appendix A) uses the all-blocker construction; the minimal-blocker membership test is formalised separately as an equivalence of predicates.

9 Certificates and algorithms

The results above suggest a small toolkit for a given finite CRS.

Elementary-core certificates. Build $\Gamma(\mathcal{Q})$ (Definition 5.3). A shortest directed cycle can be found by breadth-first search from each vertex in time $O(|V_\Gamma|(|V_\Gamma| + |E_\Gamma|))$. By Lemma 5.4 its vertex set is an elementary RAF, and it is a smallest elementary RAF by reaction count, since every elementary RAF contains a cycle [21]. By Theorem 4.1 it contains a reaction abundant in the whole system. This is a compact certificate: a small set guaranteed to contain an abundant reaction, produced without enumerating any RAF family. The certificate does not identify which member is abundant unless it is a singleton (a food-catalysed or self-catalysed food-ready reaction).

Cycle packing. Repeatedly extract a shortest cycle and delete its vertices. The cycles found are vertex-disjoint, so by Corollary 5.5 their number is a lower bound on the number of abundant reactions. The greedy packing need not be maximum.

Module diagnostics. For a module E of food-ready reactions, compute $U = \mathrm{maxRAF}_{\mathcal{Q}}(E)$ by the RAF algorithm restricted to E; if nonempty, U contains an abundant reaction of the full system (Theorem 4.1, with the isolated maxRAF as core). Run the $|E|$ deletion queries of Theorem 6.1 to decide whether the module's isolated analysis is faithful globally, and, if not, obtain an explicit RAF witnessing the failure. All of this is polynomial in the size of the system.

Benchmarks with known answers. Construction 8.1 turns any explicit union-closed family into a CRS whose RAF family, frequencies and maxRAF operator are known exactly, which is useful for validating RAF software; the marker construction of Theorem 7.5 does the same for exact-ground realisations.

10 Scope and open problems

The results are statements about finite set families. They assert nothing about concentrations, fluxes, persistence under dynamics, thermodynamic feasibility, or the chemical realisability of marker molecules, and a reaction that lies in half of all RAFs is not thereby essential to any of them. The weighted theorem allows arbitrary nonnegative weights on exterior contexts but not arbitrary weights on individual RAFs.

The proved sufficient condition for abundance is the existence of a viable elementary core, supplemented by the one- and two-exception cases, which also cover some coreless systems. Theorem 8.4 shows that no condition satisfied by the universal source can be sufficient short of the full conjecture. Between these lies a range of natural questions.

Question 10.1. Which structural hypotheses on a CRS without an elementary core, not satisfied by Construction 8.1, imply abundance? Candidates should constrain reactant requirements, since the construction shows that catalytic structure alone can be made trivial.

Question 10.2. Classify the systems in which every exterior fibre has zero slack (Remark 4.3), and find conditions forcing strictly positive slack, i.e. a frequency strictly above $N/2$ in the core.

Question 10.3. When can the witness be narrowed inside the core, or shown to be stable across a specified family of exterior contexts? Example 5.2 shows that some hypothesis beyond viability is needed.

Question 10.4. Characterise the union-closed families that admit a producer–gate, or exact-ground, realisation with polynomially many markers.

The strongest positive statement we know of for general RAF families is the general constant of the entropy method [1, 3, 4, 8, 19], which gives a reaction in a 0.38 fraction of all RAFs of any CRS with a RAF. Whether the RAF structure can push that constant to 1/2 is, by Theorem 8.4, exactly the union-closed sets conjecture.

A Formal verification

A.1 Environment and provenance

The formal development is written in Lean 4 (toolchain `leanprover/lean4:v4.30.0`) against Mathlib at commit `c5ea0035` (the `v4.30.0` tag). Each result below was compiled with `-DwarningAsError=true`; the compiler produced no output. A verification receipt records, for each root module, the SHA-256 of the root source, of every local dependency source and of the Lake manifest, together with an elaboration probe of each exported declaration. Every exported theorem depends only on the axioms `Classical.choice`, `Quot.sound` and `propext`; there are no `sorry`s and no `native_decide`. The receipts used here are:

Root module	Exported declarations	Root SHA-256 (prefix)
`RAF/Frankl/StructuralMain.lean`	11	`d23caca1b9b9705c`
`RAF/Frankl/CoreConsequences.lean`	3	`53f5cd106e1e85a5`
`RAF/Frankl/PairCore.lean`	2	`c6e1118ff58e6fc`
`RAFInteriorRealizability/Corollaries.lean`	3	`96128a929c660aa4`

The RAF semantics is defined literally as in Definition 2.1: a structure `CRS M R` with `food`, `inputs` and `outputs`, a relation `Catalysis M R`, the iterated closure `closureAt Q S k`, the predicates `FoodGenerated`, `CatalyzedFromClosure` and `IsRAF`, the finite families `rafFamily` and `fixedFamily` (the latter is $\mathcal{F}$), and `frequency`. A food-ready reaction is a `SeedReaction`. The executable maxRAF used in the projection results is `RAFQueryCompilation.evaluate`, the iterated pruning operator of the RAF algorithm, proved to compute the largest RAF inside its argument.

A.2 Correspondence

Article statement	Lean declaration (namespace `RAF.Frankl` unless stated)
Lemma 3.1	`supported_upward_average` (module `ElementaryConditioning`); the injection is `conditionedCellMap` with `conditionedCellMap_injOn`
Theorem 4.1, identity (5)	`extension_fibre_iff`
Theorem 4.1, fibre inequality (3)	`extensionFibre_average`
Theorem 4.1, global witness	`elementary_core_exists_abundant`; fibre summation is `exists_frequency_of_fibre_average`
Theorem 4.1, empty-inclusive form	`elementary_core_rafFixedFrankl`
Corollary 5.1, inequality (6)	`exterior_weighted_fibre_average` (real weights, fibre-indexed)
Corollary 5.5, cycle vertex sets	`supported_food_ready_abundant` (support predicate `ProductGraphCatalyzed`)
Corollary 5.5, disjoint cores	`disjoint_elementary_cores_distinct_abundant`
Lemma 5.6 (RAF form)	`two_hit_rafFixedFrankl`
Corollary 5.7	`one_nonfood_rafFixedFrankl`, `two_nonfood_rafFixedFrankl`
Example 5.10	`ExtensionExample.abundant_core_reaction`
Theorem 6.1	`projection_iff_deletion_queries`, `failed_query_witness`
Isolated maxRAF as core	`elementary_maxRAF_exists_abundant`
Theorem 6.2	`feedback_defect_bound`, `feedback_defect_frequency`
Layer identities	`layer_card_identity`, `layer_frequency_identity`
Lemma 7.2	`UnionClosedData.mem_interior_iff_hits_blockers`
Lemma 7.3	`UnionClosedData.mem_interior_iff_hits_minimal_blockers`, `blocker_card_bound`, `marker_output_incidence_bound`
Lemma 7.4	`RAF.foodGenerated_accessible`
Theorem 7.5	`RAFInteriorRealizability.rafInteriorOperator_realizable_iff`
Corollary 7.6	`RAFInteriorRealizability.foodGenerationFamilyRealizable_iff_antimatroid`
Theorem 8.2	`PairGadget.sparse_two_round_realization`
Proposition 8.3	`PairGadget.gate_food_ready_iff_singleton`, `PairGadget.evaluate_food_ready_eq_singletons`
Theorem 8.4, (a)⇔(b)	`PairGadget.all_rafFixedFrankl_iff_all_unionClosedFrankl`
Molecule count of $\mathcal{Q}_{\mathcal{D}}$	`PairGadget.declared_molecule_count`

A.3 What is and is not formalised

The formal statements are literal. In particular, `elementary_core_exists_abundant` takes an arbitrary `CRS`, an arbitrary catalysis relation, a finite set `U` with `IsRAF Q C U` and every element a `SeedReaction`, and returns $r \in U$ with $|\mathcal{F}| \leq 2f(r)$ for the full system; no projection or feedback hypothesis appears. The weighted export is the weighted sum of fibre inequalities over an explicit finite set of contexts; the reindexing by $W \mapsto W \setminus U$ and the weighted witness deduction in Corollary 5.1 are ordinary mathematics in this paper. The cycle corollary is connected through the food/product support predicate on a vertex set rather than through a graph-library notion of cycle; Lemma 5.4 is proved here. The formal version of Lemma 5.6 is stated for RAF families. Theorem 8.4(c) is a paper corollary of the formal (a)$\Leftrightarrow$(b) and the formal properties of the universal source. Examples 5.2, 5.8, 5.11 and 6.3, the threshold-family blocker count, the incidence formulas, and the feedback-vertex-set remark are checked by hand (the examples also by exhaustive enumeration in an independent script); Example 5.10 is compiled. The shortest-cycle and packing procedures of Section 9 are ordinary algorithms whose correctness follows from the theorems; the accompanying implementation is validated by tests against exhaustive enumeration on small systems, not by kernel certificates.

Acknowledgements

To be inserted.

References

[1] Ryan Alweiss, Brice Huang, and Mark Sellke. Improved lower bound for Frankl's union-closed sets conjecture, 2022. arXiv:2211.11731.

[2] Henning Bruhn and Oliver Schaudt. The journey of the union-closed sets conjecture. *Graphs and Combinatorics*, 31(6):2043–2074, 2015. doi: 10.1007/s00373-014-1515-0.

[3] Stijn Cambie. Better bounds for the union-closed sets conjecture using the entropy approach, 2022. arXiv:2212.12500.

[4] Zachary Chase and Shachar Lovett. Approximate union closed conjecture, 2022. arXiv:2211.11689.

[5] 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. doi: 10.1007/978-3-030-79876-5_37.

[6] Robert P. Dilworth. Lattices with unique irreducible decompositions. *Annals of Mathematics*, 41(4):771–777, 1940. doi: 10.2307/1968857.

[7] Peter Frankl. Extremal set systems. In Ronald L. Graham, Martin Grötschel, and László Lovász, editors, *Handbook of Combinatorics*, volume 2, pages 1293–1329. Elsevier, Amsterdam, 1995.

[8] Justin Gilmer. A constant lower bound for the union-closed sets conjecture, 2022. arXiv:2211.09055.

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

[10] Wim Hordijk and Mike Steel. Chasing the tail: the emergence of autocatalytic networks. *BioSystems*, 152:1–10, 2017. doi: 10.1016/j.biosystems.2016.12.002.

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

[12] Stuart A. Kauffman. *The Origins of Order: Self-Organization and Selection in Evolution*. Oxford University Press, New York, 1993.

[13] Bernhard Korte, László Lovász, and Rainer Schrader. *Greedoids*, volume 4 of *Algorithms and Combinatorics*. Springer, Berlin, 1991.

[14] Vadim Lozin and Viktor Zamaraev. Union-closed sets and Horn Boolean functions. *Journal of Combinatorial Theory, Series A*, 202:105818, 2024. doi: 10.1016/j.jcta.2023.105818.

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

[16] Bjorn Poonen. Union-closed families. *Journal of Combinatorial Theory, Series A*, 59(2): 253–268, 1992. doi: 10.1016/0097-3165(92)90068-7.

[17] David Reimer. An average set size theorem. *Combinatorics, Probability and Computing*, 12(1):89–93, 2003. doi: 10.1017/S0963548302005539.

[18] Dinesh G. Sarvate and Jean-Claude Renaud. On the union-closed sets conjecture. *Ars Combinatoria*, 27:149–154, 1989.

[19] Will Sawin. An improved lower bound for the union-closed set conjecture, 2022. arXiv:2211.11504.

[20] Mike Steel. Interior operators and their relationship to autocatalytic networks. *Acta Biotheoretica*, 71(4):21, 2023. doi: 10.1007/s10441-023-09472-8. arXiv:2305.16714.

[21] 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. doi: 10.1098/rsif.2018.0808.

[22] The mathlib Community. The Lean mathematical library. In *Proceedings of the 9th ACM SIGPLAN International Conference on Certified Programs and Proofs (CPP 2020)*, pages 367–381. ACM, 2020. doi: 10.1145/3372885.3373824.

[23] Piotr Wójcik. Union-closed families of sets. *Discrete Mathematics*, 199(1–3):173–182, 1999. doi: 10.1016/S0012-365X(98)00292-0.