

Interior operators realizable by autocatalytic networks: an antimatroid characterization

Anonymous manuscript

September 7, 2026

Abstract

A catalytic reaction system with a food set determines an interior operator on the subsets of its reaction set: each set of reactions is sent to the largest reflexively autocatalytic and food-generated (RAF) subset it contains. Steel (*Acta Biotheoretica*, 2023) showed that not every interior operator arises in this way when the reaction set is required to be the given ground set, and asked for a set-theoretic characterization of the set systems that are realizable as the RAFs of a catalytic reaction system. We answer this question for same-ground realizations. An interior operator ψ on a finite set E is the maxRAF operator of a catalytic reaction system with reaction set E if and only if its family of fixed sets is the intersection of an antimatroid on E with the family $\mathcal{C}(D)$ of vertex sets of a directed graph D on E in which every vertex has an in-neighbour. The antimatroid records food generation and the digraph records catalysis. Necessity rests on two observations: the food-generated subsets of any catalytic reaction system form an antimatroid, and on a food-generated set the catalysis condition reduces to a fixed digraph. Sufficiency is proved by an explicit construction with exactly one reaction per element of E , in which “blocker marker” molecules encode the antimatroid and private catalyst molecules encode the digraph. As consequences, finite antimatroids are exactly the food-generated families of finite catalytic reaction systems; the interior operator whose fixed sets are the empty set and all sets of size at least three has a same-ground realization if and only if $|E| \leq 4$, which sharpens Steel’s bound of twelve to the exact threshold; and every interior operator is realizable if two reactions per element are permitted. The main theorem and its corollaries have been formalized and checked in the Lean 4 proof assistant.

Keywords. autocatalytic sets; RAF theory; interior operators; union-closed families; antimatroids; catalytic reaction systems; formal verification.

MSC 2020. 05B35, 05C20, 06A15, 92E20, 68V20.

1 Introduction

Reflexively autocatalytic and food-generated sets (RAFTs) are a finite combinatorial model of collectively self-sustaining reaction networks. The concept goes back to Kauffman’s autocatalytic sets of proteins [13, 14], was formalized by Hordijk and Steel [11], and has since been developed into a structural theory with efficient algorithms [10, 19, 20] and applications ranging from the origin of metabolism [23, 12] to cognition and cultural evolution [8]. A catalytic reaction system (CRS) consists of molecule types, reactions with reactants and products, a food set of freely available molecules, and a catalysis relation. A nonempty set of reactions is a RAF if it can be built up from the food set in some order and every reaction in it is catalysed by a food molecule or by a product of the set. Unions of RAFs are RAFs, so every set of reactions contains a unique largest RAF, and the map sending a set of reactions to that largest RAF (or to the empty set) is an interior operator on the Boolean lattice of reaction subsets: it is contractive, monotone and idempotent. Steel [18] observed that most generic results of RAF

theory are consequences of this interior-operator property alone, and asked to what extent the converse holds.

Steel’s question has a precise form. Since finite interior operators are in bijection with union-closed families containing the empty set (the family of fixed sets), one may ask which such families are the fixed families of maxRAF operators. Steel defined an interior operator ψ on a finite set Y to have a *RAF-realisation* if there is a CRS whose reaction set is in bijection with Y and whose maxRAF operator agrees with ψ under that bijection; auxiliary molecules, food and catalysts are unrestricted. He proved that for $|Y| \geq 12$ the interior operator whose fixed sets are the empty set and all sets of size at least three has no RAF-realisation, and in his concluding remarks he posed two problems: a set-theoretic characterization of the union-closed families realizable by a directed graph, and the “more difficult task” of characterizing the set systems realizable as the RAFs of some CRS [18, Section 4].

This paper solves the second problem for realizations on the given ground set. The answer combines two classical objects. An *antimatroid* on E is a family of subsets of E that contains the empty set, is closed under union, and is *accessible*: every nonempty member has an element whose removal leaves a member. Antimatroids are the combinatorial abstraction of “feasible orderings” [15, 6, 16, 2]. For a directed graph D on E , with loops permitted, let $\mathcal{C}(D)$ denote the family consisting of the empty set together with the nonempty $S \subseteq E$ in which every vertex has an in-neighbour inside S ; these are Steel’s digraph families [18].

Theorem 1.1 (Main theorem). *Let E be a finite set and let ψ be an interior operator on 2^E . Then ψ has a RAF-realisation with reaction set E if and only if there are an antimatroid $\mathcal{A}$ on E and a directed graph D on E such that*

$$\text{Fix}(\psi) = \mathcal{A} \cap \mathcal{C}(D).$$

Moreover, from $(\mathcal{A}, D)$ one can write down an explicit realizing CRS with exactly $|E|$ reactions.

The two factors have transparent meanings. The antimatroid is the family of food-generated reaction sets: it records the order-sensitive way in which reactants become available. The digraph is the “product catalysis” graph: it records the order-insensitive fact that once a set is food-generated, every product of the set is present, so catalysis becomes a static requirement that each selected reaction has a selected predecessor. The content of the theorem is that these two layers are independent and that nothing else is recorded by the fixed family. The necessity direction is short. The sufficiency direction requires constructing, for an arbitrary antimatroid $\mathcal{A}$, a CRS with exactly one reaction per element whose food-generated sets are precisely the members of $\mathcal{A}$. We do this with marker molecules indexed by “blockers”: a set B blocks e if no feasible set containing e avoids B . Reaction e requires the marker $m_{e,B}$ for every blocker B of e , and every reaction in B produces $m_{e,B}$. The marker is therefore available exactly when the selected set hits B , which is exactly the feasibility test for e . A provenance argument shows that no cyclic bootstrapping is possible.

Consequences

The theorem has several corollaries, proved in Section 6.

- *Antimatroids are food-generation families.* A family of subsets of E is the family of food-generated sets of some finite CRS with reaction set E if and only if it is an antimatroid (Theorem 6.1).
- *Digraph families are the elementary case.* A union-closed family containing $\emptyset$ is of the form $\mathcal{C}(D)$ if and only if it is the fixed family of a CRS on E in which every reaction set is food-generated, for instance an elementary CRS in the sense of Steel (Proposition 6.2).

- *Steel’s obstruction, sharpened.* Let $\mathcal{C}_k^+$ be the family consisting of $\emptyset$ and all subsets of E of size at least k . Steel showed that for $k = 3$ and $|E| \geq 12$ this family is not RAF-realizable, and exhibited a realization on four reactions. Theorem 1.1 gives a short proof that $\mathcal{C}_k^+$ is not realizable on the same ground set once $|E| \geq k^2 - 2k + 3$, and for $k = 3$ the exact threshold: $\mathcal{C}_3^+$ has a RAF-realisation on E if and only if $|E| \leq 4$ (Theorem 6.6).
- *Two reactions per element suffice for everything.* Every interior operator on E is realized exactly, up to the doubling $S \mapsto S \times \{p, g\}$, by a CRS with reaction set $E \times \{p, g\}$ (Theorem 6.8). In particular Steel’s question of whether some reaction always lies in at least half of the RAFs of a CRS implies the union-closed sets conjecture up to an additive constant of one.

Formal verification

The main theorem and the corollaries in Sections 6 have been formalized in Lean 4 [3] on top of Mathlib [21]; details are given in Section 7. The formal development uses the closure-based semantics of food generation and catalysis, which agrees with the ordering-based definitions of [11, 18] (Section 2.1). The proofs in this paper are complete and self-contained; the formalization is an additional guarantee, not a substitute. Before the formal proof was attempted, the construction was tested exhaustively on all labelled antimatroids with at most four elements and all digraphs with at most three vertices; those computations are reported in Section 7 for reproducibility and play no role in the proofs.

Organization

Section 2 fixes definitions. Section 3 proves the necessity direction. Section 4 introduces blockers and the marker construction and proves that it realizes any antimatroid as a food-generation family. Section 5 adds catalysis and completes the proof of the main theorem. Section 6 derives the consequences listed above. Section 7 describes the Lean formalization and the computational checks, and Section 8 discusses limitations and open problems.

2 Preliminaries

All sets are finite. For a set E we write 2^E for its power set. For a family $\mathcal{F} \subseteq 2^E$, $\bigcup \mathcal{F}$ denotes the union of its members.

2.1 Catalytic reaction systems and RAFs

Definition 2.1 (CRS). A *catalytic reaction system* (CRS) is a tuple $\mathcal{Q} = (X, R, \rho, \pi, F, C)$ where X is a finite set of molecule types, R is a finite set of reactions, $\rho, \pi: R \rightarrow 2^X$ assign to each reaction its set of *reactants* and its set of *products*, $F \subseteq X$ is the *food set*, and $C \subseteq X \times R$ is the *catalysis relation*; $(x, r) \in C$ is read “ x catalyses r ”. For $S \subseteq R$ we write $\pi(S) = \bigcup_{r \in S} \pi(r)$.

Steel [18] writes reactions as pairs of multisets and assumes that $\rho(r)$ and $\pi(r)$ are nonempty. Multiplicities play no role in any notion considered here, so we ignore them. We do not impose nonemptiness in the definition, which only enlarges the class of systems for which the necessity direction of the main theorem is proved; the CRS constructed in the sufficiency direction has nonempty reactant and product sets for every reaction.

Definition 2.2 (Closure). Let $\mathcal{Q}$ be a CRS and $S \subseteq R$. Define $W_0(S) = F$ and

$$W_{i+1}(S) = W_i(S) \cup \bigcup \{\pi(r) : r \in S, \rho(r) \subseteq W_i(S)\} \quad (i \geq 0).$$

The sequence is increasing and stabilizes after at most $|X|$ steps; its union is the *closure* $\text{cl}_S(F)$. A reaction r is *enabled* at stage i if $\rho(r) \subseteq W_i(S)$.

Definition 2.3 (Food-generated sets and RAFs). Let $\mathcal{Q}$ be a CRS and $S \subseteq R$.

- (i) S is *food-generated* (F -generated) if $\rho(r) \subseteq \text{cl}_S(F)$ for every $r \in S$.
- (ii) S is *reflexively autocatalytic* if for every $r \in S$ there is $x \in \text{cl}_S(F)$ with $(x, r) \in C$.
- (iii) S is a *RAF* if it is nonempty, food-generated and reflexively autocatalytic.

We write $\text{RAF}(\mathcal{Q})$ for the set of RAFs of $\mathcal{Q}$ and $\mathcal{A}_{\mathcal{Q}} = \{S \subseteq R : S \text{ is food-generated}\}$.

Hordijk and Steel [11] define S to be F -generated when its reactions admit a linear ordering $r_1, \dots, r_k$ with $\rho(r_j) \subseteq F \cup \pi(\{r_1, \dots, r_{j-1}\})$ for each j , and define S to be reflexively autocatalytic when each $r \in S$ has a catalyst in $F \cup \pi(S)$. It is standard that the ordering condition is equivalent to (i) [18, Section 3.1], and Lemma 2.4(ii) below shows that for a food-generated set $\text{cl}_S(F) = F \cup \pi(S)$, so the two catalysis conditions agree on food-generated sets. Hence the RAFs of Definition 2.3 are exactly the RAFs of [11, 18].

Lemma 2.4 (Provenance and monotonicity). *Let $\mathcal{Q}$ be a CRS and $S, T \subseteq R$.*

- (i) *Every $x \in \text{cl}_S(F)$ lies in F or in $\pi(r)$ for some $r \in S$.*
- (ii) *If S is food-generated then $\text{cl}_S(F) = F \cup \pi(S)$.*
- (iii) *If $S \subseteq T$ then $W_i(S) \subseteq W_i(T)$ for all i ; hence $\text{cl}_S(F) \subseteq \text{cl}_T(F)$.*
- (iv) *If S and T are food-generated, so is $S \cup T$. If S and T are RAFs, so is $S \cup T$.*

Proof. (i) By induction on i : $W_0(S) = F$, and every element added at stage $i + 1$ is a product of a reaction in S . (ii) The inclusion $\subseteq$ is (i). Conversely, if $r \in S$ then $\rho(r) \subseteq \text{cl}_S(F) = W_i(S)$ for some i , so $\pi(r) \subseteq W_{i+1}(S)$. (iii) By induction on i , using that a reaction enabled at stage i for S is in T and enabled at stage i for T . (iv) For $r \in S$ we have $\rho(r) \subseteq \text{cl}_S(F) \subseteq \text{cl}_{S \cup T}(F)$ by (iii), and similarly for $r \in T$; catalysts are handled in the same way. $\square$

By Lemma 2.4(iv), every $T \subseteq R$ contains a unique inclusion-maximal RAF or no RAF at all. Following Steel we define the *maxRAF operator* $\varphi_{\mathcal{Q}}: 2^R \rightarrow 2^R$ by

$$\varphi_{\mathcal{Q}}(T) = \bigcup \{S \subseteq T : S \in \text{RAF}(\mathcal{Q})\}, \quad (1)$$

so that $\varphi_{\mathcal{Q}}(T)$ is the largest RAF contained in T , or $\emptyset$ if there is none. Since the closure and catalysis conditions for S refer only to reactions in S , the RAFs of the restricted system $\mathcal{Q}|_T$ are exactly the RAFs of $\mathcal{Q}$ contained in T , so (1) agrees with Steel's definition $\varphi_{\mathcal{Q}}(T) = \text{maxRAF}(\mathcal{Q}|_T)$ [18, Eq. (3)].

2.2 Interior operators and fixed families

Definition 2.5. An *interior operator* on 2^E is a map $\psi: 2^E \rightarrow 2^E$ such that for all $S, T \subseteq E$: $\psi(S) \subseteq S$; $S \subseteq T$ implies $\psi(S) \subseteq \psi(T)$; and $\psi(\psi(S)) = \psi(S)$. Its *fixed family* is $\text{Fix}(\psi) = \{S \subseteq E : \psi(S) = S\}$.

The following facts are elementary; see [18, Lemma 1] for proofs.

Lemma 2.6. *Let E be finite.*

- (i) *If ψ is an interior operator on 2^E then $\text{Fix}(\psi)$ contains $\emptyset$ and is closed under unions, and $\psi(S) = \bigcup \{U \in \text{Fix}(\psi) : U \subseteq S\}$ for all S .*
- (ii) *Conversely, if $\mathcal{D} \subseteq 2^E$ contains $\emptyset$ and is union-closed, then $\psi_{\mathcal{D}}(S) = \bigcup \{U \in \mathcal{D} : U \subseteq S\}$ is the unique interior operator with $\text{Fix}(\psi_{\mathcal{D}}) = \mathcal{D}$.*

In particular two interior operators on 2^E are equal if and only if their fixed families are equal.

For a CRS $\mathcal{Q}$ the map $\varphi_{\mathcal{Q}}$ is the interior operator $\psi_{\mathcal{D}}$ for $\mathcal{D} = \{\emptyset\} \cup \text{RAF}(\mathcal{Q})$, and

$$\text{Fix}(\varphi_{\mathcal{Q}}) = \{\emptyset\} \cup \text{RAF}(\mathcal{Q}). \quad (2)$$

2.3 RAF-realisations

Definition 2.7 (Steel [18]). An interior operator ψ on 2^E has a *RAF-realisation* if there are a CRS $\mathcal{Q} = (X, R, \rho, \pi, F, C)$ and a bijection $b: E \rightarrow R$ such that $\psi(S) = b^{-1}(\varphi_{\mathcal{Q}}(b(S)))$ for every $S \subseteq E$. Nothing is assumed about X , F or C beyond finiteness.

Relabelling reactions along b , we may and shall assume that the realizing CRS has reaction set exactly E ; we then say ψ is *same-ground realizable*. By Lemma 2.6 and (2), ψ is same-ground realizable if and only if there is a CRS $\mathcal{Q}$ with reaction set E and $\text{Fix}(\psi) = \{\emptyset\} \cup \text{RAF}(\mathcal{Q})$. The qualification “same ground” is essential: each element of E must correspond to a single reaction, not to a gadget of several reactions (compare Theorem 6.8).

2.4 Antimatroids

Definition 2.8 (Antimatroid). A family $\mathcal{A} \subseteq 2^E$ is an *antimatroid* on E if $\emptyset \in \mathcal{A}$, $\mathcal{A}$ is closed under unions, and $\mathcal{A}$ is *accessible*: for every nonempty $S \in \mathcal{A}$ there is $e \in S$ with $S \setminus \{e\} \in \mathcal{A}$.

This is the standard definition [16, 2], except that we do not require $\bigcup \mathcal{A} = E$: elements of E that belong to no member of $\mathcal{A}$ are allowed. Accessibility is equivalent, given union-closure and $\emptyset \in \mathcal{A}$, to the existence for every $S \in \mathcal{A}$ of a chain $\emptyset = S_0 \subset S_1 \subset \dots \subset S_{|S|} = S$ of members of $\mathcal{A}$ with $|S_i| = i$. Antimatroids are also known as the feasible-set systems of convex geometries [6] (by complementation), as learning spaces in the theory of knowledge spaces [5, 7], and as the closure systems described by Horn rules with singleton premises [24, 22]. For an antimatroid $\mathcal{A}$ we write $\psi_{\mathcal{A}}$ for its interior operator (Lemma 2.6(ii)).

2.5 Digraph families

Let $D = (E, A)$ be a directed graph with vertex set E ; loops (v, v) are allowed. For $v \in E$ let $N_D^-(v) = \{u \in E : (u, v) \in A\}$ be the set of in-neighbours of v (so $v \in N_D^-(v)$ if and only if v has a loop). Following [18, Section 2.1], let

$$\mathcal{C}(D) = \{\emptyset\} \cup \{S \subseteq E : S \neq \emptyset \text{ and } S \cap N_D^-(v) \neq \emptyset \text{ for all } v \in S\}.$$

Thus a nonempty S lies in $\mathcal{C}(D)$ exactly when every vertex of the induced subgraph $D|_S$ has in-degree at least one. Equivalently, writing $P(v) = N_D^-(v)$, $S \in \mathcal{C}(D)$ if and only if every $e \in S$ has a *predecessor* $u \in S \cap P(e)$; we refer to this as S being *P-supported*. The family $\mathcal{C}(D)$ contains $\emptyset$ and is union-closed [18, Proposition 2(i)], so it is the fixed family of an interior operator. It need not be accessible: for a directed 2-cycle on $\{a, b\}$, $\mathcal{C}(D) = \{\emptyset, \{a, b\}\}$.

3 The characterization and the necessity direction

We restate the main theorem in the notation just introduced.

Theorem 3.1. *Let E be a finite set and ψ an interior operator on 2^E . The following are equivalent.*

- (a) ψ has a RAF-realisation (equivalently, ψ is same-ground realizable).
- (b) There exist an antimatroid $\mathcal{A}$ on E and a directed graph D on E with $\text{Fix}(\psi) = \mathcal{A} \cap \mathcal{C}(D)$. When (b) holds, the CRS $\mathcal{Q}_{\mathcal{A}, D}$ of Construction 4.4 and Definition 5.1 has reaction set E and $\varphi_{\mathcal{Q}_{\mathcal{A}, D}} = \psi$.

The proof of (a) $\Rightarrow$ (b) occupies the rest of this section; (b) $\Rightarrow$ (a) is proved in Sections 4 and 5.

3.1 Food-generated sets form an antimatroid

Proposition 3.2. *For every CRS $\mathcal{Q}$, the family $\mathcal{A}_{\mathcal{Q}}$ of food-generated subsets of R is an antimatroid on R .*

Proof. The empty set is food-generated, and $\mathcal{A}_{\mathcal{Q}}$ is union-closed by Lemma 2.4(iv). For accessibility let $S \in \mathcal{A}_{\mathcal{Q}}$ be nonempty. For $r \in S$ let $k(r)$ be the least i with $\rho(r) \subseteq W_i(S)$; this exists because S is food-generated. Choose $r^* \in S$ with $k(r^*)$ maximal and put $S' = S \setminus \{r^*\}$. We claim that $W_j(S') = W_j(S)$ for all $j \leq k(r^*)$. This holds for $j = 0$. If it holds for some $j < k(r^*)$, then r^* is not enabled at stage j for S (by minimality of $k(r^*)$), so the reactions of S enabled at stage j all lie in S' and are enabled at stage j for S' as well, whence $W_{j+1}(S') = W_{j+1}(S)$. Now every $u \in S'$ has $k(u) \leq k(r^*)$ and therefore $\rho(u) \subseteq W_{k(u)}(S) = W_{k(u)}(S') \subseteq \text{cl}_{S'}(F)$. Thus S' is food-generated. $\square$

In words: among the reactions of a food-generated set, one whose reactants become available last can be deleted without disturbing the availability of anything the others need. This is the “last reaction of an admissible ordering” phenomenon in RAF theory, and Proposition 3.2 says that it is precisely the accessibility axiom of antimatroids.

3.2 Catalysis is static on food-generated sets

Definition 3.3 (Product catalysis graph). For a CRS $\mathcal{Q}$ let $D(\mathcal{Q})$ be the directed graph on R with an arc (u, e) whenever some product of u catalyses e , and additionally a loop (e, e) whenever some food molecule catalyses e . Equivalently,

$$N_{D(\mathcal{Q})}^-(e) = \{u \in R : \exists x \in \pi(u), (x, e) \in C\} \cup \{e : \exists x \in F, (x, e) \in C\}.$$

For an elementary CRS this is the digraph used by Steel [18, Section 3.2]; here it is defined for arbitrary CRSs.

Proposition 3.4. *Let $\mathcal{Q}$ be a CRS and let $S \subseteq R$ be food-generated. Then S is reflexively autocatalytic if and only if $S \in \mathcal{C}(D(\mathcal{Q}))$.*

Proof. For $S = \emptyset$ both statements hold. Let $S \neq \emptyset$ and $e \in S$. By Lemma 2.4(ii), $\text{cl}_S(F) = F \cup \pi(S)$. Hence e has a catalyst in $\text{cl}_S(F)$ if and only if some food molecule catalyses e (in which case $e \in S \cap N_{D(\mathcal{Q})}^-(e)$) or some product of some $u \in S$ catalyses e (in which case $u \in S \cap N_{D(\mathcal{Q})}^-(e)$), and conversely each of these alternatives yields a catalyst in $\text{cl}_S(F)$. Thus e is catalysed from $\text{cl}_S(F)$ if and only if $S \cap N_{D(\mathcal{Q})}^-(e) \neq \emptyset$. $\square$

Corollary 3.5 (Necessary factorization). *For every CRS $\mathcal{Q}$,*

$$\text{Fix}(\varphi_{\mathcal{Q}}) = \mathcal{A}_{\mathcal{Q}} \cap \mathcal{C}(D(\mathcal{Q})).$$

Consequently, every interior operator with a RAF-realisation satisfies condition (b) of Theorem 3.1.

Proof. Both sides contain $\emptyset$. A nonempty S lies in $\text{Fix}(\varphi_{\mathcal{Q}})$ if and only if it is a RAF by (2), that is, if and only if it is food-generated and reflexively autocatalytic; by Proposition 3.4 this holds if and only if $S \in \mathcal{A}_{\mathcal{Q}}$ and $S \in \mathcal{C}(D(\mathcal{Q}))$. The second statement follows from Proposition 3.2 after transporting $\mathcal{A}_{\mathcal{Q}}$ and $D(\mathcal{Q})$ along the bijection of Definition 2.7. $\square$

Remark 3.6. If every subset of R is food-generated (for instance if $\mathcal{Q}$ is *elementary*, i.e. $\rho(r) \subseteq F$ for all r), then $\mathcal{A}_{\mathcal{Q}} = 2^R$ and Corollary 3.5 reads $\text{Fix}(\varphi_{\mathcal{Q}}) = \mathcal{C}(D(\mathcal{Q}))$, which is Lemma 2 of [18].

4 Blockers and the marker construction

Throughout this section $\mathcal{D} \subseteq 2^E$ is a union-closed family containing $\emptyset$, with interior operator $\psi_{\mathcal{D}}$. The results are stated in this generality because they are used both for antimatroids (Theorem 4.6) and for arbitrary fixed families (Theorem 6.8).

4.1 Blockers

Definition 4.1 (Blocker). Let $e \in E$. A set $B \subseteq E$ is a *blocker* for e (with respect to $\mathcal{D}$) if $e \notin B$ and $e \notin \psi_{\mathcal{D}}(E \setminus B)$. Write $\text{Bl}_{\mathcal{D}}(e)$ for the set of blockers for e .

Thus B blocks e when no member of $\mathcal{D}$ containing e avoids B . The family $\text{Bl}_{\mathcal{D}}(e)$ is an up-set in $2^{E \setminus \{e\}}$: if $B \subseteq B'$, $e \notin B'$ and B blocks e , then $\psi_{\mathcal{D}}(E \setminus B') \subseteq \psi_{\mathcal{D}}(E \setminus B)$ by monotonicity, so B' blocks e .

Lemma 4.2 (Hitting lemma). Let $S \subseteq E$ and $e \in S$. Then $e \in \psi_{\mathcal{D}}(S)$ if and only if $S \cap B \neq \emptyset$ for every $B \in \text{Bl}_{\mathcal{D}}(e)$. Consequently

$$S \in \mathcal{D} \iff \text{for every } e \in S \text{ and every } B \in \text{Bl}_{\mathcal{D}}(e) : S \cap B \neq \emptyset.$$

Proof. Suppose $e \in \psi_{\mathcal{D}}(S)$ and let B be a blocker for e with $S \cap B = \emptyset$. Then $S \subseteq E \setminus B$, so $e \in \psi_{\mathcal{D}}(S) \subseteq \psi_{\mathcal{D}}(E \setminus B)$, contradicting the definition of a blocker. Conversely, suppose $e \notin \psi_{\mathcal{D}}(S)$ and put $B = E \setminus S$. Then $e \notin B$ because $e \in S$, and $\psi_{\mathcal{D}}(E \setminus B) = \psi_{\mathcal{D}}(S) \not\ni e$, so B is a blocker for e disjoint from S . The second statement follows because $S \in \mathcal{D}$ if and only if $\psi_{\mathcal{D}}(S) = S$ (Lemma 2.6), i.e. if and only if every $e \in S$ lies in $\psi_{\mathcal{D}}(S)$. $\square$

Remark 4.3. Lemma 4.2 describes every union-closed family containing $\emptyset$ by rules of the form “if $e \in S$ then S meets B ”. Under complementation $S \mapsto E \setminus S$ these rules become the implications $B \rightarrow e$ of a closure system, so the lemma is the union-closed counterpart of the implicational description of closure systems [22]. For an antimatroid $\mathcal{A}$, every rooted circuit (C, e) of $\mathcal{A}$ in the sense of [16, 4] yields the blocker $C \setminus \{e\}$ for e , and Lemma 4.2 is an all-blocker form of the rooted-circuit characterization of feasibility; see also [24]. We use all blockers rather than minimal ones. This is redundant but harmless: the proofs below use only the up-closure of $\text{Bl}_{\mathcal{D}}(e)$ and Lemma 4.2, and the number of markers is not an issue for the existence question addressed here (see Remark 4.7). The digraph families $\mathcal{C}(D)$ are exactly the families described by one such rule per element, namely $B = N_D^-(e)$; the difference between $\mathcal{C}(D)$ and a general union-closed family is that a general family may require several incomparable blockers for the same element.

4.2 The marker CRS of an antimatroid

Construction 4.4 (Marker CRS). Let $\mathcal{A}$ be an antimatroid on E and write $\text{Bl}(e) = \text{Bl}_{\mathcal{A}}(e)$. The CRS $\mathcal{Q}_{\mathcal{A}} = (X, E, \rho, \pi, F, \emptyset)$ (with empty catalysis relation for the moment) has reaction set E and molecule set

$$X = \{f\} \cup \{c_e : e \in E\} \cup \{m_{t,B} : t \in E, B \in \text{Bl}(t)\},$$

consisting of a single food molecule f , one *private catalyst* c_e per element, and one *marker* $m_{t,B}$ per pair of a target t and a blocker B of t . The food set is $F = \{f\}$. Reaction $e \in E$ has

$$\rho(e) = \{f\} \cup \{m_{e,B} : B \in \text{Bl}(e)\}, \quad \pi(e) = \{c_e\} \cup \{m_{t,B} : t \in E, B \in \text{Bl}(t), e \in B\}.$$

In words, reaction e consumes food and every marker that certifies that one of its blockers has been hit, produces its private catalyst, and produces the marker $m_{t,B}$ for every blocker B that it belongs to. Note that reaction e never produces a marker $m_{e,B}$ that it requires, because $e \notin B$ for $B \in \text{Bl}(e)$. All reactant and product sets are nonempty, and $|X| = 1 + |E| + \sum_e |\text{Bl}(e)| \leq 1 + |E| + |E| 2^{|E|-1}$.

Lemma 4.5 (Marker provenance). *Let $\mathcal{A}$ be an antimatroid on E , let $\mathcal{Q} = \mathcal{Q}_{\mathcal{A}}$, and let $S \subseteq E$.*

(i) If $m_{t,B} \in \text{cl}_S(F)$ then $B \in \text{Bl}(t)$ and $S \cap B \neq \emptyset$.

(ii) If $c_e \in \text{cl}_S(F)$ then $e \in S$.

(iii) If S is food-generated, $B \in \text{Bl}(t)$ and $S \cap B \neq \emptyset$, then $m_{t,B} \in \text{cl}_S(F)$.

Proof. (i) and (ii): markers and catalysts are not food, so by Lemma 2.4(i) they occur in $\text{cl}_S(F)$ only as products of reactions in S ; a marker $m_{t,B}$ is produced only by reactions $e \in B$ with $B \in \text{Bl}(t)$, and c_e only by e . (iii): pick $e \in S \cap B$; then $m_{t,B} \in \pi(e)$ and $\pi(e) \subseteq \text{cl}_S(F)$ by Lemma 2.4(ii). $\square$

Theorem 4.6 (Antimatroids are food-generation families). *Let $\mathcal{A}$ be an antimatroid on E . For every $S \subseteq E$,*

$$S \text{ is food-generated in } \mathcal{Q}_{\mathcal{A}} \iff S \in \mathcal{A}.$$

That is, $\mathcal{A}_{\mathcal{Q}_{\mathcal{A}}} = \mathcal{A}$.

Proof. Write $\mathcal{Q} = \mathcal{Q}_{\mathcal{A}}$ and $\psi = \psi_{\mathcal{A}}$. Both directions are proved by strong induction on $|S|$; the case $S = \emptyset$ is trivial in both.

$(\Leftarrow)$ Let $S \in \mathcal{A}$ be nonempty. By accessibility choose $e \in S$ with $S' = S \setminus \{e\} \in \mathcal{A}$; by induction S' is food-generated. We show that every reactant of e lies in $\text{cl}_{S'}(F)$. Certainly $f \in F$. Let $B \in \text{Bl}(e)$. Since $S \in \mathcal{A}$ we have $\psi(S) = S \ni e$, so by Lemma 4.2 there is $j \in S \cap B$. As $e \notin B$, $j \neq e$, so $j \in S' \cap B$, and Lemma 4.5(iii) applied to the food-generated set S' gives $m_{e,B} \in \text{cl}_{S'}(F)$. Hence $\rho(e) \subseteq \text{cl}_{S'}(F) \subseteq \text{cl}_S(F)$ by Lemma 2.4(iii). For $r \in S'$ we have $\rho(r) \subseteq \text{cl}_{S'}(F) \subseteq \text{cl}_S(F)$ as well. Therefore S is food-generated.

$(\Rightarrow)$ Let S be nonempty and food-generated. By Proposition 3.2 there is $e \in S$ such that $S' = S \setminus \{e\}$ is food-generated; by induction $S' \in \mathcal{A}$. Let $B \in \text{Bl}(e)$. Since $m_{e,B} \in \rho(e) \subseteq \text{cl}_S(F)$, Lemma 4.5(i) gives $S \cap B \neq \emptyset$. By Lemma 4.2, $e \in \psi(S)$. For $x \in S'$ we have $x \in S' = \psi(S') \subseteq \psi(S)$ by monotonicity. Hence $\psi(S) = S$, i.e. $S \in \mathcal{A}$. $\square$

The provenance lemma is the decisive safeguard in the $(\Rightarrow)$ direction. A reaction cannot create a marker that it itself requires, and more generally a set of mutually blocked reactions cannot manufacture the markers they need unless some selected reaction genuinely lies in each blocker. Note also that the $(\Rightarrow)$ direction does not assume any admissible ordering of S in advance: it uses the accessibility of food generation (Proposition 3.2) to peel off one reaction at a time.

Remark 4.7 (Minimal blockers suffice). Construction 4.4 remains correct if $\text{Bl}(e)$ is replaced throughout by the set of inclusion-minimal blockers of e . The only place where the full family is used is the “ $\Leftarrow$ ” half of Lemma 4.2, and there the blocker $E \setminus S$ contains a minimal blocker that is still disjoint from S ; the “ $\Rightarrow$ ” half persists because a set that hits all minimal blockers hits all blockers. With this modification the number of markers equals the number of rooted circuits of $\mathcal{A}$. We do not use this reduction, and it is not part of the formal development.

5 Adding catalysis and completing the proof

Definition 5.1 (Catalysis from a digraph). Let $\mathcal{A}$ be an antimatroid on E and D a directed graph on E . The CRS $\mathcal{Q}_{\mathcal{A},D}$ is $\mathcal{Q}_{\mathcal{A}}$ equipped with the catalysis relation

$$C_D = \{(c_u, e) : (u, e) \text{ is an arc of } D\}.$$

Food and markers catalyse nothing.

Lemma 5.2. $D(\mathcal{Q}_{\mathcal{A},D}) = D$.

Proof. The products of u are c_u and markers, and markers catalyse nothing, so some product of u catalyses e if and only if $(c_u, e) \in C_D$, i.e. if and only if (u, e) is an arc of D . No food molecule catalyses anything, so no additional loops arise. $\square$

Proof of Theorem 3.1. (a) $\Rightarrow$ (b) is Corollary 3.5. For (b) $\Rightarrow$ (a), let $\text{Fix}(\psi) = \mathcal{A} \cap \mathcal{C}(D)$ and let $\mathcal{Q} = \mathcal{Q}_{\mathcal{A}, D}$, a CRS with reaction set E . By Corollary 3.5, Theorem 4.6 and Lemma 5.2,

$$\text{Fix}(\varphi_{\mathcal{Q}}) = \mathcal{A}_{\mathcal{Q}} \cap \mathcal{C}(D(\mathcal{Q})) = \mathcal{A} \cap \mathcal{C}(D) = \text{Fix}(\psi).$$

By Lemma 2.6, $\varphi_{\mathcal{Q}} = \psi$. $\square$

Remark 5.3 (What the construction uses). The realizing CRS $\mathcal{Q}_{\mathcal{A}, D}$ has exactly $|E|$ reactions, each with the common food reactant f and at least one product, a finite molecule set, ordinary catalysis, and no inhibition, kinetics or thermodynamic assumptions. Each reaction has a single molecule that catalyses other reactions, namely its private product c_e , so catalysis “by a product of u ” and catalysis “by reaction u ” coincide, and the digraph D is recovered verbatim as the product catalysis graph.

Example 5.4. Steel observes that the union-closed family $\mathcal{D} = \{\emptyset, \{a\}, \{a, b, c\}\}$ on $E = \{a, b, c\}$ is not of the form $\mathcal{C}(D')$ for any digraph D' , since $\{a\} \subsetneq \{a, b, c\}$ violates the interpolation property of digraph families [18, Proposition 2(ii)]. It is nevertheless same-ground realizable. Take the chain antimatroid $\mathcal{A} = \{\emptyset, \{a\}, \{a, b\}, \{a, b, c\}\}$ and the digraph D with a loop at a and arcs $a \rightarrow c$, $c \rightarrow b$. Then $\mathcal{C}(D)$ contains $\{a\}$ and $\{a, b, c\}$ but not $\{a, b\}$ (the vertex b has no in-neighbour in $\{a, b\}$), so $\mathcal{A} \cap \mathcal{C}(D) = \mathcal{D}$. The blockers with respect to $\mathcal{A}$ are $\text{Bl}(a) = \emptyset$, $\text{Bl}(b) = \{\{a\}, \{a, c\}\}$ and $\text{Bl}(c) = \{\{a\}, \{b\}, \{a, b\}\}$, and Construction 4.4 gives

$$\begin{aligned} \rho(a) &= \{f\}, & \pi(a) &= \{c_a, m_{b,\{a\}}, m_{b,\{a,c\}}, m_{c,\{a\}}, m_{c,\{a,b\}}\}, \\ \rho(b) &= \{f, m_{b,\{a\}}, m_{b,\{a,c\}}\}, & \pi(b) &= \{c_b, m_{c,\{b\}}, m_{c,\{a,b\}}\}, \\ \rho(c) &= \{f, m_{c,\{a\}}, m_{c,\{b\}}, m_{c,\{a,b\}}\}, & \pi(c) &= \{c_c, m_{b,\{a,c\}}\}, \end{aligned}$$

with c_a catalysing a and c , and c_c catalysing b . The RAFs of this system are $\{a\}$ (food-enabled and self-catalysed) and $\{a, b, c\}$ (reaction a fires first, then b , then c , and every reaction is catalysed). The set $\{a, b\}$ is food-generated but b lacks its catalyst c_c , and $\{a, c\}$ is not food-generated because the marker $m_{c,\{b\}}$ is produced only by b .

6 Consequences

6.1 Antimatroids are exactly the food-generation families

Theorem 6.1. *A family $\mathcal{A} \subseteq 2^E$ is the family of food-generated subsets of some finite CRS with reaction set E if and only if $\mathcal{A}$ is an antimatroid on E .*

Proof. Necessity is Proposition 3.2; sufficiency is Theorem 4.6. $\square$

Food generation is the “temporal” half of RAF semantics, and Theorem 6.1 identifies it exactly with the classical notion of feasibility in antimatroids. We are not aware of an earlier statement of this identification, although the union-closure of food-generated sets and the existence of admissible orderings are standard [11, 19].

6.2 Digraph families are the elementary case

Proposition 6.2. *For a union-closed family $\mathcal{D} \subseteq 2^E$ containing $\emptyset$ the following are equivalent.*

- (a) $\mathcal{D} = \mathcal{C}(D)$ for some directed graph D on E .
- (b) $\mathcal{D} = \text{Fix}(\varphi_Q)$ for some elementary CRS Q with reaction set E .
- (c) $\mathcal{D} = \text{Fix}(\varphi_Q)$ for some CRS Q with reaction set E in which every subset of E is food-generated.

Proof. (a) $\Rightarrow$ (b): take $\mathcal{A} = 2^E$. Then $\psi_{\mathcal{A}}(E \setminus B) = E \setminus B$, so $e \notin B$ and $e \notin \psi_{\mathcal{A}}(E \setminus B)$ cannot both hold, and no element has a blocker. Hence $\rho(e) = \{f\} \subseteq F$ for all e , so $\mathcal{Q}_{2^E, D}$ is elementary, and $\text{Fix}(\varphi_{\mathcal{Q}_{2^E, D}}) = 2^E \cap \mathcal{C}(D) = \mathcal{C}(D)$ by Theorem 3.1. (b) $\Rightarrow$ (c) is immediate, and (c) $\Rightarrow$ (a) is Corollary 3.5 with $\mathcal{A}_Q = 2^E$. $\square$

Thus Steel's first concluding question, the intrinsic characterization of digraph-realizable families, is exactly the question of characterizing the fixed families of elementary CRSs, and Theorem 3.1 reduces the general same-ground question to it modulo an antimatroid factor.

6.3 Steel's obstruction, sharpened

For $k \geq 1$ let $\mathcal{C}_k^+ = \{\emptyset\} \cup \{S \subseteq E : |S| \geq k\}$; this family is union-closed and contains $\emptyset$. Steel proved that for $k \geq 3$ and $|E| \geq k^3 - 3k^2 + 4k$ (so $|E| \geq 12$ when $k = 3$) the interior operator $\psi_{\mathcal{C}_k^+}$ has no RAF-realisation, and gave a four-reaction CRS realizing $\psi_{\mathcal{C}_3^+}$ on four reactions [18, Proposition 5 and the subsequent remarks]. The key ingredient of his proof is the following counting result.

Lemma 6.3 ([18, Proposition 2(iii)]). *Let $k \geq 3$ and let D be a directed graph on a set Y such that $\mathcal{C}(D)$ contains every k -subset of Y and no nonempty subset of size less than k . Then $|Y| \leq 1 + (k - 1)(k - 2)$.*

The next proposition extracts from the main theorem the structure that Steel's argument exploits: the reactions whose reactants are all food.

Proposition 6.4 (Seeds). *Suppose $\text{Fix}(\psi) = \mathcal{A} \cap \mathcal{C}(D)$ with $\mathcal{A}$ an antimatroid and D a digraph on E , and let $E_0 = \{e \in E : \{e\} \in \mathcal{A}\}$ be the set of seeds. Then:*

- (i) *every nonempty member of $\mathcal{A}$, hence every nonempty member of $\text{Fix}(\psi)$, contains a seed;*
- (ii) $2^{E_0} \subseteq \mathcal{A}$, and consequently $\text{Fix}(\psi) \cap 2^{E_0} = \mathcal{C}(D|_{E_0})$.

Proof. (i) Accessibility applied repeatedly to a nonempty $S \in \mathcal{A}$ yields a chain of members of $\mathcal{A}$ ending in a singleton $\{e\} \subseteq S$. (ii) A subset of E_0 is a union of singletons in $\mathcal{A}$. Hence for $T \subseteq E_0$, $T \in \text{Fix}(\psi)$ if and only if $T \in \mathcal{C}(D)$, and $T \in \mathcal{C}(D)$ if and only if $T \in \mathcal{C}(D|_{E_0})$ since in-neighbours inside T are the same in D and in $D|_{E_0}$. $\square$

Corollary 6.5. *Let $k \geq 3$. If $\psi_{\mathcal{C}_k^+}$ has a RAF-realisation on E then $|E| \leq k^2 - 2k + 2$.*

Proof. By Theorem 3.1 write $\mathcal{C}_k^+ = \mathcal{A} \cap \mathcal{C}(D)$ and let E_0 be the seeds. Every k -subset of E is a nonempty member of $\text{Fix}(\psi)$ and so contains a seed by Proposition 6.4(i); hence $|E \setminus E_0| \leq k - 1$. By Proposition 6.4(ii), $\mathcal{C}(D|_{E_0}) = \mathcal{C}_k^+ \cap 2^{E_0}$. If $|E_0| \geq k$ this family contains every k -subset of E_0 and no nonempty smaller set, so Lemma 6.3 gives $|E_0| \leq 1 + (k - 1)(k - 2)$; if $|E_0| < k$ the same bound holds because $1 + (k - 1)(k - 2) \geq k$ for $k \geq 3$. Altogether $|E| \leq (k - 1) + 1 + (k - 1)(k - 2) = k^2 - 2k + 2$. $\square$

For $k = 3$ the bound is $|E| \leq 5$. The remaining case $|E| = 5$ can be excluded by a direct argument, which yields the exact threshold.

Theorem 6.6. *The interior operator $\psi_{\mathcal{C}_3^+}$ on 2^E has a RAF-realisation if and only if $|E| \leq 4$.*

Proof. Realizations for $|E| \leq 4$. If $|E| \leq 2$ then $\mathcal{C}_3^+ = \{\emptyset\} = 2^E \cap \mathcal{C}(D)$ for the digraph D with no arcs. If $|E| = 3$ then $\mathcal{C}_3^+ = \{\emptyset, E\} = 2^E \cap \mathcal{C}(D)$ for a directed 3-cycle D . For $E = \{a, b, c, d\}$ let

$$\mathcal{A} = \{\emptyset, \{a\}, \{b\}, \{a, b\}, \{a, c\}, \{b, c\}\} \cup \{S : |S| \geq 3\}$$

and let D have the arcs $b \rightarrow a, d \rightarrow a, c \rightarrow b, d \rightarrow b, a \rightarrow c, d \rightarrow c, a \rightarrow d, b \rightarrow d$. One checks directly that $\mathcal{A}$ is union-closed and accessible, that every 3-subset of E lies in $\mathcal{C}(D)$ (in $\{a, b, c\}$ the arcs $b \rightarrow a, c \rightarrow b, a \rightarrow c$ form a cycle; in $\{a, b, d\}$ use $b \rightarrow a, d \rightarrow b, a \rightarrow d$; in $\{a, c, d\}$ use $d \rightarrow a, a \rightarrow c, a \rightarrow d$; in $\{b, c, d\}$ use $c \rightarrow b, d \rightarrow c, b \rightarrow d$), and that none of $\{a\}, \{b\}, \{a, b\}, \{a, c\}, \{b, c\}$ lies in $\mathcal{C}(D)$ (a and b have no loops, and b, a, c respectively have no in-neighbour in $\{a, b\}, \{a, c\}, \{b, c\}$). Since $\mathcal{C}(D)$ is union-closed it contains all sets of size at least three, so $\mathcal{A} \cap \mathcal{C}(D) = \mathcal{C}_3^+$, and Theorem 3.1 applies. (Steel's explicit four-reaction CRS [18, Section 3.3] is another realization.)

Non-realizability for $|E| \geq 5$. By Corollary 6.5 it suffices to treat $|E| = 5$. Suppose $\mathcal{C}_3^+ = \mathcal{A} \cap \mathcal{C}(D)$ with $\mathcal{A}$ an antimatroid and D a digraph on E , $|E| = 5$, and let E_0 be the seeds. As in the proof of Corollary 6.5, $|E \setminus E_0| \leq 2$ and $|E_0| \leq 3$, so $|E_0| = 3$; write $E_0 = \{a, b, c\}$ and $E \setminus E_0 = \{d, e\}$.

Step 1: $D|_{E_0}$ is a directed 3-cycle. By Proposition 6.4(ii), $\mathcal{C}(D|_{E_0}) = \mathcal{C}_3^+ \cap 2^{E_0} = \{\emptyset, E_0\}$. Hence no seed has a loop (singletons are excluded), no two seeds form a 2-cycle (pairs are excluded), and every seed has an in-neighbour in E_0 (because $E_0 \in \mathcal{C}(D)$). If some seed y had both other seeds as in-neighbours, each of those two would need an in-neighbour in E_0 other than itself, and any choice creates a 2-cycle. So every seed has exactly one in-neighbour in E_0 .

Step 2: $d \rightarrow y$ and $e \rightarrow y$ for every seed y . Let y be a seed and $\{u, u'\}$ any two of the other four vertices. The 3-set $\{y, u, u'\}$ lies in $\mathcal{C}(D)$ and y has no loop, so u or u' is an in-neighbour of y . Thus at most one of the four other vertices fails to be an in-neighbour of y . By Step 1 one of the two other seeds already fails, so both d and e are in-neighbours of y .

Step 3: partners. For a seed y the set $\{y, d, e\}$ lies in $\mathcal{C}_3^+ \subseteq \mathcal{A}$, so by accessibility one of $\{y, d\}, \{y, e\}, \{d, e\}$ lies in $\mathcal{A}$. If $\{d, e\} \in \mathcal{A}$ then by accessibility $\{d\}$ or $\{e\}$ would be in $\mathcal{A}$, contradicting that d, e are not seeds. Hence each seed y has a *partner* $p(y) \in \{d, e\}$ with $\{y, p(y)\} \in \mathcal{A}$. Since there are three seeds and two candidates, two distinct seeds y, z share a partner x .

Step 4: contradiction. The pairs $\{y, x\}$ and $\{z, x\}$ lie in $\mathcal{A}$ but not in $\mathcal{C}_3^+$, hence not in $\mathcal{C}(D)$. By Step 2, $x \rightarrow y$, so y has an in-neighbour in $\{y, x\}$; therefore x has none: x has no loop and $y \nrightarrow x$. Likewise $z \nrightarrow x$. But $\{x, y, z\} \in \mathcal{C}(D)$ requires x to have an in-neighbour among x, y, z , which is impossible. This contradiction completes the proof. $\square$

Remark 6.7. Steel notes that within elementary CRSs the threshold for $\mathcal{C}_3^+$ is four reactions. In the language of Theorem 3.1 the elementary case is $\mathcal{A} = 2^E$, so all elements are seeds and Lemma 6.3 gives $|E| \leq 3$ directly. The general case is different precisely because non-seed reactions, which need products of other reactions, can exist; Theorem 6.6 shows that this freedom buys exactly one additional reaction for $k = 3$. For $k \geq 4$ we do not know whether the bound $k^2 - 2k + 2$ of Corollary 6.5 is attained.

6.4 Two reactions per element suffice

Theorem 3.1 shows that the same-ground requirement is a genuine restriction. If each element may be represented by two reactions, every interior operator becomes realizable, with the fixed family recovered exactly.

Theorem 6.8 (Paired realization). *Let $\mathcal{D} \subseteq 2^E$ be union-closed with $\emptyset \in \mathcal{D}$, and let $\psi = \psi_{\mathcal{D}}$. There is a CRS $\mathcal{Q}$ with reaction set $E \times \{p, g\}$ such that*

$$\text{RAF}(\mathcal{Q}) = \{S \times \{p, g\} : S \in \mathcal{D}, S \neq \emptyset\}.$$

In particular $S \in \mathcal{D}$ if and only if $S \times \{p, g\} \in \text{Fix}(\varphi_{\mathcal{Q}})$, and every fixed set of $\varphi_{\mathcal{Q}}$ is of this form.

Proof. Write $\text{Bl}(e) = \text{Bl}_{\mathcal{D}}(e)$. The molecules are a food molecule f , molecules α_e, β_e for $e \in E$, and markers $m_{t,B}$ for $t \in E, B \in \text{Bl}(t)$; the food set is $\{f\}$. For each $e \in E$ there are a *producer* (e, p) and a *gate* (e, g) :

$$\begin{aligned} \rho(e, p) &= \{f\}, & \pi(e, p) &= \{\alpha_e\} \cup \{m_{t,B} : B \in \text{Bl}(t), e \in B\}, \\ \rho(e, g) &= \{f\} \cup \{m_{e,B} : B \in \text{Bl}(e)\}, & \pi(e, g) &= \{\beta_e\}. \end{aligned}$$

The catalysis relation consists of the pairs $(\beta_e, (e, p))$ and $(\alpha_e, (e, g))$: each reaction is catalysed only by the product of its partner. Let T be a RAF of this system and $S = \{e : (e, p) \in T\}$. Since β_e is produced only by (e, g) and α_e only by (e, p) , Lemma 2.4(i) shows that $(e, p) \in T$ if and only if $(e, g) \in T$, so $T = S \times \{p, g\}$ with $S \neq \emptyset$. Producers are enabled at stage 0, so $\text{cl}_T(F)$ contains α_e for $e \in S$ and exactly the markers $m_{t,B}$ with $S \cap B \neq \emptyset$ (Lemma 2.4(i) for “only these”, and enabledness of producers for “all these”). Hence T is food-generated if and only if for every $e \in S$ and $B \in \text{Bl}(e)$, $S \cap B \neq \emptyset$, which by Lemma 4.2 holds if and only if $S \in \mathcal{D}$. Conversely, if $S \in \mathcal{D}$ is nonempty then $T = S \times \{p, g\}$ is food-generated by the same computation, and each (e, p) is catalysed by $\beta_e \in \pi(e, g) \subseteq \text{cl}_T(F)$ and each (e, g) by $\alpha_e \in \text{cl}_T(F)$; so T is a RAF. $\square$

Corollary 6.9 (A reaction-fusion dichotomy). *For every interior operator ψ on 2^E , either ψ is realized by a CRS with one reaction per element (exactly when $\text{Fix}(\psi) = \mathcal{A} \cap \mathcal{C}(D)$ for some antimatroid $\mathcal{A}$ and digraph D), or else two reactions per element are necessary and, by Theorem 6.8, sufficient.*

Remark 6.10 (Steel’s half-frequency question). Steel asked whether every CRS with at least one RAF has a reaction lying in at least half of its RAFs [18, Section 4], a RAF analogue of the union-closed sets conjecture [1, 9]. Theorem 6.8 shows that this question is at least as hard as the union-closed conjecture in the following form: if $\mathcal{D}$ is union-closed with $\emptyset \in \mathcal{D}$ and $|\mathcal{D}| \geq 2$, then the paired CRS of $\mathcal{D}$ has $|\mathcal{D}| - 1$ RAFs, and the reaction (i, p) (or (i, g)) lies in exactly $|\{S \in \mathcal{D} : i \in S\}|$ of them. A positive answer to Steel’s question for all CRSs would therefore give, for every such $\mathcal{D}$, an element i with $2|\{S \in \mathcal{D} : i \in S\}| \geq |\mathcal{D}| - 1$.

7 Formal verification and computation

7.1 Lean formalization

The definitions and results of Sections 2–6 have been formalized in Lean 4 (toolchain v4.30.0) against a pinned version of Mathlib (tag v4.30.0, commit c5ea003). The formalization follows the closure semantics of Definitions 2.2 and 2.3 verbatim: a CRS is a structure with `inputs`, `outputs` and `food`, closure is defined by recursion on the stage, and `FoodGenerated`, `ReflexivelyAutocatalytic` and `IsRAF` are the predicates of Definition 2.3. An interior operator is represented by its fixed family, as a structure `UnionClosedData` carrying $\emptyset \in \mathcal{D}$ and union-closure, with the operator recovered by Lemma 2.6(ii). A same-ground realization is a structure `LiteralRealization` consisting of a finite molecule type, a CRS whose reaction type is exactly the ground type E , and a catalysis relation. The public statement, rendered in ASCII, is

```

theorem RAFInteriorRealizability.rafInteriorOperator_realizable_iff
  {E : Type u} [Fintype E] [DecidableEq E] (psi : InteriorOperator E) :
  SameGroundRAFRealizable psi <=>
    exists (A : AntimatroidData E) (P : E -> Finset E),
      forall S : Finset E,
        S in psi.family <=> (S in A.family /\ PredSupported P S)

```

where `AntimatroidData` is Definition 2.8, P is the in-neighbourhood map of a digraph, and `PredSupported P S` is the condition $S \in \mathcal{C}(D)$. The correspondence between the sections of this paper and the Lean modules is given in Table 1. The consequences of Section 6 are formalized in a further module `Corollaries.lean` as the following declarations:

- `foodGenerationFamilyRealizable_iff_antimatroid`: Theorem 6.1;
- `elementary_fixedFamily_iff_predSupported`: Proposition 6.2, (c) $\Rightarrow$ (a);
- `pairGadget_fixedFamily_exact` and `reactionFusionDichotomy`: Theorem 6.8 and the reaction-fusion dichotomy;
- `rafFixedFrankl_implies_unionClosedFrankl` (in `PairGadget.lean`): the implication from Steel’s half-frequency question to the union-closed conjecture.

The bound of Corollary 6.5 and Theorem 6.6 are not formalized.

Lean module	Content
RAF/Core/CRS.lean, RAF/Core/Closure.lean, RAF/Core/RAF.lean, RAF/Frankl/Semantics.lean RAF/Frankl/Antimatroid.lean	CRS, closure, RAFs, fixed family, maxRAF operator (Section 2.1) Lemma 2.4, Proposition 3.2, Proposition 3.4
RAF/Frankl/PairGadget.lean Basic.lean, Realizable.lean	Lemma 2.6, blockers, Lemma 4.2, Theorem 6.8 interior operators, same-ground realizability
PredSupport.lean, Necessity.lean, FoodSupportCertificate.lean	Definition 3.3, Corollary 3.5
MarkerSource.lean, MarkerProvenance.lean	Construction 4.4, Lemma 4.5
AntimatroidRealization.lean	Theorem 4.6
CatalysisRealization.lean, Characterization.lean	Definition 5.1, Lemma 5.2, Theorem 3.1
Corollaries.lean	Theorem 6.1, Proposition 6.2, Theorem 6.8

Table 1: Correspondence between the paper and the Lean development. Modules are in the directory `proofs/RAFInteriorRealizability` unless a path is indicated, in which case they are relative to `proofs/`.

The aggregate module was compiled with warnings promoted to errors, using the compiler option `-DwarningAsError=true`. The build reports exit code 0, no warnings and no use of `sorry`. The axioms on which the main theorem depends are exactly `propext`, `Classical.choice` and `Quot.sound`, the standard axioms of Lean’s classical logic. The verified source of the main module has SHA-256 hash

858de7bf29f4f1bcf0422229ac7942c89d89e801638337e6988be3a450f7996b,

and the hash of the elaborated statement (the declaration interface) is

effd320f43a7ae418eacce6f063c46cf9c9051265c9fc15e2b5c9ccb819d4c83.

The three corollary declarations were verified separately in the same way, with the same axiom set.

7.2 Exhaustive finite checks

Before the formalization, the marker construction was tested exhaustively by a short program. For $n = 0, \dots, 4$ the program enumerated all labelled antimatroids on an n -set (in the sense of Definition 2.8, allowing elements in no member), obtaining 1, 2, 6, 35 and 596 families, built the marker CRS of Construction 4.4, and compared food generation with membership in the antimatroid for every subset; this gave 9 845 comparisons with no discrepancy. For $n \leq 3$ it additionally ranged over all 2^{n^2} digraphs, computed the RAFs of $\mathcal{Q}_{\mathcal{A},D}$ and compared the fixed family with $\mathcal{A} \cap \mathcal{C}(D)$; this gave 143 753 comparisons with no discrepancy. A second program, written after the fact, enumerates for $n \leq 5$ all pairs $(\mathcal{A}, D)$ with $\mathcal{A} \cap \mathcal{C}(D) = \mathcal{C}_3^+$; it finds certificates for $n \leq 4$ (the one displayed in the proof of Theorem 6.6 is among them) and none for $n = 5$, in agreement with Theorem 6.6. These computations informed the proofs and are reported for reproducibility; they are not used as evidence in any proof.

8 Discussion

Theorem 3.1 gives finite RAF fixed families a two-layer semantics: an antimatroid that records emergence from the food set, and a digraph that records catalysis among the selected reactions. The fixed family of a CRS on E is realizable on the same ground set exactly when it admits such a decomposition, and the marker construction shows that the decomposition is complete rather than merely necessary. The construction is explicit and uses one reaction per element; it may use exponentially many auxiliary molecules, although Remark 4.7 reduces their number to the number of rooted circuits of the antimatroid, and we have not investigated how many molecules are necessary.

Several questions remain.

Question 8.1. Is there an intrinsic (digraph-free) set-theoretic characterization of the families $\mathcal{C}(D)$? By Proposition 6.2 this is Steel’s first concluding question and is equivalent to characterizing the fixed families of elementary CRSs. Together with Theorem 3.1 it would yield an intrinsic characterization of same-ground RAF-realizable interior operators.

Question 8.2. What is the complexity of deciding whether an interior operator given succinctly (for example by its minimal nonempty fixed sets, or by a CRS on a different ground set) is same-ground realizable? The decomposition $\text{Fix}(\psi) = \mathcal{A} \cap \mathcal{C}(D)$ is far from unique, and the search for one is not obviously polynomial.

Question 8.3. For $k \geq 4$, what is the largest $|E|$ for which $\psi_{\mathcal{C}_k^+}$ has a RAF-realisation? Corollary 6.5 gives $|E| \leq k^2 - 2k + 2$; the elementary threshold is $1 + (k - 1)(k - 2)$.

Finally, the theorem is a representation theorem and does not address kinetic or thermodynamic feasibility of the constructed systems, nor variants of RAF theory with inhibition, with pseudo-RAFTs, or with the algebraic semantics of [17]. Whether the antimatroid factor survives in those settings is a natural direction for further work.

Availability

The Lean sources, the verification receipts recording the hashes and axioms quoted in Section 7, and the enumeration programs are available in the repository <https://github.com/mmislan/ErDOS>, in the directories `proofs/RAFInteriorRealizability` and `problem_workspaces/RAF_interior_operator_realizability`.

References

- [1] Igor Balla, Béla Bollobás, and Tom Eccles. Union-closed families of sets. *Journal of Combinatorial Theory, Series A*, 120(3):531–544, 2013.
- [2] Anders Björner and Günter M. Ziegler. Introduction to greedoids. In Neil White, editor, *Matroid Applications*, volume 40 of *Encyclopedia of Mathematics and its Applications*, pages 284–357. Cambridge University Press, 1992.
- [3] 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.
- [4] Brenda L. Dietrich. A circuit set characterization of antimatroids. *Journal of Combinatorial Theory, Series B*, 43(3):314–321, 1987.
- [5] Jean-Paul Doignon and Jean-Claude Falmagne. *Knowledge Spaces*. Springer, Berlin, 1999.
- [6] Paul H. Edelman and Robert E. Jamison. The theory of convex geometries. *Geometriae Dedicata*, 19(3):247–270, 1985.
- [7] Jean-Claude Falmagne and Jean-Paul Doignon. *Learning Spaces: Interdisciplinary Applied Mathematics*. Springer, Berlin, 2011.
- [8] Liane Gabora and Mike Steel. Autocatalytic networks in cognition and the origin of culture. *Journal of Theoretical Biology*, 431:87–95, 2017.
- [9] Justin Gilmer. A constant lower bound for the union-closed sets conjecture. arXiv:2211.09055, 2022.
- [10] Wim Hordijk, Joshua I. Smith, and Mike Steel. Algorithms for detecting and analysing autocatalytic sets. *Algorithms for Molecular Biology*, 10:15, 2015.
- [11] Wim Hordijk and Mike Steel. Detecting autocatalytic, self-sustaining sets in chemical reaction systems. *Journal of Theoretical Biology*, 227(4):451–461, 2004.
- [12] Wim Hordijk and Mike Steel. Autocatalytic networks at the basis of life’s origin and organization. *Life*, 8(4):62, 2018.
- [13] Stuart A. Kauffman. Autocatalytic sets of proteins. *Journal of Theoretical Biology*, 119(1):1–24, 1986.
- [14] Stuart A. Kauffman. *The Origins of Order: Self-Organization and Selection in Evolution*. Oxford University Press, New York, 1993.
- [15] Bernhard Korte and László Lovász. Shelling structures, convexity and a happy end. In Béla Bollobás, editor, *Graph Theory and Combinatorics*, pages 219–232. Academic Press, London, 1984.
- [16] Bernhard Korte, László Lovász, and Rainer Schrader. *Greedoids*, volume 4 of *Algorithms and Combinatorics*. Springer, Berlin, 1991.
- [17] Dimitri Loutchko. An algebraic characterization of self-generating chemical reaction networks using semigroup models. *Journal of Mathematical Biology*, 86:76, 2023.
- [18] Mike Steel. Interior operators and their relationship to autocatalytic networks. *Acta Biotheoretica*, 71(4):21, 2023.

- [19] 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.
- [20] Mike Steel, Joana C. Xavier, and Daniel H. Huson. The structure of autocatalytic networks, with application to early biochemistry. *Journal of the Royal Society Interface*, 17(171):20200488, 2020.
- [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] Marcel Wild. The joy of implications, aka pure Horn formulas: Mainly a survey. *Theoretical Computer Science*, 658:264–292, 2017.
- [23] Joana C. Xavier, Wim Hordijk, Stuart Kauffman, Mike Steel, and William F. Martin. Autocatalytic chemical networks at the origin of metabolism. *Proceedings of the Royal Society B*, 287(1922):20192377, 2020.
- [24] Hiyori Yoshikawa, Hiroshi Hirai, and Kazuhisa Makino. A representation of antimatroids by Horn rules and its application to educational systems. *Journal of Mathematical Psychology*, 77:82–93, 2017.