Representatives of Box-Closure Systems on Downsets of Products of Chains

Luca Blanchi

June 2026

Abstract

We introduce box-closure systems on finite downsets of products of chains. The closure is generated by reflexivity, transitivity, and a coordinate-mixture rule: if two elements lie below the same upper endpoint, then every coordinatewise mixture of them also lies below that endpoint. For a finite downset D, we completely characterize all representative sets whose box-closure is the product order on D. The classification decomposes by upper endpoint: axis elements have forced incoming cover generators, while every element with at least two active coordinates requires at least one incoming generator from its strict lower ideal. As consequences, inclusion-minimal representatives coincide with cardinality-minimal representatives, all minimum representatives have size $|D| - 1$, and after removing the fixed forced axis generators the variable parts of the minimum representatives are precisely the bases of an explicit partition matroid. We derive exact formulae for the number of minimum representatives, Hasse-supported minimum representatives, and all representatives by size, and we solve the weighted representative problem by independent local choices. We also obtain optimal compression ratios for product orders under box-closure, a factorized model of random minimum representatives, and a product-of-simplices description of the minimum-representative polytope. A two-dimensional restricted LCA-like system on designated Ferrers cross-pairs is recovered as a motivating special case; in that case we spell out the projected LCA-closure, the rooted-DAG realization, exact cover-contained counts, jump-length enumerators, and sampling formulae. The box-closure studied here is not the full LCA (+)-closure, nor in general the projection of that closure.

1 Introduction

Closure systems generated by local inference rules arise naturally in combinatorics, order theory, database theory, and phylogenetics. This paper studies a simple but structured closure system on finite downsets of products of chains. The basic rule is a coordinate-mixture rule: if two lower elements lie below a common upper element, then all coordinatewise mixtures of the two lower elements also lie below the same upper element.

The construction is motivated by restricted cross-consistency rules that appear in LCA-like constraint systems, but the object studied here is not the full LCA (+)-closure. This distinction is essential. In the full LCA setting one works with all unordered pairs of leaves, possible non-uniqueness or non-existence of LCAs in directed acyclic graphs, same-side pairs, non-designated cross-pairs, and additional interactions. The projection of a full LCA (+)-closure to a designated subsystem need not coincide with the box-closure of that projected subsystem. In this paper, box-closure is treated as an autonomous combinatorial closure operator.

Let

$$D \subseteq [n_1] \times \cdots \times [n_d]$$

be a finite downset. For each $x \in D$, introduce a symbol p_x. The target relation is the product order

$$R_D = \{p_x \preceq p_y : x \leq y\}.$$

A representative is a subset $T \subseteq R_D$ whose box-closure is all of R_D. The central question is: which subsets T represent R_D, and how small can such a representative be?

We solve this problem completely. If $y \in D$ lies on a coordinate axis, then the immediate predecessor relation into y is forced. If y has at least two active coordinates, then any one incoming relation

$$p_x \preceq p_y, \qquad x < y,$$

is sufficient. Thus the representative problem decomposes over upper endpoints.

The main consequences are:

- $$\min\{|T| : T^{\square} = R_D\} = |D| - 1;$$

- every inclusion-minimal representative is cardinality-minimal;

- after removing the fixed forced axis generators, the minimum representatives are bases of a partition matroid;

- the number of representatives, the number of minimum representatives, and the size enumerator of all representatives have closed product formulae;

- the weighted representative problem decomposes independently over upper endpoints;

- for full boxes, the product order has exponentially larger strict relation set than its minimum box-closure representative.

- in dimension two, projected LCA-closures on designated Ferrers cross-pairs give an explicit LCA-motivated realization with exact jump-length enumerators and cover-contained counts.

The two-dimensional case $D \subseteq [m] \times [n]$ recovers a restricted Ferrers LCA-like system on designated cross-pairs $a_i b_j$. We discuss that connection near the end as motivation, not as the main theorem.

2 Downsets in products of chains

Let

$$[n] = \{1, \ldots, n\}$$

with its usual order. Fix positive integers $n_1, \ldots, n_d$, and write

$$\mathbf{n} = [n_1] \times \cdots \times [n_d].$$

We order $\mathbf{n}$ coordinatewise:

$$x \leq y \quad \Longleftrightarrow \quad x_t \leq y_t \text{ for every } t = 1, \ldots, d.$$

A subset

$$D \subseteq \mathbf{n}$$

is a *downset* if

$$y \in D, \quad x \le y \quad \Longrightarrow \quad x \in D.$$

Throughout, D is a nonempty finite downset. Hence D contains the bottom element

$$\hat{0} = (1, \dots, 1).$$

For $y = (y_1, \dots, y_d) \in D$, define its active support by

$$\mathrm{supp}(y) = \{t : y_t > 1\}.$$

Then

$$y = \hat{0} \quad \Longleftrightarrow \quad \mathrm{supp}(y) = \varnothing.$$

A non-minimal element $y \ne \hat{0}$ is called an *axis element* if

$$|\,\mathrm{supp}(y)| = 1.$$

If y is an axis element, then it has a unique lower cover, denoted

$$y^-.$$

Explicitly, if t is the unique active coordinate, then

$$y^- = (1, \dots, 1, y_t - 1, 1, \dots, 1).$$

An element $y \in D$ is called *multi-coordinate* if

$$|\,\mathrm{supp}(y)| \ge 2.$$

For $y \in D$, write

$$D_{<y} = \{x \in D : x < y\},$$

and

$$D_{\le y} = \{x \in D : x \le y\}.$$

3 Box-closure systems

For each $x \in D$, introduce a formal symbol p_x. Let

$$P_D = \{p_x : x \in D\}.$$

The target relation is the product order on these symbols:

$$R_D = \{p_x \preceq p_y : x \le y\}.$$

Its strict part is

$$R_D^< = \{p_x \preceq p_y : x < y\}.$$

A generating set will always mean a subset

$$T \subseteq R_D^<.$$

Reflexive relations are built into the closure and are not counted as generators.

Definition 3.1 (Box-closure)**.** *For $T \subseteq R_D^{<}$, the* box-closure $T^{\square}$ *is the smallest relation on P_D containing T and satisfying the following rules.*
(B0) Reflexivity. *For every $x \in D$,*

$$p_x \preceq p_x.$$

(B1) Transitivity. *If*

$$p_x \preceq p_y \quad \textit{and} \quad p_y \preceq p_z,$$

then

$$p_x \preceq p_z.$$

(B2) Box rule. *If*

$$p_x \preceq p_z \quad \textit{and} \quad p_y \preceq p_z,$$

then for every coordinatewise mixture w of x and y, meaning

$$w_t \in \{x_t, y_t\} \qquad (t = 1, \dots, d),$$

with $w \in D$, we add

$$p_w \preceq p_z.$$

The box rule preserves the upper endpoint z.

Definition 3.2 (Representatives)**.** *A subset $T \subseteq R_D^{<}$ is a* representative *of R_D if*

$$T^{\square} = R_D.$$

The main problem is to characterize all representatives.

Lemma 3.3 (The product order is box-closed)**.** *The relation R_D is closed under the box-closure rules.*

Proof. Reflexivity and transitivity are immediate.
Suppose

$$p_x \preceq p_z \quad \text{and} \quad p_y \preceq p_z$$

belong to R_D. Then

$$x \leq z, \qquad y \leq z.$$

Let w be a coordinatewise mixture of x and y. For every coordinate t,

$$w_t \in \{x_t, y_t\},$$

so

$$w_t \leq z_t.$$

Thus

$$w \leq z.$$

If $w \in D$, then

$$p_w \preceq p_z \in R_D.$$

Hence R_D is box-closed. □

4 Endpoint preservation

The first structural observation is that non-reflexive relations cannot be generated without a generator having the same upper endpoint.

Lemma 4.1 (Endpoint preservation)**.** *Let $T \subseteq R_D^{<}$. Suppose*

$$p_x \preceq p_y \in T^{\Box}$$

with $x < y$. Then T contains at least one generator of the form

$$p_u \preceq p_y$$

with $u < y$.

Proof. Let
$$U(T) = \{y \in D : \exists u < y \text{ such that } p_u \preceq p_y \in T\}.$$
We prove by induction over derivations in $T^{\Box}$ that every non-reflexive relation $p_x \preceq p_y$ has $y \in U(T)$.

If the relation is a generator in T, the claim is immediate.

Reflexivity produces only relations $p_y \preceq p_y$, so it produces no non-reflexive relation.

For transitivity, suppose
$$p_x \preceq p_z$$
is obtained from
$$p_x \preceq p_y \quad \text{and} \quad p_y \preceq p_z.$$
If $y < z$, then by induction applied to $p_y \preceq p_z$, we have $z \in U(T)$. If $y = z$, then the conclusion is the first premise, and the claim follows from induction applied to that premise.

For the box rule, the conclusion has the same upper endpoint as the two premises. If the conclusion is non-reflexive, then at least one premise is non-reflexive with the same upper endpoint, unless the conclusion coincides with a premise. In either case, induction gives that the common upper endpoint belongs to $U(T)$.

Thus every non-reflexive relation in $T^{\Box}$ has an upper endpoint that already appears as the upper endpoint of a non-reflexive generator in T. $\Box$

Corollary 4.2. *If $T^{\Box} = R_D$, then for every $y \neq \hat{0}$, the set T contains at least one generator with upper endpoint y.*

Proof. For every $y \neq \hat{0}$, choose a lower cover $x \lessdot y$. Then

$$p_x \preceq p_y \in R_D = T^{\Box}$$

is non-reflexive. Apply Lemma 4.1. $\Box$

Lemma 4.3 (Axis covers are forced)**.** *Let y be an axis element. If*

$$p_{y^-} \preceq p_y \in T^{\Box},$$

then

$$p_{y^-} \preceq p_y \in T.$$

Proof. Since y is an axis element, the principal ideal $D_{\le y}$ is a chain:

$$\hat{0} < \cdots < y^- < y.$$

In a chain, the box rule cannot create a new lower endpoint: a coordinatewise mixture of two elements of the chain is one of the two elements.

Consider the fiber of relations with upper endpoint y. By Lemma 4.1, any non-reflexive relation ending at y ultimately depends on some generator ending at y. If that generator is

$$p_u \preceq p_y$$

with $u < y^-$, then transitivity can only produce relations

$$p_z \preceq p_y$$

with $z \le u$, not the immediate cover relation $p_{y^-} \preceq p_y$. The box rule also cannot create y^- as a new lower endpoint inside this chain.

Therefore the relation

$$p_{y^-} \preceq p_y$$

can belong to $T^\square$ only if it is itself a generator in T. $\square$

5 Complete representative theorem

We now prove the main classification.

Theorem 5.1 (Complete representative classification)**.** *Let $T \subseteq R_D^<$. Then*

$$T^\square = R_D$$

if and only if the following two conditions hold:

1. *for every axis element $y \ne \hat{0}$,*
$$p_{y^-} \preceq p_y \in T;$$
2. *for every multi-coordinate element y, there exists at least one $x \in D_{<y}$ such that*
$$p_x \preceq p_y \in T.$$

Thus axis elements have forced incoming cover generators, while each multi-coordinate element requires at least one incoming generator from its strict lower ideal.

Proof. We prove necessity and sufficiency.
Necessity. Assume

$$T^\square = R_D.$$

By Corollary 4.2, every $y \ne \hat{0}$ occurs as the upper endpoint of at least one generator

$$p_x \preceq p_y \in T$$

with $x < y$. This gives condition (2) for all multi-coordinate y.

Now let y be an axis element. Since

$$p_{y^-} \preceq p_y \in R_D = T^{\square},$$

Lemma 4.3 implies

$$p_{y^-} \preceq p_y \in T.$$

Thus condition (1) is necessary.
Sufficiency. Assume conditions (1) and (2).

For every $y \neq \hat{0}$, choose a parent $\pi(y) \in D_{<y}$ as follows:

- if y is an axis element, set
$$\pi(y) = y^-;$$
- if y is multi-coordinate, choose any $x < y$ such that
$$p_x \preceq p_y \in T.$$

Define

$$T_\pi = \{p_{\pi(y)} \preceq p_y : y \in D,\ y \neq \hat{0}\}.$$

Then

$$T_\pi \subseteq T.$$

It suffices to prove

$$T_\pi^{\square} = R_D.$$

We prove by induction on

$$\rho(y) = y_1 + \cdots + y_d$$

that for every $y \in D$,

$$p_z \preceq p_y \in T_\pi^{\square} \qquad \text{for all } z \leq y.$$

If $y = \hat{0}$, this is reflexivity.

First suppose y is an axis element. Then

$$p_{y^-} \preceq p_y \in T_\pi.$$

By induction, every $z \leq y^-$ satisfies

$$p_z \preceq p_{y^-} \in T_\pi^{\square}.$$

By transitivity,

$$p_z \preceq p_y.$$

Together with reflexivity $p_y \preceq p_y$, this gives every $z \leq y$.

Now suppose y is multi-coordinate. Let

$$x = \pi(y).$$

The generator

$$p_x \preceq p_y$$

belongs to T_π. Since $x < y$, the induction hypothesis gives

$$p_{\hat{0}} \preceq p_x.$$

By transitivity,

$$p_{\hat{0}} \preceq p_y.$$

Apply the box rule to

$$p_{\hat{0}} \preceq p_y \quad \text{and} \quad p_y \preceq p_y.$$

We obtain

$$p_c \preceq p_y$$

for every corner c of the box between $\hat{0}$ and y, i.e. every $c \in D$ satisfying

$$c_t \in \{1, y_t\} \qquad (t = 1, \dots, d).$$

In particular, for every active coordinate $t \in \operatorname{supp}(y)$, let

$$e_t(y_t) = (1, \dots, 1, y_t, 1, \dots, 1).$$

Then

$$p_{e_t(y_t)} \preceq p_y.$$

Let $z \leq y$ be arbitrary. For each coordinate t, define

$$e_t(z_t) = (1, \dots, 1, z_t, 1, \dots, 1).$$

If $z_t = 1$, then $e_t(z_t) = \hat{0}$, and we already know

$$p_{\hat{0}} \preceq p_y.$$

If $z_t > 1$, then

$$e_t(z_t) \leq e_t(y_t).$$

By the axis case and transitivity,

$$p_{e_t(z_t)} \preceq p_y.$$

Finally,

$$z = \bigvee_{t \in \operatorname{supp}(z)} e_t(z_t).$$

Repeated applications of the box rule to the relations

$$p_{e_t(z_t)} \preceq p_y$$

produce

$$p_z \preceq p_y.$$

Thus the entire principal ideal below y is generated.

The induction is complete. Hence

$$T_\pi^\square = R_D.$$

Since

$$T_\pi \subseteq T,$$

we obtain

$$T^\square = R_D.$$

$\square$

6 Minimum and inclusion-minimal representatives

The complete representative theorem immediately identifies all minimal representatives.

Theorem 6.1 (Inclusion-minimal equals cardinality-minimal)**.** *Let $T \subseteq R_D^{<}$ be a representative. Then T is inclusion-minimal if and only if:*

1. *for every axis element $y \neq \hat{0}$, T contains the forced generator*

$$p_{y^-} \preceq p_y;$$

2. *for every multi-coordinate element y, T contains exactly one generator*

$$p_x \preceq p_y$$

with $x < y$.

Consequently, every inclusion-minimal representative has cardinality

$$|D| - 1.$$

In particular,

$$\min\{|T| : T^{\square} = R_D\} = |D| - 1.$$

Proof. By Theorem 5.1, every representative must contain all forced axis generators and at least one incoming generator for every multi-coordinate element.

If a representative contains two or more incoming generators with the same multi-coordinate upper endpoint y, deleting all but one still leaves a representative by Theorem 5.1. Hence an inclusion-minimal representative contains exactly one incoming generator for each multi-coordinate y.

Conversely, any set satisfying the two stated conditions is a representative by Theorem 5.1. Removing any generator violates one of the necessary conditions, so the representative is inclusion-minimal.

Such a representative contains exactly one generator for each non-minimal element of D. Therefore its cardinality is

$$|D| - 1.$$

□

7 Partition matroid structure

Let

$$F_{\mathrm{ax}} = \{p_{y^-} \preceq p_y : y \neq \hat{0},\ |\operatorname{supp}(y)| = 1\}$$

be the set of forced axis generators.

For every multi-coordinate element y, define

$$E_y = \{p_x \preceq p_y : x < y\}.$$

Then Theorem 6.1 says that every inclusion-minimal representative is exactly

$$F_{\mathrm{ax}} \cup \{e_y : y \text{ multi-coordinate},\ e_y \in E_y\}.$$

Theorem 7.1 (Partition matroid theorem)**.** *After removing the fixed forced set F_{ax}, the variable parts of the inclusion-minimal representatives are precisely the bases of the partition matroid*

$$\mathcal{M}_D = \bigoplus_{\substack{y \in D \\ |\operatorname{supp}(y)| \geq 2}} U_{1,E_y},$$

where U_{1,E_y} is the rank-one uniform matroid on E_y.

Proof. The variable ground set is

$$E_{\mathrm{multi}} = \bigcup_{\substack{y \in D \\ |\operatorname{supp}(y)| \geq 2}} E_y.$$

The sets E_y are pairwise disjoint because their elements have distinct upper endpoints. By Theorem 6.1, choosing the variable part of an inclusion-minimal representative is exactly choosing one element from each block E_y. This is precisely the basis family of the direct sum of the rank-one uniform matroids U_{1,E_y}. □

Remark 7.2. *Equivalently, one may build a matroid on the full generating set by adding the elements of F_{ax} as fixed coloops and taking the direct sum with $\mathcal{M}_D$. We avoid this extra formalism and simply remove the fixed forced set.*

8 Enumerators and compression

The preceding results yield exact enumeration formulae and optimal compression ratios.

8.1 Inclusion-minimal representatives

The number of inclusion-minimal representatives is

$$N_{\min}(D) = \prod_{\substack{y \in D \\ |\operatorname{supp}(y)| \geq 2}} |D_{<y}|.$$

Indeed, for each multi-coordinate y, one chooses an arbitrary parent $x \in D_{<y}$. Axis generators are forced.

For a full box

$$D = [n_1] \times \cdots \times [n_d],$$

we have

$$|D_{<y}| = \prod_{t=1}^{d} y_t - 1.$$

Therefore

$$N_{\min}([n_1] \times \cdots \times [n_d]) = \prod_{\substack{1 \leq y_t \leq n_t \\ |\operatorname{supp}(y)| \geq 2}} \left(\prod_{t=1}^{d} y_t - 1 \right).$$

8.2 Hasse-supported minimum representatives

A representative is *Hasse-supported* if all its generators are cover relations.

For a multi-coordinate element y, the number of lower covers is

$$|\operatorname{supp}(y)|.$$

Therefore the number of Hasse-supported minimum representatives is

$$N_{\mathrm{Hasse}}(D) = \prod_{\substack{y\in D \\ |\operatorname{supp}(y)|\geq 2}} |\operatorname{supp}(y)|.$$

For a full box,

$$N_{\mathrm{Hasse}}([n_1] \times \cdots \times [n_d]) = \prod_{\substack{1\leq y_t\leq n_t \\ |\operatorname{supp}(y)|\geq 2}} |\operatorname{supp}(y)|.$$

In dimension 2, this specializes to

$$N_{\mathrm{Hasse}}([m] \times [n]) = 2^{(m-1)(n-1)}.$$

8.3 All representatives

Theorem 5.1 also counts all representatives.

For an axis element y, the forced generator

$$p_{y^-} \preceq p_y$$

must be included, while any other incoming generator from

$$D_{<y} \setminus \{y^-\}$$

may be included or omitted. Thus the number of choices in that endpoint fiber is

$$2^{|D_{<y}|-1}.$$

For a multi-coordinate element y, any nonempty subset of $D_{<y}$ may be chosen as the set of incoming generators. Thus the number of choices in that endpoint fiber is

$$2^{|D_{<y}|} - 1.$$

Therefore the total number of representatives is

$$N_{\mathrm{all}}(D) = \prod_{\substack{y\in D \\ |\operatorname{supp}(y)|=1}} 2^{|D_{<y}|-1} \cdot \prod_{\substack{y\in D \\ |\operatorname{supp}(y)|\geq 2}} \left(2^{|D_{<y}|} - 1\right).$$

The ordinary generating function by cardinality is

$$F_D(t) = \prod_{\substack{y\in D \\ |\operatorname{supp}(y)|=1}} t(1+t)^{|D_{<y}|-1} \cdot \prod_{\substack{y\in D \\ |\operatorname{supp}(y)|\geq 2}} \left((1+t)^{|D_{<y}|} - 1\right).$$

The coefficient of t^k is the number of representatives of size k.

8.4 Optimal compression of the product order

The strict product-order relation has size

$$|R_D^{<}| = \sum_{y \in D} (|D_{\le y}| - 1).$$

The minimum representative has size

$$|D| - 1.$$

Thus the optimal strict-order compression ratio is

$$\frac{|R_D^{<}|}{|D|-1} = \frac{\sum_{y \in D}(|D_{\le y}| - 1)}{|D|-1}.$$

For a full box

$$D = [n_1] \times \cdots \times [n_d],$$

we have

$$|D| = \prod_{t=1}^{d} n_t$$

and

$$|R_D| = \prod_{t=1}^{d} \frac{n_t(n_t+1)}{2}.$$

Hence

$$|R_D^{<}| = \prod_{t=1}^{d} \frac{n_t(n_t+1)}{2} - \prod_{t=1}^{d} n_t,$$

while

$$\min |T| = \prod_{t=1}^{d} n_t - 1.$$

For the cube $D = [n]^d$,

$$|R_D^{<}| = \left(\frac{n(n+1)}{2}\right)^d - n^d, \qquad \min |T| = n^d - 1.$$

As $n \to \infty$,

$$\frac{|R_D^{<}|}{\min |T|} \sim \left(\frac{n}{2}\right)^d.$$

Thus box-closure gives an optimal representative that is polynomially smaller, by a factor of order n^d, than the full strict product order in fixed dimension d.

9 Weighted optimization and random minimum representatives

Let a weight

$$w(x, y)$$

be assigned to every strict relation

$$p_x \preceq p_y, \qquad x < y.$$

9.1 Minimum-weight inclusion-minimal representatives

Among inclusion-minimal representatives, the minimum weight is

$$\sum_{\substack{y\in D\\ |\operatorname{supp}(y)|=1}} w(y^-,y) + \sum_{\substack{y\in D\\ |\operatorname{supp}(y)|\geq 2}} \min_{x<y} w(x,y).$$

The axis terms are forced, and each multi-coordinate endpoint fiber independently chooses one minimum-weight incoming generator.

9.2 Minimum-weight representatives with nonnegative weights

If all weights are nonnegative, then every minimum-weight representative may be chosen inclusion-minimal. Hence the same formula gives the minimum weight among all representatives.

9.3 Arbitrary real weights

If negative weights are allowed, extra generators may reduce the total weight. The problem still decomposes by upper endpoint.

For an axis element y, the forced generator $p_{y^-} \preceq p_y$ must be included, and any additional negative-weight incoming generator should be included. The optimal axis contribution is

$$w(y^-,y) + \sum_{\substack{x<y\\ x\neq y^-}} \min\{0, w(x,y)\}.$$

For a multi-coordinate element y, one must choose a nonempty subset of $D_{<y}$. The optimal contribution is

$$\min_{\varnothing\neq A\subseteq D_{<y}} \sum_{x\in A} w(x,y).$$

Equivalently, include all negative-weight incoming generators; if there are no negative-weight incoming generators, include one generator of minimum weight.

Thus the arbitrary-weight representative problem decomposes into independent endpoint-fiber optimizations.

9.4 Random minimum representatives

A uniformly random inclusion-minimal representative is sampled as follows:

1. include every forced axis generator;
2. for each multi-coordinate element y, choose one parent $x \in D_{<y}$ uniformly and independently.

Thus the probability that a generator $p_x \preceq p_y$ appears in a uniformly random minimum representative is

$$\mathbb{P}(p_x \preceq p_y \in T) = \begin{cases} 1, & y \text{ is an axis element and } x = y^-, \\ 0, & y \text{ is an axis element and } x \neq y^-, \\ \dfrac{1}{|D_{<y}|}, & y \text{ is multi-coordinate and } x < y. \end{cases}$$

The entropy of the uniform distribution on minimum representatives is

$$H_{\min}(D) = \sum_{\substack{y\in D\\ |\operatorname{supp}(y)|\geq 2}} \log|D_{<y}| = \log N_{\min}(D).$$

9.5 Minimum-representative polytope and Boltzmann sampling

The endpoint factorization gives a polyhedral form of the minimum-representative space.

For $y \in D \setminus \{\hat{0}\}$, define the admissible parent set

$$A_D(y) = \begin{cases} \{y^-\}, & |\operatorname{supp}(y)| = 1, \\ D_{<y}, & |\operatorname{supp}(y)| \geq 2. \end{cases}$$

Let

$$E_D = \{(x, y) : y \neq \hat{0},\ x \in A_D(y)\}.$$

Minimum representatives are exactly the sets

$$T_\pi = \{p_{\pi(y)} \preceq p_y : y \neq \hat{0}\}, \qquad \pi(y) \in A_D(y).$$

Therefore the convex hull of incidence vectors of minimum representatives is

$$\mathcal{P}_D = \left\{ z \in \mathbb{R}_{\geq 0}^{E_D} : \sum_{x \in A_D(y)} z_{(x,y)} = 1 \text{ for every } y \neq \hat{0} \right\}.$$

Equivalently,

$$\mathcal{P}_D \cong \prod_{y \neq \hat{0}} \Delta_{A_D(y)}.$$

This is the base polytope of the partition matroid from Theorem 7.1, with singleton blocks for the forced axis endpoints.

For an additive energy w, the partition function over minimum representatives is

$$Z_D(\beta) = \prod_{y \neq \hat{0}} \left(\sum_{x \in A_D(y)} e^{-\beta w(x,y)} \right).$$

Thus exact Boltzmann sampling is obtained by choosing each parent independently with probability

$$\mathbb{P}_\beta(\pi(y) = x) = \frac{e^{-\beta w(x,y)}}{\sum_{x' \in A_D(y)} e^{-\beta w(x',y)}}.$$

Uniform sampling is the case $\beta = 0$.

10 The two-dimensional projected LCA case

The box-closure system was motivated by a restricted cross-consistency rule for designated LCA-like pairs. We now describe this connection in dimension 2 in a more explicit form.

10.1 Ferrers pair systems and projected closure

Let

$$D \subseteq [m] \times [n]$$

be a Ferrers downset. Introduce leaf labels a_i for represented rows and b_j for represented columns, and write

$$p_{ij} = a_i b_j.$$

The designated universe is

$$P_D = \{a_i b_j : (i,j) \in D\}.$$

We define the projected LCA closure $+_D$ on relations in $P_D \times P_D$ by reflexivity, transitivity, and the following restricted cross-consistency rule. If

$$p_{i\ell} \preceq p_{km} \quad \text{and} \quad p_{rj} \preceq p_{km},$$

then, whenever the displayed pair symbols belong to P_D, one obtains

$$p_{ij} \preceq p_{km} \quad \text{and} \quad p_{r\ell} \preceq p_{km}.$$

This is exactly the two-dimensional box rule for the coordinate pairs (i,ℓ) and (r,j) with common upper endpoint (k,m).

This projected closure is not the full LCA (+)-closure on all pairs of leaves. It ignores same-side pairs $a_i a_{i'}$, $b_j b_{j'}$, singleton symbols, non-designated cross-pairs, and other interactions present in the full LCA universe. Moreover, the projection of a full LCA (+)-closure to the designated universe need not coincide with the box-closure of the designated subsystem.

10.2 A rooted DAG realizing the designated product order

For completeness, we recall a realization of the designated product order by a rooted DAG.

Proposition 10.1 (Ferrers grids are realized by rooted DAGs)**.** *For every finite Ferrers downset D, there is a rooted DAG N_D with leaf set containing the labels a_i, b_j and internal vertices v_{ij}, $(i,j) \in D$, such that*

$$\mathrm{lca}_{N_D}(a_i, b_j) = v_{ij}$$

for every $(i,j) \in D$, and

$$v_{ij} \preceq v_{k\ell} \iff i \le k,\ j \le \ell.$$

Equivalently, the LCA order of N_D, restricted to the designated pair set P_D, is the product order on D.

Proof. For each $(i,j) \in D$, introduce an internal vertex v_{ij}. If (i,j) is covered by (k,ℓ) in the product order, add a directed edge

$$v_{k\ell} \to v_{ij}.$$

If D has several maximal cells, add a root ρ with edges to the maximal vertices. If D has a unique maximal cell, it may serve as the root.

For each represented row i, add a leaf a_i below v_{i1}. For each represented column j, add a leaf b_j below v_{1j}.

The graph is acyclic because every internal edge decreases at least one coordinate. The internal ancestors of a_i are exactly the vertices $v_{k\ell}$ with $k \ge i$, and the internal ancestors of b_j are exactly the vertices $v_{k\ell}$ with $\ell \ge j$. Hence their common internal ancestors are precisely

$$\{v_{k\ell} : (k,\ell) \in D,\ k \ge i,\ \ell \ge j\}.$$

Because D is a downset and $(i,j) \in D$, this set has unique minimal element v_{ij}. Thus

$$\mathrm{lca}_{N_D}(a_i, b_j) = v_{ij}.$$

Finally, there is a directed path from $v_{k\ell}$ to v_{ij} if and only if one can decrease the first coordinate from k to i and the second coordinate from ℓ to j, namely if and only if

$$i \le k, \qquad j \le \ell.$$

This proves the claim. ∎

This statement is only about designated cross-pairs. The full LCA relation of the DAG may contain additional comparisons involving other pairs of leaves. If one wants the DAG to satisfy a particular convention for phylogenetic networks, degree-$(1,1)$ vertices may be suppressed or the construction may be modified in the standard way. The representative theorems above do not depend on these network conventions.

10.3 Projected closure equals the product order

Let H_D be the Hasse-cover relation of the product order on D:

$$H_D = \{p_x \preceq p_y : x \lessdot y\}.$$

Here $x \lessdot y$ means that x is an immediate predecessor of y in the product order.

Theorem 10.2 (Projected Ferrers closure)**.** *For every finite Ferrers downset D,*

$$H_D^{+_D} = \{p_{ij} \preceq p_{k\ell} : (i,j),(k,\ell) \in D,\ i \le k,\ j \le \ell\}.$$

Equivalently, the projected closure of the Hasse-cover relation is exactly the product order.

Proof. Let

$$Q_D = \{p_{ij} \preceq p_{k\ell} : (i,j),(k,\ell) \in D,\ i \le k,\ j \le \ell\}.$$

The relation Q_D contains H_D, is reflexive, and is transitive.

We check projected cross-consistency. Suppose two premises with common upper endpoint $p_{k\ell}$ belong to Q_D. Their lower endpoints have the form

$$p_{rc} = a_r b_c \quad \text{and} \quad p_{r'c'} = a_{r'} b_{c'}$$

with

$$r, r' \le k, \qquad c, c' \le \ell.$$

Any conclusion obtained by projected cross-consistency combines one a-index from one premise and one b-index from the other. Thus its lower endpoint has the form $p_{\alpha\beta}$, where

$$\alpha \le k, \qquad \beta \le \ell.$$

If $p_{\alpha\beta} \in P_D$, then

$$p_{\alpha\beta} \preceq p_{k\ell}$$

belongs to Q_D. Therefore Q_D is $+_D$-closed.

Since Q_D is $+_D$-closed and contains H_D,

$$H_D^{+_D} \subseteq Q_D.$$

Conversely, suppose $(i,j) \le (k,\ell)$ in the product order. Since D is a downset and both endpoints lie in D, there is a monotone path in the Hasse graph of D from (i,j) to (k,ℓ). Each step of this path is a cover in H_D. By transitivity,

$$p_{ij} \preceq p_{k\ell}$$

belongs to $H_D^{+_D}$. Hence

$$Q_D \subseteq H_D^{+_D}.$$

Thus equality holds. □

10.4 Minimum representatives in Ferrers grids

The two-dimensional representative results are immediate specializations of the general theory, but it is useful to spell them out in LCA notation.

The axis elements are the first row and first column. Their immediate predecessor generators are forced. Every interior cell (i, j), $i, j > 1$, requires exactly one incoming generator from its strict lower rectangle

$$[1, i] \times [1, j] \setminus \{(i, j)\}.$$

Thus

$$\min\{|T| : T^{+_D} = H_D^{+_D}\} = |D| - 1.$$

Every minimum representative is determined by an admissible parent map

$$\pi : D \setminus \{(1, 1)\} \to D$$

such that

$$\pi(1, j) = (1, j - 1), \qquad \pi(i, 1) = (i - 1, 1),$$

and, for $i, j > 1$,

$$\pi(i, j) \in [1, i] \times [1, j] \setminus \{(i, j)\}.$$

The representative is

$$T_\pi = \{p_{\pi(i,j)} \preceq p_{ij} : (i, j) \in D \setminus \{(1, 1)\}\}.$$

Consequently,

$$\#\{\text{minimum representatives}\} = \prod_{\substack{(i,j)\in D \\ i>1,\ j>1}} (ij - 1).$$

For a rectangle $D = [m] \times [n]$, this becomes

$$\prod_{i=2}^{m} \prod_{j=2}^{n} (ij - 1).$$

For the square $D = [s] \times [s]$,

$$\log \#\{\text{minimum representatives}\} = 2s^2 \log s - 2s^2 + O(s \log s).$$

10.5 Cover-contained representatives

A minimum representative is *cover-contained* if it is a subset of the Hasse-cover relation H_D.

The cover-contained minimum representatives are exactly the parent maps for which each interior cell chooses one of its two Hasse parents:

$$(i - 1, j) \quad \text{or} \quad (i, j - 1).$$

Therefore

$$\#\{\text{cover-contained minimum representatives}\} = 2^{\mathrm{int}(D)},$$

where

$$\mathrm{int}(D) = |\{(i, j) \in D : i > 1,\ j > 1\}|.$$

For $D = [m] \times [n]$, this is

$$2^{(m-1)(n-1)}.$$

The Hasse-cover presentation itself has exact redundancy

$$|H_D| - (|D| - 1) = \mathrm{int}(D).$$

Indeed, if r is the number of nonempty rows and c the number of nonempty columns, then the number of vertical covers is $|D| - c$ and the number of horizontal covers is $|D| - r$. Hence

$$|H_D| = 2|D| - r - c,$$

and

$$|H_D| - (|D| - 1) = |D| - r - c + 1 = \mathrm{int}(D).$$

For $D = [s] \times [s]$, the probability that a uniformly random minimum representative is cover-contained is

$$\prod_{i=2}^{s} \prod_{j=2}^{s} \frac{2}{ij - 1}.$$

Its logarithm is

$$-2s^2 \log s + O(s^2).$$

Thus local, cover-contained representatives form an asymptotically negligible subfamily of all minimum representatives.

10.6 Jump-length enumerator

The classification also yields an exact enumerator by locality. For a parent choice

$$(u, v) \to (i, j),$$

define its jump length as

$$\lambda((u, v), (i, j)) = (i - u) + (j - v).$$

For a minimum representative T_π, define its total jump length by

$$\Lambda(T_\pi) = \sum_{(i,j) \neq (1,1)} \lambda(\pi(i, j), (i, j)).$$

Let

$$b(D) = |\{(i, j) \in D \setminus \{(1, 1)\} : i = 1 \text{ or } j = 1\}|$$

be the number of non-minimum boundary cells.

Theorem 10.3 (Jump-length generating function)**.** *The generating function*

$$F_D(q) = \sum_{T \text{ minimum}} q^{\Lambda(T)}$$

is

$$F_D(q) = q^{b(D)} \prod_{\substack{(i,j) \in D \\ i>1,\ j>1}} \left(\sum_{\substack{0 \le a \le i-1 \\ 0 \le b \le j-1 \\ (a,b) \neq (0,0)}} q^{a+b} \right).$$

Equivalently,

$$F_D(q) = q^{b(D)} \prod_{\substack{(i,j) \in D \\ i>1,\ j>1}} \left(\frac{(1 - q^i)(1 - q^j)}{(1 - q)^2} - 1 \right),$$

where the second expression is interpreted as the corresponding polynomial.

Proof. Boundary cells have one forced parent at jump length 1, giving the factor $q^{b(D)}$.

For an interior cell (i,j), parent choices are strict lower cells

$$(u,v) \in [1,i] \times [1,j] \setminus \{(i,j)\}.$$

Writing

$$a = i - u, \qquad b = j - v,$$

the possible jumps are all pairs

$$0 \le a \le i-1, \qquad 0 \le b \le j-1, \qquad (a,b) \ne (0,0).$$

Since parent choices are independent over cells, the total generating function factors. The closed form follows from the product of two finite geometric sums. □

Setting $q = 1$ recovers the product formula for the number of minimum representatives.

For a uniformly random minimum representative T,

$$\mathbb{E}[\# \text{ cover edges in } T] = b(D) + \sum_{\substack{(i,j)\in D \\ i>1,\ j>1}} \frac{2}{ij-1}.$$

For $D = [s] \times [s]$, this is

$$2(s-1) + \sum_{i=2}^{s}\sum_{j=2}^{s} \frac{2}{ij-1} = 2s + O((\log s)^2).$$

Thus a uniformly random minimum representative has only $O(s + (\log s)^2)$ cover edges out of $s^2 - 1$ total generators.

Similarly,

$$\mathbb{E}[\Lambda(T)] = b(D) + \sum_{\substack{(i,j)\in D \\ i>1,\ j>1}} \frac{ij(i+j-2)}{2(ij-1)}.$$

For $D = [s] \times [s]$,

$$\mathbb{E}[\Lambda(T)] = \frac{1}{2}s^3 + O(s^2).$$

10.7 Small rectangles

For $D = [2] \times [3]$, there are six designated pairs and the minimum representative size is 5. Boundary parents are forced:

$$p_{11} \preceq p_{12}, \qquad p_{12} \preceq p_{13}, \qquad p_{11} \preceq p_{21}.$$

The interior cells are $(2,2)$ and $(2,3)$. They have respectively 3 and 5 possible parents, so the total number of minimum representatives is

$$3 \cdot 5 = 15.$$

The cover-contained subfamily has

$$2^2 = 4$$

members.

For $D = [3] \times [3]$, the minimum representative size is 8. The four interior cells

$$(2,2), \quad (2,3), \quad (3,2), \quad (3,3)$$

have respectively

$$3, \quad 5, \quad 5, \quad 8$$

possible parents. Thus the total number of minimum representatives is

$$3 \cdot 5 \cdot 5 \cdot 8 = 600.$$

The cover-contained subfamily has size

$$2^4 = 16.$$

11 Relationship with LCA constraints and triple representatives

The classical LCA-constraint problem concerns rooted trees and constraints comparing lowest common ancestors. Recent work extends LCA constraints to directed acyclic graphs and phylogenetic networks, where LCA existence and uniqueness become nontrivial and where closure operations involve the full universe of leaf pairs [ASSU81, EH26, HLM26, LH25, LAMSH25].

The present paper studies a different restricted closure system. Its relation to LCA theory is motivational and occurs explicitly in the two-dimensional designated cross-pair case. The main results should not be read as theorems about the full LCA (+)-closure.

Representative triple sets in rooted phylogenetic trees have a known matroidal structure. The box-closure systems studied here also yield a matroidal structure, but of a different and simpler kind: after removing the forced axis generators, the variable parts of the inclusion-minimal representatives form the bases of a partition matroid. This follows from endpoint-fiber decomposition, not from triple-closure theory.

12 Conclusion

We introduced box-closure systems on finite downsets of products of chains and solved their representative problem completely.

The main results are:

1. a complete characterization of all representatives;
2. the equivalence of inclusion-minimality and cardinality-minimality;
3. the minimum size formula $|D| - 1$;
4. a partition-matroid description of the variable parts of minimum representatives after removing forced axis generators;
5. exact enumerators for minimum, Hasse-supported, and all representatives;
6. optimal compression ratios for product orders under box-closure;
7. a decomposed solution of the weighted representative problem;
8. a factorized model of random minimum representatives and a product-of-simplices description of their polytope;

9. recovery of the two-dimensional projected Ferrers LCA-like system as a special case, including exact jump-length enumerators and cover-contained counts.

The contribution is a complete combinatorial analysis of a coordinate-mixture closure system. Its connection to LCA constraints is motivational and restricted, not a claim about the full LCA (+)-closure.

References

[ASSU81] A. V. Aho, Y. Sagiv, T. G. Szymanski, and J. D. Ullman, *Inferring a tree from lowest common ancestors with an application to the optimization of relational expressions*, SIAM Journal on Computing 10 (1981), 405–421.

[EH26] P. A. Ebert and M. Hellmuth, *Inferring phylogenetic networks from allowed and forbidden LCA-constraints*, arXiv:2605.03827, 2026.

[HLM26] M. Hellmuth, A. Lindeberg, and V. Moulton, *Encoding phylogenetic networks with least common ancestor constraints*, arXiv:2606.16963, 2026.

[HS18] M. Hellmuth and C. R. Seemann, *The matroid structure of representative triple sets and triple-closure computation*, European Journal of Combinatorics 70 (2018), 384–407.

[LAMSH25] A. Lindeberg, A. Alfonsson, V. Moulton, G. E. Scholz, and M. Hellmuth, *Inferring DAGs and phylogenetic networks from least common ancestors*, arXiv:2511.07965, 2025.

[LH25] A. Lindeberg and M. Hellmuth, *Simplifying and characterizing DAGs and phylogenetic networks via least common ancestor constraints*, Bulletin of Mathematical Biology 87 (2025), 44.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work, ChatGPT, by OpenAI, was used to assist with mathematical drafting, formalization, review, and editing. This work is shared as a preliminary AI-assisted mathematical note. The mathematical content may have been only partially reviewed and may contain errors; it should not be treated as peer-reviewed or as a fully verified manuscript.