UNDERDETERMINED POLYTOPIC MATRIX FACTORIZATION: IDENTIFIABILITY VIA DETERMINANT MAXIMIZATION AND FACE-SPAN GEOMETRY

Anonymous authors
Paper under double-blind review

ABSTRACT

Structured matrix factorization supports analytical representation learning by recovering latent vectors from unknown linear mixtures through structural constraints. Nonnegativity, sparsity, and antisparsity motivate nonnegative matrix factorization, sparse component analysis, and bounded component analysis, respectively. Polytopic matrix factorization (PMF) unifies these approaches through a prescribed polytopic domain, allowing different combinations of structural properties, e.g., nonnegative sparse coefficients with signed, amplitude bounded coefficients in the same representation. Existing PMF identifiability, however, requires at least as many measurements as latent components. We extend PMF to the underdetermined regime, with fewer measurements than latent components, while retaining its source correlation determinant-maximization (Det-Max) criterion. Our approach combines PMF's maximum-volume inscribed ellipsoid based sufficient scattering condition with a face-span skeleton formed by intersecting the polytope with linear spans of low-dimensional faces. Under explicit assumptions on cell population, relations among face spans, and local injectivity of the mixing matrix, we establish finite-sample exact recovery from noiseless observations: every global maximizer recovers the mixing matrix and all latent samples up to linear automorphism of the polytope. We derive sufficient measurement dimension conditions determined by the face-span geometry, demonstrate that Det-Max can exclude coordinate-mixing ambiguities, and establish a vanishing-regularization characterization of Det-Max through a source-entropy and reconstruction objective. These results provide a geometric foundation for underdetermined PMF, with implications for overcomplete dictionary learning with structural constraints tailored to different groups of latent coefficients.

1 INTRODUCTION

Structured matrix factorization provides an analytical approach to representation learning by recovering latent representations from unknown linear mixtures through explicit structural constraints. Nonnegativity, sparsity, and boundedness motivate established frameworks such as nonnegative matrix factorization (NMF), sparse component analysis (SCA), and bounded component analysis (BCA), respectively. Polytopic matrix factorization (PMF) unifies these approaches within a common geometric framework by constraining latent vectors to a prescribed polytope $\mathcal{P}$ (Tatli & Erdogan, 2021a;b). Beyond encompassing familiar models as special cases, PMF allows different structural properties to coexist within the same representation. For example, a nonnegative subvector with a shared ℓ_1 budget can coexist with signed, amplitude-bounded components, while coupled inequalities can impose overlapping group constraints. The choice of polytope thus provides a unified way to express and combine structural assumptions on the learned representation.

The value of this modeling flexibility depends on *identifiability*: whether the prescribed structure distinguishes the underlying latent representation and mixing matrix, rather than merely permitting an accurate reconstruction of the observations. In $\mathbf{Y} = \mathbf{H}_g\mathbf{S}_g$, the unknown mixing matrix $\mathbf{H}_g \in \mathbb{R}^{M\times r}$ maps unknown r-dimensional latent vectors to M measurements. Existing PMF theory assumes that $\mathbf{H}_g$ has full column rank, requiring $M \geq r$. Under sufficient scattering, its source-correlation

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

determinant-maximization (Det-Max) criterion identifies both factors up to *linear automorphisms* of $\mathcal{P}$, the invertible linear maps preserving the source domain.

This work develops a general framework for *underdetermined PMF*, extending PMF's unified treatment of latent structure to $M < r$ while retaining its Det-Max criterion. A principal application is *overcomplete dictionary learning*, where $\mathbf{H}_g$ contains more atoms than observation dimensions. Our goal is joint recovery of the dictionary and its coefficients, with structural constraints tailored to different coefficient groups rather than a representation restricted to coordinate sparsity.

The main obstacle is that underdetermination introduces nullspace ambiguities: latent vectors can change without affecting the observations. Boundedness alone cannot resolve these ambiguities. We therefore constrain samples to a *face-span skeleton*, a union of cells $\mathcal{P} \cap \mathrm{span}(F)$ associated with low-dimensional faces F. These cells are linear slices, not merely boundary faces. For an ℓ_1 ball, they recover coordinate-sparse models; for a cube, they admit dense vectors in which many components share a common magnitude and only a few vary independently.

Our identification argument separates the roles of low-dimensional structure and bounded source geometry. Under explicit cell-population, local-injectivity, and face-span compatibility assumptions, every exact feasible factorization is related to the true factors through a single invertible transformation. PMF's sufficient-scattering condition, based on the polytope's maximum-volume inscribed ellipsoid and polar geometry, and Det-Max then restrict this transformation to a polytope automorphism. Together, these ingredients yield finite-sample joint identifiability from noiseless observations for every global maximizer, extending PMF's common identification principle to the underdetermined setting.

1.1 Related Work

NMF (Lee & Seung, 1999), SCA (Zibulevsky & Pearlmutter, 2001), and BCA (Cruces, 2010; Erdogan, 2013) motivate PMF's unified treatment of source geometry. Log-determinant information criteria, including correlative information maximization (CorInfoMax), further develop PMF's optimization perspective (Erdogan, 2022; Bozkurt et al., 2023). Democratic representations instead control peak coefficient magnitude for a *known* underdetermined operator (Studer et al., 2014). Sparse coding and dictionary-learning algorithms (Olshausen & Field, 1997; Aharon et al., 2006a; Mairal et al., 2010), together with sparse underdetermined separation (Bofill & Zibulevsky, 2001), establish the main application context. Importantly, Cruces (2015) studies underdetermined BCA mixing-matrix identification while distinguishing it from source separability; our additional cell structure supports joint recovery.

Classical results address uniqueness and local identifiability for sparse codes (Aharon et al., 2006b; Gribonval & Schnass, 2010). Finite-sample joint recovery up to permutation and scaling is established by Hillar & Sommer (2015); Garfinkle & Hillar (2019). We extend their cell-matching argument to non-coordinate face spans, using span relations to obtain one ambient invertible transformation to which PMF's Det-Max principle applies. Sufficient-scattering NMF (Fu et al., 2018) and dictionary-volume criteria provide complementary principles. Sun & Huang (2025) combine rowwise ℓ_1 coefficient penalties with a dictionary log-volume term under strong scattering; we instead couple prescribed face-span geometry with the source-correlation determinant. Appendix A provides detailed comparisons.

1.2 Contributions

- **Unified underdetermined PMF framework.** We combine a polytope with a face-span skeleton for $M < r$, extending analytical representation learning to mixed sparse, nonnegative, bounded, and non-coordinate latent structures.
- **Finite-sample joint identifiability.** Under structural assumptions, each exact feasible factorization is a single invertible transform of the truth. Scattering and Det-Max select polytope automorphisms; we derive $d_q(\mathcal{L}) \leq M < r$, generic mixing conditions, and canonical certificates.
- **Skeleton versus polytope symmetries.** We show that Det-Max reduces $G_{\mathcal{L}}$ to $\mathcal{A}_{\mathcal{P}}$, can eliminate coordinate mixing beyond gain calibration, and is recovered as the vanishing-regularization limit of a source-entropy and reconstruction objective.

- **Online overcomplete dictionary learning with bio-plausible networks.** We derive online ER-CorInfoMax as locally learned, dimension-expanding Hebbian/anti-Hebbian networks. For codes combining nonnegativity and sparsity, a natural pairing since firing rates are nonnegative, interneurons enforce the activity budget and p-winner constraint; experiments use $10\times$ overcomplete factor.

2 Background on Polytopic Matrix Factorization

In this section, we aim to provide a brief background on PMF, which has so far been restricted to the overdetermined setting. In the following section, we will propose its extension to underdetermined setting.

2.1 Polytopic Matrix Factorization Problem

We study PMF through the generative model summarized from Tatli & Erdogan (2021b):

$$\mathbf{Y} = \mathbf{H}_g \mathbf{S}_g, \tag{1}$$

where $\mathbf{Y} \in \mathbb{R}^{M\times N}$ is the observation matrix with M as the observation vector dimension and N as the number of samples. $\mathbf{H}_g \in \mathbb{R}^{M\times r}$ is the unknown mixing matrix, and the columns of the $\mathbf{S}_g \in \mathbb{R}^{r\times N}$ matrix are the unknown latent vectors with r as the latent space dimension. The basic PMF assumption is that each column of $\mathbf{S}_g$ lies in a convex polytope $\mathcal{P} \subset \mathbb{R}^r$ (see Appendix B.1 for standard polytope representations), i.e.,

$$\mathbf{S}_{g_{:,j}} \in \mathcal{P}, \qquad j = 1, \dots, N. \tag{2}$$

Given $\mathbf{Y}$ and the structural prior encoded by $\mathcal{P}$, whose selection corresponds to latent features (see Appendix B.4 for examples), the goal is to recover both factors on the right side of Equation (1) up to a sign and permutation ambiguity:

$$\mathbf{H} = \mathbf{H}_g \mathbf{Q}^{-1}, \qquad \mathbf{S} = \mathbf{Q}\mathbf{S}_g, \tag{3}$$

where $\mathbf{Q} \in \mathrm{SP}(r)$ is an $r \times r$ signed permutation matrix.

2.2 Determinant Maximization Based Overdetermined PMF

The original PMF framework of Tatli & Erdogan (2021b) assumes an overdetermined setting, i.e., $M \geq r$, and that $\mathbf{H}_g$ has full column rank. Furthermore, it assumes that the latent vectors are "sufficiently scattered" with respect to a precise notion defined based on the containment of the maximum volume inscribed ellipsoid (MVIE) of $\mathcal{P}$ by the convex hull of the latent vectors of the generative model (see Appendix B for formal definitions of the MVIE, polar sets, and sufficient scattering). The same reference proves that, under these assumptions, the global maxima of the determinant maximization problem

$$\begin{aligned} \underset{\mathbf{H}\in\mathbb{R}^{M\times r}, \mathbf{S}\in\mathbb{R}^{r\times N}, \mathbf{S}_{:,k}\in\mathcal{P}, k=1,\dots,N}{\text{maximize}} \ & \frac{1}{2}\log\det \mathbf{R}_{\mathbf{S}}, \qquad \mathbf{R}_{\mathbf{S}} := \frac{1}{N}\mathbf{S}\mathbf{S}^\top, \\ \text{subject to } & \mathbf{Y} = \mathbf{H}\mathbf{S}, \end{aligned} \tag{4}$$

satisfy Equation (3) with $\mathbf{Q}$ restricted to the linear automorphism set of $\mathcal{P}$, defined as

$$\mathcal{A}_{\mathcal{P}} := \{\mathbf{Q} \in \mathbb{R}^{r\times r} : \mathbf{Q}\mathcal{P} = \mathcal{P}\}. \tag{5}$$

Furthermore, if $\mathcal{A}_{\mathcal{P}}$ is a subset of signed permutation matrices, (such a $\mathcal{P}$ is called "Det-Max identifiable"), then the original factors in Equation (1) are obtained up to sign and permutation ambiguity as desired. Background on PMF is provided in Appendix B.

2.3 Information maximization perspective

The reference Erdogan (2022) proposes an information theoretic lens for the Det-Max approach, where the objective of the optimization in Equation (4) is interpreted as the (correlative) entropy of the latent samples. Specifically, for $\epsilon > 0$, the empirical correlative information is

$$\begin{aligned} \widehat{I}_{\mathrm{c}}^{(\epsilon)}(\mathbf{Y}, \mathbf{S}) &:= \frac{1}{2}\log\det(\widehat{\mathbf{R}}_{\mathbf{S}} + \epsilon\mathbf{I}_r) - \frac{1}{2}\log\det(\widehat{\mathbf{R}}^{(\epsilon)}_{\mathbf{S}|\mathbf{Y}} + \epsilon\mathbf{I}_r), \\ \widehat{\mathbf{R}}^{(\epsilon)}_{\mathbf{S}|\mathbf{Y}} &:= \widehat{\mathbf{R}}_{\mathbf{S}} - \widehat{\mathbf{R}}_{\mathbf{S}\mathbf{Y}}(\widehat{\mathbf{R}}_{\mathbf{Y}} + \epsilon\mathbf{I}_M)^{-1}\widehat{\mathbf{R}}_{\mathbf{S}\mathbf{Y}}^\top, \end{aligned} \tag{6}$$

where $\widehat{\mathbf{R}}_{\mathbf{S}}$, $\widehat{\mathbf{R}}_{\mathbf{Y}}$, and $\widehat{\mathbf{R}}_{\mathbf{SY}}$ denote centered sample covariance and cross-covariance matrices. Using this measure, Erdogan (2022) proposes Correlative Information Maximization (CorInfoMax)

$$\underset{\mathbf{S}\in\mathbb{R}^{r\times N}}{\text{maximize}}\ \widehat{I}_{\text{c}}^{(\epsilon)}(\mathbf{Y},\mathbf{S}) \quad \text{subject to} \quad \mathbf{S}_{:,j}\in\mathcal{P}, \quad j=1,\dots,N, \tag{7}$$

as the replacement for the Det-Max problem. The first term is a centered, regularized counterpart of the Det-Max objective in Equation (4), promoting the spread of the latent samples. The second penalizes uncertainty in linearly predicting $\mathbf{S}$ from $\mathbf{Y}$, thereby softly encouraging the linear consistency imposed by exact factorization in the full-column-rank setting. This provides an information-maximization perspective on Det-Max, combining latent entropy maximization with a linear-predictability penalty. The reference Bozkurt et al. (2023) offers the online solution to CorInfoMax optimization where biologically plausible neural networks with piecewise-linear activations emerges as the natural solutions. Appendix C investigates the connection between the CorInfoMax and Det-Max optimizations, and proposes an alternative entropy based formulation defined in Equation (15) whose global maximizers approach Det-Max solutions as the regularization vanishes.

3 Underdetermined PMF on Face-Span Skeletons

The identifiability guarantee for overdetermined PMF relies on the mixing matrix $\mathbf{H}_g$ having full column rank. Moving to the underdetermined regime ($M < r$) introduces a genuine obstacle: the non-trivial nullspace $\ker(\mathbf{H}_g)$ allows changes in the latent space that leave the observations entirely unchanged. Consequently, the invertible change-of-coordinates argument that resolves identifiability in standard PMF no longer follows from exact reconstruction and bounded polytope geometry alone. To resolve this structural ambiguity, we constrain the latent vectors to a low-dimensional *face-span skeleton*, constructed from the linear spans of the polytope's faces.

3.1 Data Model

In the underdetermined setting of the generative model in Equation (1), the observation dimension M is strictly less than the latent dimension r. To allow identifiability, we constrain the latent vectors to a subset of $\mathcal{P}$ based on the subspaces generated by its k_0-dimensional faces.

First, we fix a face dimension $0 \le k_0 < r$ and set $p := k_0 + 1$. For any k_0-dimensional face F of the source polytope $\mathcal{P}$, we define the associated face-span subspace and polytope cell by

$$L_F := \operatorname{span}(F), \qquad C_F := L_F \cap \mathcal{P}. \tag{8}$$

A k_0-face has a p-dimensional linear span when its affine hull does not contain the origin ($\mathbf{0} \notin \operatorname{aff}(F)$). Retaining these specific faces and removing duplicate spans yields a unique face-span arrangement:

$$\mathcal{L}(\mathcal{P}, k_0) = \{\operatorname{span}(F) : F \in \mathcal{F}_{k_0}(\mathcal{P}),\ 0 \notin \operatorname{aff}(F)\} = \{L_a\}_{a=1}^{J}. \tag{9}$$

As shown in Proposition F.2, cells corresponding to omitted faces are already entirely contained within the retained cells, so omitting these faces does not change the skeleton. Here, $[J] := \{1,\dots,J\}$ denotes the index set of the distinct face spans. We define the resulting "face-span skeleton" as the union of these linear slices:

$$\operatorname{Skel}_{k_0}(\mathcal{P}) := \bigcup_{a=1}^{J} C_a, \qquad C_a := L_a \cap \mathcal{P}. \tag{10}$$

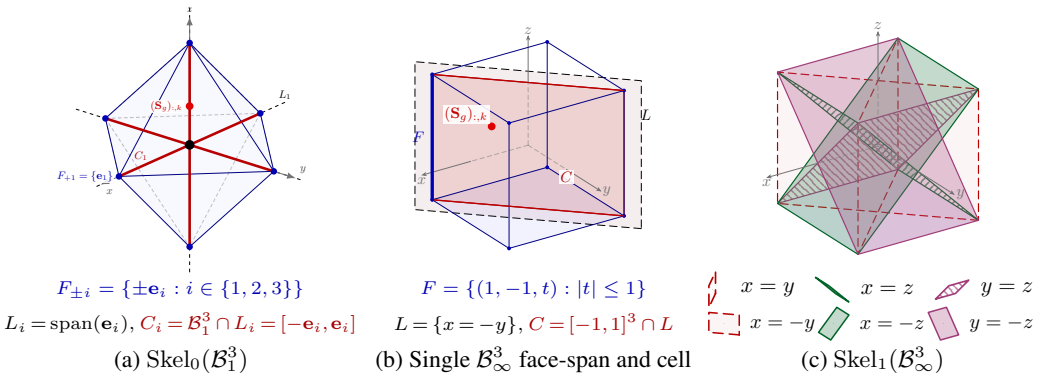
Unlike an ordinary union of boundary faces, this skeleton consists of full linear slices passing through the polytope. Each true source vector is assigned a cell label $a_g(k)$ satisfying

$$\mathbf{S}_{g_{:,k}} \in C_{a_g(k)}. \tag{11}$$

If we consider the canonical polytopes in Appendix B.4: for the ℓ_1 norm ball, the distinct spans are the coordinate p-subspaces, so

$$\operatorname{Skel}_{p-1}(\mathcal{B}_1^r) = \{\mathbf{s} : \|\mathbf{s}\|_1 \le 1,\ \|\mathbf{s}\|_0 \le p\}.$$

For other polytopes, the spans are not restricted to coordinate subspaces. Sources can therefore be dense while varying within a low-dimensional cell. For instance, in the ℓ_∞ norm ball, each face span



(a) $\mathrm{Skel}_0(\mathcal{B}_1^3)$ (b) Single $\mathcal{B}_\infty^3$ face-span and cell (c) $\mathrm{Skel}_1(\mathcal{B}_\infty^3)$

Figure 1: Face-span geometry of $\mathcal{B}_1^3$ and $\mathcal{B}_\infty^3$.

has $p-1$ free coordinate directions and one direction with fixed signs on the remaining coordinates. We illustrate the face-span skeletons of $\mathcal{B}_1^3$ and $\mathcal{B}_\infty^3$ in Figure 1.

Finally, we retain the sufficient scattering condition of Tatli & Erdogan (2021b). This condition quantifies how well the latent samples fill the polytope by utilizing its maximum volume inscribed ellipsoid (MVIE). For brevity, the formal parameterization of the MVIE, its optimization problem, and the precise polar set definitions are deferred to Appendix B.

3.2 Feasible Set, Det-Max Identifiability and Source-Entropy Objective

With the latent vectors constrained to the face-span skeleton, our goal is to isolate the true factors from all feasible factorizations. Given the data $\mathbf{Y}$, the structural prior $\mathcal{P}$, and the face dimension k_0, we formalize this search space as the face-span feasible set:

$$\mathcal{F}_{k_0}^{\mathrm{span}}(\mathbf{Y}, \mathcal{P}) := \{(\mathbf{H}, \mathbf{S}) : \mathbf{Y} = \mathbf{H}\mathbf{S},\ \mathbf{S}_{:,k} \in \mathrm{Skel}_{k_0}(\mathcal{P}) \text{ for every } k\}. \tag{12}$$

Because the problem is underdetermined, this feasible set contains structural ambiguities. To resolve them, we must distinguish transformations that merely preserve the face-span arrangement from those that preserve the entire polytope. We define the set of matrices preserving the face-span arrangement as $G_{\mathcal{L}} := \{\mathbf{B} \in \mathbb{R}^{r \times r} : \mathbf{B}\mathcal{L} = \mathcal{L}\}$. In contrast, the target ambiguities we accept are restricted to $\mathcal{A}_{\mathcal{P}}$, as defined in Equation (5).

While every polytope automorphism inherently permutes its face spans (meaning $\mathcal{A}_{\mathcal{P}} \subseteq G_{\mathcal{L}}$), the reverse is not true. For example, a uniform scaling $c\mathbf{I}_r$ belongs to $G_{\mathcal{L}}$ for any $c \neq 0$, whereas preserving the full polytope requires $|c| = 1$ by volume preservation. To filter out these excess transformations in $G_{\mathcal{L}}$, we search within the feasible set to maximize the volume of the source correlation:

$$\max_{(\mathbf{H},\mathbf{S}) \in \mathcal{F}_{k_0}^{\mathrm{span}}(\mathbf{Y},\mathcal{P})} \frac{1}{2} \log\det \mathbf{R}_{\mathbf{S}}, \qquad \mathbf{R}_{\mathbf{S}} := \frac{1}{N}\mathbf{S}\mathbf{S}^\top. \tag{13}$$

We say that the underdetermined face-span model is Det-Max identifiable if every global optimizer of Equation (13) successfully eliminates all excess arrangement ambiguities, yielding factors of the form:

$$\mathbf{S}^\star = \mathbf{Q}\mathbf{S}_g, \qquad \mathbf{H}^\star = \mathbf{H}_g\mathbf{Q}^{-1}, \qquad \mathbf{Q} \in \mathcal{A}_{\mathcal{P}}. \tag{14}$$

When $\mathcal{A}_{\mathcal{P}} \subseteq \mathrm{SP}(r)$, this optimal recovery reduces to the standard signed-permutation ambiguity.

Finally, for practical optimization, we relax the strict exact-reconstruction constraint and formulate a regularized source-entropy and reconstruction objective from Appendix C:

$$K_{\epsilon,\lambda}(\mathbf{H}, \mathbf{S}) := \frac{\lambda}{2} \log\det(\mathbf{R}_{\mathbf{S}} + \epsilon\mathbf{I}_r) - \frac{1}{2} \log\det(\mathbf{R}_{\mathbf{E}}(\mathbf{H}, \mathbf{S}) + \epsilon\mathbf{I}_M), \tag{15}$$

where $\epsilon > 0$, $\lambda > 0$, $\mathbf{R}_{\mathbf{S}} := \frac{1}{N}\mathbf{S}\mathbf{S}^\top$, and $\mathbf{R}_{\mathbf{E}}(\mathbf{H}, \mathbf{S}) := \frac{1}{N}(\mathbf{Y} - \mathbf{H}\mathbf{S})(\mathbf{Y} - \mathbf{H}\mathbf{S})^\top$. We maximize this objective while keeping the source columns within $\mathrm{Skel}_{k_0}(\mathcal{P})$. We call this entropy-regularized CorInfoMax (ER-CorInfoMax); unlike symmetric CorInfoMax, its global maximizers approach Det-Max solutions as $\epsilon \to 0$ (Lemma C.1).

3.3 CONDITIONS FOR JOINT IDENTIFIABILITY

To guarantee that global maxima of the Det-Max program recover only the true factors, two requirements must hold. First, every exact feasible factorization must be a single global linear transform of the truth. In the underdetermined regime this is not automatic: since $\ker(\mathbf{H}_g) \neq \{\mathbf{0}\}$, a candidate may assign samples to different cells, or alter them within their cells, without changing $\mathbf{Y}$. We call this the *face-arrangement ambiguity*. We resolve it by matching true and candidate samples cell by cell, which yields a linear map on each face span L_a. These local maps define one global change of coordinates only if they preserve the linear relations among the spans: whenever $\sum_a \mathbf{z}_a = \mathbf{0}$ with $\mathbf{z}_a \in L_a$, the mapped vectors must also sum to zero. Second, among the resulting transforms, determinant maximization must select exactly the polytope automorphisms $\mathcal{A}_\mathcal{P}$. Assumptions 3.1–3.3 below address the first requirement and Assumption 3.4 the second. Because the local maps can only be compared on small groups of spans, the first requirement is governed by two quantities determined by $\mathcal{P}$ and k_0 alone.

Definition 3.1 (Relation completeness and local dimension budget)**.** *For an integer $q \geq 2$, the arrangement $\mathcal{L} = \mathcal{L}(\mathcal{P}, k_0)$ is* q-relation-complete *if every relation $\sum_{a=1}^{J} \mathbf{z}_a = \mathbf{0}$ with $\mathbf{z}_a \in L_a$ is a finite sum of relations, each involving at most q spans, with contributions added separately within each L_a. Its* local dimension budget *is*

$$d_q(\mathcal{L}) := \max_{\substack{I \subseteq [J] \\ 1 \leq |I| \leq q}} \rho(I), \qquad \rho(I) := \dim\Big(\sum_{a \in I} L_a\Big). \tag{16}$$

Thus q counts the spans needed to check relations, while d_q is the dimension that the mixing matrix must preserve on their sums. Appendix E illustrates both notions, and Definition F.2 gives the equivalent block-sum formulation used in the proofs.

Assumption 3.1 (Polytope and face spans)**.** *$\mathcal{P} \subset \mathbb{R}^r$ is full-dimensional, and $\mathcal{L}(\mathcal{P}, k_0)$ is q-relation-complete for a known integer $q \geq 2$ with $d_q(\mathcal{L}(\mathcal{P}, k_0)) \leq M$.*

Assumption 3.2 (q-GP)**.** *$\mathbf{H}_g$ is injective on every sum of at most q face spans: $\ker(\mathbf{H}_g) \cap \sum_{a \in I} L_a = \{\mathbf{0}\}$ for every $I \subseteq [J]$ with $1 \leq |I| \leq q$.*

Assumption 3.3 (Cell population)**.** *For every $a \in [J]$, the samples with $a_g(k) = a$ include at least $J(p-1)+1$ vectors in linear general position in L_a (see Appendix B.6 for a formal definition and genericity properties).*

Assumption 3.4 (Sufficient scattering (Tatli & Erdogan, 2021b))**.** *With $\mathrm{conv}(\mathbf{S}_g) = \mathrm{conv}\{\mathbf{S}_{g_{:,1}}, \ldots, \mathbf{S}_{g_{:,N}}\}$ and $\mathcal{E}_\mathcal{P}$ as the MVIE of $\mathcal{P}$, we have $\mathcal{E}_\mathcal{P} \subseteq \mathrm{conv}(\mathbf{S}_g) \subseteq \mathcal{P}$ and*

$$\mathrm{conv}(\mathbf{S}_g)^{*,\mathbf{g}_\mathcal{P}} \cap \partial\big(\mathcal{E}_\mathcal{P}^{*,\mathbf{g}_\mathcal{P}}\big) = \mathrm{ext}\big(\mathcal{P}^{*,\mathbf{g}_\mathcal{P}}\big), \tag{17}$$

where $A^{*,\mathbf{g}_\mathcal{P}}$ is the $g_\mathcal{P}$-centered polar of the set A (see Appendix B.3).

Assumption 3.2 holds almost surely for absolutely continuously distributed $\mathbf{H}_g$ whenever $d_q \leq M$ (see Proposition D.1 in Appendix D). When $q = 2$, it reduces to $\mathrm{spark}(\mathbf{H}_g) > 2p$ for p-sparse models (Donoho & Elad, 2003), and for general subspaces, it characterizes injectivity on their union for a known linear operator (Lu & Do, 2008; Nam et al., 2013). To extend this to a polytope geometry, we require injectivity on sums up to q spans. Assumption 3.3 follows the population and subspace matching arguments used in sparse component analysis and dictionary identifiability (Georgiev et al., 2005; Aharon et al., 2006b; Hillar & Sommer, 2015; Garfinkle & Hillar, 2019), applied here to general face spans. Proposition B.4 gives a sufficient sampling condition for linear general position. Relation completeness ensures that preservation of relations involving at most q spans is sufficient for the local maps to extend to one linear map (Proposition F.1). The coordinate face-span arrangements satisfy it with $q = 2$. For one-dimensional spans and $3 \leq q \leq J$, it coincides with $(q-1)$-generation of the corresponding hyperplane arrangement (Brandt & Terao, 1994; Le & Mohammadi, 2015). Proposition G.2 verifies Assumption 3.1 for the canonical polytopes. Assumption 3.4 is the PMF sufficient scattering condition, which parallels scattering conditions for nonnegative factorization (Fu et al., 2018). Further discussion is deferred to Appendix E.

3.4 IDENTIFIABILITY

The underdetermined face-span model is Det-Max identifiable: its global maximizers recover the true mixing matrix and latent samples up to a linear automorphism of the source polytope. Following the

structural conditions outlined in Section 3.3, the proof proceeds in two distinct stages. First, exact reconstruction on the polytope skeleton reduces every feasible factorization to a single global linear transformation of the ground truth. Second, determinant maximization isolates the true polytope automorphisms from among these feasible transformations. The first step relies on cell population matching to pair each true cell with a candidate cell, local injectivity to induce an invertible linear map on each face span L_a, and relation completeness to glue these local maps into a unified transformation on $\mathbb{R}^r$.

Theorem 3.1 (Face-arrangement ambiguity)**.** *Assume Assumptions 3.1 to 3.3. For every* $(\mathbf{H}, \mathbf{S}) \in \mathcal{F}_{k_0}^{\mathrm{span}}(\mathbf{Y}, \mathcal{P})$*, there exists* $\mathbf{B} \in G_{\mathcal{L}}$ *such that*

$$\mathbf{S} = \mathbf{B}\mathbf{S}_g, \qquad \mathbf{H} = \mathbf{H}_g\mathbf{B}^{-1}. \tag{18}$$

Because the global transformation $\mathbf{B}$ acts uniformly across all source columns, the kernel of $\mathbf{H}_g$ cannot conceal sample-wise deviations. However, feasibility alone is insufficient to guarantee full recovery: the transformation group $G_{\mathcal{L}}$ is typically broader than the automorphism group $\mathcal{A}_{\mathcal{P}}$ (for instance, $G_{\mathcal{L}}$ contains isotropic scalings $c\mathbf{I}_r$ for all $c \neq 0$).

To execute the second step, observe that for any feasible factorization, the Det-Max objective decomposes as

$$\tfrac{1}{2}\log\det \mathbf{R}_{\mathbf{B}\mathbf{S}_g} = \tfrac{1}{2}\log\det \mathbf{R}_{\mathbf{S}_g} + \log|\det \mathbf{B}|. \tag{19}$$

Maximizing the objective is therefore mathematically equivalent to maximizing $|\det \mathbf{B}|$ over all feasible transformations. This optimization problem is governed by the PMF scattering argument (Tatli & Erdogan, 2021b), which operates directly on the polytope geometry and sample points, independently of the underlying face-span relations.

Lemma 3.2 (Det-Max and polar tightness)**.** *Let* $\mathcal{P}$ *be full-dimensional and assume Assumption 3.4. If an invertible* $\mathbf{B} \in \mathbb{R}^{r\times r}$ *satisfies* $\mathbf{B}\mathbf{S}_{g_{:,k}} \in \mathcal{P}$ *for every* k*, then* $|\det \mathbf{B}| \leq 1$*, with equality if and only if* $\mathbf{B} \in \mathcal{A}_{\mathcal{P}}$*.*

Because the identity transformation $\mathbf{B} = \mathbf{I}_r$ is always feasible, the upper bound $|\det \mathbf{B}| = 1$ is achievable. Combining Theorem 3.1 and Lemma 3.2 yields our primary recovery result.

Theorem 3.3 (Det-Max identifiability)**.** *Assume Assumptions 3.1 to 3.4. For every global maximizer* $(\mathbf{H}^\star, \mathbf{S}^\star)$ *of Equation* (13)*, there exists* $\mathbf{Q} \in \mathcal{A}_{\mathcal{P}}$ *such that*

$$\mathbf{H}^\star = \mathbf{H}_g\mathbf{Q}^{-1}, \qquad \mathbf{S}^\star = \mathbf{Q}\mathbf{S}_g. \tag{20}$$

Moreover, as $\epsilon \to 0$*, every source accumulation point of approximate global maximizers of* $K_{\epsilon,\lambda}$ *in Equation* (15)*, with optimization error tending to zero, is of the form* $\mathbf{Q}\mathbf{S}_g$ *for some* $\mathbf{Q} \in \mathcal{A}_{\mathcal{P}}$*. If* $\mathcal{A}_{\mathcal{P}} \subseteq \mathrm{SP}(r)$*, recovery is exact up to signed permutation; for nonnegative canonical polytopes, recovery is achieved up to permutation.*

The theorem therefore separates two layers of identifiability. The face-span constraints restrict exact feasible factorizations to arrangement-preserving transformations in $G_{\mathcal{L}}$, whereas the bounded polytope geometry and Det-Max reduce this residual class to $\mathcal{A}_{\mathcal{P}}$. How much the second layer contributes depends on the source domain. Section 4 makes this distinction concrete for canonical PMF polytopes and for a noncanonical example with a genuine coordinate-mixing ambiguity. Formal proofs are deferred to Appendix F.

4 Canonical Models and Residual Ambiguities

The abstract conditions above admit explicit certificates for four canonical polytopes in Appendix B.4:

$$\mathcal{B}_1^r := \{\mathbf{x} \in \mathbb{R}^r : \|\mathbf{x}\|_1 \leq 1\} \text{ (sparse)}, \qquad \mathcal{B}_{1,+}^r := \{\mathbf{x} \in \mathbb{R}_+^r : \|\mathbf{x}\|_1 \leq 1\} \text{ (nonnegative-sparse)},$$
$$\mathcal{B}_\infty^r := [-1,1]^r \text{ (anti-sparse)}, \qquad \mathcal{B}_{\infty,+}^r := [0,1]^r \text{ (nonnegative anti-sparse)}.$$

With $p = k_0 + 1$, the ℓ_1 skeletons impose bounded p-sparsity, whereas cube cells can be dense and non-coordinate. Table 1 in Appendix G and Proposition D.1 give generic local injectivity for $2p \leq M < r$ in the ℓ_1 and positive-dimensional cube cases. Cube vertex lines have sufficient budgets $4 \leq M < r$ (symmetric) and $3 \leq M < r$ (nonnegative). Polytope automorphisms are signed permutations for symmetric domains and permutations otherwise.

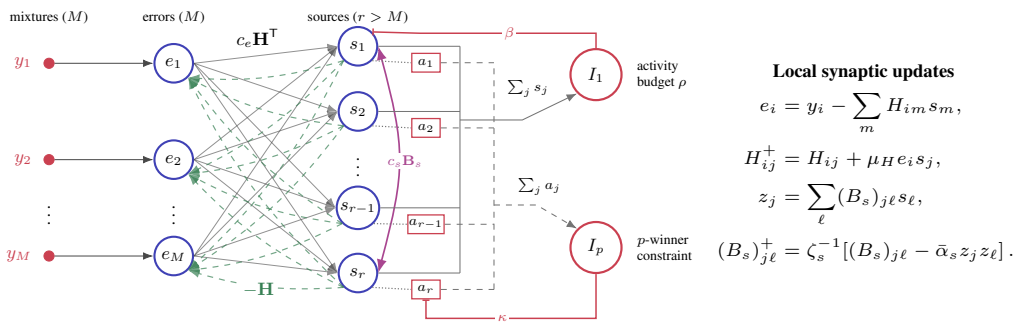


Figure 2: Online network for the nonnegative ℓ_1 face-span skeleton. The source layer expands M inputs to $r > M$ outputs; I_p imposes the p-winner constraint and I_1 the activity budget. The dictionary update is error–source Hebbian, while the lateral update is anti-Hebbian; $\bar{\alpha}_s$ is a constant approximation to $\alpha_s(k)$. See Appendix J for more details.

Under the structural hypotheses, Proposition H.1 leaves component symmetries and individual gains for the ℓ_1 models, but only a common gain for cubes. With sufficient scattering, Det-Max calibrates these gains; for cubes, one boundary-attaining sample replaces full scattering without relaxing cell population (Appendix H). To exhibit ambiguity reduction beyond calibration, let $r = 2m$, $m \geq 3$, and consider

$$\mathcal{P}_m = \mathrm{conv}\big(\{\pm\tfrac{6}{5}\mathbf{e}_i : i \in [2m]\} \cup \{\sigma_1\mathbf{e}_{2b-1} + \sigma_2\mathbf{e}_{2b} : b \in [m],\ \sigma_1, \sigma_2 \in \{\pm 1\}\}\big)\,.$$

For $k_0 = 1$ and $4 \leq M < r$, Appendix I constructs an exact alternative factorization by rotating and contracting one coordinate pair. This transformation preserves the face-span arrangement but not $\mathcal{P}_m$, and Det-Max excludes it. Thus Det-Max can remove genuine coordinate mixing left unresolved by the structural constraints, rather than merely calibrating gains.

5 Batch and Online Algorithms

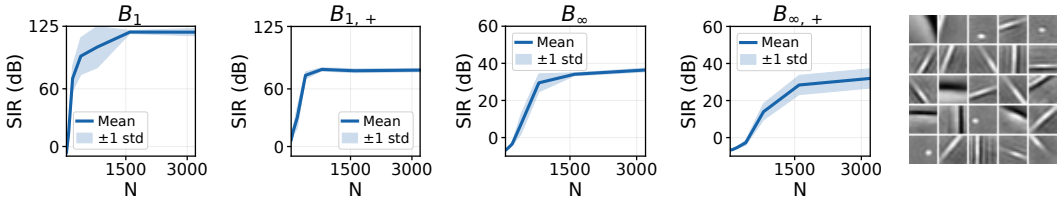
In order to provide numerical experiments supporting identifiability analysis, we use a full-data batch solver and an online neural approximation derived from $K_{\epsilon,\lambda}$ in Equation (15).

Batch algorithm. At each iteration, the batch method forms the source and reconstruction-error moments from all N samples, alternates updates of $\mathbf{H}$ and $\mathbf{S}$, and projects every source column onto $\mathrm{Skel}_{k_0}(\mathcal{P})$. Thus, the source step enforces both the polytopic domain and the face-span constraint. The complete procedure, including the gradients and exact projections for the canonical polytopes, is given in Appendix K.1. Synthetic experiments for canonical polytopes use this full-data method.

Online neural algorithm. The overdetermined CorInfoMax algorithm of Bozkurt et al. (2023) admits a biologically plausible recurrent-network interpretation with local Hebbian/anti-Hebbian updates and polytopic piecewise-linear activations. Appendix J adapts this construction to an underdetermined expansion network with $M < r$, where the face-span skeleton supplies an additional structural constraint. Error neurons form $\mathbf{e} = \mathbf{y} - \mathbf{H}\mathbf{s}$, while source neurons combine the error-correction drive $\mathbf{H}^\top\mathbf{e}$ with the lateral source-precision drive $\mathbf{B}_s\mathbf{s}$. The dictionary $\mathbf{H}$ therefore serves as the source-to-mixture backward predictor, and $\mathbf{B}_s$ supplies lateral recurrence boosting representation entropy. For the special case of nonnegative sparse polytope, $\mathcal{B}_{1,+}^r$ in Figure 2, the interneuron I_p implements the p-winner for skeleton constraint, whereas I_1 enforces the ℓ_1 activity budget. Figure 2 summarizes the circuit and its local synaptic updates. See Appendix J for more details.

6 Numerical Experiments

The impact of sample size: Figure 3a reports the signal-to-interference power ratio (SIR) versus N for batch recovery in four canonical models with $r = 25$, $M = 20$, and $k_0 = 3$. High SIR occurs far below the sufficient population count $J[J(p-1)+1]$. Appendices K.3 and K.4 explain this fitted-map separation metric, its scope, and the N- and M-sweeps.



(a) (b)

Figure 3: (a) Synthetic source recovery SIR versus sample count N for $r = 25$, $M = 20$, and $k_0 = 3$ for the $\mathcal{B}_1^r$ (sparse), $\mathcal{B}_{1,+}^r$ (nonnegative sparse), $\mathcal{B}_\infty^r$ (antisparse), and $\mathcal{B}_{\infty,+}^r$ (nonnegative antisparse) polytopes. (b) Example learned atoms with $\mathcal{B}_{1,+}^r$ codes. For the full dictionary, see Figure 12.

Octagon experiment: This experiment illustrates ambiguity reduction beyond gain calibration. Reconstruction-only fitting produced 76 near-exact fits in 80 trials, 58 with coordinate mixing. Rescaling and local projected-gradient updates did not remove this mixing, whereas adding block-reorientation proposals selected using $K_{\epsilon,1}$ recovered all 76 near-exact fits. Fixed-candidate comparisons isolate Det-Max's role: reconstruction scores tie, but the source-volume term favors the correct orientation (Appendix K.6).

Overcomplete nonnegative-sparse dictionary learning: We use 16×16 natural-image patches (van Hateren & van der Schaaf, 1998), compressed to $M = 100$ by PCA. Following the sparse-coding setting of Olshausen & Field (1997); Olshausen et al. (2009), we constrain $r = 1000$ coefficients to $\mathcal{B}_{1,+}^r$ with the face-span skeleton. The localized, oriented atoms qualitatively resemble V1 simple-cell receptive fields (Figure 3b). Appendix K provides details and online dictionary-learning results using the same $\mathcal{B}_{1,+}^r$ face-span model, illustrating PMF's flexibility in combining sparsity and nonnegativity to learn representations compatible with biological neuronal firing rates.

7 Conclusions and Limitations

This work extends PMF as an analytical framework for representation learning to the underdetermined regime, where the latent dimension exceeds the observation dimension. Through the choice of source polytope, PMF can flexibly encode sparse, nonnegative, bounded, antisparse, and heterogeneous combinations of such structures. The additional face-span skeleton supplies the low-dimensional organization needed for identifiability, recovering coordinate sparsity for the ℓ_1 ball while also allowing non-coordinate and dense structured representations for other polytopes.

Our analysis separates two ambiguity reductions. The face-arrangement theorem shows that finite cell populations, local injectivity, and compatibility among face spans force all cellwise changes of coordinates to arise from one ambient transformation in $G_{\mathcal{L}}$. This is stronger than a block-sparse interpretation, which alone permits unrelated basis changes across blocks. The polytope constraint and Det-Max then reduce this skeleton symmetry to the polytope automorphism group $\mathcal{A}_{\mathcal{P}}$. For canonical models, Det-Max mainly calibrates residual gains; the octagon example shows that it can also eliminate genuine coordinate mixing. Its role is therefore determined by the gap between skeleton and polytope symmetries.

The source-entropy and reconstruction objective provides a Det-Max-consistent alternative, with batch projected updates and exact projections for canonical polytopes. Its online approximation yields a biologically plausible, dimension-expanding feature learner with local Hebbian and anti-Hebbian updates, lateral source-precision connections, polytopic piecewise-linear responses, and, for the nonnegative ℓ_1 model, separate interneurons enforcing the activity budget and p-winner skeleton constraint.

Limitations. The all-cell population assumption entails $N \geq J[J(p-1)+1]$, a worst-case requirement of our theorem rather than a necessary recovery bound (Appendix E). High SIR below this count demonstrates empirical separation, not theorem-certified exact recovery. The theory concerns noiseless global optima; noise stability and solver convergence are not established. We assume known polytope and face dimension; general skeletons may lack efficient projections or simple neural circuits.

AI USE STATEMENT

We used generative AI tools, including large language models, to assist with implementing the numerical methods and preparing the manuscript. For the numerical experiments, these tools helped generate initial code for the optimization algorithms, synthetic-data generation, performance evaluation, and plotting, and assisted with exploring optimization hyperparameters. We also used these tools to check mathematical proofs and suggest corrections and revisions to improve their correctness and rigor. Additional uses included grammatical editing, refinement of the exposition, and identification of potentially relevant literature.

We did not use generative AI tools to propose research hypotheses, design the experiments, translate text, or clean or reformat existing datasets. The implementation assistance described above included code for generating the synthetic datasets used in the numerical experiments.

The authors reviewed all AI-assisted material, including text, code, and proposed proof revisions. We reviewed the generated initial codes for correctness and manually further extended them, checked suggested proof corrections and their connections to existing results against the relevant literature, and incorporated selected revisions with appropriate citations. References identified with AI assistance were examined manually before inclusion. We take full responsibility for the final content of this work, including its text, mathematical claims and proofs, code, and reported results.

REPRODUCIBILITY STATEMENT

All details of the numerical experiments including hyperparameters are provided in the main text and in Appendix K. The proofs and justifications of theoretical claims and statements are provided in Appendices C,D,F,G,H,I.1-I.3, and J. We also share all programs used for the numerical experiments as the supplementary material.

REFERENCES

Michal Aharon, Michael Elad, and Alfred M. Bruckstein. K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Transactions on Signal Processing*, 54 (11):4311–4322, 2006a. doi: 10.1109/TSP.2006.881199.

Michal Aharon, Michael Elad, and Alfred M. Bruckstein. On the uniqueness of overcomplete dictionaries, and a practical way to retrieve them. *Linear Algebra and its Applications*, 416(1): 48–67, 2006b. doi: 10.1016/j.laa.2005.06.035.

Sanjeev Arora, Rong Ge, Tengyu Ma, and Ankur Moitra. Simple, efficient, and neural algorithms for sparse coding. In *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pp. 113–149. PMLR, 2015. URL https://proceedings.mlr.press/v40/Arora15.html.

Alexander Barvinok. *A Course in Convexity*, volume 54 of *Graduate Studies in Mathematics*. American Mathematical Society, 2002. doi: 10.1090/gsm/054.

Thomas Blumensath and Mike E. Davies. Iterative thresholding for sparse approximations. *Journal of Fourier Analysis and Applications*, 14:629–654, 2008. doi: 10.1007/s00041-008-9035-z.

Pau Bofill and Michael Zibulevsky. Underdetermined blind source separation using sparse representations. *Signal Processing*, 81(11):2353–2362, 2001. doi: 10.1016/S0165-1684(01)00120-7.

Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004. URL https://web.stanford.edu/~boyd/cvxbook/.

Bariscan Bozkurt and Alper T. Erdogan. On identifiable polytope characterization for polytopic matrix factorization. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 3343–3347. IEEE, 2022. doi: 10.1109/ICASSP43922.2022.9746370.

Bariscan Bozkurt, Cengiz Pehlevan, and Alper T. Erdogan. Biologically-plausible determinant maximization neural networks for blind separation of correlated sources. In *Advances in Neural Information Processing Systems*, volume 35, pp. 13704–13717, 2022. doi: 10.52202/068431-0996.

Bariscan Bozkurt, Ateş İsfendiyaroğlu, Cengiz Pehlevan, and Alper T. Erdogan. Correlative information maximization based biologically plausible neural networks for correlated source separation. In *International Conference on Learning Representations*, 2023. URL https://openreview.net/forum?id=8JsaP7j1cL0.

K. A. Brandt and H. Terao. Free arrangements and relation spaces. *Discrete & Computational Geometry*, 12(1):49–63, 1994. doi: 10.1007/BF02574365.

Thomas M. Cover. Geometrical and statistical properties of systems of linear inequalities with applications in pattern recognition. *IEEE Transactions on Electronic Computers*, EC-14(3): 326–334, 1965. doi: 10.1109/PGEC.1965.264137.

Sergio Cruces. Bounded component analysis of linear mixtures: A criterion of minimum convex perimeter. *IEEE Transactions on Signal Processing*, 58(4):2141–2154, 2010. doi: 10.1109/TSP.2010.2040685.

Sergio Cruces. Bounded component analysis of noisy underdetermined and overdetermined mixtures. *IEEE Transactions on Signal Processing*, 63(9):2279–2294, 2015. doi: 10.1109/TSP.2015.2404312.

David L. Donoho. For most large underdetermined systems of linear equations the minimal ℓ_1-norm solution is also the sparsest solution. *Communications on Pure and Applied Mathematics*, 59(6): 797–829, 2006. doi: 10.1002/cpa.20132.

David L. Donoho and Michael Elad. Optimally sparse representation in general (nonorthogonal) dictionaries via ℓ_1 minimization. *Proceedings of the National Academy of Sciences*, 100(5): 2197–2202, 2003. doi: 10.1073/pnas.0437847100.

Will Dorrell, Peter Latham, and James Whittington. Convex efficient coding. In *International Conference on Learning Representations*, pp. 26478–26514, 2026. URL https://proceedings.iclr.cc/paper_files/paper/2026/hash/2c74f005aabbf90a8f1747d99f387321-Abstract-Conference.html.

John Duchi, Shai Shalev-Shwartz, Yoram Singer, and Tushar Chandra. Efficient projections onto the ℓ_1-ball for learning in high dimensions. In *Proceedings of the 25th International Conference on Machine Learning*, pp. 272–279. ACM, 2008. doi: 10.1145/1390156.1390191.

Alper T. Erdogan. A class of bounded component analysis algorithms for the separation of both independent and dependent sources. *IEEE Transactions on Signal Processing*, 61(22):5730–5743, 2013. doi: 10.1109/TSP.2013.2280115.

Alper T. Erdogan. An information maximization based blind source separation approach for dependent and independent sources. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 4378–4382. IEEE, 2022. doi: 10.1109/ICASSP43922.2022.9746099.

Xiao Fu, Kejun Huang, and Nicholas D. Sidiropoulos. On identifiability of nonnegative matrix factorization. *IEEE Signal Processing Letters*, 25(3):328–332, 2018. doi: 10.1109/LSP.2018.2789405.

Charles J. Garfinkle and Christopher J. Hillar. On the uniqueness and stability of dictionaries for sparse representation of noisy signals. *IEEE Transactions on Signal Processing*, 67(23):5884–5892, 2019. URL https://arxiv.org/abs/1606.06997.

Pando Georgiev, Fabian Theis, and Andrzej Cichocki. Sparse component analysis and blind source separation of underdetermined mixtures. *IEEE Transactions on Neural Networks*, 16(4):992–996, 2005.

Rémi Gribonval and Karin Schnass. Dictionary identification - sparse matrix-factorization via ℓ_1-minimization. *IEEE Transactions on Information Theory*, 56(7):3523–3539, 2010. doi: 10.1109/TIT.2010.2048466.

Christopher J. Hillar and Friedrich T. Sommer. When can dictionary learning uniquely recover sparse data from subsamples? *IEEE Transactions on Information Theory*, 61(11):6290–6297, 2015. doi: 10.1109/TIT.2015.2460238.

Jingzhou Hu and Kejun Huang. Global identifiability of ℓ_1-based dictionary learning via matrix volume optimization. In *Advances in Neural Information Processing Systems*, volume 36, pp. 36165–36186, 2023. doi: 10.52202/075280-1568.

Anastasios Kyrillidis, Stephen Becker, Volkan Cevher, and Christoph Koch. Sparse projections onto the simplex. In *Proceedings of the 30th International Conference on Machine Learning*, number 2 in Proceedings of Machine Learning Research, pp. 235–243. PMLR, 2013. URL https://proceedings.mlr.press/v28/kyrillidis13.html.

Dinh Van Le and Fatemeh Mohammadi. On the Orlik–Terao ideal and the relation space of a hyperplane arrangement. *Advances in Applied Mathematics*, 71:34–51, 2015. doi: 10.1016/j.aam.2015.08.003.

Daniel D. Lee and H. Sebastian Seung. Learning the parts of objects by non-negative matrix factorization. *Nature*, 401(6755):788–791, 1999. doi: 10.1038/44565.

Yue M. Lu and Minh N. Do. A theory for sampling signals from a union of subspaces. *IEEE Transactions on Signal Processing*, 56(6):2334–2345, 2008.

Julien Mairal, Francis Bach, Jean Ponce, and Guillermo Sapiro. Online learning for matrix factorization and sparse coding. *Journal of Machine Learning Research*, 11:19–60, 2010. URL https://www.jmlr.org/papers/v11/mairal10a.html.

Sangnam Nam, Mike E. Davies, Michael Elad, and Rémi Gribonval. The cosparse analysis model and algorithms. *Applied and Computational Harmonic Analysis*, 34(1):30–56, 2013. doi: 10.1016/j.acha.2012.03.006.

Bruno A. Olshausen and David J. Field. Sparse coding with an overcomplete basis set: A strategy employed by V1? *Vision Research*, 37(23):3311–3325, 1997. doi: 10.1016/S0042-6989(97)00169-7.

Bruno A. Olshausen, Charles F. Cadieu, and David K. Warland. Learning real and complex overcomplete representations from the statistics of natural images. In Vivek K. Goyal, Manos Papadakis, and Dimitri Van De Ville (eds.), *Wavelets XIII*, volume 7446, pp. 74460S. International Society for Optics and Photonics, SPIE, 2009. doi: 10.1117/12.825882.

Pentti Paatero and Unto Tapper. Positive matrix factorization: A non-negative factor model with optimal utilization of error estimates of data values. *Environmetrics*, 5(2):111–126, 1994. doi: 10.1002/env.3170050203.

Jorge Silva, Minhua Chen, Yonina C. Eldar, Guillermo Sapiro, and Lawrence Carin. Blind compressed sensing over a structured union of subspaces. arXiv preprint arXiv:1103.2469, 2011. URL https://arxiv.org/abs/1103.2469.

Meera Sitharam, Mohamad Tarifi, and Menghan Wang. Combinatorial rigidity of incidence systems and application to dictionary learning. *Journal of Symbolic Computation*, 88:21–46, 2018. doi: 10.1016/j.jsc.2018.02.003.

Daniel A. Spielman, Huan Wang, and John Wright. Exact recovery of sparsely-used dictionaries. In *Proceedings of the 25th Annual Conference on Learning Theory*, volume 23 of *Proceedings of Machine Learning Research*, pp. 37.1–37.18. PMLR, 2012. URL https://proceedings.mlr.press/v23/spielman12.html.

Christoph Studer, Tom Goldstein, Wotao Yin, and Richard G. Baraniuk. Democratic representations. arXiv preprint arXiv:1401.3420, 2014. URL https://arxiv.org/abs/1401.3420.

Yuchen Sun and Kejun Huang. Global identifiability of overcomplete dictionary learning via ℓ_1 and volume minimization. In *International Conference on Learning Representations*, 2025. URL https://proceedings.iclr.cc/paper_files/paper/2025/hash/84409c45a0defe347c895b004b1c675b-Abstract-Conference.html.

Gokcan Tatli and Alper T. Erdogan. Generalized polytopic matrix factorization. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 3235–3239. IEEE, 2021a. doi: 10.1109/ICASSP39728.2021.9413709.

Gokcan Tatli and Alper T. Erdogan. Polytopic matrix factorization: Determinant maximization based criterion and identifiability. *IEEE Transactions on Signal Processing*, 69:5431–5447, 2021b. doi: 10.1109/TSP.2021.3112918.

J. H. van Hateren and A. van der Schaaf. Independent component filters of natural images compared with simple cells in primary visual cortex. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 265(1394):359–366, 1998. doi: 10.1098/rspb.1998.0303. URL https://doi.org/10.1098/rspb.1998.0303.

James C. R. Whittington, Will Dorrell, Surya Ganguli, and Timothy E. J. Behrens. Disentanglement with biological constraints: A theory of functional cell types. In *International Conference on Learning Representations*, 2023. URL https://arxiv.org/abs/2210.01768.

Michael Zibulevsky and Barak A. Pearlmutter. Blind source separation by sparse decomposition in a signal dictionary. *Neural Computation*, 13(4):863–882, 2001. doi: 10.1162/089976601300014385.

Günter M. Ziegler. *Lectures on Polytopes*, volume 152 of *Graduate Texts in Mathematics*. Springer, 1995. doi: 10.1007/978-1-4613-8431-1.

APPENDIX CONTENTS

A Related Work: Detailed Discussion **15**
- A.1 Structured factorizations and polytopic latent domains . . . 15
- A.2 Dictionary learning and underdetermined separation . . . 15
- A.3 Sparse identifiability and general-position arguments . . . 16
- A.4 Subspace geometry and ambient-factor recovery . . . 16
- A.5 Volume criteria and ambiguity selection . . . 17
- A.6 Information objectives and optimization . . . 17

B Background **17**
- B.1 Polytopes . . . 17
- B.2 Maximum Volume Inscribed Ellipsoid . . . 18
- B.3 Polar Sets . . . 18
- B.4 Canonical and Heterogeneous Feature Polytopes . . . 19
- B.5 Sufficiently Scattered Sets . . . 19
- B.6 Linear general position . . . 19

C CorInfoMax and a Det-Max-Consistent Reconstruction Objective **21**
- C.1 The CorInfoMax Objective and Its Optimization Profile . . . 22
- C.2 The Rank-Deficiency Pitfall of Standard CorInfoMax . . . 22
- C.3 A Weighted Det-Max-Consistent Reconstruction Objective . . . 23

D Genericity of q-GP **26**

E Discussion of the Identifiability Conditions **27**

F Proof of The Identifiability Theorem **28**
- F.1 Preliminary Work . . . 28
- F.2 Proof of the Face-Arrangement Ambiguity Theorem . . . 35
- F.3 Proof of the Det-Max and Polar-Tightness Lemma . . . 42
- F.4 Proof of the Main Theorem . . . 45

G Canonical Polytope Certificates **46**
- G.1 Computational verification of finite instances . . . 50

H Structural ambiguity in the canonical models **52**

I A polytope with coordinate-mixing ambiguities **53**
- I.1 Polytope and face spans . . . 53
- I.2 Polytope symmetries and feasibility of scattering . . . 55
- I.3 A coordinate-mixing alternative for k=1 . . . 55
- I.4 Euclidean projection onto the skeleton . . . 56

J Online ER-CorInfoMax for Underdetermined PMF **58**
- J.1 Online objective and neural drive . . . 58
- J.2 Fast inference and slow synaptic learning for a general polytope . . . 59
- J.3 Nonnegative ℓ_1-ball: a p-winner, activity-budget circuit . . . 60

K Algorithm and Details on Numerical Experiments **62**
- K.1 Full-data projected optimization for synthetic recovery . . . 62
- K.2 Synthetic experiments for canonical polytopes . . . 63
- K.3 Source recovery metric and interpretation . . . 64
- K.4 Discussion of synthetic SIR results . . . 65
- K.5 $B_{1,+}$ dictionary learning from image patches . . . 66
- K.6 Synthetic Experiments for Octagon . . . 67
- K.7 Online nonnegative sparse ER-CorInfoMax for image-patch dictionary learning . . . 69

A Related Work: Detailed Discussion

This appendix expands the comparisons in Section 1.1. We distinguish the structure imposed on latent vectors, the factorization object to be identified, and the role of the optimization criterion. These distinctions separate the modeling flexibility inherited from PMF from the additional geometric conditions used here to obtain underdetermined joint identifiability.

A.1 Structured factorizations and polytopic latent domains

Early nonnegative factorizations and parts-based representation learning were developed by Paatero & Tapper (1994); Lee & Seung (1999). Sparse component analysis exploits sparse source descriptions (Zibulevsky & Pearlmutter, 2001), whereas bounded component analysis uses bounded-support geometry and can accommodate dependent sources (Cruces, 2010; Erdogan, 2013). These approaches illustrate how assumptions on latent representations can distinguish otherwise ambiguous matrix factorizations.

PMF combines these geometric viewpoints by constraining sources to a prescribed polytope and maximizing their sample-correlation determinant (Tatli & Erdogan, 2021a;b). Bozkurt & Erdogan (2022) further characterize identifiable polytopes. The modeling advantage is not merely a choice between a sparse model and a bounded model. A single polytope can encode different combinations of sign restrictions, componentwise amplitude bounds, and shared activity budgets across groups of components, including coupled or overlapping groups. For example, a nonnegative subvector subject to a shared ℓ_1 budget can coexist with a signed subvector subject to individual amplitude bounds. An ℓ_1 budget alone does not impose an exact cardinality constraint; in our underdetermined model, the additional low-dimensional structure is supplied by the face-span skeleton.

PMF's theoretical appeal is a common identification principle across these domains. Sufficient scattering combines containment of the polytope's maximum-volume inscribed ellipsoid within the source convex hull with polar tightness. Together with determinant optimality, this geometry restricts the remaining ambiguity to linear automorphisms of the prescribed polytope. Our work retains this selection principle rather than proposing a replacement for determined PMF. The missing step addressed here is to establish that competing underdetermined factorizations are related through a single invertible transformation to which that principle can be applied. The resulting guarantee remains conditional on the selected polytope, its face-span arrangement, and the samples satisfying the stated assumptions; modeling flexibility is not an unconditional identifiability claim for every polytope.

A.2 Dictionary learning and underdetermined separation

Overcomplete sparse coding (Olshausen & Field, 1997), K-SVD (Aharon et al., 2006a), and online matrix factorization (Mairal et al., 2010) established influential representation-learning formulations and algorithms. Sparse representations also support blind separation with fewer mixtures than sources (Bofill & Zibulevsky, 2001). These lines emphasize sparse coefficients, but dictionary learning as an application is not inherently restricted to sparsity or nonnegativity.

In our notation, $\mathbf{H}_g \in \mathbb{R}^{M\times r}$ with $M < r$ is both an underdetermined mixing matrix and an overcomplete synthesis dictionary: its r columns are atoms in the M-dimensional observation space. Joint dictionary–code recovery up to permutation and scaling is already established for sparse models (Hillar & Sommer, 2015; Garfinkle & Hillar, 2019). Our extension concerns the prescribed polytopic face-span model and the characterization of its residual ambiguities. Accordingly, the application-level extension is dictionary learning with prescribed combinations of coefficient properties, while the mathematical result concerns joint identifiability of the generating factorization.

Underdetermined identification is not confined to sparse source models. In particular, Cruces (2015) establishes underdetermined BCA mixing-matrix identifiability from Cartesian convex-support geometry while distinguishing mixing-system identification from source separability. This distinction is important: identifying the mixing matrix does not, by itself, resolve the source ambiguity created by its nontrivial kernel. Our additional face-span structure supports a finite-sample joint recovery result with a prescribed polytopic domain and a source determinant objective. We therefore do not claim the first underdetermined identification result for bounded sources.

Democratic representations (Studer et al., 2014) provide a complementary reason to consider coefficient properties beyond sparsity: minimizing the largest coefficient magnitude distributes a signal across many coordinates. Their representation-design problem uses a *known* underdetermined operator. Our problem instead identifies an unknown operator and its generating latent samples. Moreover, the face-span restrictions used here are additional assumptions; they do not follow from bounded coefficients or peak control alone.

A.3 Sparse identifiability and general-position arguments

For a known dictionary, spark-based uniqueness and ℓ_1-recovery theory give conditions for recovering sparse representations (Donoho & Elad, 2003; Donoho, 2006). With an unknown dictionary, early overcomplete uniqueness results (Aharon et al., 2006b), local ℓ_1 identifiability (Gribonval & Schnass, 2010), and provable sparse-coding algorithms (Spielman et al., 2012; Arora et al., 2015) address complementary assumptions and recovery regimes. Our global-optimizer characterization is distinct from a guarantee for a particular dictionary-learning algorithm.

The closest starting point for our structural argument is the finite-sample general-position approach of Hillar & Sommer (2015); Garfinkle & Hillar (2019). We use their cell-matching logic: sufficiently many samples in general position in one true image subspace cannot all be covered by proper intersections with candidate image subspaces. Our sufficient per-cell count is $J(p-1)+1$, where J is the number of distinct cells and p their span dimension. For coordinate-sparse sources, $J = \binom{r}{p}$, and this recovers the existing full-support sufficient count. We do not claim an improved sparse-case sample bound or a new origin for this counting argument.

The extension concerns non-coordinate face spans and consistency across them. Sources in $\mathcal{P} \cap \operatorname{span}(F)$ can be dense in their original coordinates while varying in a low-dimensional cell. Identifying the collection of observed cell images is therefore only an intermediate step: their local correspondences must arise from one map on the original latent space. The relation conditions in our lifting theorem establish this consistency and relate all source columns, not only the designated general-position samples.

A.4 Subspace geometry and ambient-factor recovery

Union-of-subspaces and cosparse models provide related geometric perspectives (Silva et al., 2011; Nam et al., 2013). Incidence geometry connects dictionary learning to rigidity (Sitharam et al., 2018), while generation of relation spaces by relations of bounded support is established in arrangement theory (Le & Mohammadi, 2015; Brandt & Terao, 1994). We use a block formulation for face spans, together with local injectivity, to establish a common ambient transformation between exact feasible factorizations.

This explains the geometry-dependent sufficient measurement budget $d_q(\mathcal{L}) \le M < r$: $d_q(\mathcal{L})$ is the largest dimension spanned by at most q cells, and the chosen relation order q must generate the arrangement's block-sum relations. The budget applies to the general arrangement, with sparse and cube models providing explicit specializations. Local injectivity holds almost surely for absolutely continuously distributed mixing matrices under this budget, but this genericity does not supply the population or scattering assumptions, nor does the budget assert an optimal recovery threshold.

The model can also be expressed as block-sparse coding. If $\mathbf{U}$ concatenates bases of the known cell spans, then

$$\mathbf{S} = \mathbf{U}\mathbf{Z}, \qquad \mathbf{Y} = (\mathbf{H}\mathbf{U})\mathbf{Z},$$

where each column of $\mathbf{Z}$ is supported on one block. The original polytopic constraints remain necessary: each column of $\mathbf{U}\mathbf{Z}$ must still lie in $\mathcal{P}$. This is a useful reformulation, not an alternative recovery theorem for the original factors. Its effective dictionary is $\mathbf{H}\mathbf{U}$, rather than $\mathbf{H}$. Recovering effective blocks up to unrelated blockwise changes does not automatically identify the original mixing matrix or latent vectors. The common ambient transformation enforced by the face-span relations is the additional consistency statement required here.

A.5 VOLUME CRITERIA AND AMBIGUITY SELECTION

Sufficient scattering supports nonnegative-factor identification (Fu et al., 2018), and Hu & Huang (2023) study dictionary-volume optimization for nonsingular dictionaries. Sun & Huang (2025) propose an overcomplete dictionary-learning formulation combining rowwise ℓ_1 penalties on the coefficient matrix with a dictionary log-volume term, and analyze its identifiability under a strong-scattering condition. Our comparison concerns the objective, prescribed latent domain, and recovery target: we retain PMF's *source-correlation determinant* and combine it with face-span constraints, rather than adopting a dictionary-volume formulation.

The structural and objective-based steps have distinct roles. Writing $\mathcal{L}$ for the face-span arrangement, the lifting theorem relates exact feasible factorizations through a matrix in the arrangement-preserving group

$$G_{\mathcal{L}} = \{\mathbf{B} \in \mathbb{R}^{r\times r} : \mathbf{B}\mathcal{L} = \mathcal{L}\}.$$

Under sufficient scattering, determinant maximization then selects the smaller polytope automorphism group

$$\mathcal{A}_{\mathcal{P}} = \{\mathbf{Q} \in \mathbb{R}^{r\times r} : \mathbf{Q}\mathcal{P} = \mathcal{P}\}.$$

The additional constraints $\mathbf{BS}_{g_{:,j}} \in \mathcal{P}$ depend on the samples, whereas $G_{\mathcal{L}}$ and $\mathcal{A}_{\mathcal{P}}$ depend only on the arrangement and the polytope, respectively. The substantive conclusion is that its determinant-maximizing transformations are exactly $\mathcal{A}_{\mathcal{P}}$, not simply that $\mathcal{A}_{\mathcal{P}} \subseteq G_{\mathcal{L}}$.

The role of determinant maximization also depends on the source geometry. For the canonical sparse models, the structural step already reduces ambiguity to component gains and permutations; for the canonical cube models, it leaves a common gain and component symmetries. The objective then calibrates the gains. Our octagon example establishes a different possibility: a coordinate-mixing transformation remains feasible after the structural step but is excluded by determinant maximization. The theory distinguishes this additional ambiguity reduction from gain calibration and does not make determinant maximization or full scattering necessary in every special case.

A.6 INFORMATION OBJECTIVES AND OPTIMIZATION

Log-determinant information criteria extend the PMF perspective (Erdogan, 2022), with neural realizations and CorInfoMax formulations developed by Bozkurt et al. (2022; 2023). Related nonnegative representation theories study disentanglement and convex efficient coding (Whittington et al., 2023; Dorrell et al., 2026). These works provide complementary motivations for structured representations; our contribution here is the underdetermined face-span factorization and its conditional joint-identifiability guarantee.

Our vanishing-regularization result concerns a source-entropy and reconstruction objective, which is distinct from the symmetric CorInfoMax criterion of Section 2. Source accumulation points of approximate global maximizers of its least-squares profile solve exact Det-Max as regularization and optimization error vanish. Appendix C proves this result and explains why it is not a general equivalence between CorInfoMax and Det-Max. This is a global-optimization statement, not a recovery guarantee for finite-regularization stationary points or online iterations.

Finally, exact skeleton projections provide implementable constraint-handling operations. Sparse support selection and projection onto ℓ_1 balls or simplices have established antecedents (Blumensath & Davies, 2008; Duchi et al., 2008; Kyrillidis et al., 2013); our discussion credits those ingredients alongside the sorting-based cube rules. Solving a projection subproblem exactly does not solve the full nonconvex matrix-factorization problem. The identifiability theorem is therefore kept separate from convergence guarantees for numerical solvers and from noise-stability claims.

B BACKGROUND

B.1 POLYTOPES

A convex polytope in $\mathbb{R}^r$ admits two standard descriptions. In H-form, it is written as the intersection of finitely many half-spaces,

$$\mathcal{P} = \{\mathbf{x} \mid \langle \mathbf{a}_i, \mathbf{x}\rangle \le b_i,\ i = 1, \dots, f\}. \tag{21}$$

In V-form, it is written as the convex hull of its vertices,

$$\mathcal{P} = \mathrm{conv}(\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}). \tag{22}$$

These two descriptions encode the same object through its faces and its extreme points.

B.2 Maximum Volume Inscribed Ellipsoid

The maximum volume inscribed ellipsoid (MVIE) of a full-dimensional polytope $\mathcal{P} \subset \mathbb{R}^r$ is the ellipsoid of largest volume contained in $\mathcal{P}$. Following the notation in Boyd & Vandenberghe (2004), we write it as

$$\mathcal{E}_{\mathcal{P}} = \{\mathbf{C}_{\mathcal{P}}\mathbf{u} + \mathbf{g}_{\mathcal{P}} \mid \|\mathbf{u}\|_2 \leq 1\}, \tag{23}$$

where $\mathbf{g}_{\mathcal{P}} \in \mathbb{R}^r$ is the center and the symmetric positive-definite matrix $\mathbf{C}_{\mathcal{P}} \in \mathbb{R}^{r\times r}$ determines the principal axes and their lengths. For the half-space representation $\mathcal{P} = \{\mathbf{x} \in \mathbb{R}^r : \mathbf{a}_i^\top \mathbf{x} \leq b_i,\ i = 1, \ldots, f\}$, the pair $(\mathbf{C}_{\mathcal{P}}, \mathbf{g}_{\mathcal{P}})$ is obtained by solving the convex optimization problem (Boyd & Vandenberghe, 2004, Section 8.4.2)

$$\begin{aligned} \underset{\mathbf{C}\in\mathbb{R}^{r\times r},\, \mathbf{g}\in\mathbb{R}^r}{\text{minimize}} \quad & -\log\det \mathbf{C} \\ \text{subject to} \quad & \|\mathbf{C}\mathbf{a}_i\|_2 + \mathbf{a}_i^\top \mathbf{g} \leq b_i, \qquad i = 1, \ldots, f, \\ & \mathbf{C} = \mathbf{C}^\top \succ \mathbf{0}. \end{aligned} \tag{24}$$

In the PMF identifiability analysis, the MVIE is the reference object used to quantify how well the latent samples fill the polytope.

B.3 Polar Sets

Let $K \subset \mathbb{R}^r$ and let $\mathbf{d}$ be an interior point. The centralized polar of K with respect to $\mathbf{d}$ is defined as (see Tatli & Erdogan (2021b))

$$K^{*,\mathbf{d}} = \{\mathbf{x} \in \mathbb{R}^r \mid \langle \mathbf{x}, \mathbf{y} - \mathbf{d}\rangle \leq 1 \text{ for all } \mathbf{y} \in K\}. \tag{25}$$

When $\mathbf{d} = \mathbf{0}$, we simply write K^*. A key fact is reversal of inclusion: if $A \subset B$, then $B^{*,\mathbf{d}} \subset A^{*,\mathbf{d}}$.

For the ellipsoid in Equation (23), the polar with respect to $\mathbf{g}_{\mathcal{P}}$ is

$$\mathcal{E}^{*,\mathbf{g}_{\mathcal{P}}} = \{\mathbf{x} \in \mathbb{R}^r \mid \langle \mathbf{x}, \mathbf{C}_{\mathcal{P}}\mathbf{u}\rangle \leq 1 \text{ for all } \mathbf{u} \in \mathbb{R}^r \text{ with } \|\mathbf{u}\|_2 \leq 1\}. \tag{26}$$

Since $\max_{\|\mathbf{u}\|_2\leq 1}\langle \mathbf{x}, \mathbf{C}_{\mathcal{P}}\mathbf{u}\rangle = \|\mathbf{C}_{\mathcal{P}}^\top \mathbf{x}\|_2 = \|\mathbf{C}_{\mathcal{P}}\mathbf{x}\|_2$, this simplifies to

$$\mathcal{E}^{*,\mathbf{g}_{\mathcal{P}}} = \{\mathbf{x} \in \mathbb{R}^r \mid \|\mathbf{C}_{\mathcal{P}}\mathbf{x}\|_2 \leq 1\} = \{\mathbf{C}_{\mathcal{P}}^{-1}\mathbf{u} \mid \mathbf{u} \in \mathbb{R}^r,\ \|\mathbf{u}\|_2 \leq 1\}. \tag{27}$$

We also recall the bipolar theorem (Barvinok, 2002, Chapter IV, 1.2): for any closed convex set $K \subseteq \mathbb{R}^r$ containing the origin,

$$(K^*)^* = K. \tag{28}$$

For compact convex $K \subset \mathbb{R}^r$ and $\mathbf{g} \in \mathrm{int}(K)$, we have $K^{*,\mathbf{g}} = (K - \mathbf{g})^*$, so the bipolar theorem also gives

$$(K^{*,\mathbf{g}})^* = ((K - \mathbf{g})^*)^* = K - \mathbf{g}, \tag{29}$$

where $K - \mathbf{g} := \{\mathbf{x} - \mathbf{g} : \mathbf{x} \in K\}$ and the outer polar is taken with respect to the origin.

The following two results are Lemmas 1 and 2 of Tatli & Erdogan (2021a;b) which are used in the later proofs.

Lemma B.1 (Linear invariance of an ellipsoid)**.** *Let $\mathcal{E} \subset \mathbb{R}^r$ be a full-dimensional ellipsoid with center $\mathbf{g}$, and let $\mathbf{A} \in \mathbb{R}^{r\times r}$. If $\mathbf{A}\mathcal{E} = \mathcal{E}$, then*

$$\mathbf{A}(\partial\mathcal{E}) = \partial\mathcal{E} \qquad \textit{and} \qquad \mathbf{A}\mathbf{g} = \mathbf{g}. \tag{30}$$

Lemma B.2 (Polar of a linearly transformed set)**.** *Let $K \subset \mathbb{R}^r$, let $\mathbf{d} \in \mathbb{R}^r$, and let a full rank matrix $\mathbf{A} \in \mathbb{R}^{r\times r}$ satisfy $\mathbf{A}\mathbf{d} = \mathbf{d}$. Then*

$$(\mathbf{A}K)^{*,\mathbf{d}} = \mathbf{A}^{-\top} K^{*,\mathbf{d}}. \tag{31}$$

In particular, for the origin-centered polar,

$$(\mathbf{A}K)^* = \mathbf{A}^{-\top} K^*. \tag{32}$$

B.4 Canonical and Heterogeneous Feature Polytopes

According to the PMF formulation, the choice of the latent source polytope determines the specific structural attributes of the recovered representations, enabling different configurations of sparsity, nonnegativity, and amplitude boundedness (Tatli & Erdogan, 2021b). The reference Tatli & Erdogan (2021b) categorizes four canonical domains alongside an example polytope to illustrate the PMF framework's capability for assigning heterogeneous features to subvectors:

- **Antisparse Polytope ($\mathcal{B}_\infty^r$)**: Defined as the unit ℓ_∞-norm ball, $[-1,1]^r$. Enforcing this domain encourages antisparse components, meaning that the latent vectors tend to achieve near-maximum magnitude values simultaneously across all their coordinates. This case has been covered under bounded component analysis (Cruces, 2010; Erdogan, 2013), democratic/dense (Studer et al., 2014) and anti-sparse (Bozkurt et al., 2022) representations.
- **Sparse Polytope ($\mathcal{B}_1^r$)**: Defined as the unit ℓ_1-norm ball. This polytope induces sparsity, restricting the recovered representations to have few non-zero elements.
- **Nonnegative Antisparse Polytope ($\mathcal{B}_{\infty,+}^r$)**: Formed by the intersection of the ℓ_∞-norm ball with the nonnegative orthant, yielding $\mathcal{B}_{\infty,+}^r = [0,1]^r$. It produces components that are strictly nonnegative and antisparse.
- **Nonnegative Sparse Polytope ($\mathcal{B}_{1,+}^r$)**: Formed by the intersection of the ℓ_1-norm ball with the nonnegative orthant. This creates a standard unit simplex augmented with the origin, yielding components that are both nonnegative and sparse.
- **Heterogeneous Feature Polytope ($\mathcal{P}_{\text{ex}}$)**: The PMF framework permits assigning different structural attributes at the subvector level to model heterogeneous features without relying on statistical independence assumptions. For example, a polytope $\mathcal{P}_{\text{ex}} \subset \mathbb{R}^3$ can be defined by the constraints $x_1, x_2 \in [-1,1]$ and $x_3 \in [0,1]$, coupled with overlapping sparsity budgets $\|[x_1,x_2]^\top\|_1 \leq 1$ and $\|[x_2,x_3]^\top\|_1 \leq 1$. In this configuration, x_3 is nonnegative, x_1 and x_2 are signed, and sparsity is enforced locally on the intersecting subvectors. The vertices for this specific mixed-feature polytope are given by the columns of the matrix

$$\mathcal{V}_{\mathcal{P}_{\text{ex}}} = \begin{bmatrix} 1 & -1 & 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}.$$

  These polytope examples are illustrated in Figure 4.

B.5 Sufficiently Scattered Sets

The overdetermined PMF theory replaces vertex-inclusion assumptions in bounded component analysis and sparse component analysis with an MVIE-based sufficient scattering condition (Tatli & Erdogan, 2021b):

Definition B.1 (Sufficiently scattered factor)**.** *A matrix $\mathbf{S} \in \mathbb{R}^{r\times N}$ is called* sufficiently scattered *with respect to $\mathcal{P}$ if*

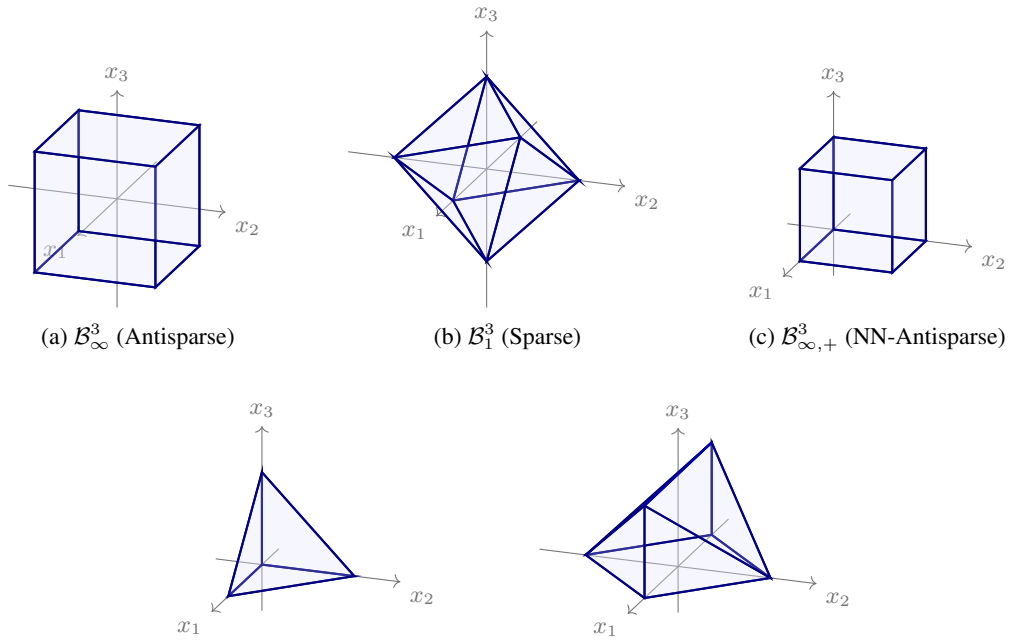
(PMF.SS.i) $\mathcal{P} \supseteq \operatorname{conv}(\mathbf{S}) \supset \mathcal{E}_{\mathcal{P}}$*, and*

(PMF.SS.ii) $\operatorname{conv}(\mathbf{S})^{*,\mathbf{g}_{\mathcal{P}}} \cap \operatorname{bd}(\mathcal{E}_{\mathcal{P}}^{*,\mathbf{g}_{\mathcal{P}}}) = \operatorname{ext}(\mathcal{P}^{*,\mathbf{g}_{\mathcal{P}}})$

where $\mathcal{E}_{\mathcal{P}}$ is the MVIE of $\mathcal{P}$ with the center $\mathbf{g}_{\mathcal{P}}$ and $\operatorname{conv}(\mathbf{S}) = \operatorname{conv}\{\mathbf{s}(1),\ldots,\mathbf{s}(N)\}$. Condition (PMF.SS.i) requires the convex hull of the latent samples to contain the MVIE of the polytope, while (PMF.SS.ii) controls how tightly that convex hull can touch the MVIE of the polytope through the corresponding polar geometry. This is the standard PMF condition that turns determinant optimality into the preservation of the MVIE and of the polytope.

B.6 Linear general position

Identifiability in this work is established from finitely many samples, so we need a condition ensuring that a finite set of vectors is rich enough to reveal the whole subspace containing it, rather than only a lower-dimensional part of it. Linear general position is a convenient tool for this purpose: it



(a) $\mathcal{B}^3_\infty$ (Antisparse) (b) $\mathcal{B}^3_1$ (Sparse) (c) $\mathcal{B}^3_{\infty,+}$ (NN-Antisparse)

(d) $\mathcal{B}^3_{1,+}$ (NN-Sparse) (e) $\mathcal{P}_{\text{ex}}$ (Heterogeneous)

Figure 4: Geometric domains in $\mathbb{R}^3$ for the four canonical PMF polytopes and the heterogeneous feature polytope $\mathcal{P}_{\text{ex}}$.

guarantees that any collection of the vectors as large as the dimension of the space already spans the entire space, so no lower-dimensional subspace can contain more than a few of them. It is a deterministic, geometric condition on the samples rather than a statistical assumption such as independence of the source components, and it holds almost surely under generic random sampling (Proposition B.4). We use the vector-space notion of general position from Cover (1965).

Definition B.2 (Linear general position)**.** *Let V be a p-dimensional vector space. An indexed family of vectors $(\mathbf{t}_k)_{k\in K}$ in V is in* linear general position *if, for every set of p distinct indices $k_1, \ldots, k_p \in K$, the vectors $\mathbf{t}_{k_1}, \ldots, \mathbf{t}_{k_p}$ are linearly independent, or equivalently, span V.*

Geometrically, linear general position means that the vectors point in sufficiently many different directions: no proper linear subspace of V, that is, no line, plane, or higher-dimensional subspace through the origin other than V itself, contains p of them. For instance, in a plane no two of the vectors lie on a common line through the origin, and in three dimensions no three lie on a common plane through the origin.

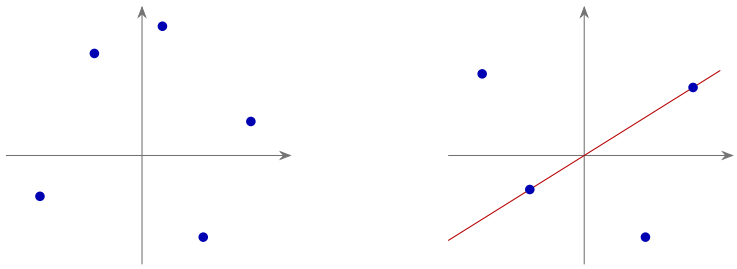
The next proposition makes this precise; it is the formulation used in the finite-sample counting argument.

Proposition B.3 (Equivalent characterizations of linear general position)**.** *Let V be a p-dimensional vector space and let $(\mathbf{t}_k)_{k\in K}$ be an indexed family in V. The following are equivalent:*

(i) The family $(\mathbf{t}_k)_{k\in K}$ is in linear general position.

(ii) Every proper linear subspace of V contains at most $p-1$ members of the family, counted by their indices.

Proof. Suppose the family is in linear general position. If a proper linear subspace contained p members, those p vectors could not span V, which would contradict Definition B.2. Conversely, suppose the family is not in linear general position. Then some p members fail to span V, so their span is a proper linear subspace containing all p members. This proves the equivalence. □



(a) general position in $V = \mathbb{R}^2$ ($p = 2$): no two points on a common line through $\mathbf{0}$

(b) *not* in general position: two points lie on one line through $\mathbf{0}$

Figure 5: Linear general position in a 2-dimensional space. Proper linear subspaces of $\mathbb{R}^2$ are lines through the origin; general position means that each such line contains at most $p - 1 = 1$ member of the indexed family.

Remark B.1 (Linear versus affine general position)**.** *The term "general position" is also used in an affine sense, concerning affine subspaces such as lines that need not pass through the origin. In contrast, linear general position concerns linear subspaces, all of which contain the origin. In Figure 5(b), the four points are in affine general position because no three are collinear, but they are not in linear general position because two lie on the same line through the origin. Since the article is based on the linear mixing model $\mathbf{Y} = \mathbf{H}_g\mathbf{S}_g$ with no offset, linear general position is the relevant notion used in the identifiability arguments.*

Linear general position is a mild requirement: the next proposition shows that it holds almost surely whenever the vectors are drawn independently from distributions that assign zero probability to every proper linear subspace, as does any distribution with a density on V.

Proposition B.4 (Genericity of linear general position)**.** *Let K be finite, and let $(\mathbf{t}_k)_{k\in K}$ be independent random vectors in a p-dimensional vector space V. Suppose that, for every $k \in K$ and every proper linear subspace $U \subsetneq V$,*

$$\mathbb{P}(\mathbf{t}_k \in U) = 0.$$

Then $(\mathbf{t}_k)_{k\in K}$ is in linear general position almost surely.

Proof. Fix any p distinct indices $k_1, \ldots, k_p \in K$. We show successively that the corresponding vectors are linearly independent almost surely.

For $j = 1, \ldots, p$, condition on $\mathbf{t}_{k_1}, \ldots, \mathbf{t}_{k_{j-1}}$, and consider

$$U := \operatorname{span}\{\mathbf{t}_{k_1}, \ldots, \mathbf{t}_{k_{j-1}}\}.$$

For $j = 1$, take $U = \{\mathbf{0}\}$. Since U is generated by at most $j - 1 < p$ vectors, it is a proper linear subspace of V. By independence, the conditional distribution of $\mathbf{t}_{k_j}$ is its original distribution. The hypothesis therefore gives

$$\mathbb{P}\big(\mathbf{t}_{k_j} \in U \mid \mathbf{t}_{k_1}, \ldots, \mathbf{t}_{k_{j-1}}\big) = 0 \quad \text{almost surely.}$$

Thus each successive vector lies outside the span of the preceding vectors almost surely, and these p vectors are linearly independent.

There are only finitely many selections of p distinct indices. Taking a finite union of the corresponding probability-zero events proves the assertion. □

C CorInfoMax and a Det-Max-Consistent Reconstruction Objective

This appendix examines the relationship between correlative information maximization (CorInfoMax) (Erdogan, 2022; Bozkurt et al., 2023) and determinant maximization (Det-Max) (Fu et al., 2018; Tatli & Erdogan, 2021b) in the underdetermined PMF setting. We first characterize the exact mixing-matrix profile of symmetric CorInfoMax and explain why it need not select Det-Max

solutions. We then introduce a weighted source-entropy and reconstruction objective, identify its connection to a reconstruction-form CorInfoMax surrogate, and prove Det-Max consistency for every fixed source-entropy weight $\lambda > 0$ as regularization vanishes.

C.1 The CorInfoMax Objective and Its Optimization Profile

Let $\mathbf{C}_N = \mathbf{I}_N - N^{-1}\mathbf{1}\mathbf{1}^\top$, $\overline{\mathbf{Y}} = \mathbf{Y}\mathbf{C}_N$, and $\overline{\mathbf{S}} = \mathbf{S}\mathbf{C}_N$. Define the centered empirical covariances

$$\widehat{\mathbf{R}}_{\mathbf{S}} = \frac{1}{N}\overline{\mathbf{S}}\overline{\mathbf{S}}^\top, \qquad \widehat{\mathbf{R}}_{\mathbf{Y}} = \frac{1}{N}\overline{\mathbf{Y}}\overline{\mathbf{Y}}^\top, \qquad \widehat{\mathbf{R}}_{\mathbf{YS}} = \frac{1}{N}\overline{\mathbf{Y}}\overline{\mathbf{S}}^\top,$$

with $\widehat{\mathbf{R}}_{\mathbf{SY}} = \widehat{\mathbf{R}}_{\mathbf{YS}}^\top$. By symmetry, the regularized empirical correlative information defined in Section 2 also admits the observation-prediction form

$$\begin{aligned} \widehat{I}_{\mathrm{c}}^{(\epsilon)}(\mathbf{Y},\mathbf{S}) &= c_\epsilon(\mathbf{Y}) - \frac{1}{2}\log\det\left(\widehat{\mathbf{R}}_{\mathbf{Y}|\mathbf{S}}^{(\epsilon)} + \epsilon\mathbf{I}_M\right), \\ c_\epsilon(\mathbf{Y}) &:= \frac{1}{2}\log\det\left(\widehat{\mathbf{R}}_{\mathbf{Y}} + \epsilon\mathbf{I}_M\right), \\ \widehat{\mathbf{R}}_{\mathbf{Y}|\mathbf{S}}^{(\epsilon)} &:= \widehat{\mathbf{R}}_{\mathbf{Y}} - \widehat{\mathbf{R}}_{\mathbf{YS}}(\widehat{\mathbf{R}}_{\mathbf{S}} + \epsilon\mathbf{I}_r)^{-1}\widehat{\mathbf{R}}_{\mathbf{SY}}. \end{aligned} \tag{33}$$

The corresponding exact variational objective is

$$J_\epsilon(\mathbf{H},\mathbf{S}) := c_\epsilon(\mathbf{Y}) - \frac{1}{2}\log\det\left(\widehat{\mathbf{R}}_{\mathbf{E}}(\mathbf{H},\mathbf{S}) + \epsilon\mathbf{H}\mathbf{H}^\top + \epsilon\mathbf{I}_M\right), \tag{34}$$

where

$$\widehat{\mathbf{R}}_{\mathbf{E}}(\mathbf{H},\mathbf{S}) := \frac{1}{N}(\overline{\mathbf{Y}} - \mathbf{H}\overline{\mathbf{S}})(\overline{\mathbf{Y}} - \mathbf{H}\overline{\mathbf{S}})^\top.$$

Indeed, with

$$\widehat{\mathbf{H}}_\epsilon(\mathbf{S}) := \widehat{\mathbf{R}}_{\mathbf{YS}}(\widehat{\mathbf{R}}_{\mathbf{S}} + \epsilon\mathbf{I}_r)^{-1},$$

completion of the square gives

$$\begin{aligned} &\widehat{\mathbf{R}}_{\mathbf{E}}(\mathbf{H},\mathbf{S}) + \epsilon\mathbf{H}\mathbf{H}^\top \\ &\quad = \widehat{\mathbf{R}}_{\mathbf{Y}|\mathbf{S}}^{(\epsilon)} + (\mathbf{H} - \widehat{\mathbf{H}}_\epsilon(\mathbf{S}))(\widehat{\mathbf{R}}_{\mathbf{S}} + \epsilon\mathbf{I}_r)(\mathbf{H} - \widehat{\mathbf{H}}_\epsilon(\mathbf{S}))^\top. \end{aligned} \tag{35}$$

The final term is positive semidefinite. Monotonicity of the log-determinant therefore implies

$$\max_{\mathbf{H}} J_\epsilon(\mathbf{H},\mathbf{S}) = J_\epsilon(\widehat{\mathbf{H}}_\epsilon(\mathbf{S}),\mathbf{S}) = \widehat{I}_{\mathrm{c}}^{(\epsilon)}(\mathbf{Y},\mathbf{S}).$$

Thus, the ridge term $\epsilon\mathbf{H}\mathbf{H}^\top$ is required for this exact finite-ϵ representation of symmetric CorInfoMax.

Removing terms independent of the optimization variables yields the normalized profile

$$\begin{aligned} \Psi_\epsilon(\mathbf{S}) &:= \widehat{I}_{\mathrm{c}}^{(\epsilon)}(\mathbf{Y},\mathbf{S}) - c_\epsilon(\mathbf{Y}) + \frac{M}{2}\log\epsilon \\ &= \max_{\mathbf{H}}\left\{-\frac{1}{2}\log\det\left(\mathbf{I}_M + \mathbf{H}\mathbf{H}^\top + \frac{\widehat{\mathbf{R}}_{\mathbf{E}}(\mathbf{H},\mathbf{S})}{\epsilon}\right)\right\}. \end{aligned} \tag{36}$$

This profile penalizes reconstruction error and dictionary gain. Even at exact reconstruction, however, it retains a dictionary-gain penalty rather than an explicit source-volume objective. Consequently, reconstruction alone does not ensure full-row-rank source estimates or Det-Max optimality.

C.2 The Rank-Deficiency Pitfall of Standard CorInfoMax

The following example illustrates the distinction between reconstructing the observations and maximizing source volume.

Example C.1 (A rank-deficient alternative preferred by CorInfoMax)**.** *Let $M = 1$, $r = 2$, $N = 4$, and $\mathcal{P} = [-1,1]^2$. Consider*

$$\mathbf{S}_g = \begin{bmatrix} 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \end{bmatrix}, \qquad \mathbf{Y} = \begin{bmatrix} 1 & 1 & -1 & -1 \end{bmatrix}.$$

Then $\mathbf{Y} = [1\ \ 0]\mathbf{S}_g$ *and* $\mathbf{R}_{\mathbf{S}_g} = N^{-1}\mathbf{S}_g\mathbf{S}_g^\top = \mathbf{I}_2$. *The alternative*

$$\mathbf{S}_{\text{col}} = \begin{bmatrix} 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & -1 \end{bmatrix}$$

also reconstructs the observations exactly, with $\mathbf{H}_{\text{col}} = [1/2\ \ 1/2]$, *but has* $\det \mathbf{R}_{\mathbf{S}_{\text{col}}} = 0$. *Both source matrices have columns in* $\text{Skel}_0(\mathcal{P})$, *and both have zero sample mean. Direct substitution gives*

$$\widehat{I}_{\text{c}}^{(\epsilon)}(\mathbf{Y}, \mathbf{S}_g) = \frac{1}{2}\log\frac{(1+\epsilon)^2}{\epsilon(2+\epsilon)}, \qquad \widehat{I}_{\text{c}}^{(\epsilon)}(\mathbf{Y}, \mathbf{S}_{\text{col}}) = \frac{1}{2}\log\frac{(1+\epsilon)(2+\epsilon)}{\epsilon(3+\epsilon)}.$$

Therefore,

$$\widehat{I}_{\text{c}}^{(\epsilon)}(\mathbf{Y}, \mathbf{S}_{\text{col}}) - \widehat{I}_{\text{c}}^{(\epsilon)}(\mathbf{Y}, \mathbf{S}_g) = \frac{1}{2}\log\frac{(2+\epsilon)^2}{(1+\epsilon)(3+\epsilon)} > 0 \qquad (\epsilon > 0).$$

For source columns in this cube, Hadamard's inequality gives $\det \mathbf{R}_{\mathbf{S}} \leq 1$; hence $\mathbf{S}_g$ is Det-Max optimal. Nevertheless, symmetric CorInfoMax assigns a strictly larger value to the collapsed alternative. Thus the two criteria do not generally have the same maximizers, even when a full-row-rank exact factorization is feasible. This example distinguishes the objectives; it does not invoke the complete identifiability assumptions of the main theorem.

C.3 A Weighted Det-Max-Consistent Reconstruction Objective

To preserve the source-volume criterion while relaxing exact reconstruction, we use the uncentered second moments appearing in the Det-Max formulation:

$$\mathbf{R}_{\mathbf{S}} := \frac{1}{N}\mathbf{S}\mathbf{S}^\top, \qquad \mathbf{R}_{\mathbf{E}}(\mathbf{H}, \mathbf{S}) := \frac{1}{N}(\mathbf{Y} - \mathbf{H}\mathbf{S})(\mathbf{Y} - \mathbf{H}\mathbf{S})^\top.$$

For $\epsilon > 0$ and a fixed $\lambda > 0$, define the weighted *source-entropy and reconstruction objective*

$$K_{\epsilon,\lambda}(\mathbf{H}, \mathbf{S}) := \frac{\lambda}{2}\log\det(\mathbf{R}_{\mathbf{S}} + \epsilon\mathbf{I}_r) - \frac{1}{2}\log\det(\mathbf{R}_{\mathbf{E}}(\mathbf{H}, \mathbf{S}) + \epsilon\mathbf{I}_M)\,. \tag{37}$$

Its connection to CorInfoMax can be made explicit by introducing the *reconstruction-form CorInfoMax surrogate*

$$J_\epsilon^{\text{rec}}(\mathbf{H}, \mathbf{S}) := \frac{1}{2}\log\det\Big(\widehat{\mathbf{R}}_{\mathbf{Y}} + \epsilon\mathbf{I}_M\Big) - \frac{1}{2}\log\det(\mathbf{R}_{\mathbf{E}}(\mathbf{H}, \mathbf{S}) + \epsilon\mathbf{I}_M)\,. \tag{38}$$

Adding and subtracting the observation term gives the exact identity

$$\begin{aligned} K_{\epsilon,\lambda}(\mathbf{H}, \mathbf{S}) = J_\epsilon^{\text{rec}}(\mathbf{H}, \mathbf{S}) &+ \frac{\lambda}{2}\log\det(\mathbf{R}_{\mathbf{S}} + \epsilon\mathbf{I}_r) \\ &- \frac{1}{2}\log\det\Big(\widehat{\mathbf{R}}_{\mathbf{Y}} + \epsilon\mathbf{I}_M\Big)\,. \end{aligned} \tag{39}$$

Because the input data $\mathbf{Y}$ are fixed, the last term is constant with respect to both $\mathbf{H}$ and $\mathbf{S}$. Maximizing $K_{\epsilon,\lambda}$ is therefore equivalent to maximizing this reconstruction-form CorInfoMax surrogate augmented by the weighted source-entropy term.

We refer to the special case of $\lambda = 1$ as,

$$K_\epsilon(\mathbf{H}, \mathbf{S}) := \frac{1}{2}\log\det(\mathbf{R}_{\mathbf{S}} + \epsilon\mathbf{I}_r) - \frac{1}{2}\log\det(\mathbf{R}_{\mathbf{E}}(\mathbf{H}, \mathbf{S}) + \epsilon\mathbf{I}_M)\,. \tag{40}$$

Distinction from symmetric CorInfoMax. The qualifier "reconstruction-form" is essential: J_ϵ^{rec} uses the uncentered reconstruction residual and omits the ridge term in Equation (34). It is not the exact variational objective J_ϵ of symmetric CorInfoMax. More precisely,

$$\begin{aligned} J_\epsilon^{\text{rec}}(\mathbf{H}, \mathbf{S}) &= J_\epsilon(\mathbf{H}, \mathbf{S}) + \Delta_\epsilon(\mathbf{H}, \mathbf{S}), \\ \Delta_\epsilon(\mathbf{H}, \mathbf{S}) &:= \frac{1}{2}\log\frac{\det\Big(\widehat{\mathbf{R}}_{\mathbf{E}}(\mathbf{H}, \mathbf{S}) + \epsilon\mathbf{H}\mathbf{H}^\top + \epsilon\mathbf{I}_M\Big)}{\det(\mathbf{R}_{\mathbf{E}}(\mathbf{H}, \mathbf{S}) + \epsilon\mathbf{I}_M)}. \end{aligned} \tag{41}$$

Thus, replacing J_ϵ^{rec} by standard J_ϵ in Equation (39) requires the additional Δ_ϵ term. Even when both data matrices have zero sample mean, so that the centered and uncentered residual moments agree, $\Delta_\epsilon \geq 0$ is generally nonzero because of the ridge term. At centered exact reconstruction, $\Delta_\epsilon = \frac{1}{2}\log\det(\mathbf{I}_M + \mathbf{H}\mathbf{H}^\top)$, which need not vanish as $\epsilon \downarrow 0$. No equivalence with unmodified symmetric CorInfoMax is asserted.

Role of the source-entropy weight. The case $\lambda = 1$ recovers the objective in Equation (15). For finite regularization, λ controls the relative emphasis on source spread and reconstruction; it is not merely an overall scaling of the objective. For any exact factorization,

$$K_{\epsilon,\lambda}(\mathbf{H}, \mathbf{S}) + \frac{M}{2} \log \epsilon = \frac{\lambda}{2} \log \det(\mathbf{R}_{\mathbf{S}} + \epsilon \mathbf{I}_r) \,. \tag{42}$$

Consequently, every fixed positive λ preserves the limiting Det-Max ordering among full-row-rank exact factorizations. The result below also controls sequences that are not exactly reconstructing at finite ϵ. Positivity is essential: when $\lambda = 0$, the criterion does not distinguish between exact factorizations with different source volumes.

Least-squares profiling. For fixed $\mathbf{S}$, let

$$\widehat{\mathbf{H}}_0(\mathbf{S}) := \mathbf{Y}\mathbf{S}^\dagger, \qquad \mathbf{R}_{\mathbf{E}}(\mathbf{S}) := \mathbf{R}_{\mathbf{E}}(\widehat{\mathbf{H}}_0(\mathbf{S}), \mathbf{S}),$$

where $\mathbf{S}^\dagger$ is the Moore–Penrose pseudoinverse. The least-squares orthogonality relation

$$(\mathbf{Y} - \widehat{\mathbf{H}}_0(\mathbf{S})\mathbf{S})\mathbf{S}^\top = \mathbf{0}$$

gives

$$\mathbf{R}_{\mathbf{E}}(\mathbf{H}, \mathbf{S}) = \mathbf{R}_{\mathbf{E}}(\mathbf{S}) + (\mathbf{H} - \widehat{\mathbf{H}}_0(\mathbf{S}))\mathbf{R}_{\mathbf{S}}(\mathbf{H} - \widehat{\mathbf{H}}_0(\mathbf{S}))^\top. \tag{43}$$

Hence $\widehat{\mathbf{H}}_0(\mathbf{S})$ maximizes $K_{\epsilon,\lambda}$ over $\mathbf{H}$ for every $\epsilon > 0$ and $\lambda > 0$. The normalized profile is

$$\begin{aligned} \Phi_{\epsilon,\lambda}(\mathbf{S}) &:= \max_{\mathbf{H}} K_{\epsilon,\lambda}(\mathbf{H}, \mathbf{S}) + \frac{M}{2} \log \epsilon \\ &= \frac{\lambda}{2} \log \det(\mathbf{R}_{\mathbf{S}} + \epsilon \mathbf{I}_r) - \frac{1}{2} \log \det\left(\mathbf{I}_M + \frac{\mathbf{R}_{\mathbf{E}}(\mathbf{S})}{\epsilon} \right). \end{aligned} \tag{44}$$

Unlike Equation (36), this profile retains an explicit source-volume term without a dictionary-gain penalty at exact reconstruction.

Lemma C.1 (Det-Max limit of the weighted reconstruction profile)**.** *Let $\mathcal{K} \subset \mathbb{R}^{r \times N}$ be compact, for example $\mathcal{K} = (\mathrm{Skel}_{k_0}(\mathcal{P}))^N$. Assume that an exact feasible pair $\mathbf{Y} = \mathbf{H}_g \mathbf{S}_g$ exists with $\mathbf{S}_g \in \mathcal{K}$ and $\mathbf{R}_{\mathbf{S}_g} \succ \mathbf{0}$. Fix $\lambda > 0$. Let $\epsilon_n \downarrow 0$, $\delta_n \geq 0$ with $\delta_n \to 0$, and suppose*

$$\Phi_{\epsilon_n,\lambda}(\mathbf{S}_n) \geq \sup_{\mathbf{S} \in \mathcal{K}} \Phi_{\epsilon_n,\lambda}(\mathbf{S}) - \delta_n, \qquad \mathbf{S}_n \in \mathcal{K}. \tag{45}$$

Every accumulation point $\mathbf{S}^\star$ has full row rank and is the source factor of a global maximizer of

$$v_{\det} := \sup_{\substack{\mathbf{S} \in \mathcal{K},\, \mathbf{H} \in \mathbb{R}^{M \times r} \\ \mathbf{Y} = \mathbf{H}\mathbf{S}}} \frac{1}{2} \log \det \mathbf{R}_{\mathbf{S}}, \tag{46}$$

where singular source moments are assigned objective value $-\infty$. Moreover,

$$\sup_{\mathbf{S} \in \mathcal{K}} \Phi_{\epsilon_n,\lambda}(\mathbf{S}) \longrightarrow \lambda v_{\det}, \qquad \|\mathbf{R}_{\mathbf{E}}(\mathbf{S}_n)\|_2 = o(\epsilon_n). \tag{47}$$

Along every subsequence with $\mathbf{S}_n \to \mathbf{S}^\star$,

$$\widehat{\mathbf{H}}_0(\mathbf{S}_n) \longrightarrow \mathbf{H}^\star := \mathbf{Y}(\mathbf{S}^\star)^\top \big(\mathbf{S}^\star(\mathbf{S}^\star)^\top\big)^{-1}, \qquad \mathbf{Y} = \mathbf{H}^\star \mathbf{S}^\star.$$

Proof. **Uniform residual and rank bounds.** Compactness gives a constant $C \geq 1$ such that $\lambda_{\max}(\mathbf{R}_{\mathbf{S}}) \leq C$ for all $\mathbf{S} \in \mathcal{K}$. Write

$$\begin{aligned} A_n &:= \frac{\lambda}{2} \log \det(\mathbf{R}_{\mathbf{S}_n} + \epsilon_n \mathbf{I}_r) \,, \\ P_n &:= \frac{1}{2} \log \det\left(\mathbf{I}_M + \frac{\mathbf{R}_{\mathbf{E}}(\mathbf{S}_n)}{\epsilon_n} \right), \qquad a := \frac{\lambda}{2} \log \det \mathbf{R}_{\mathbf{S}_g}. \end{aligned}$$

Here $P_n \geq 0$ and a is finite. Comparing Equation (45) with $\mathbf{S}_g$, whose profiled residual vanishes, yields

$$A_n - P_n \geq \frac{\lambda}{2} \log \det\big(\mathbf{R}_{\mathbf{S}_g} + \epsilon_n \mathbf{I}_r\big) - \delta_n \geq a - \delta_n.$$

For all sufficiently large n, $\epsilon_n \le 1$, $\delta_n \le 1$, and

$$A_n \le B := \frac{\lambda r}{2} \log(C+1).$$

Consequently,

$$0 \le P_n \le B - a + 1 =: D, \qquad \|\mathbf{R}_{\mathbf{E}}(\mathbf{S}_n)\|_2 \le (e^{2D} - 1)\epsilon_n.$$

Also $A_n \ge a - 1$. Since $\lambda > 0$,

$$\det(\mathbf{R}_{\mathbf{S}_n} + \epsilon_n \mathbf{I}_r) \ge \exp\left(\frac{2(a-1)}{\lambda}\right).$$

Bounding the other $r - 1$ eigenvalues by $C + 1$ gives

$$\lambda_{\min}(\mathbf{R}_{\mathbf{S}_n} + \epsilon_n \mathbf{I}_r) \ge \frac{\exp(2(a-1)/\lambda)}{(C+1)^{r-1}} =: c > 0.$$

Thus $\mathbf{R}_{\mathbf{S}_n} \succeq (c/2)\mathbf{I}_r$ for all sufficiently large n.

Exact feasibility of accumulation points. Take any convergent subsequence $\mathbf{S}_n \to \mathbf{S}^\star$. Compactness and the preceding rank bound imply $\mathbf{S}^\star \in \mathcal{K}$ and $\mathbf{R}_{\mathbf{S}^\star} \succ \mathbf{0}$. Continuity of matrix inversion on this uniformly full-row-rank set gives

$$\widehat{\mathbf{H}}_0(\mathbf{S}_n) = \mathbf{Y}\mathbf{S}_n^\top (\mathbf{S}_n \mathbf{S}_n^\top)^{-1} \longrightarrow \mathbf{H}^\star.$$

Furthermore,

$$\|\mathbf{Y} - \widehat{\mathbf{H}}_0(\mathbf{S}_n)\mathbf{S}_n\|_F^2 = N \operatorname{tr} \mathbf{R}_{\mathbf{E}}(\mathbf{S}_n) \longrightarrow 0.$$

Hence $\mathbf{Y} = \mathbf{H}^\star \mathbf{S}^\star$.

Det-Max optimality. For any exact feasible pair $(\mathbf{H}, \mathbf{T})$ with $\mathbf{R}_{\mathbf{T}} \succ \mathbf{0}$, the profiled residual at $\mathbf{T}$ is zero, so

$$A_n \ge \Phi_{\epsilon_n,\lambda}(\mathbf{S}_n) \ge \frac{\lambda}{2} \log\det(\mathbf{R}_{\mathbf{T}} + \epsilon_n \mathbf{I}_r) - \delta_n.$$

Taking the limit along the chosen subsequence and dividing by the fixed positive λ yields

$$\frac{1}{2} \log\det \mathbf{R}_{\mathbf{S}^\star} \ge \frac{1}{2} \log\det \mathbf{R}_{\mathbf{T}}.$$

Singular competitors have value $-\infty$. Thus $(\mathbf{H}^\star, \mathbf{S}^\star)$ is a global Det-Max maximizer, and the supremum in Equation (46) is attained.

Optimal-value convergence and the sharper residual bound. Every source accumulation point is Det-Max optimal, and the source moments are uniformly positive definite for all large n. Compactness and continuity therefore give $A_n \to \lambda v_{\det}$ for the entire sequence. Fix an optimal source matrix $\mathbf{S}^\star$ and set

$$w_n := \sup_{\mathbf{S} \in \mathcal{K}} \Phi_{\epsilon_n,\lambda}(\mathbf{S}).$$

Then

$$\frac{\lambda}{2} \log\det(\mathbf{R}_{\mathbf{S}^\star} + \epsilon_n \mathbf{I}_r) \le w_n \le \Phi_{\epsilon_n,\lambda}(\mathbf{S}_n) + \delta_n \le A_n + \delta_n.$$

Both endpoints tend to $\lambda v_{\det}$. Hence $w_n \to \lambda v_{\det}$ and, by near-optimality, $\Phi_{\epsilon_n,\lambda}(\mathbf{S}_n) \to \lambda v_{\det}$. It follows that

$$P_n = A_n - \Phi_{\epsilon_n,\lambda}(\mathbf{S}_n) \longrightarrow 0.$$

Finally, positive semidefiniteness of the residual moment gives

$$0 \le \frac{\|\mathbf{R}_{\mathbf{E}}(\mathbf{S}_n)\|_2}{\epsilon_n} \le e^{2P_n} - 1 \longrightarrow 0,$$

which proves Equation (47). □

Scope of the conclusion. The weight λ changes the finite-regularization trade-off but not the limiting Det-Max solution set when it is held fixed and strictly positive. Under the identifiability assumptions of the main theorem, the limiting factors therefore have the stated polytope-automorphism ambiguity. Approximate global maximizers of the joint objective inherit the source conclusion, since profiling can only improve their objective value. These statements concern vanishing regularization and vanishing global optimization error; they do not establish convergence of finite-regularization stationary points, batch iterations, or online neural dynamics.

D GENERICITY OF Q-GP

Proposition D.1 (Genericity of q-GP)**.** *Let $\mathcal{L} = \{L_a\}_{a=1}^{J}$ be a finite subspace arrangement in $\mathbb{R}^r$. If $d_q(\mathcal{L}) \le M$, and if $\mathbf{H}_g \in \mathbb{R}^{M\times r}$ is drawn from a distribution with a density with respect to Lebesgue measure on $\mathbb{R}^{M\times r}$, then $\mathbf{H}_g$ satisfies q-GP almost surely.*

Proof of Proposition D.1. For each subset $I \subseteq [J]$ with $|I| \le q$, define the local sum subspace

$$\mathcal{U}_I := \sum_{a\in I} L_a, \qquad d_I := \dim(\mathcal{U}_I) = \rho(I).$$

By the definition of $d_q(\mathcal{L})$,

$$d_I \le d_q(\mathcal{L}) \le M.$$

Choose a basis $\mathbf{u}_1, \dots, \mathbf{u}_{d_I}$ of $\mathcal{U}_I$, and collect these basis vectors in the matrix

$$U_I = [\mathbf{u}_1 \ \cdots \ \mathbf{u}_{d_I}] \in \mathbb{R}^{r\times d_I}.$$

$\mathbf{H}_g|_{\mathcal{U}_I}$ is injective if and only if the vectors

$$\mathbf{H}_g\mathbf{u}_1, \dots, \mathbf{H}_g\mathbf{u}_{d_I}$$

are linearly independent in $\mathbb{R}^M$. Equivalently,

$$\operatorname{rank}(\mathbf{H}_g U_I) = d_I.$$

Let

$$\mathcal{G}_I := \left\{\mathbf{H} \in \mathbb{R}^{M\times r} : \operatorname{rank}(\mathbf{H}U_I) < d_I\right\}$$

be the failure set for this fixed I. We will first show that $\mathcal{G}_I$ has Lebesgue measure zero.

Since $d_I \le M$, we can choose a linear map $\widetilde{\mathbf{H}} : \mathbb{R}^r \to \mathbb{R}^M$ that sends a basis $\mathbf{u}_1, \dots, \mathbf{u}_{d_I}$ of $\mathcal{U}_I$ to d_I linearly independent vectors in $\mathbb{R}^M$. Its values on a complementary subspace of $\mathcal{U}_I$ can be chosen arbitrarily. Hence $\widetilde{\mathbf{H}}$ is injective on $\mathcal{U}_I$. For this choice,

$$\operatorname{rank}(\widetilde{\mathbf{H}}U_I) = d_I.$$

Therefore at least one $d_I \times d_I$ minor of $\mathbf{H}U_I$, viewed as a polynomial in the entries of $\mathbf{H}$, is not the zero polynomial. Denote such a minor by $\Delta_I(\mathbf{H})$.

If $\mathbf{H} \in \mathcal{G}_I$, then $\operatorname{rank}(\mathbf{H}U_I) < d_I$, so every $d_I \times d_I$ minor of $\mathbf{H}U_I$ vanishes. In particular,

$$\mathcal{G}_I \subseteq \left\{\mathbf{H} \in \mathbb{R}^{M\times r} : \Delta_I(\mathbf{H}) = 0\right\}.$$

The set on the right is the zero set of a nonzero polynomial, and hence has Lebesgue measure zero. Therefore

$$\lambda(\mathcal{G}_I) = 0,$$

where λ denotes Lebesgue measure on $\mathbb{R}^{M\times r}$.

Now take the union over all local subsets:

$$\mathcal{G} := \bigcup_{\substack{I\subseteq[J]\\ 1\le|I|\le q}} \mathcal{G}_I.$$

There are only finitely many such subsets I because J is finite. Hence by subadditivity

$$0 \le \lambda(\mathcal{G}) = \lambda\Big(\bigcup_{\substack{I\subseteq[J]\\ 1\le|I|\le q}} \mathcal{G}_I\Big) \le \sum_{\substack{I\subseteq[J]\\ 1\le|I|\le q}} \lambda(\mathcal{G}_I) = 0.$$

Thus the total failure set $\mathcal{G}$ has Lebesgue measure zero.

Finally, since $\mathbf{H}_g$ has a density f with respect to Lebesgue measure, its distribution is absolutely continuous with respect to Lebesgue measure. Therefore, for any Lebesgue-measure-zero set A,

$$\mathbb{P}(\mathbf{H}_g \in A) = \int_A f(\mathbf{H})\, d\lambda(\mathbf{H}) = 0.$$

Applying this with $A = \mathcal{G}$, we obtain

$$\mathbb{P}(\mathbf{H}_g \in \mathcal{G}) = 0.$$

Equivalently,

$$\mathbb{P}(\mathbf{H}_g \notin \mathcal{G}) = 1.$$

Being outside $\mathcal{G}$ means that for every $I \subseteq [J]$ with $1 \le |I| \le q$,

$$\operatorname{rank}(\mathbf{H}_g U_I) = d_I,$$

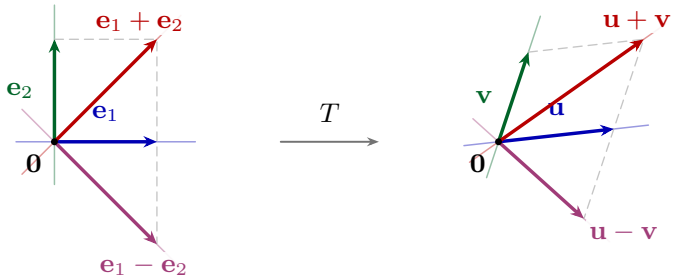
or equivalently $\mathbf{H}_g|_{\mathcal{U}_I}$ is injective almost surely. □

E Discussion of the Identifiability Conditions

This appendix illustrates the conditions of Section 3.3 and relates them to existing results.

Relation completeness. Figure 6 shows a small arrangement and why preserving relations among few spans forces the local maps to agree with one common linear map. For one-dimensional spans, generation of relations by relations involving few spans is studied in Le & Mohammadi (2015); the block-sum formulation in Definition F.2 and the linear extension criterion in Proposition F.1 treat higher-dimensional face spans.

(a) Four spans, three-span relations (b) One common linear map



$q = 3 < J = 4$ $T(x_1, x_2) = x_1\mathbf{u} + x_2\mathbf{v}$

Figure 6: Relations among face spans and a common linear map. (a) Let L_1, L_2, L_3, L_4 be the lines generated by $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_1 + \mathbf{e}_2$, and $\mathbf{e}_1 - \mathbf{e}_2$, respectively. Every relation $a\mathbf{e}_1 + b\mathbf{e}_2 + c(\mathbf{e}_1 + \mathbf{e}_2) + d(\mathbf{e}_1 - \mathbf{e}_2) = \mathbf{0}$ satisfies $a = -c - d$ and $b = -c + d$. It is therefore the sum of c times $-\mathbf{e}_1 - \mathbf{e}_2 + (\mathbf{e}_1 + \mathbf{e}_2) = \mathbf{0}$ and d times $-\mathbf{e}_1 + \mathbf{e}_2 + (\mathbf{e}_1 - \mathbf{e}_2) = \mathbf{0}$. Adding the contributions in each span recovers the original relation. Each of these two relations involves only three spans, so the four spans are 3-relation-complete. (b) Suppose linear maps T_a on the individual lines preserve these two relations. Writing $\mathbf{u} = T_1\mathbf{e}_1$ and $\mathbf{v} = T_2\mathbf{e}_2$, preservation forces $T_3(\mathbf{e}_1 + \mathbf{e}_2) = \mathbf{u} + \mathbf{v}$ and $T_4(\mathbf{e}_1 - \mathbf{e}_2) = \mathbf{u} - \mathbf{v}$. Thus all four maps agree with the single linear map $T(x_1, x_2) = x_1\mathbf{u} + x_2\mathbf{v}$. The image of each diagonal generator is determined by the images of the two coordinate generators.

Local injectivity (q-GP). The budget $d_q(\mathcal{L}) \leq M$ allows every sum of at most q spans to fit in $\mathbb{R}^M$, and Assumption 3.2 requires that the true mixing matrix loses no information on any of them. For the p-sparse model in the underdetermined setting, $q = 2$ gives the usual condition $\operatorname{spark}(\mathbf{H}_g) > 2p$ and generic budget $2p \leq M$ (Donoho & Elad, 2003). Injectivity on pairwise sums also characterizes uniqueness for a known linear map on a union of subspaces (Nam et al., 2013, Proposition 2). Here, sums of up to q spans additionally preserve the relations needed when the mixing map is unknown. The budget is sufficient, not claimed optimal.

Cell population and the sample budget. Assumption 3.3 requires at least $J(p-1)+1$ samples in linear general position in each cell. The factor J comes from comparing one true image subspace with all J candidate image subspaces: each proper intersection contains at most $p-1$ designated samples. Requiring this population in every true cell gives the total requirement $N \geq N_{\text{pop}} := J[J(p-1)+1]$, since each sample has one assigned cell label. This uses the general-position and counting arguments from dictionary identifiability (Hillar & Sommer, 2015; Garfinkle & Hillar, 2019);for coordinate-sparse models, $J = \binom{r}{p}$. The total count alone does not ensure the required per-cell distribution, general position, or sufficient scattering. Linear general position is a geometric condition, not statistical independence of source components; Proposition B.4 gives a sufficient sampling construction.

Sufficient scattering. Assumption 3.4 is the PMF condition of Section B.5 applied to $\mathbf{S}_g$: the sample convex hull must contain the MVIE and satisfy the polar contact condition. It uses the bounded polytope geometry to reduce the remaining transformations to polytope symmetries through Det-Max. General position alone does not ensure it.

F Proof of The Identifiability Theorem

Proof overview. The preliminary results establish the linear extension criterion and show that retaining the selected face spans preserves the skeleton. The identifiability argument then proceeds as follows.

1. **Match the cell images.** The population and general-position assumptions prevent the samples from one true cell image from being covered by proper intersections with the candidate cell images. This gives a matching between true and candidate image subspaces (Lemmas F.3 and F.4).
2. **Obtain one transformation for all sources.** The matching transfers local injectivity to the candidate mixing matrix (Lemma F.5). The resulting maps on individual face spans preserve relations involving at most q spans. Relation completeness and Proposition F.1 therefore extend them to one invertible matrix $\mathbf{B}$ satisfying $\mathbf{HB} = \mathbf{H}_g$ (Lemma F.6). A separate application of injectivity on two-span sums then gives $\mathbf{S} = \mathbf{BS}_g$ for every source column, completing Theorem 3.1.
3. **Restrict the transformation to a polytope symmetry.** Scattering and feasibility place the transformed MVIE inside $\mathcal{P}$, giving $|\det \mathbf{B}| \leq 1$. The polar scattering condition makes equality possible only for $\mathbf{B} \in \mathcal{A}_{\mathcal{P}}$ (Lemma 3.2). Since the true factorization is feasible, every Det-Max maximizer attains equality. This proves Theorem 3.3; its regularized-objective conclusion follows from Lemma C.1.

F.1 Preliminary Work

We record the block-sum notation used in the proof and the formal version of Definition 3.1.

Definition F.1 (Block-sum relation space)**.** *For an arrangement $\mathcal{L} = \{L_1, \ldots, L_J\}$, define*

$$\mathcal{E}_{\mathcal{L}} := \bigoplus_{a=1}^{J} L_a = \{Z = (\mathbf{z}_1, \ldots, \mathbf{z}_J) : \mathbf{z}_a \in L_a \text{ for every } a \in \{1, \ldots, J\}\}. \tag{48}$$

Thus an element $Z \in \mathcal{E}_{\mathcal{L}}$ consists of one labeled vector component $\mathbf{z}_a \in L_a$ for each cell L_a. Define the block-sum map by

$$\Sigma_{\mathcal{L}} : \mathcal{E}_{\mathcal{L}} \to \mathbb{R}^r, \qquad \Sigma_{\mathcal{L}}(Z) = \sum_{a=1}^{J} \mathbf{z}_a, \tag{49}$$

and define the block support of Z by

$$\operatorname{bsupp}(Z = (\mathbf{z}_1, \dots, \mathbf{z}_J)) := \{a \in [J] : \mathbf{z}_a \neq \mathbf{0}\}. \tag{50}$$

Thus $\operatorname{bsupp}(Z)$ *is an index set: it selects exactly those indices $a \in [J]$ for which the corresponding vector component $\mathbf{z}_a$ of $Z = (\mathbf{z}_1, \dots, \mathbf{z}_J)$ is nonzero. We define the block-sum relation space as*

$$\ker \Sigma_{\mathcal{L}} = \left\{ (\mathbf{z}_1, \dots, \mathbf{z}_J) \in \mathcal{E}_{\mathcal{L}} : \sum_{a=1}^{J} \mathbf{z}_a = \mathbf{0} \right\}. \tag{51}$$

The block index records the cell from which a vector comes. Thus a vector in $L_a \cap L_b$ can occupy either block, while $\Sigma_{\mathcal{L}}$ forgets the labels and adds the vectors.

Definition F.2 (q-relation-completeness)**.** *For an integer $q \geq 2$, the arrangement $\mathcal{L}$ is q-relation-complete if*

$$\ker \Sigma_{\mathcal{L}} = \operatorname{span} \{ Z \in \ker \Sigma_{\mathcal{L}} : |\operatorname{bsupp}(Z)| \leq q \}. \tag{52}$$

Equivalently, every block-sum relation can be written as a finite linear combination of block-sum relations, each involving at most q nonzero blocks. This does not require every element of $\ker \Sigma_{\mathcal{L}}$ *itself to have block support of size at most q.*

Generation of relation spaces by small-support relations is established in arrangement theory. For one-dimensional $L_a = \operatorname{span}(\mathbf{v}_a)$, write $\mathbf{z}_a = c_a \mathbf{v}_a$ and associate the hyperplane $\ker(\mathbf{v}_a^\top)$ to each span. For $3 \leq q \leq J$, the condition above is exactly $(q-1)$-generation (Brandt & Terao, 1994, Definitions 3.2–3.3); see also Le & Mohammadi (2015, Definition 2.1). In particular, $q = 3$ corresponds to formality, also called 2-formality. For higher-dimensional spans, support counts nonzero blocks.

The block-sum formulation describes relations directly through vectors in the face spans. An equivalent matrix form is obtained by choosing a basis matrix $\mathbf{U}_a$ for each L_a and setting $\mathbf{U} = [\, \mathbf{U}_1 \ \cdots \ \mathbf{U}_J \,]$. Writing $\mathbf{z}_a = \mathbf{U}_a \mathbf{c}_a$, we have

$$\sum_{a=1}^{J} \mathbf{z}_a = \mathbf{0} \quad \Longleftrightarrow \quad \mathbf{U}\mathbf{c} = \mathbf{0}, \qquad \mathbf{c} = (\mathbf{c}_1^\top, \dots, \mathbf{c}_J^\top)^\top.$$

Thus q-relation completeness means that $\ker \mathbf{U}$ is spanned by its vectors having at most q nonzero coefficient blocks. The count concerns blocks, not individual coefficients, and is independent of the bases chosen within the spans. We give the geometric formulation to keep the connection with the face spans explicit. The following elementary linear-algebra criterion explains the role of this condition in our proof.

Proposition F.1 (Linear extension criterion)**.** *Let $\mathcal{L} = \{L_a\}_{a=1}^{J}$ be a finite arrangement of linear subspaces of $\mathbb{R}^r$, let $q \geq 2$ be an integer, and set $V := \sum_{a=1}^{J} L_a$. The following statements are equivalent.*

(i) The arrangement $\mathcal{L}$ is q-relation-complete.

(ii) For every real vector space W and every family of linear maps $T_a : L_a \to W$ satisfying

$$\sum_{a \in I} \mathbf{z}_a = \mathbf{0} \quad \Longrightarrow \quad \sum_{a \in I} T_a \mathbf{z}_a = \mathbf{0} \tag{53}$$

for every $I \subseteq [J]$ with $1 \leq |I| \leq q$ and every choice of vectors $\mathbf{z}_a \in L_a$, $a \in I$, there exists a unique linear map $T : V \to W$ such that

$$T|_{L_a} = T_a \qquad \text{for every } a \in [J].$$

Proof. We use the block-sum notation introduced in Definition F.1:

$$\begin{aligned} \mathcal{E}_{\mathcal{L}} &:= \bigoplus_{a=1}^{J} L_a = \{ Z = (\mathbf{z}_1, \dots, \mathbf{z}_J) : \mathbf{z}_a \in L_a \text{ for every } a \in [J] \}, \\ \Sigma_{\mathcal{L}} &: \mathcal{E}_{\mathcal{L}} \to \mathbb{R}^r, \qquad \Sigma_{\mathcal{L}}(Z) := \sum_{a=1}^{J} \mathbf{z}_a. \end{aligned} \tag{54}$$

Recall that

$$V := \sum_{a=1}^{J} L_a.$$

Thus the image of $\Sigma_{\mathcal{L}}$ is V.

Suppose first that $\mathcal{L}$ is q-relation-complete, and let $T_a : L_a \to W$ satisfy Equation (53).

Construction and properties of the blockwise map $\widetilde{T}$. Define

$$\widetilde{T} : \mathcal{E}_{\mathcal{L}} \to W, \qquad \widetilde{T}(Z) := \sum_{a=1}^{J} T_a \mathbf{z}_a. \tag{55}$$

This map is well-defined and linear because each $T_a : L_a \to W$ is linear.

Let $Z^\star \in \ker \Sigma_{\mathcal{L}}$ be arbitrary. By q-relation-completeness, $\ker \Sigma_{\mathcal{L}}$ is spanned by relations supported on at most q blocks. Consequently, there exist relations

$$Z^{(1)}, \dots, Z^{(t)} \in \ker \Sigma_{\mathcal{L}}, \qquad |\operatorname{bsupp}(Z^{(\ell)})| \le q,$$

and scalars $\alpha_1, \dots, \alpha_t$ such that

$$Z^\star = \sum_{\ell=1}^{t} \alpha_\ell Z^{(\ell)}. \tag{56}$$

The case $Z^\star = \mathbf{0}$ is immediate, so we may discard any zero relations from this representation.

Fix $\ell \in [t]$, write

$$Z^{(\ell)} = (\mathbf{z}_1^{(\ell)}, \dots, \mathbf{z}_J^{(\ell)}),$$

and set

$$I_\ell := \operatorname{bsupp}(Z^{(\ell)}).$$

Then $1 \le |I_\ell| \le q$, and $\mathbf{z}_a^{(\ell)} = \mathbf{0}$ whenever $a \notin I_\ell$. Since $Z^{(\ell)} \in \ker \Sigma_{\mathcal{L}}$, we have

$$\sum_{a \in I_\ell} \mathbf{z}_a^{(\ell)} = \mathbf{0}. \tag{57}$$

Applying Equation (53) gives

$$\sum_{a \in I_\ell} T_a \mathbf{z}_a^{(\ell)} = \mathbf{0}. \tag{58}$$

Therefore,

$$\widetilde{T}(Z^{(\ell)}) = \sum_{a=1}^{J} T_a \mathbf{z}_a^{(\ell)} = \sum_{a \in I_\ell} T_a \mathbf{z}_a^{(\ell)} = \mathbf{0}. \tag{59}$$

Because ℓ was arbitrary, linearity yields

$$\widetilde{T}(Z^\star) = \sum_{\ell=1}^{t} \alpha_\ell \widetilde{T}(Z^{(\ell)}) = \mathbf{0}. \tag{60}$$

Thus,

$$\ker \Sigma_{\mathcal{L}} \subseteq \ker \widetilde{T}. \tag{61}$$

Construction and properties of the global map T. Every $\mathbf{x} \in V$ can be represented as

$$\mathbf{x} = \Sigma_{\mathcal{L}}(Z) = \sum_{a=1}^{J} \mathbf{z}_a \tag{62}$$

for some $Z = (\mathbf{z}_1, \dots, \mathbf{z}_J) \in \mathcal{E}_{\mathcal{L}}$. Choose such a representation and define

$$T(\mathbf{x}) := \widetilde{T}(Z). \tag{63}$$

Now choose another $Z' \in \mathcal{E}_{\mathcal{L}}$ satisfying

$$\Sigma_{\mathcal{L}}(Z') = \Sigma_{\mathcal{L}}(Z) = \mathbf{x}. \tag{64}$$

Since $\Sigma_{\mathcal{L}}$ is linear,

$$\Sigma_{\mathcal{L}}(Z' - Z) = \mathbf{0},$$

and hence

$$Z' - Z \in \ker \Sigma_{\mathcal{L}}.$$

By Equation (61),

$$Z' - Z \in \ker \widetilde{T}.$$

Therefore,

$$\widetilde{T}(Z') - \widetilde{T}(Z) = \widetilde{T}(Z' - Z) = \mathbf{0}. \tag{65}$$

Thus the value assigned to $T(\mathbf{x})$ does not depend on the chosen representation of $\mathbf{x}$. Hence the definition gives a well-defined map

$$T : V \to W.$$

Moreover, let

$$\mathbf{x} = \Sigma_{\mathcal{L}}(Z), \qquad \mathbf{y} = \Sigma_{\mathcal{L}}(Z')$$

be arbitrary vectors in V, and let $\alpha, \beta \in \mathbb{R}$. Using the linearity of $\Sigma_{\mathcal{L}}$ and $\widetilde{T}$, we obtain

$$T(\alpha\mathbf{x} + \beta\mathbf{y}) = T\big(\Sigma_{\mathcal{L}}(\alpha Z + \beta Z')\big) \tag{66}$$

$$= \widetilde{T}(\alpha Z + \beta Z') \tag{67}$$

$$= \alpha\widetilde{T}(Z) + \beta\widetilde{T}(Z') \tag{68}$$

$$= \alpha T(\mathbf{x}) + \beta T(\mathbf{y}). \tag{69}$$

Therefore, T is linear.

To identify the restriction of T to a single face span L_a, fix $a \in [J]$ and let $\mathbf{x}_a \in L_a$. Consider the block vector

$$Z = (\mathbf{0}, \ldots, \mathbf{0}, \mathbf{x}_a, \mathbf{0}, \ldots, \mathbf{0}) \in \mathcal{E}_{\mathcal{L}}, \tag{70}$$

whose a-th component is $\mathbf{x}_a$. Then

$$\Sigma_{\mathcal{L}}(Z) = \mathbf{x}_a, \qquad \widetilde{T}(Z) = T_a\mathbf{x}_a.$$

Consequently,

$$T(\mathbf{x}_a) = \widetilde{T}(Z) = T_a\mathbf{x}_a. \tag{71}$$

Since this holds for every $\mathbf{x}_a \in L_a$,

$$T|_{L_a} = T_a.$$

Finally, suppose that $T' : V \to W$ is another linear map with the same restrictions. For any $\mathbf{x} \in V$, choose a representation $\mathbf{x} = \sum_{a=1}^{J} \mathbf{z}_a$, where $\mathbf{z}_a \in L_a$. Then

$$T'(\mathbf{x}) = \sum_{a=1}^{J} T'(\mathbf{z}_a) = \sum_{a=1}^{J} T_a\mathbf{z}_a = T(\mathbf{x}). \tag{72}$$

Thus $T' = T$, proving uniqueness and establishing (ii).

Proof of the converse. Suppose now that (ii) holds. Define

$$\mathcal{R}_q := \operatorname{span}\left\{Z \in \ker \Sigma_{\mathcal{L}} : |\operatorname{bsupp}(Z)| \le q\right\}. \tag{73}$$

By definition,

$$\mathcal{R}_q \subseteq \ker \Sigma_{\mathcal{L}}.$$

We must show that the reverse inclusion also holds.

Assume for a contradiction, that there exists

$$Z^\star = (\mathbf{z}_1^\star, \ldots, \mathbf{z}_J^\star) \in \ker \Sigma_{\mathcal{L}} \setminus \mathcal{R}_q.$$

Because $\mathcal{E}_{\mathcal{L}}$ is finite-dimensional, there exists a linear map

$$\widetilde{T} : \mathcal{E}_{\mathcal{L}} \to \mathbb{R}$$

such that

$$\widetilde{T}(Z) = 0 \quad \text{for every } Z \in \mathcal{R}_q, \qquad \widetilde{T}(Z^\star) = 1. \tag{74}$$

To construct this map explicitly, let $U_1, \dots, U_m$ be a basis of $\mathcal{R}_q$. Since $Z^\star \notin \mathcal{R}_q$, it cannot be written as a linear combination of these basis vectors. Thus adjoining $Z^\star$ gives a linearly independent family. We can add further vectors $V_1, \dots, V_s$, if needed, to obtain a basis

$$U_1, \dots, U_m,\ Z^\star,\ V_1, \dots, V_s$$

of the whole space $\mathcal{E}_\mathcal{L}$. Every $Z \in \mathcal{E}_\mathcal{L}$ therefore has a unique representation

$$Z = \sum_{i=1}^{m} \alpha_i U_i + \beta Z^\star + \sum_{j=1}^{s} \gamma_j V_j.$$

Define $\widetilde{T}(Z) := \beta$; that is, the map simply returns the coefficient of $Z^\star$ in this basis. Uniqueness of the representation makes this definition well-defined, and taking a coordinate in a fixed basis is linear. If $Z \in \mathcal{R}_q$, its representation uses only the vectors $U_1, \dots, U_m$, so $\beta = 0$ and $\widetilde{T}(Z) = 0$. For $Z = Z^\star$, the coefficient β is one, so $\widetilde{T}(Z^\star) = 1$. So we can assume such map Equation (74).

For each $a \in [J]$, define

$$T_a : L_a \to \mathbb{R}, \qquad T_a \mathbf{x}_a := \widetilde{T}(\mathbf{0}, \dots, \mathbf{0}, \mathbf{x}_a, \mathbf{0}, \dots, \mathbf{0}), \tag{75}$$

where $\mathbf{x}_a$ occupies the a-th block. Each T_a is linear. Moreover, by decomposing a block vector into its individual blocks, we have

$$\widetilde{T}(Z) = \sum_{a=1}^{J} T_a \mathbf{z}_a \qquad \text{for every } Z = (\mathbf{z}_a)_{a=1}^{J} \in \mathcal{E}_\mathcal{L}. \tag{76}$$

We now verify Equation (53). Let $I \subseteq [J]$ satisfy $1 \le |I| \le q$, and choose $\mathbf{z}_a \in L_a$, $a \in I$, such that

$$\sum_{a \in I} \mathbf{z}_a = \mathbf{0}.$$

Form the block vector Z with these components in the blocks indexed by I and zero components elsewhere. Then

$$Z \in \ker \Sigma_\mathcal{L}, \qquad |\operatorname{bsupp}(Z)| \le q,$$

so $Z \in \mathcal{R}_q$. By Equation (74) and Equation (76),

$$\sum_{a \in I} T_a \mathbf{z}_a = \widetilde{T}(Z) = 0. \tag{77}$$

Thus the maps T_a satisfy Equation (53).

Applying (ii) with $W = \mathbb{R}$, there exists a linear map $T : V \to \mathbb{R}$ whose restriction to each L_a is T_a. Since $Z^\star \in \ker \Sigma_\mathcal{L}$,

$$\sum_{a=1}^{J} \mathbf{z}_a^\star = \mathbf{0}.$$

Consequently,

$$\begin{aligned} 1 &= \widetilde{T}(Z^\star) && (78)\\ &= \sum_{a=1}^{J} T_a \mathbf{z}_a^\star && (79)\\ &= \sum_{a=1}^{J} T(\mathbf{z}_a^\star) && (80)\\ &= T\left(\sum_{a=1}^{J} \mathbf{z}_a^\star \right) && (81)\\ &= T(\mathbf{0}) = 0, && (82) \end{aligned}$$

a contradiction. Therefore,

$$\ker \Sigma_{\mathcal{L}} \subseteq \mathcal{R}_q.$$

Combining the two inclusions gives

$$\ker \Sigma_{\mathcal{L}} = \mathcal{R}_q.$$

Hence $\mathcal{L}$ is q-relation-complete. □

We next state a proposition regarding the faces of a polytope which are containing the origin in their affine hull.

Proposition F.2 (Coverage by (k_0+1)-dimensional face spans)**.** *Let $\mathcal{P} \subset \mathbb{R}^r$ be a full-dimensional polytope, and let $0 \le k_0 < r$. Denote the collection of all k_0-dimensional faces of $\mathcal{P}$ by $\mathcal{F}_{k_0}(\mathcal{P})$.*

Every $G \in \mathcal{F}_{k_0}(\mathcal{P})$ satisfying

$$\mathbf{0} \notin \operatorname{aff}(G) \tag{83}$$

has

$$\dim \operatorname{span}(G) = k_0 + 1. \tag{84}$$

Moreover, for every $F \in \mathcal{F}_{k_0}(\mathcal{P})$, there exists $G \in \mathcal{F}_{k_0}(\mathcal{P})$ satisfying

$$\mathbf{0} \notin \operatorname{aff}(G) \qquad \textit{and} \qquad \operatorname{span}(F) \subseteq \operatorname{span}(G). \tag{85}$$

Consequently,

$$\bigcup_{F \in \mathcal{F}_{k_0}(\mathcal{P})} \big(\mathcal{P} \cap \operatorname{span}(F)\big) = \bigcup_{\substack{G \in \mathcal{F}_{k_0}(\mathcal{P}) \\ \mathbf{0} \notin \operatorname{aff}(G)}} \big(\mathcal{P} \cap \operatorname{span}(G)\big). \tag{86}$$

Proof. **Dimension of the face spans.** We first show that a face has one additional dimension in its linear span exactly when its affine hull does not contain the origin.

Let Q be a k-dimensional face of $\mathcal{P}$, fix $\mathbf{x}_0 \in Q$, and let

$$D_Q := \operatorname{span}(Q - Q)$$

be the direction space of its affine hull. Then

$$\dim D_Q = k, \qquad \operatorname{aff}(Q) = \mathbf{x}_0 + D_Q,$$

and

$$\operatorname{span}(Q) = D_Q + \operatorname{span}\{\mathbf{x}_0\}. \tag{87}$$

If $\mathbf{0} \in \operatorname{aff}(Q)$, then $\mathbf{x}_0 \in D_Q$, and therefore

$$\operatorname{span}(Q) = D_Q, \qquad \dim \operatorname{span}(Q) = k.$$

If $\mathbf{0} \notin \operatorname{aff}(Q)$, then $\mathbf{x}_0 \notin D_Q$. Thus adding $\mathbf{x}_0$ to D_Q increases the dimension by one, and hence

$$\dim \operatorname{span}(Q) = k + 1.$$

We have therefore established the general identity

$$\dim \operatorname{span}(Q) = \begin{cases} \dim Q, & \mathbf{0} \in \operatorname{aff}(Q), \\ \dim Q + 1, & \mathbf{0} \notin \operatorname{aff}(Q). \end{cases} \tag{88}$$

Applying this identity to a k_0-dimensional face G satisfying $\mathbf{0} \notin \operatorname{aff}(G)$ proves Equation (84).

Coverage of the omitted face spans. We next prove Equation (85). Let $F \in \mathcal{F}_{k_0}(\mathcal{P})$ be arbitrary. If

$$\mathbf{0} \notin \operatorname{aff}(F),$$

we may simply take $G = F$. It remains to consider a k_0-dimensional face F satisfying

$$\mathbf{0} \in \operatorname{aff}(F). \tag{89}$$

Suppose first that $k_0 = 0$. Since F is zero-dimensional, $F = \{\mathbf{v}\}$ for some vertex $\mathbf{v}$ of $\mathcal{P}$. Condition Equation (89) forces $\mathbf{v} = \mathbf{0}$, and therefore

$$\operatorname{span}(F) = \{\mathbf{0}\}.$$

Since $\mathcal{P}$ is full-dimensional and is the convex hull of its vertices, it has a nonzero vertex $\mathbf{w}$. Taking $G = \{\mathbf{w}\}$ gives

$$\mathbf{0} \notin \operatorname{aff}(G) \qquad \text{and} \qquad \operatorname{span}(F) = \{\mathbf{0}\} \subseteq \operatorname{span}(G).$$

Now suppose that $k_0 \geq 1$. We will find a facet R of F with the same linear span as F, and then a different k_0-dimensional face G containing R. Because $\mathbf{0} \in \operatorname{aff}(F)$, the affine hull of F is itself a k_0-dimensional linear subspace. In particular,

$$\operatorname{aff}(F) = \operatorname{span}(F), \qquad \dim \operatorname{span}(F) = k_0. \tag{90}$$

We claim that F has a facet R satisfying

$$\mathbf{0} \notin \operatorname{aff}(R). \tag{91}$$

To see this, consider an irredundant facet description of F within the linear space $\operatorname{aff}(F)$:

$$F = \{\mathbf{x} \in \operatorname{aff}(F) : \langle \mathbf{a}_i, \mathbf{x} \rangle \leq \beta_i, \quad i = 1, \ldots, m\}. \tag{92}$$

The affine hull of the facet corresponding to the ith inequality is contained in

$$\{\mathbf{x} \in \operatorname{aff}(F) : \langle \mathbf{a}_i, \mathbf{x} \rangle = \beta_i\}.$$

If the affine hull of every facet of F contained the origin, then $\beta_i = 0$ for every i. The description Equation (92) would consequently consist entirely of homogeneous inequalities:

$$F = \{\mathbf{x} \in \operatorname{aff}(F) : \langle \mathbf{a}_i, \mathbf{x} \rangle \leq 0, \quad i = 1, \ldots, m\}.$$

For every $\mathbf{x} \in F$ and every $t \geq 0$, we would then have

$$\langle \mathbf{a}_i, t\mathbf{x} \rangle = t \langle \mathbf{a}_i, \mathbf{x} \rangle \leq 0, \qquad i = 1, \ldots, m.$$

Thus $t\mathbf{x} \in F$ for every $t \geq 0$. Since $\dim F = k_0 \geq 1$, we may choose $\mathbf{x} \neq \mathbf{0}$ in F. The points $t\mathbf{x}$ would then make F unbounded. This contradicts the fact that a face of a polytope is itself a polytope and hence is bounded (Ziegler, 1995, Proposition 2.3(i), p. 53). We conclude that F has a facet R satisfying Equation (91).

Since R is a facet of F,

$$\dim R = k_0 - 1.$$

A face of a face is also a face of the ambient polytope, so R is a face of $\mathcal{P}$ by Ziegler (1995, Proposition 2.3(iii), pp. 53–54). Applying Equation (88) to R gives

$$\dim \operatorname{span}(R) = \dim R + 1 = k_0. \tag{93}$$

Since $R \subseteq F$,

$$\operatorname{span}(R) \subseteq \operatorname{span}(F).$$

Equations (90) and (93) show that both spaces have dimension k_0. Therefore,

$$\operatorname{span}(R) = \operatorname{span}(F). \tag{94}$$

We next show that R belongs to another k_0-dimensional face. Since $k_0 < r$, there is a $(k_0 + 1)$-dimensional face Q of $\mathcal{P}$ containing F; when $k_0 = r - 1$, take $Q = \mathcal{P}$. The existence of such a face follows from Ziegler (1995, Theorem 2.7(i), p. 57). Since $R \subseteq F \subseteq Q$ and $\dim R = \dim Q - 2$, R is contained in exactly two facets of Q, by Ziegler (1995, Theorem 2.7(iii), p. 57). One is F; denote the other by G. Both facets have dimension k_0, and a face of Q is also a face of $\mathcal{P}$. Thus $G \neq F$ is a k_0-dimensional face of $\mathcal{P}$ with

$$R \subseteq G. \tag{95}$$

It remains to show that

$$\mathbf{0} \notin \operatorname{aff}(G). \tag{96}$$

Suppose, to the contrary, that

$$\mathbf{0} \in \operatorname{aff}(G).$$

Then, just as for F, the affine hull of G is a linear space and hence

$$\operatorname{aff}(G) = \operatorname{span}(G), \qquad \dim \operatorname{span}(G) = k_0.$$

Because $R \subseteq G$, we have

$$\operatorname{span}(R) \subseteq \operatorname{span}(G).$$

Both spaces have dimension k_0 by Equation (93), so

$$\operatorname{span}(R) = \operatorname{span}(G).$$

Combining this equality with Equation (94) yields

$$\operatorname{aff}(F) = \operatorname{span}(F) = \operatorname{span}(R) = \operatorname{span}(G) = \operatorname{aff}(G).$$

Every face Q of a polytope satisfies

$$Q = \mathcal{P} \cap \operatorname{aff}(Q).$$

Therefore,

$$F = \mathcal{P} \cap \operatorname{aff}(F) = \mathcal{P} \cap \operatorname{aff}(G) = G,$$

contradicting the choice of $G \neq F$. This proves Equation (96).

Finally, Equation (94) and Equation (95) give

$$\operatorname{span}(F) = \operatorname{span}(R) \subseteq \operatorname{span}(G).$$

Thus Equation (85) holds for every $F \in \mathcal{F}_{k_0}(\mathcal{P})$.

Equality of the cell unions. Finally, we prove Equation (86). The right-hand side is contained in the left-hand side because it takes the union over a subcollection of $\mathcal{F}_{k_0}(\mathcal{P})$. Conversely, let

$$\mathbf{x} \in \mathcal{P} \cap \operatorname{span}(F)$$

for some $F \in \mathcal{F}_{k_0}(\mathcal{P})$. By Equation (85), there exists $G \in \mathcal{F}_{k_0}(\mathcal{P})$ such that

$$\mathbf{0} \notin \operatorname{aff}(G) \qquad \text{and} \qquad \operatorname{span}(F) \subseteq \operatorname{span}(G).$$

It follows that

$$\mathbf{x} \in \mathcal{P} \cap \operatorname{span}(G),$$

which places $\mathbf{x}$ in the right-hand side of Equation (86). This proves the reverse inclusion and completes the proof. □

F.2 Proof of the Face-Arrangement Ambiguity Theorem

We now state a lemma that allows us to utilize Assumption 3.3.

Lemma F.3 (Why the full-cell count works)**.** *Let V be a p-dimensional vector space, let K be a finite index set, and let*

$$T = (\mathbf{t}_k)_{k \in K}, \qquad \mathbf{t}_k \in V \quad \textit{for every } k \in K, \tag{97}$$

be an indexed family of points in V. Suppose that every proper linear subspace of V contains at most $p-1$ members of T, counted by their indices. Then, for any proper linear subspaces $U_1, \dots, U_J \subsetneq V$, their union contains at most $J(p-1)$ members of T, counted by their indices; equivalently,

$$\left| \left\{ k \in K : \mathbf{t}_k \in \bigcup_{i=1}^{J} U_i \right\} \right| \leq J(p-1). \tag{98}$$

Proof of Lemma F.3. We adapt the general-position and pigeonhole argument in Hillar & Sommer (2015, proof of Theorem 1) to the face-span setting. Let $U_1, \dots, U_J \subsetneq V$ be proper linear subspaces of V. For each $i \in [J]$, let K_i be the set of indices of the points U_i cover in the family T, explicitly,

$$K_i := \{k \in K : \mathbf{t}_k \in U_i\}. \tag{99}$$

By the hypothesis, K_i contains less than p elements, i.e.,

$$|K_i| \leq p - 1 \qquad \text{for every } i \in [J]. \tag{100}$$

Summing these inequalities over $i \in [J]$ gives

$$\sum_{i=1}^{J} |K_i| \leq J(p-1). \tag{101}$$

Moreover,

$$\left|\bigcup_{i=1}^{J} K_i\right| \le \sum_{i=1}^{J} |K_i|. \tag{102}$$

Therefore,

$$\left|\bigcup_{i=1}^{J} K_i\right| \le J(p-1). \tag{103}$$

Thus the union $\bigcup_{i=1}^{J} U_i$ contains at most $J(p-1)$ members of the indexed family T, counted by their indices. □

Throughout for the rest of this this subsection, assume the hypotheses of Theorem 3.1, the Assumptions 3.1, 3.2 and 3.3, write

$$\mathcal{L} := \mathcal{L}(\mathcal{P}, k_0) = \{L_1, \dots, L_J\}, \qquad p := k_0 + 1, \tag{104}$$

and fix an arbitrary exact feasible factorization

$$(\mathbf{H}, \mathbf{S}) \in \mathcal{F}_{k_0}^{\mathrm{span}}(\mathbf{Y}, \mathcal{P}). \tag{105}$$

For every true-cell index $a \in [J]$ and candidate-cell index $b \in [J]$, define the corresponding image subspaces by

$$V_a := \mathbf{H}_g L_a, \qquad W_b := \mathbf{H} L_b. \tag{106}$$

Lemma F.4 (Cell matching)**.** *There exists a permutation $c : [J] \to [J]$ such that*

$$V_a = W_{c(a)} \qquad \textit{for every } a \in [J]. \tag{107}$$

Proof. Fix an arbitrary $a \in [J]$. By the source-side full-cell-population condition in Assumption 3.3, we may choose an index set

$$\mathcal{K}_a \subseteq \{k : a_g(k) = a\} \tag{108}$$

such that

$$|\mathcal{K}_a| \ge J(p-1) + 1, \tag{109}$$

and the indexed family $\{\mathbf{S}_{g_{:,k}} : k \in \mathcal{K}_a\}$ is in linear general position in L_a.

By Proposition B.3, every proper linear subspace of L_a contains at most $p-1$ members of this indexed family.

Recall that $V_a := \mathbf{H}_g L_a$,. By the singleton case of Assumption 3.2, $\mathbf{H}_g$ is injective on L_a, and hence

$$\dim V_a = \dim L_a = p. \tag{110}$$

We now show that the observations indexed by $\mathcal{K}_a$ are also in linear general position in V_a. Let $T \subsetneq V_a$ be a proper linear subspace. Suppose, toward a contradiction, that T contains p observations $\mathbf{Y}_{:,k_1}, \dots, \mathbf{Y}_{:,k_p}$, where $k_1, \dots, k_p \in \mathcal{K}_a$. By the source-side general-position condition,

$$\mathrm{span}\{\mathbf{S}_{g_{:,k_1}}, \dots, \mathbf{S}_{g_{:,k_p}}\} = L_a. \tag{111}$$

Applying $\mathbf{H}_g$ gives

$$\begin{aligned} V_a &= \mathbf{H}_g L_a \\ &= \mathbf{H}_g \,\mathrm{span}\{\mathbf{S}_{g_{:,k_1}}, \dots, \mathbf{S}_{g_{:,k_p}}\} \\ &= \mathrm{span}\{\mathbf{Y}_{:,k_1}, \dots, \mathbf{Y}_{:,k_p}\} \\ &\subseteq T, \end{aligned}$$

contradicting $T \subsetneq V_a$. Therefore, every proper linear subspace of V_a contains at most $p-1$ observations indexed by $\mathcal{K}_a$. By Proposition B.3, these observations are therefore in linear general position in V_a.

By the definition of $\mathcal{K}_a$, it collects the indices of the selected ground-truth samples assigned to the true cell L_a. Therefore, if $k \in \mathcal{K}_a$, then the k-th ground-truth source column satisfies $\mathbf{S}_{g_{:,k}} \in L_a$. Using $\mathbf{Y} = \mathbf{H}_g \mathbf{S}_g$ and $V_a = \mathbf{H}_g L_a$, we obtain

$$\mathbf{Y}_{:,k} = \mathbf{H}_g \mathbf{S}_{g_{:,k}} \in V_a. \tag{112}$$

On the other hand, feasibility of the candidate factorization $(\mathbf{H}, \mathbf{S})$ implies that there exists an index $b(k) \in [J]$ such that $\mathbf{S}_{:,k} \in L_{b(k)}$. Since $\mathbf{Y} = \mathbf{HS}$ and $W_{b(k)} = \mathbf{H}L_{b(k)}$, it follows that

$$\mathbf{Y}_{:,k} = \mathbf{H}\mathbf{S}_{:,k} \in W_{b(k)}. \tag{113}$$

Combining the two inclusions for $\mathbf{Y}_{:,k}$, we obtain

$$\mathbf{Y}_{:,k} \in V_a \cap W_{b(k)}. \tag{114}$$

Since $b(k) \in [J]$ for every $k \in \mathcal{K}_a$, all observations indexed by $\mathcal{K}_a$ satisfy

$$\{\mathbf{Y}_{:,k} : k \in \mathcal{K}_a\} \subseteq \bigcup_{b=1}^{J} (V_a \cap W_b). \tag{115}$$

Claim: There exists at least one index $b_a \in [J]$ such that $V_a \subseteq W_{b_a}$.

To prove the claim, suppose that

$$V_a \nsubseteq W_b \qquad \text{for every } b \in [J]. \tag{116}$$

Then, for every $b \in [J]$, the intersection $V_a \cap W_b$ is a proper linear subspace of V_a. By the observation side general position conclusion established above, each proper linear subspace of V_a contains at most $p-1$ observations indexed by $\mathcal{K}_a$. Lemma F.3 implies that the union of the J proper subspaces $V_a \cap W_1, \ldots, V_a \cap W_J \subsetneq V_a$ contains at most $J(p-1)$ such observations. However, our choice of $\mathcal{K}_a$ gives

$$|\mathcal{K}_a| \geq J(p-1) + 1. \tag{117}$$

This contradicts Equation (115). Thus, the claim holds, and there exists at least one index $b_a \in [J]$ such that

$$V_a \subseteq W_{b_a}. \tag{118}$$

For each $a \in [J]$, choose one such index b_a, and define the mapping

$$c : [J] \to [J], \qquad c(a) := b_a. \tag{119}$$

By the definition of $W_{c(a)}$, $W_{c(a)} = \mathbf{H}L_{c(a)}$. A linear map cannot increase dimension, so $\dim W_{c(a)} \leq \dim L_{c(a)} = p$.

On the other hand, since $V_a \subseteq W_{c(a)}$, $\dim V_a \leq \dim W_{c(a)}$.

To determine $\dim V_a$, apply the q-GP condition in Assumption 3.2 with $I = \{a\}$. This choice is admissible because $|I| = 1 \leq q$, where $q \geq 2$ by Assumption 3.2. Hence $\mathbf{H}_g$ is injective on L_a, and therefore $\dim V_a = \dim(\mathbf{H}_g L_a) = \dim L_a = p$.

Combining these relations gives

$$p = \dim V_a \leq \dim W_{c(a)} \leq \dim L_{c(a)} = p. \tag{120}$$

Thus $\dim V_a = \dim W_{c(a)} = p$. Since $V_a \subseteq W_{c(a)}$ and the two subspaces have the same dimension,

$$V_a = W_{c(a)}. \tag{121}$$

Because $a \in [J]$ was arbitrary, this proves the asserted cell-matching identity for every $a \in [J]$.

Claim: The mapping $c : [J] \to [J]$ is a permutation.

We first show that the subspaces $V_1, \ldots, V_J$ are pairwise distinct. Suppose that

$$V_a = V_{a'} \tag{122}$$

for some $a, a' \in [J]$. Let $\mathbf{x} \in L_a$. Since $\mathbf{H}_g\mathbf{x} \in V_a = V_{a'} = \mathbf{H}_g L_{a'}$, there exists $\mathbf{x}' \in L_{a'}$ such that

$$\mathbf{H}_g\mathbf{x} = \mathbf{H}_g\mathbf{x}'. \tag{123}$$

Consequently,

$$\mathbf{H}_g(\mathbf{x} - \mathbf{x}') = 0, \qquad \mathbf{x} - \mathbf{x}' \in L_a + L_{a'}. \tag{124}$$

By Assumption 3.2 with $I = \{a, a'\}$, we have $|I| \le 2 \le q$ because $q \ge 2$ by Assumption 3.2. Hence $\mathbf{H}_g$ is injective on $L_a + L_{a'}$, which implies $\mathbf{x} - \mathbf{x}' = 0$, and therefore $\mathbf{x} = \mathbf{x}' \in L_{a'}$. Since $\mathbf{x} \in L_a$ was arbitrary, this proves

$$L_a \subseteq L_{a'}. \tag{125}$$

Interchanging a and a' gives the reverse inclusion, so

$$L_a = L_{a'}. \tag{126}$$

The arrangement $\mathcal{L} = \{L_1, \dots, L_J\}$ contains no duplicate subspaces; hence $a = a'$. Thus $V_1, \dots, V_J$ are pairwise distinct.

Now suppose that $c(a) = c(a')$. By the cell-matching identity established above,

$$V_a = W_{c(a)} = W_{c(a')} = V_{a'}. \tag{127}$$

Since the subspaces $V_1, \dots, V_J$ are pairwise distinct, it follows that $a = a'$. Therefore c is injective. Because c is an injective map from the finite set $[J]$ to itself, it is also surjective and hence is a permutation. This proves the claim.

□

Lemma F.5 (Candidate local injectivity)**.** *Let c be the permutation supplied by Lemma F.4. Then the candidate mixing matrix is injective on every local sum of at most q arrangement subspaces. Equivalently, for every $K \subseteq [J]$ with $1 \le |K| \le q$,*

$$\ker(\mathbf{H}) \cap \sum_{b \in K} L_b = \{\mathbf{0}\}. \tag{128}$$

Proof. Recall from Definition 3.1 that, for every $K \subseteq [J]$,

$$\rho(K) := \dim \left(\sum_{a \in K} L_a \right). \tag{129}$$

Let $I \subseteq [J]$, and define

$$c(I) := \{c(a) : a \in I\}. \tag{130}$$

If $1 \le |I| \le q$, then, using Lemma F.4,

$$\mathbf{H} \left(\sum_{a \in I} L_{c(a)} \right) = \sum_{a \in I} \mathbf{H} L_{c(a)} = \sum_{a \in I} W_{c(a)} \tag{131}$$

$$= \sum_{a \in I} V_a = \sum_{a \in I} \mathbf{H}_g L_a = \mathbf{H}_g \left(\sum_{a \in I} L_a \right). \tag{132}$$

By the q-GP condition in Assumption 3.2, $\mathbf{H}_g$ is injective on $\sum_{a \in I} L_a$. Therefore, using the equality of image subspaces above,

$$\dim \left(\operatorname{im} \left(\mathbf{H} \Big|_{\sum_{a \in I} L_{c(a)}} \right) \right) = \dim \left(\mathbf{H}_g \left(\sum_{a \in I} L_a \right) \right) \tag{133}$$

$$= \dim \left(\sum_{a \in I} L_a \right) = \rho(I). \tag{134}$$

Also, by the definition of ρ and the definition of $c(I)$, $\dim \left(\sum_{a \in I} L_{c(a)} \right) = \rho(c(I))$ and since a linear map cannot increase dimension, it follows that

$$\dim \left(\mathbf{H} \left(\sum_{a \in I} L_{c(a)} \right) \right) \le \dim \left(\sum_{a \in I} L_{c(a)} \right) = \rho(c(I)). \tag{135}$$

Combining this inequality with Equation (134), where the dimension of the same image subspace was shown to equal $\rho(I)$, gives

$$\rho(I) \le \rho(c(I)). \tag{136}$$

Now, since c is a permutation of the finite set $[J]$, its induced action on subsets of $[J]$ is also a permutation. Therefore, $c(I)$ is periodic, and there exists an integer $m \geq 1$ such that $c^m(I) = I$. If we apply the preceding arguments successively to the subsets $K = I, c(I), \ldots, c^{m-1}(I)$ with $\rho(K) \leq \rho(c(K))$ consecutively,

$$\rho(I) \leq \rho(c(I)) \leq \cdots \leq \rho(c^m(I)) = \rho(I). \tag{137}$$

Since this chain begins and ends with the same quantity, every inequality in the chain must be an equality. In particular,

$$\rho(I) = \rho(c(I)). \tag{138}$$

Thus the restriction $\mathbf{H}|_{\sum_{a\in I} L_{c(a)}}$ has image dimension equal to the dimension of its domain. By the rank-nullity theorem, $\dim \ker(\mathbf{H}|_{\sum_{a\in I} L_{c(a)}}) = 0$ and $\mathbf{H}$ is injective on $\sum_{a\in I} L_{c(a)}$, i.e., $\ker(\mathbf{H}) \cap \sum_{a\in I} L_{c(a)} = \{0\}$.

Now let $K \subseteq [J]$ satisfy $1 \leq |K| \leq q$, and set $I = c^{-1}(K)$. Since c is a permutation, $|I| = |K| \leq q$ and $\sum_{a\in I} L_{c(a)} = \sum_{b\in K} L_b$.

Therefore

$$\ker(\mathbf{H}) \cap \sum_{b\in K} L_b = \{0\}. \tag{139}$$

□

Lemma F.6 (Extending cellwise maps)**.** *Let c be the permutation supplied by Lemma F.4. Then there exists an invertible $\mathbf{B} \in \mathbb{R}^{r\times r}$ such that*

$$\mathbf{H}\mathbf{B} = \mathbf{H}_g, \qquad \mathbf{B}L_a = L_{c(a)} \quad \textit{for every } a \in [J]. \tag{140}$$

Proof. **Construction and properties of the local maps $\mathbf{B}_a$.** Let $c : [J] \to [J]$ be the permutation given by Lemma F.4. Then

$$V_a = W_{c(a)} \qquad \text{for every } a \in [J]. \tag{141}$$

By the definitions of V_a and $W_{c(a)}$, this is equivalent to

$$\mathbf{H}_g L_a = \mathbf{H} L_{c(a)} \qquad \text{for every } a \in [J]. \tag{142}$$

Fix an arbitrary $a \in [J]$ and $\mathbf{x} \in L_a$. By Equation (142), there exists $\mathbf{y} \in L_{c(a)}$ such that

$$\mathbf{H}_g\mathbf{x} = \mathbf{H}\mathbf{y}. \tag{143}$$

Moreover, $\mathbf{H}|_{L_{c(a)}}$ is injective by Lemma F.5, applied with $K = \{c(a)\}$, so this vector $\mathbf{y}$ is unique.

We define a local linear map

$$\mathbf{B}_a\mathbf{x} := \mathbf{y}. \tag{144}$$

Equivalently, the local linear map $\mathbf{B}_a : L_a \to L_{c(a)}$ is

$$\mathbf{B}_a := \left(\mathbf{H}|_{L_{c(a)}}\right)^{-1} \left(\mathbf{H}_g|_{L_a}\right) \qquad \text{for every } a \in [J]. \tag{145}$$

We next show that each $\mathbf{B}_a$ is an isomorphism. Let $\mathbf{x}, \mathbf{x}' \in L_a$ and suppose that

$$\mathbf{B}_a\mathbf{x} = \mathbf{B}_a\mathbf{x}'. \tag{146}$$

Multiplying both sides from left with $\mathbf{H}$ and using the definition of $\mathbf{B}_a$ gives

$$\mathbf{H}_g\mathbf{x} = \mathbf{H}\mathbf{B}_a\mathbf{x} = \mathbf{H}\mathbf{B}_a\mathbf{x}' = \mathbf{H}_g\mathbf{x}'. \tag{147}$$

By Assumption 3.2, with $I = \{a\}$, $|I| = 1 \leq q$ as $q \geq 2$ the restriction $\mathbf{H}_g|_{L_a}$ is injective. Therefore $\mathbf{x} = \mathbf{x}'$, and hence $\mathbf{B}_a$ is injective. Since

$$\dim L_a = \dim L_{c(a)} = p, \tag{148}$$

the injective linear map $\mathbf{B}_a : L_a \to L_{c(a)}$ is also surjective. Hence, $\mathbf{B}_a$ is an isomorphism.

Verification of the condition in Proposition F.1. Let $I \subseteq [J]$ satisfy $1 \le |I| \le q$, and let $\mathbf{z}_a \in L_a$, $a \in I$, satisfy

$$\sum_{a \in I} \mathbf{z}_a = \mathbf{0}. \tag{149}$$

We will show that

$$\sum_{a \in I} \mathbf{B}_a \mathbf{z}_a = \mathbf{0}. \tag{150}$$

By the definition of the local maps $\mathbf{B}_a$,

$$\mathbf{H}\mathbf{B}_a \mathbf{z}_a = \mathbf{H}_g \mathbf{z}_a \qquad \text{for every } a \in I.$$

Therefore,

$$\mathbf{H}\left(\sum_{a \in I} \mathbf{B}_a \mathbf{z}_a\right) = \sum_{a \in I} \mathbf{H}\mathbf{B}_a \mathbf{z}_a \tag{151}$$

$$= \sum_{a \in I} \mathbf{H}_g \mathbf{z}_a \tag{152}$$

$$= \mathbf{H}_g \left(\sum_{a \in I} \mathbf{z}_a\right) \tag{153}$$

$$= \mathbf{0}. \tag{154}$$

Hence,

$$\sum_{a \in I} \mathbf{B}_a \mathbf{z}_a \in \ker(\mathbf{H}). \tag{155}$$

On the other hand, since $\mathbf{B}_a \mathbf{z}_a \in L_{c(a)}$ for every $a \in I$, we also have

$$\sum_{a \in I} \mathbf{B}_a \mathbf{z}_a \in \sum_{a \in I} L_{c(a)}. \tag{156}$$

Combining these memberships gives

$$\sum_{a \in I} \mathbf{B}_a \mathbf{z}_a \in \ker(\mathbf{H}) \cap \left(\sum_{a \in I} L_{c(a)}\right). \tag{157}$$

Since c is a permutation, the set $c(I) = \{c(a) : a \in I\}$ has cardinality $|c(I)| = |I| \le q$. Applying Lemma F.5 to $c(I)$, we obtain

$$\ker(\mathbf{H}) \cap \left(\sum_{a \in I} L_{c(a)}\right) = \{\mathbf{0}\}. \tag{158}$$

Consequently,

$$\sum_{a \in I} \mathbf{B}_a \mathbf{z}_a = \mathbf{0}. \tag{159}$$

Because I and the vectors $\mathbf{z}_a$ were arbitrary, the maps $\mathbf{B}_a$ satisfy Equation (53).

Construction and properties of the global map B. Since the polytope $\mathcal{P}$ is full-dimensional in $\mathbb{R}^r$, its vertices linearly span $\mathbb{R}^r$. Every vertex belongs to a k_0-dimensional face, and Proposition F.2 shows that the span of that face is contained in one of the retained face spans L_a. Thus every vertex belongs to at least one L_a. Since the vertices span $\mathbb{R}^r$, so do the subspaces in the arrangement:

$$\sum_{a=1}^{J} L_a = \mathbb{R}^r. \tag{160}$$

Thus $\Sigma_{\mathcal{L}}$ is surjective, and every $\mathbf{x} \in \mathbb{R}^r$ can be represented as

$$\mathbf{x} = \Sigma_{\mathcal{L}}(Z) = \sum_{a=1}^{J} \mathbf{z}_a \tag{161}$$

for some $Z \in \mathcal{E}_{\mathcal{L}}$.

By Assumption 3.1, the arrangement $\mathcal{L}$ is q-relation-complete. We have also shown that the maps $\mathbf{B}_a$ satisfy Equation (53). Therefore, Proposition F.1, applied with $W = \mathbb{R}^r$ and $T_a = \mathbf{B}_a$, gives a unique linear map

$$\mathbf{B} : \mathbb{R}^r \to \mathbb{R}^r \tag{162}$$

whose restriction to each L_a is $\mathbf{B}_a$. In particular,

$$\mathbf{B}|_{L_a} = \mathbf{B}_a \qquad \text{for every } a \in [J]. \tag{163}$$

Since $\mathbf{B}_a : L_a \to L_{c(a)}$ is surjective, it follows that

$$\mathbf{B}L_a = \mathbf{B}_a L_a = L_{c(a)}. \tag{164}$$

This proves the second part of the claim.

Now for the first part, we start by inspecting the range of $\mathbf{B}$. Using Equation (160), Equation (164), and the fact that c is a permutation of $[J]$, we obtain

$$\operatorname{range}(\mathbf{B}) = \mathbf{B}(\mathbb{R}^r) = \mathbf{B}\left(\sum_{a=1}^{J} L_a\right) = \sum_{a=1}^{J} \mathbf{B}L_a \tag{165}$$

$$= \sum_{a=1}^{J} \mathbf{B}_a L_a = \sum_{a=1}^{J} L_{c(a)} = \sum_{a=1}^{J} L_a = \mathbb{R}^r. \tag{166}$$

Thus $\mathbf{B}$ is surjective. Since $\mathbf{B} : \mathbb{R}^r \to \mathbb{R}^r$ is a surjective linear map between finite-dimensional vector spaces of the same dimension, it is invertible.

Finally, for every $a \in [J]$ and every $\mathbf{z}_a \in L_a$, Equation (163) and Equation (147) give

$$\mathbf{H}\mathbf{B}\mathbf{z}_a = \mathbf{H}\mathbf{B}_a\mathbf{z}_a = \mathbf{H}_g\mathbf{z}_a. \tag{167}$$

Now let $\mathbf{x} \in \mathbb{R}^r$ be arbitrary. By Equation (161), there exist vectors $\mathbf{z}_a \in L_a$ such that $\mathbf{x} = \sum_{a=1}^{J} \mathbf{z}_a$. Using linearity and the preceding identity Equation (167), we obtain

$$\mathbf{H}\mathbf{B}\mathbf{x} = \mathbf{H}\mathbf{B}\left(\sum_{a=1}^{J} \mathbf{z}_a\right) = \sum_{a=1}^{J} \mathbf{H}\mathbf{B}\mathbf{z}_a \tag{168}$$

$$= \sum_{a=1}^{J} \mathbf{H}_g\mathbf{z}_a = \mathbf{H}_g\left(\sum_{a=1}^{J} \mathbf{z}_a\right) = \mathbf{H}_g\mathbf{x}. \tag{169}$$

Since this equality holds for every $\mathbf{x} \in \mathbb{R}^r$, the two linear maps are equal. Therefore,

$$\mathbf{H}\mathbf{B} = \mathbf{H}_g. \tag{170}$$

□

Proof of Theorem 3.1. For the exact feasible factorization $(\mathbf{H}, \mathbf{S})$ fixed at the beginning of this section, Lemma F.6 supplies an invertible matrix $\mathbf{B} \in \mathbb{R}^{r\times r}$ and the permutation $c : [J] \to [J]$ satisfying

$$\mathbf{H}\mathbf{B} = \mathbf{H}_g, \qquad \mathbf{B}L_a = L_{c(a)} \quad \text{for every } a \in [J]. \tag{171}$$

Since $c : [J] \to [J]$ is a permutation, it follows immediately that

$$\mathbf{B}\mathcal{L} = \{\mathbf{B}L_a : a \in [J]\} \tag{172}$$

$$= \{L_{c(a)} : a \in [J]\} \tag{173}$$

$$= \{L_b : b \in [J]\} = \mathcal{L}. \tag{174}$$

It remains to show that $\mathbf{S} = \mathbf{B}\mathbf{S}_g$. Fix an arbitrary sample index $k \in [N]$, and let $\mathbf{S}_{g_{:,k}}$ denote the corresponding ground-truth source column. Let $a = a_g(k) \in [J]$ be its true cell index, so that

$$\mathbf{S}_{g_{:,k}} \in C_a \subseteq L_a. \tag{175}$$

Likewise, choose a candidate cell index $b = b(k) \in [J]$ such that the corresponding candidate source column satisfies

$$\mathbf{S}_{:,k} \in C_b \subseteq L_b, \qquad \mathbf{H}\mathbf{S}_{:,k} = \mathbf{Y}_{:,k}. \tag{176}$$

Since $\mathbf{S}_{g_{:,k}} \in L_a$ and $\mathbf{B}L_a = L_{c(a)}$, we have

$$\mathbf{B}\mathbf{S}_{g_{:,k}} \in L_{c(a)}. \tag{177}$$

Define the difference vector

$$\mathbf{d}_k := \mathbf{S}_{:,k} - \mathbf{B}\mathbf{S}_{g_{:,k}}. \tag{178}$$

By Equation (176) and Equation (177),

$$\mathbf{d}_k \in L_b + L_{c(a)}. \tag{179}$$

Moreover, using $\mathbf{H}\mathbf{B} = \mathbf{H}_g$ and the two exact factorizations of $\mathbf{Y}_{:,k}$,

$$\begin{aligned} \mathbf{H}\mathbf{d}_k &= \mathbf{H}\big(\mathbf{S}_{:,k} - \mathbf{B}\mathbf{S}_{g_{:,k}}\big) && (180)\\ &= \mathbf{H}\mathbf{S}_{:,k} - \mathbf{H}\mathbf{B}\mathbf{S}_{g_{:,k}} && (181)\\ &= \mathbf{Y}_{:,k} - \mathbf{H}_g\mathbf{S}_{g_{:,k}} && (182)\\ &= 0. && (183) \end{aligned}$$

Therefore,

$$\mathbf{d}_k \in \ker(\mathbf{H}). \tag{184}$$

Combining Equation (179) and Equation (184), we obtain

$$\mathbf{d}_k \in \ker(\mathbf{H}) \cap \big(L_b + L_{c(a)}\big). \tag{185}$$

Set $I := \{b, c(a)\}$. Then $1 \le |I| \le q$, as by definition $q \ge 2$. Therefore, Lemma F.5 gives

$$\ker(\mathbf{H}) \cap \big(L_b + L_{c(a)}\big) = \{0\}. \tag{186}$$

Together with Equation (185), this implies

$$\mathbf{d}_k = 0. \tag{187}$$

By the definition of $\mathbf{d}_k$, we conclude that

$$\mathbf{S}_{:,k} = \mathbf{B}\mathbf{S}_{g_{:,k}}. \tag{188}$$

Since $k \in [N]$ was arbitrary, this identity holds for every $k \in [N]$. Therefore,

$$\mathbf{S} = \mathbf{B}\mathbf{S}_g. \tag{189}$$

Finally, since $\mathbf{H}\mathbf{B} = \mathbf{H}_g$ and $\mathbf{B}$ is invertible,

$$\mathbf{H} = \mathbf{H}_g\mathbf{B}^{-1}. \tag{190}$$

□

F.3 Proof of the Det-Max and Polar-Tightness Lemma

Proof of Lemma 3.2. Let an invertible $\mathbf{B} \in \mathbb{R}^{r\times r}$ satisfy $\mathbf{B}\mathbf{S}_{g_{:,k}} \in \mathcal{P}$ for every k. Since $\mathcal{P}$ is convex, taking the convex hull of these transformed source points gives

$$\operatorname{conv}\{\mathbf{B}\mathbf{S}_{g_{:,1}}, \dots, \mathbf{B}\mathbf{S}_{g_{:,N}}\} \subseteq \mathcal{P}. \tag{191}$$

By the definition of $\operatorname{conv}(\mathbf{S}_g)$ in Assumption 3.4 and the fact that a linear map commutes with taking convex hulls, we have

$$\begin{aligned} \mathbf{B}\operatorname{conv}(\mathbf{S}_g) &= \mathbf{B}\operatorname{conv}\{\mathbf{S}_{g_{:,1}}, \dots, \mathbf{S}_{g_{:,N}}\} && (192)\\ &= \operatorname{conv}\{\mathbf{B}\mathbf{S}_{g_{:,1}}, \dots, \mathbf{B}\mathbf{S}_{g_{:,N}}\} \subseteq \mathcal{P}. && (193) \end{aligned}$$

Next, the left-hand inclusion in Assumption 3.4 gives

$$\mathcal{E}_{\mathcal{P}} \subseteq \operatorname{conv}(\mathbf{S}_g). \tag{194}$$

Applying the linear map $\mathbf{B}$ to both sides preserves this inclusion, so

$$\mathbf{B}\mathcal{E}_{\mathcal{P}} \subseteq \mathbf{B}\,\mathrm{conv}(\mathbf{S}_g). \tag{195}$$

Combining this inclusion with $\mathbf{B}\,\mathrm{conv}(\mathbf{S}_g) \subseteq \mathcal{P}$, obtained above, yields

$$\mathbf{B}\mathcal{E}_{\mathcal{P}} \subseteq \mathbf{B}\,\mathrm{conv}(\mathbf{S}_g) \subseteq \mathcal{P}. \tag{196}$$

In particular,

$$\mathbf{B}\mathcal{E}_{\mathcal{P}} \subseteq \mathcal{P}. \tag{197}$$

Because $\mathbf{B}$ is invertible, the set $\mathbf{B}\mathcal{E}_{\mathcal{P}}$ is a full-dimensional ellipsoid. By construction, $\mathcal{E}_{\mathcal{P}}$ is the maximum-volume inscribed ellipsoid of $\mathcal{P}$. Therefore no ellipsoid contained in $\mathcal{P}$, including $\mathbf{B}\mathcal{E}_{\mathcal{P}}$, can have volume larger than that of $\mathcal{E}_{\mathcal{P}}$. Hence

$$\mathrm{vol}(\mathbf{B}\mathcal{E}_{\mathcal{P}}) \leq \mathrm{vol}(\mathcal{E}_{\mathcal{P}}). \tag{198}$$

Under an invertible linear map, volume scales by the absolute value of the determinant. Consequently,

$$\mathrm{vol}(\mathbf{B}\mathcal{E}_{\mathcal{P}}) = |\det \mathbf{B}|\ \mathrm{vol}(\mathcal{E}_{\mathcal{P}}). \tag{199}$$

Substituting this identity into the preceding volume inequality gives

$$|\det \mathbf{B}|\ \mathrm{vol}(\mathcal{E}_{\mathcal{P}}) \leq \mathrm{vol}(\mathcal{E}_{\mathcal{P}}). \tag{200}$$

Since $\mathrm{vol}(\mathcal{E}_{\mathcal{P}}) > 0$, division by $\mathrm{vol}(\mathcal{E}_{\mathcal{P}})$ yields

$$|\det \mathbf{B}| \leq 1. \tag{201}$$

We now prove the "only if" part of the equality case. Suppose that

$$|\det \mathbf{B}| = 1. \tag{202}$$

Then, volume of the scaled MVIE is

$$\mathrm{vol}(\mathbf{B}\mathcal{E}_{\mathcal{P}}) = |\det \mathbf{B}|\ \mathrm{vol}(\mathcal{E}_{\mathcal{P}}) = \mathrm{vol}(\mathcal{E}_{\mathcal{P}}). \tag{203}$$

Since $\mathbf{B}\mathcal{E}_{\mathcal{P}} \subseteq \mathcal{P}$, the ellipsoid $\mathbf{B}\mathcal{E}_{\mathcal{P}}$ is also a maximum-volume inscribed ellipsoid of $\mathcal{P}$. The MVIE of $\mathcal{P}$ is unique, so

$$\mathbf{B}\mathcal{E}_{\mathcal{P}} = \mathcal{E}_{\mathcal{P}}. \tag{204}$$

Since $\mathbf{B}$ preserves this ellipsoid, Lemma B.1 gives

$$\mathbf{B}(\partial\mathcal{E}_{\mathcal{P}}) = \partial\mathcal{E}_{\mathcal{P}}, \tag{205}$$

and

$$\mathbf{B}\mathbf{g}_{\mathcal{P}} = \mathbf{g}_{\mathcal{P}}. \tag{206}$$

We can therefore apply Lemma B.2 with $K = \mathcal{E}_{\mathcal{P}}$, $\mathbf{A} = \mathbf{B}$, and $\mathbf{d} = \mathbf{g}_{\mathcal{P}}$. Since $\mathbf{B}\mathcal{E}_{\mathcal{P}} = \mathcal{E}_{\mathcal{P}}$, we obtain

$$\mathbf{B}^{-\top}\mathcal{E}_{\mathcal{P}}^{*,\mathbf{g}_{\mathcal{P}}} = (\mathbf{B}\mathcal{E}_{\mathcal{P}})^{*,\mathbf{g}_{\mathcal{P}}} = \mathcal{E}_{\mathcal{P}}^{*,\mathbf{g}_{\mathcal{P}}}. \tag{207}$$

Applying Lemma B.1 to this polar ellipsoid shows that its boundary is preserved by $\mathbf{B}^{-\top}$, and hence also by its inverse $\mathbf{B}^{\top}$:

$$\mathbf{B}^{-\top}\big(\partial(\mathcal{E}_{\mathcal{P}}^{*,\mathbf{g}_{\mathcal{P}}})\big) = \partial(\mathcal{E}_{\mathcal{P}}^{*,\mathbf{g}_{\mathcal{P}}}) = \mathbf{B}^{\top}\big(\partial(\mathcal{E}_{\mathcal{P}}^{*,\mathbf{g}_{\mathcal{P}}})\big). \tag{208}$$

Recall that $\mathbf{B}\,\mathrm{conv}(\mathbf{S}_g) \subseteq \mathcal{P}$. Taking polars reverses the inclusion, and also applying Lemma B.2 gives

$$\mathcal{P}^{*,\mathbf{g}_{\mathcal{P}}} \subseteq (\mathbf{B}\,\mathrm{conv}(\mathbf{S}_g))^{*,\mathbf{g}_{\mathcal{P}}} = \mathbf{B}^{-\top}\,\mathrm{conv}(\mathbf{S}_g)^{*,\mathbf{g}_{\mathcal{P}}}. \tag{209}$$

Now let $\mathbf{u} \in \mathrm{ext}(\mathcal{P}^{*,\mathbf{g}_{\mathcal{P}}})$. By the polar part of Assumption 3.4,

$$\mathrm{ext}\big(\mathcal{P}^{*,\mathbf{g}_{\mathcal{P}}}\big) = \mathrm{conv}(\mathbf{S}_g)^{*,\mathbf{g}_{\mathcal{P}}} \cap \partial\big(\mathcal{E}_{\mathcal{P}}^{*,\mathbf{g}_{\mathcal{P}}}\big), \tag{210}$$

so $\mathbf{u} \in \partial\big(\mathcal{E}_{\mathcal{P}}^{*,\mathbf{g}_{\mathcal{P}}}\big)$. Moreover, equation 209 implies that $\mathbf{u} \in \mathbf{B}^{-\top} \operatorname{conv}(\mathbf{S}_g)^{*,\mathbf{g}_{\mathcal{P}}}$. Applying $\mathbf{B}^{\top}$ therefore gives

$$\mathbf{B}^{\top}\mathbf{u} \in \operatorname{conv}(\mathbf{S}_g)^{*,\mathbf{g}_{\mathcal{P}}}. \tag{211}$$

Since $\mathbf{u} \in \operatorname{ext}(\mathcal{P}^{*,\mathbf{g}_{\mathcal{P}}})$, equation 210 also gives

$$\mathbf{u} \in \partial\big(\mathcal{E}_{\mathcal{P}}^{*,\mathbf{g}_{\mathcal{P}}}\big). \tag{212}$$

By equation 208, we therefore have

$$\mathbf{u} \in \mathbf{B}^{-\top}\big(\partial(\mathcal{E}_{\mathcal{P}}^{*,\mathbf{g}_{\mathcal{P}}})\big). \tag{213}$$

Applying $\mathbf{B}^{\top}$ to both sides gives

$$\mathbf{B}^{\top}\mathbf{u} \in \partial\big(\mathcal{E}_{\mathcal{P}}^{*,\mathbf{g}_{\mathcal{P}}}\big). \tag{214}$$

Combining the memberships in equations 211 and 214 with equation 210 yields

$$\mathbf{B}^{\top} \operatorname{ext}\big(\mathcal{P}^{*,\mathbf{g}_{\mathcal{P}}}\big) \subseteq \operatorname{ext}\big(\mathcal{P}^{*,\mathbf{g}_{\mathcal{P}}}\big). \tag{215}$$

The set $\operatorname{ext}(\mathcal{P}^{*,\mathbf{g}_{\mathcal{P}}})$ is finite, and $\mathbf{B}^{\top}$ is invertible, so the inclusion in equation 215 is an equality:

$$\mathbf{B}^{\top} \operatorname{ext}\big(\mathcal{P}^{*,\mathbf{g}_{\mathcal{P}}}\big) = \operatorname{ext}\big(\mathcal{P}^{*,\mathbf{g}_{\mathcal{P}}}\big). \tag{216}$$

Taking convex hulls gives

$$\mathbf{B}^{\top}\mathcal{P}^{*,\mathbf{g}_{\mathcal{P}}} = \mathcal{P}^{*,\mathbf{g}_{\mathcal{P}}}. \tag{217}$$

Taking polars with respect to the origin on both sides gives

$$\big(\mathbf{B}^{\top}\mathcal{P}^{*,\mathbf{g}_{\mathcal{P}}}\big)^{*} = \big(\mathcal{P}^{*,\mathbf{g}_{\mathcal{P}}}\big)^{*}. \tag{218}$$

By equation 32 and shifted bipolarity 29, this becomes

$$\mathbf{B}^{-1}(\mathcal{P} - \mathbf{g}_{\mathcal{P}}) = \mathcal{P} - \mathbf{g}_{\mathcal{P}}. \tag{219}$$

Applying $\mathbf{B}$ and using $\mathbf{B}\mathbf{g}_{\mathcal{P}} = \mathbf{g}_{\mathcal{P}}$ from equation 206, we obtain

$$\mathbf{B}\mathcal{P} - \mathbf{g}_{\mathcal{P}} = \mathcal{P} - \mathbf{g}_{\mathcal{P}}, \tag{220}$$

and hence

$$\mathbf{B}\mathcal{P} = \mathcal{P}. \tag{221}$$

Therefore

$$\mathbf{B} \in \mathcal{A}_{\mathcal{P}}. \tag{222}$$

For the "if" part, suppose that

$$\mathbf{B} \in \mathcal{A}_{\mathcal{P}}. \tag{223}$$

By definition of the automorphism group,

$$\mathbf{B}\mathcal{P} = \mathcal{P}. \tag{224}$$

Taking volumes on both sides of Equation (224) gives

$$\operatorname{vol}(\mathcal{P}) = \operatorname{vol}(\mathbf{B}\mathcal{P}) = |\det \mathbf{B}| \operatorname{vol}(\mathcal{P}). \tag{225}$$

Since $\operatorname{vol}(\mathcal{P}) > 0$, we conclude that

$$|\det \mathbf{B}| = 1. \tag{226}$$

This proves the converse and completes the equality characterization.

□

F.4 Proof of the Main Theorem

Proof of Theorem 3.3. Let $(\mathbf{H}^\star, \mathbf{S}^\star)$ be a global maximizer of Equation (13). By Theorem 3.1, there exists an invertible matrix $\mathbf{B} \in \mathbb{R}^{r\times r}$ such that

$$\mathbf{S}^\star = \mathbf{B}\mathbf{S}_g, \qquad \mathbf{H}^\star = \mathbf{H}_g\mathbf{B}^{-1}. \tag{227}$$

Since $\mathbf{S}^\star$ is feasible, every column of $\mathbf{S}^\star$ belongs to the face-span skeleton, and therefore

$$\mathbf{B}\mathbf{S}_{g_{:,k}} = \mathbf{S}^\star_{:,k} \in \mathrm{Skel}_{k_0}(\mathcal{P}) \subseteq \mathcal{P} \qquad \text{for every } k. \tag{228}$$

Then by Lemma 3.2 we have $|\det \mathbf{B}| \le 1$.

Now, using the definition of the sample correlation matrix and substituting $\mathbf{S}^\star = \mathbf{B}\mathbf{S}_g$, we obtain the following for the Det-Max objective Equation (13)

$$\begin{aligned}
\frac{1}{2}\log\det \mathbf{R}_{\mathbf{S}^\star} &= \frac{1}{2}\log\det\left(\frac{1}{N}\mathbf{S}^\star(\mathbf{S}^\star)^\top\right) && (229)\\
&= \frac{1}{2}\log\det\left(\frac{1}{N}(\mathbf{B}\mathbf{S}_g)(\mathbf{B}\mathbf{S}_g)^\top\right) && (230)\\
&= \frac{1}{2}\log\det\left(\mathbf{B}\left(\frac{1}{N}\mathbf{S}_g\mathbf{S}_g^\top\right)\mathbf{B}^\top\right) && (231)\\
&= \frac{1}{2}\log\det\left(\mathbf{B}\mathbf{R}_{\mathbf{S}_g}\mathbf{B}^\top\right) && (232)\\
&= \frac{1}{2}\log\left(\det(\mathbf{B})\det(\mathbf{R}_{\mathbf{S}_g})\det(\mathbf{B}^\top)\right) && (233)\\
&= \frac{1}{2}\log\left(|\det\mathbf{B}|^2\det\mathbf{R}_{\mathbf{S}_g}\right) && (234)\\
&= \frac{1}{2}\log\det\mathbf{R}_{\mathbf{S}_g} + \log|\det\mathbf{B}|. && (235)
\end{aligned}$$

On the other hand, the ground-truth factorization $(\mathbf{H}_g, \mathbf{S}_g)$ is feasible, so the global optimality of $(\mathbf{H}^\star, \mathbf{S}^\star)$ implies

$$\frac{1}{2}\log\det\mathbf{R}_{\mathbf{S}^\star} \ge \frac{1}{2}\log\det\mathbf{R}_{\mathbf{S}_g}. \tag{236}$$

Combining equations 235 and 236 gives $\log|\det\mathbf{B}| \ge 0$, which implies

$$|\det\mathbf{B}| \ge 1. \tag{237}$$

Together with the upper bound $|\det\mathbf{B}| \le 1$ obtained earlier, we have

$$|\det\mathbf{B}| = 1. \tag{238}$$

The equality case of Lemma 3.2 then gives $\mathbf{B} \in \mathcal{A}_\mathcal{P}$. Taking $\mathbf{Q} = \mathbf{B}$ proves the stated Det-Max identifiability result.

The source-entropy and reconstruction conclusion follows from Lemma C.1: as regularization and optimization error vanish, every source accumulation point of approximate global maximizers of Φ_ϵ is an exact Det-Max optimizer and hence has the same form.

Finally, suppose that $\mathcal{A}_\mathcal{P} \subseteq \mathrm{SP}(r)$. Then $\mathbf{Q} = \mathbf{\Pi}\mathbf{D}$ for some permutation matrix $\mathbf{\Pi}$ and diagonal sign matrix $\mathbf{D}$. Since $\mathbf{D}^{-1} = \mathbf{D}$ and $\mathbf{\Pi}^{-1} = \mathbf{\Pi}^\top$, we have

$$\mathbf{Q}^{-1} = \mathbf{D}\mathbf{\Pi}^\top. \tag{239}$$

Substituting these expressions for $\mathbf{Q}$ and $\mathbf{Q}^{-1}$ into Equation (20) gives the stated signed-permutation form.

□

G CANONICAL POLYTOPE CERTIFICATES

We now verify the polytope and arrangement conditions for four canonical source domains of Tatli & Erdogan (2021b) (see Appendix B.4). Throughout this section, let

$$\mathcal{B}_1^r := \{\mathbf{x} \in \mathbb{R}^r : \|\mathbf{x}\|_1 \le 1\}, \qquad \mathcal{B}_\infty^r := [-1,1]^r, \tag{240}$$

$$\mathcal{B}_{1,+}^r := \{\mathbf{x} \in \mathbb{R}^r : \mathbf{x} \ge \mathbf{0},\ \mathbf{1}^\top \mathbf{x} \le 1\}, \qquad \mathcal{B}_{\infty,+}^r := [0,1]^r. \tag{241}$$

No affine normalization is applied to these sets. For $T, A \subseteq [r]$, write

$$E_T := \operatorname{span}\{\mathbf{e}_i : i \in T\}, \qquad \mathbf{1}_A := \sum_{i \in A} \mathbf{e}_i, \tag{242}$$

and let

$$\mathrm{P}(r) := \{\mathbf{\Pi} \in \mathbb{R}^{r \times r} : \mathbf{\Pi} \text{ is a permutation matrix}\}. \tag{243}$$

We first record a general criterion that will be convenient for proving two-relation-completeness.

Corollary G.1 (of Proposition F.1 for q=2)**.** *An arrangement $\mathcal{L} = \{L_a\}_{a=1}^J$ is 2-relation-complete if and only if, for every real vector space W and every collection of linear maps $T_a : L_a \to W$ that agree on pairwise intersections,*

$$T_a \mathbf{x} = T_b \mathbf{x} \qquad \textit{for all } \mathbf{x} \in L_a \cap L_b,$$

there exists a linear map

$$T : \sum_{a=1}^{J} L_a \to W$$

such that $T|_{L_a} = T_a$ for every $a \in [J]$. This extension is unique.

Proof. Apply Proposition F.1 with $q = 2$. A relation supported on a single block is necessarily zero. For two distinct blocks a, b, a relation $\mathbf{z}_a + \mathbf{z}_b = \mathbf{0}$ has the form

$$\mathbf{z}_a = \mathbf{x}, \qquad \mathbf{z}_b = -\mathbf{x}, \qquad \mathbf{x} \in L_a \cap L_b.$$

Thus preservation of every such relation is equivalent to

$$T_a \mathbf{x} - T_b \mathbf{x} = \mathbf{0} \qquad \text{for every } \mathbf{x} \in L_a \cap L_b.$$

Consequently, Equation (53) for $q = 2$ is exactly agreement on pairwise intersections. The result follows from Proposition F.1. □

Proposition G.2 (Certificates for the canonical source polytopes)**.** *Let $r \ge 2$, fix $0 \le k_0 < r$, and set $p := k_0 + 1$. After faces whose affine hull contains the origin and duplicate face spans are removed, the following statements hold.*

(i) For both $\mathcal{B}_1^r$ and $\mathcal{B}_{1,+}^r$,

$$\mathcal{L}(\mathcal{B}_1^r, k_0) = \mathcal{L}(\mathcal{B}_{1,+}^r, k_0) = \{E_T : T \subseteq [r],\ |T| = p\}. \tag{244}$$

This arrangement has $J = \binom{r}{p}$, is 2-relation-complete, and satisfies

$$d_2 = \min\{r, 2p\}. \tag{245}$$

(ii) For $\mathcal{B}_\infty^r$ and $k_0 = 0$, choose the representatives

$$\mathbf{c}_\mathbf{x} := (1, \mathbf{x}) \in \mathbb{R}^r, \qquad \mathbf{x} \in \{\pm 1\}^{r-1}. \tag{246}$$

Then

$$\mathcal{L}(\mathcal{B}_\infty^r, 0) = \{\mathbb{R}\mathbf{c}_\mathbf{x} : \mathbf{x} \in \{\pm 1\}^{r-1}\}. \tag{247}$$

Thus $J = 2^{r-1}$. For $r \ge 3$, this arrangement is 4-relation-complete and

$$d_4 = \min\{r, 4\}. \tag{248}$$

(iii) For $\mathcal{B}^r_\infty$ and $1 \le k_0 \le r-2$, a face span is indexed by a free set $I \subset [r]$, $|I| = k_0$, and an unoriented sign pattern $\bar{\sigma} = \{\sigma, -\sigma\}$, where $\sigma \in \{\pm 1\}^{I^c}$. Specifically,

$$L_{I,\bar{\sigma}} = E_I + \operatorname{span}\{\mathbf{a}_{I,\sigma}\}, \qquad \mathbf{a}_{I,\sigma} := \sum_{j \notin I} \sigma_j \mathbf{e}_j. \tag{249}$$

These spans are distinct, so

$$J = \binom{r}{k_0} 2^{r-k_0-1}. \tag{250}$$

The arrangement is 2-relation-complete and satisfies

$$d_2 \le \min\{r, 2p\}. \tag{251}$$

When $r \ge 2p$, equality $d_2 = 2p$ holds. If $k_0 = r-1$, all facet spans equal $\mathbb{R}^r$; after duplicate removal the arrangement is $\{\mathbb{R}^r\}$, which is trivially 2-relation-complete and has $d_2 = r$.

(iv) For $\mathcal{B}^r_{\infty,+}$, a retained face has a free set $I \subset [r]$, $|I| = k_0$, and a nonempty set $A \subseteq I^c$ of coordinates fixed to one. Its span is

$$L_{I,A} = E_I + \operatorname{span}\{\mathbf{1}_A\}. \tag{252}$$

The number of distinct spans is

$$J_+(r, k_0) = \binom{r}{p} + \binom{r}{k_0}\left(2^{r-k_0} - 1 - (r-k_0)\right). \tag{253}$$

For $k_0 = 0$, the arrangement is

$$\mathcal{L}(\mathcal{B}^r_{\infty,+}, 0) = \{\mathbb{R}\mathbf{1}_A : \varnothing \ne A \subseteq [r]\}. \tag{254}$$

It is 3-relation-complete and satisfies

$$d_3 = \min\{r, 3\}. \tag{255}$$

For $1 \le k_0 < r$, the arrangement is 2-relation-complete and

$$d_2 = \min\{r, 2p\}. \tag{256}$$

(v) The linear automorphism groups of the four polytopes are

$$\mathcal{A}_{\mathcal{B}^r_1} = \mathcal{A}_{\mathcal{B}^r_\infty} = \mathrm{SP}(r), \tag{257}$$

$$\mathcal{A}_{\mathcal{B}^r_{1,+}} = \mathcal{A}_{\mathcal{B}^r_{\infty,+}} = \mathrm{P}(r). \tag{258}$$

Consequently, each listed polytope and arrangement satisfies Assumption 3.1 whenever the corresponding displayed dimension budget is at most M.

Proof. We divide the proof into the face-span descriptions, the relation certificates, the dimension budgets, and the automorphism groups.

The two ℓ_1 arrangements. A k_0-dimensional face of $\mathcal{B}^r_1$ is obtained by choosing a support $T \subseteq [r]$, $|T| = p$, and one sign for each coordinate in T. Its linear span is E_T, independently of the chosen sign pattern. Hence the distinct face spans are precisely the spaces in Equation (244).

The vertices of $\mathcal{B}^r_{1,+}$ are $\mathbf{0}, \mathbf{e}_1, \dots, \mathbf{e}_r$. A k_0-dimensional face is the convex hull of p of these vertices. If it contains $\mathbf{0}$, then its affine hull contains the origin and the face is omitted from the defining list. Every retained face therefore has the form

$$\operatorname{conv}\{\mathbf{e}_i : i \in T\}, \qquad |T| = p, \tag{259}$$

and its span is E_T. This proves Equation (244) and $J = \binom{r}{p}$.

To prove 2-relation-completeness, use Corollary G.1. Let $T_T : E_T \to W$ be linear maps that agree on pairwise intersections. For each $i \in [r]$, choose any p-subset T containing i and define

$$\mathbf{w}_i := T_T \mathbf{e}_i. \tag{260}$$

This definition is independent of the selected T, because two coordinate subspaces containing $\mathbf{e}_i$ intersect on at least the line $\mathbb{R}\mathbf{e}_i$. Define

$$T : \mathbb{R}^r \to W, \qquad T\left(\sum_{i=1}^r x_i \mathbf{e}_i\right) := \sum_{i=1}^r x_i \mathbf{w}_i. \tag{261}$$

For every p-subset T, this map agrees with T_T on the basis $\{\mathbf{e}_i : i \in T\}$, and hence on all of E_T. The pairwise compatibility criterion proves 2-relation-completeness.

Face spans of the centered cube. A k_0-face of $[-1,1]^r$ has a set I of k_0 free coordinates and a sign vector $\sigma \in \{\pm 1\}^{I^c}$ on its fixed coordinates. Its span is the space in Equation (249). Replacing σ by $-\sigma$ does not change this span. When $|I^c| \geq 2$, the coordinate axes contained in $L_{I,\bar{\sigma}}$ are exactly those indexed by I: if $j \notin I$, the vector $\mathbf{e}_j$ cannot be written as an element of E_I plus a multiple of $\mathbf{a}_{I,\sigma}$, since the latter vector has a nonzero coordinate at every index in I^c. Thus the span determines I. After I is known, its projection onto E_{I^c} determines the line $\mathbb{R}\sigma$, and hence $\bar{\sigma}$. This proves distinctness and the stated value of J for $1 \leq k_0 \leq r-2$. The cases $k_0 = 0$ and $k_0 = r-1$ follow directly from the same formula.

We next prove 2-relation-completeness when $1 \leq k_0 \leq r-2$. Let

$$T_{I,\bar{\sigma}} : L_{I,\bar{\sigma}} \to W \tag{262}$$

be pairwise-compatible linear maps. For every coordinate i, choose a cell for which $i \in I$ and define

$$\mathbf{w}_i := T_{I,\bar{\sigma}} \mathbf{e}_i. \tag{263}$$

Any two such cells intersect on $\mathbb{R}\mathbf{e}_i$, so $\mathbf{w}_i$ is well-defined. Let $T : \mathbb{R}^r \to W$ be the linear map satisfying $T\mathbf{e}_i = \mathbf{w}_i$.

Choose an orientation σ of $\bar{\sigma}$ and define the defect

$$\boldsymbol{\delta}_{I,\sigma} := T_{I,\bar{\sigma}} \mathbf{a}_{I,\sigma} - \sum_{j \notin I} \sigma_j \mathbf{w}_j. \tag{264}$$

We show that this defect is invariant under elementary exchanges. Fix $i \in I$ and $j \notin I$, set

$$I' := (I \setminus \{i\}) \cup \{j\}, \tag{265}$$

and choose $\tau \in \{\pm 1\}^{(I')^c}$ so that

$$\tau_\ell = \sigma_\ell \quad \text{for } \ell \in I^c \setminus \{j\}, \tag{266}$$

while $\tau_i \in \{\pm 1\}$ is arbitrary. Then

$$\mathbf{a}_{I,\sigma} + \tau_i \mathbf{e}_i = \mathbf{a}_{I',\tau} + \sigma_j \mathbf{e}_j \in L_{I,\bar{\sigma}} \cap L_{I',\bar{\tau}}. \tag{267}$$

Pairwise compatibility on this common vector, followed by the definition of the $\mathbf{w}_i$, gives

$$\boldsymbol{\delta}_{I,\sigma} = \boldsymbol{\delta}_{I',\tau}. \tag{268}$$

Elementary exchanges connect any two free sets of size k_0. They also connect all oriented sign patterns. Indeed, to change the sign of a fixed coordinate $j \notin I$, exchange j with any free coordinate $i \in I$, and then exchange them back; the sign assigned to j in the second exchange can be chosen arbitrarily. It follows from Equation (268) that all oriented defects have one common value. On the other hand, $L_{I,\bar{\sigma}} = L_{I,\overline{-\sigma}}$, and linearity gives

$$\boldsymbol{\delta}_{I,-\sigma} = -\boldsymbol{\delta}_{I,\sigma}. \tag{269}$$

The common defect must therefore be zero. Hence

$$T_{I,\bar{\sigma}} \mathbf{a}_{I,\sigma} = T \mathbf{a}_{I,\sigma}. \tag{270}$$

The two maps also agree on every $\mathbf{e}_i$, $i \in I$, so $T|_{L_{I,\bar{\sigma}}} = T_{I,\bar{\sigma}}$. The pairwise compatibility criterion now proves 2-relation-completeness.

For $k_0 = 0$, distinct vertex lines intersect only at the origin, so two-supported relations are trivial. We instead show that the kernel is generated by four-supported square relations. Let $n := r-1$, let $\varepsilon_{\mathbf{x}}$ denote the formal basis vector for the block indexed by $\mathbf{x} \in \{\pm 1\}^n$, and let $\mathbf{x}^i$ denote $\mathbf{x}$ with coordinate i flipped. For distinct $i, j \in [n]$,

$$R_{\mathbf{x},i,j} := \varepsilon_{\mathbf{x}} + \varepsilon_{\mathbf{x}^{ij}} - \varepsilon_{\mathbf{x}^i} - \varepsilon_{\mathbf{x}^j} \tag{271}$$

is a block-sum relation, because

$$\mathbf{c}_{\mathbf{x}} + \mathbf{c}_{\mathbf{x}^{ij}} - \mathbf{c}_{\mathbf{x}^i} - \mathbf{c}_{\mathbf{x}^j} = \mathbf{0}. \tag{272}$$

Let $\mathcal{R}_4$ be the span of these square relations. A scalar family $(f_{\mathbf{x}})_{\mathbf{x} \in \{\pm 1\}^n}$ annihilates every square relation if and only if

$$f_{\mathbf{x}} + f_{\mathbf{x}^{ij}} - f_{\mathbf{x}^i} - f_{\mathbf{x}^j} = 0 \qquad \text{for all } \mathbf{x} \text{ and } i \neq j. \tag{273}$$

Expand f in the multilinear basis on $\{\pm 1\}^n$:

$$f_{\mathbf{x}} = \sum_{A \subseteq [n]} \widehat{f}_A \prod_{i \in A} x_i. \tag{274}$$

Condition Equation (273) says that every mixed second difference vanishes. It therefore forces $\widehat{f}_A = 0$ for $|A| \geq 2$. Hence

$$f_{\mathbf{x}} = \alpha_0 + \sum_{i=1}^{n} \alpha_i x_i. \tag{275}$$

Let $\mathbf{C} \in \mathbb{R}^{r \times 2^n}$ be the matrix whose columns are the vectors $\mathbf{c}_{\mathbf{x}} = (1, \mathbf{x})$. These affine scalar families are exactly the elements of $\mathrm{row}(\mathbf{C})$. Thus

$$\mathcal{R}_4^{\perp} = \mathrm{row}(\mathbf{C}) = (\ker \Sigma_{\mathcal{L}})^{\perp}, \tag{276}$$

and consequently $\mathcal{R}_4 = \ker \Sigma_{\mathcal{L}}$. This proves 4-relation-completeness.

Face spans of the positive cube. A face of $[0,1]^r$ has a free-coordinate set I, a set A of coordinates fixed to one, and all remaining coordinates fixed to zero. Its affine hull contains the origin exactly when $A = \varnothing$. Thus the retained faces are precisely those with $A \neq \varnothing$, and their spans are the spaces in Equation (252).

If $|A| = 1$, then $L_{I,A}$ is the coordinate subspace $E_{I \cup A}$. Every coordinate p-subspace is obtained in this way, and its p possible representations collapse to one span after duplicate removal. If $|A| \geq 2$, the coordinate axes contained in $L_{I,A}$ are exactly those indexed by I, so the subspace determines I. Its projection onto E_{I^c} then determines the line $\mathbb{R}\mathbf{1}_A$, and the positivity of $\mathbf{1}_A$ determines A. Therefore all pairs (I, A) with $|A| \geq 2$ produce distinct non-coordinate spans. Counting the coordinate and non-coordinate cases gives

$$J_+(r, k_0) = \binom{r}{p} + \binom{r}{k_0} \sum_{s=2}^{r-k_0} \binom{r-k_0}{s} \tag{277}$$

$$= \binom{r}{p} + \binom{r}{k_0} \left(2^{r-k_0} - 1 - (r-k_0)\right), \tag{278}$$

which proves Equation (253).

Suppose first that $k_0 = 0$. The blocks are indexed by the nonempty subsets of $[r]$, with block A mapped to $\mathbf{1}_A$. For disjoint nonempty sets A, B, the identity

$$\mathbf{1}_A + \mathbf{1}_B - \mathbf{1}_{A \cup B} = \mathbf{0} \tag{279}$$

gives a three-supported block-sum relation. Let $\mathcal{R}_3$ be the span of these relations. Modulo $\mathcal{R}_3$, repeated use of Equation (279) gives

$$[\varepsilon_A] = \sum_{i \in A} [\varepsilon_{\{i\}}] \qquad \text{for every nonempty } A \subseteq [r]. \tag{280}$$

Hence the quotient of the block space by $\mathcal{R}_3$ is spanned by the r singleton classes. The block-sum map sends these singleton classes to $\mathbf{e}_1, \ldots, \mathbf{e}_r$, which are linearly independent. The induced map from the quotient to $\mathbb{R}^r$ is therefore an isomorphism. Its kernel is zero, and consequently

$$\ker \Sigma_{\mathcal{L}} = \mathcal{R}_3. \tag{281}$$

This proves 3-relation-completeness.

Now suppose $k_0 \geq 1$. Let $T_{I,A} : L_{I,A} \to W$ be pairwise-compatible linear maps. The arrangement contains every coordinate p-subspace: for $|T| = p$, choose any $a \in T$, set $I = T \setminus \{a\}$, and set $A = \{a\}$. By the coordinate-subspace argument already proved, the maps on these coordinate cells are the restrictions of one linear map $T : \mathbb{R}^r \to W$. In particular, all the maps $T_{I,A}$ agree with T on E_I.

It remains to check the generator $\mathbf{1}_A$. We proceed by induction on $|A|$. The result holds for $|A| = 1$, because the cell is then a coordinate p-subspace. Suppose $|A| \geq 2$, choose $i \in I$ and $a \in A$, and define

$$I' := (I \setminus \{i\}) \cup \{a\}, \qquad A' := A \setminus \{a\}. \tag{282}$$

Then $|I'| = k_0$, $A' \neq \varnothing$, and $A' \subseteq (I')^c$. Moreover,

$$\mathbf{1}_A = \mathbf{e}_a + \mathbf{1}_{A'} \in L_{I,A} \cap L_{I',A'}. \tag{283}$$

Pairwise compatibility and the induction hypothesis give

$$T_{I,A}\mathbf{1}_A = T_{I',A'}\mathbf{1}_A \tag{284}$$
$$= T\mathbf{e}_a + T_{I',A'}\mathbf{1}_{A'} \tag{285}$$
$$= T\mathbf{e}_a + T\mathbf{1}_{A'} = T\mathbf{1}_A. \tag{286}$$

Thus every local map is the restriction of T. By Corollary G.1, the arrangement is 2-relation-complete.

Dimension budgets. Every arrangement subspace considered here has dimension p. Therefore any sum of two such spaces has dimension at most $\min\{r, 2p\}$. For the coordinate arrangement, this bound is attained by choosing two p-subsets whose union has cardinality $\min\{r, 2p\}$. This proves Equation (245). Since the positive-cube arrangement with $k_0 \geq 1$ contains all coordinate p-subspaces, the same argument proves Equation (256).

For the centered cube with $1 \leq k_0 \leq r-2$, the general upper bound gives Equation (251). Suppose $r \geq 2p$ and write $k_0 = p - 1$. Choose disjoint free sets

$$I_1 = \{1, \ldots, k_0\}, \qquad I_2 = \{k_0 + 1, \ldots, 2k_0\}. \tag{287}$$

Choose the first dense sign vector to have all positive signs. Choose the second to have all positive signs except for one sign among the at least two coordinates outside $I_1 \cup I_2$. Modulo $E_{I_1} + E_{I_2}$, the two dense sign vectors are linearly independent. Hence the two face spans have a $2p$-dimensional sum, and $d_2 = 2p$.

For the centered-cube vertex lines, a sum of at most four lines has dimension at most $\min\{r, 4\}$. The arrangement spans $\mathbb{R}^r$. For $r \geq 4$, four independent representatives are obtained by taking the all-positive vector and three vectors formed by flipping, respectively, the first, second, and third coordinates after the leading coordinate in $\mathbf{c}_\mathbf{x} = (1, \mathbf{x})$. Thus $d_4 = \min\{r, 4\}$. Similarly, for the positive-cube vertex lines, the singleton lines $\mathbb{R}\mathbf{e}_1, \mathbb{R}\mathbf{e}_2, \mathbb{R}\mathbf{e}_3$ show, whenever available, that the general upper bound is attained. Hence $d_3 = \min\{r, 3\}$.

Linear automorphism groups. The extreme points of $\mathcal{B}_1^r$ are $\{\pm\mathbf{e}_i : i \in [r]\}$. An invertible linear map preserving $\mathcal{B}_1^r$ must permute these vertices, so it is a signed permutation. Conversely, every signed permutation preserves $\mathcal{B}_1^r$. Therefore $\mathcal{A}_{\mathcal{B}_1^r} = \mathrm{SP}(r)$.

The ordinary polar of $\mathcal{B}_\infty^r$ is $\mathcal{B}_1^r$. If $\mathbf{Q}\mathcal{B}_\infty^r = \mathcal{B}_\infty^r$, origin-centered polar duality gives

$$\mathbf{Q}^{-\top}\mathcal{B}_1^r = \mathcal{B}_1^r. \tag{288}$$

The preceding result implies $\mathbf{Q}^{-\top} \in \mathrm{SP}(r)$, and hence $\mathbf{Q} \in \mathrm{SP}(r)$. The converse is immediate, proving $\mathcal{A}_{\mathcal{B}_\infty^r} = \mathrm{SP}(r)$.

The extreme points of $\mathcal{B}_{1,+}^r$ are $\mathbf{0}, \mathbf{e}_1, \ldots, \mathbf{e}_r$. A linear automorphism fixes $\mathbf{0}$, so it must permute $\mathbf{e}_1, \ldots, \mathbf{e}_r$. Hence $\mathcal{A}_{\mathcal{B}_{1,+}^r} = \mathrm{P}(r)$.

Finally, a linear automorphism of $[0,1]^r$ fixes the vertex $\mathbf{0}$ and maps the face lattice to itself. The one-dimensional faces incident to $\mathbf{0}$ are exactly

$$[\mathbf{0}, \mathbf{e}_1], \ldots, [\mathbf{0}, \mathbf{e}_r]. \tag{289}$$

They must be permuted. Since an automorphism also maps endpoints to endpoints, it follows that the vectors $\mathbf{e}_1, \ldots, \mathbf{e}_r$ are permuted. Thus $\mathcal{A}_{\mathcal{B}_{\infty,+}^r} = \mathrm{P}(r)$, completing the proof. □

The proposition may be summarized as follows.

G.1 Computational verification of finite instances

The relation certificates can also be checked directly for fixed (r, k_0). Let $\mathbf{U}_a \in \mathbb{R}^{r \times p}$ contain a basis of L_a, and form the block-sum matrix

$$\mathbf{U} := [\mathbf{U}_1 \quad \cdots \quad \mathbf{U}_J]. \tag{290}$$

Table 1: Face-span arrangement certificates for the canonical source polytopes. Here $p = k_0 + 1$, and $J_+(r, k_0)$ is defined in Equation (253).

Polytope and skeleton	Number of spans	Relation order	Budget
$\mathcal{B}_1^r$, any k_0	$\binom{r}{p}$	$q = 2$	$d_2 = \min\{r, 2p\}$
$\mathcal{B}_{1,+}^r$, any k_0	$\binom{r}{p}$	$q = 2$	$d_2 = \min\{r, 2p\}$
$\mathcal{B}_\infty^r$, $k_0 = 0$	2^{r-1}	$q = 4$	$d_4 = \min\{r, 4\}$
$\mathcal{B}_\infty^r$, $1 \le k_0 \le r-2$	$\binom{r}{k_0} 2^{r-k_0-1}$	$q = 2$	$d_2 \le \min\{r, 2p\}$
$\mathcal{B}_{\infty,+}^r$, $k_0 = 0$	$2^r - 1$	$q = 3$	$d_3 = \min\{r, 3\}$
$\mathcal{B}_{\infty,+}^r$, $1 \le k_0 < r$	$J_+(r, k_0)$	$q = 2$	$d_2 = \min\{r, 2p\}$

In these coordinates,

$$\dim \ker \Sigma_{\mathcal{L}} = \sum_{a=1}^{J} \dim L_a - \operatorname{rank}(\mathbf{U}). \tag{291}$$

For every $I \subseteq [J]$ with $|I| \le q$, compute a basis for the nullspace of the local block matrix

$$\mathbf{U}_I := [\, \mathbf{U}_a : a \in I \,], \tag{292}$$

and embed those nullspace vectors into the full block-coordinate space by inserting zeros outside I. If the span of all embedded local nullspaces has the dimension in Equation (291), then the arrangement is q-relation-complete. The same basis matrices give

$$d_q(\mathcal{L}) = \max_{\substack{I \subseteq [J] \\ 1 \le |I| \le q}} \operatorname{rank}(\mathbf{U}_I). \tag{293}$$

This provides a finite exact-arithmetic check of the theoretical certificates and a convenient way to tabulate small-dimensional examples.

Corollary G.3 (Det-Max identifiability for the canonical polytopes)**.** *Let $\mathcal{P}$ be one of the four polytopes in Equation* (240)*–Equation* (241)*, and suppose that Assumptions 3.3 and 3.4 hold. Assume also that $\mathbf{H}_g$ satisfies q-GP for the relation order in Proposition G.2. It is sufficient to impose the following observation-dimension conditions:*

(i) $2p \le M < r$ *for* $\mathcal{B}_1^r$ *and* $\mathcal{B}_{1,+}^r$*;*

(ii) $4 \le M < r$ *for the vertex-line skeleton of* $\mathcal{B}_\infty^r$*;*

(iii) $2p \le M < r$ *for the positive-dimensional face skeleton of* $\mathcal{B}_\infty^r$*;*

(iv) $3 \le M < r$ *for the vertex-line skeleton of* $\mathcal{B}_{\infty,+}^r$*;*

(v) $2p \le M < r$ *for the positive-dimensional face skeleton of* $\mathcal{B}_{\infty,+}^r$*.*

If $\mathbf{H}_g$ is drawn from an absolutely continuous distribution, the required q-GP condition holds almost surely under these dimension bounds by Proposition D.1.

Then every global maximizer of Equation (13) *recovers the ground-truth factors up to the corresponding linear automorphism. For $\mathcal{P} \in \{\mathcal{B}_1^r, \mathcal{B}_\infty^r\}$, there exist a permutation matrix $\mathbf{\Pi}$ and a diagonal sign matrix $\mathbf{D}$ such that*

$$\mathbf{S}^\star = \mathbf{\Pi}\mathbf{D}\mathbf{S}_g, \qquad \mathbf{H}^\star = \mathbf{H}_g\mathbf{D}\mathbf{\Pi}^\top. \tag{294}$$

For $\mathcal{P} \in \{\mathcal{B}_{1,+}^r, \mathcal{B}_{\infty,+}^r\}$, there exists a permutation matrix $\mathbf{\Pi}$ such that

$$\mathbf{S}^\star = \mathbf{\Pi}\mathbf{S}_g, \qquad \mathbf{H}^\star = \mathbf{H}_g\mathbf{\Pi}^\top. \tag{295}$$

Proof. The relation certificates and dimension bounds in Proposition G.2 verify Assumption 3.1. Together with q-GP, full cell population, and Det-Max scattering, all the hypotheses of Theorem 3.3 hold. The conclusion follows from Theorem 3.3 and the automorphism-group identities Equation (257)–Equation (258). □

H STRUCTURAL AMBIGUITY IN THE CANONICAL MODELS

Proposition H.1 (Canonical structural ambiguity)**.** *Assume the hypotheses of Theorem 3.1. For the sparse models suppose* $2p \le M < r$*. For positive-dimensional cube spans suppose* $2(k_0+1) \le M < r$*; for the symmetric and nonnegative cube vertex lines suppose respectively* $4 \le M < r$ *and* $3 \le M < r$*. The map* $\mathbf{B}$ *of every exact feasible factorization has the following form:*

$\mathcal{P}$	$\mathbf{B}$
$\mathcal{B}_1^r$	$\mathbf{\Pi D}$, $\mathbf{D}$ *invertible diagonal*
$\mathcal{B}_{1,+}^r$	$\mathbf{\Pi D}$, $\mathbf{D}$ *positive diagonal*
$\mathcal{B}_\infty^r$	$c\mathbf{Q}$, $c>0,\ \mathbf{Q}\in \mathrm{SP}(r)$
$\mathcal{B}_{\infty,+}^r$	$c\mathbf{\Pi}$, $c>0$.

Here $\mathbf{\Pi}$ *is a permutation matrix. These are restrictions on feasible transformations, whose gains must also respect the sampled amplitudes. No scattering assumption is used.*

Proof. For coordinate p-subspaces with $p < r$, the one-dimensional intersections of arrangement members are exactly the coordinate axes. Each axis can be obtained by intersecting all p-subspaces containing it. Thus an arrangement automorphism permutes axes and has the form $\mathbf{\Pi D}$. Every source coordinate is nonzero somewhere because $\mathbf{S}_g$ has full row rank. In the nonnegative model this makes each diagonal gain positive.

For symmetric cube vertex lines, write $\mathbf{b}_i^\top$ for row i of $\mathbf{B}$. Since $\mathbf{B}\sigma$ has equal absolute coordinates for every sign vector σ, expanding $(\mathbf{b}_i^\top\sigma)^2 = (\mathbf{b}_j^\top\sigma)^2$ gives

$$\|\mathbf{b}_i\|_2^2 = \|\mathbf{b}_j\|_2^2, \qquad b_{iu}b_{iv} = b_{ju}b_{jv} \quad (u \ne v).$$

If one row has at least three nonzero entries, these pair products force every row to be that row or its negative, contradicting invertibility. If a row has exactly two nonzero entries, all rows must be supported on those two positions, again contradicting invertibility when $r \ge 3$. Thus each row has one nonzero entry; the equal row norms give $\mathbf{B} = c\mathbf{Q}$.

For positive-cube vertex lines, every line has a designated nonzero nonnegative sample. Feasibility and the bijective line matching imply that $\mathbf{B}$ permutes the positive indicator rays. These rays generate the nonnegative orthant, so $\mathbf{B}$ preserves that orthant and permutes its extreme rays. It is therefore a positive monomial matrix. Preservation of the pair rays $\mathbb{R}(\mathbf{e}_i + \mathbf{e}_j)$ makes all diagonal gains equal.

Now let $k = k_0 \ge 1$. For a nonzero vector $\mathbf{x}$, write $\nu(\mathbf{x})$ for the number of arrangement subspaces containing it. This count is preserved by $\mathbf{B}$. For the symmetric cube, an axis has incidence

$$\nu(\mathbf{e}_i) = \binom{r-1}{k-1}2^{r-k-1}.$$

A vector with support size $s \le k$ has incidence $\binom{r-s}{k-s}2^{r-k-1}$, strictly smaller for $s \ge 2$. If $s > k$, membership requires at least $r-k$ coordinates sharing a nonzero absolute value. Since $r \ge 2k+3$, at most one magnitude group has that size, and $\nu(\mathbf{x}) \le \binom{r}{k}$. But

$$\frac{\binom{r-1}{k-1}2^{r-k-1}}{\binom{r}{k}} = \frac{k}{r}2^{r-k-1} > 1 \qquad (r \ge 2k+3).$$

Thus the axes are exactly the directions of largest incidence, and $\mathbf{B}$ is monomial. A span with free set I contains exactly the coordinate axes in I, so its image has free set $\pi(I)$. Its remaining tied coordinates force all diagonal magnitudes outside I to agree. Any two coordinates can occur outside one such I, proving $\mathbf{B} = c\mathbf{Q}$.

For the positive cube, put $n = r-k$ and use zero for out-of-range binomial coefficients. If S is the support of $\mathbf{x}$, every cell containing $\mathbf{x}$ also contains $\mathbf{1}_S$. For $s = |S|$, counting the distinct coordinate and non-coordinate spans gives

$$\nu(\mathbf{1}_S) = \binom{r}{k} + (2^n - 2 - n)\binom{r-s}{k-s} - (s-1)\binom{r-s}{k-s+1}.$$

Indeed, a free set containing S permits $2^n - 1 - n$ tied sets of size at least two. A free set omitting at least two elements of S permits exactly the tied set $S \setminus I$. Adding the coordinate spans gives the

displayed formula. Since $n \geq 3$, the coefficient $2^n - 2 - n$ is positive. Comparing with $s = 1$ shows that every support of size at least two has strictly smaller incidence. Thus $\mathbf{B}$ is again monomial. Preservation of spans with tied set $\{i, j\}$ forces equal signed gains for every pair, so $\mathbf{B} = c\mathbf{\Pi}$ with $c \neq 0$. Nonnegative nonzero samples force $c > 0$. □

For either cube model, let $a_g := \max_j \|\mathbf{S}_{g,:,j}\|_\infty > 0$. The proposition implies that feasible gains satisfy $0 < c \leq 1/a_g$, and every such gain with the appropriate component symmetry is feasible. Since the objective changes by $r \log c$, Det-Max selects $c = 1/a_g$. Consequently, one boundary-attaining source, $a_g = 1$, suffices for calibrated cube recovery under the structural hypotheses. Full scattering is sufficient but unnecessary in this special case. If $a_g < 1$, Det-Max expands the sources to the known boundary and does not recover their original absolute gain.

I A POLYTOPE WITH COORDINATE-MIXING AMBIGUITIES

This example gives a family of source polytopes for which preserving the face-span arrangement allows transformations that mix coordinates, whereas preserving the polytope excludes them. The construction works in every even dimension $r = 2m$. We first describe its face spans and give certificates for all k, then construct an incorrect exact feasible factorization for $k = 1$. The final part gives the corresponding projection formulas. Here $k = k_0$ is the face dimension used in the main text.

I.1 POLYTOPE AND FACE SPANS

Let $m \geq 2$, set $a = 6/5$, and partition the coordinates into pairs $\{2b-1, 2b\}$, $b \in [m]$. Define

$$\mathcal{P}_m = \operatorname{conv}\left(\{\pm a\mathbf{e}_i : i \in [2m]\} \cup \{s\mathbf{e}_{2b-1} + t\mathbf{e}_{2b} : b \in [m],\ s, t \in \{-1, 1\}\}\right). \tag{296}$$

Thus every listed vertex is supported on one coordinate pair. Equivalently, let

$$Q = \operatorname{conv}\{(\pm a, 0), (0, \pm a), (\pm 1, \pm 1)\} \subset \mathbb{R}^2, \tag{297}$$

and let $Q^{(b)}$ be its copy in the coordinate plane $E_b = \operatorname{span}\{\mathbf{e}_{2b-1}, \mathbf{e}_{2b}\}$. Then $\mathcal{P}_m = \operatorname{conv}(\bigcup_{b=1}^m Q^{(b)})$. The $8m$ points in Equation (296) are all vertices: signed coordinate functionals expose the coordinate vertices because $a > 1$, and signed sums of the two coordinates in a block expose its diagonal vertices because $2 > a$. Also, 0 is an interior point of $\mathcal{P}_m$.

For a two-dimensional block vector $\mathbf{x} = (x_1, x_2)$, define

$$\gamma_Q(\mathbf{x}) = \frac{1}{a} \max\{|x_1| + (a-1)|x_2|, (a-1)|x_1| + |x_2|\}. \tag{298}$$

The eight inequalities implicit in $\gamma_Q(\mathbf{x}) \leq 1$ are exactly the edge inequalities of Q. Writing $\mathbf{x}^{(b)}$ for block b,

$$\mathcal{P}_m = \left\{\mathbf{x} \in \mathbb{R}^{2m} : \sum_{b=1}^m \gamma_Q(\mathbf{x}^{(b)}) \leq 1\right\}. \tag{299}$$

Indeed, a convex combination of embedded block points satisfies this bound. Conversely, take weights $\gamma_Q(\mathbf{x}^{(b)})$ for the nonzero blocks and use the remaining weight on the origin to obtain a convex combination.

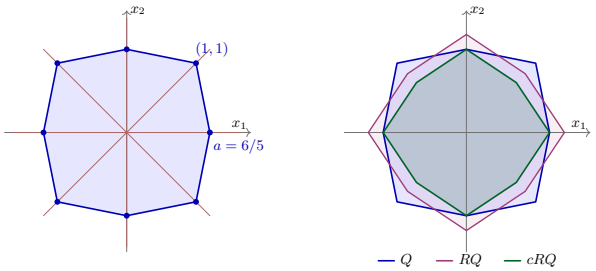
Each block has four unoriented vertex lines. Unit representatives and their radial bounds are

$$\begin{aligned} &\mathbf{d}_1 = (1, 0)^\top, && \mathbf{d}_2 = (0, 1)^\top, && \rho_1 = \rho_2 = a, \\ &\mathbf{d}_3 = 2^{-1/2}(1, 1)^\top, && \mathbf{d}_4 = 2^{-1/2}(1, -1)^\top, && \rho_3 = \rho_4 = \sqrt{2}. \end{aligned} \tag{300}$$

Write $\mathbf{d}_{b,j}$ for the embedding of $\mathbf{d}_j$ in block b.

Proposition I.1 (Face-span arrangement and certificate)**.** *Let $\mathcal{L}_{m,k}$ be the arrangement of distinct spans of the k-faces of $\mathcal{P}_m$, where $0 \leq k \leq 2m - 1$, and put $p = k + 1$. Its members are exactly the spaces*

$$L = \bigoplus_{b=1}^m L_b, \qquad L_b \in \{\{0\}, E_b, \mathbb{R}\mathbf{d}_{b,1}, \ldots, \mathbb{R}\mathbf{d}_{b,4}\}, \qquad \sum_{b=1}^m \dim L_b = p. \tag{301}$$



(a) Octagon and vertex lines (b) Rotation and feasible contraction

Figure 7: The two-dimensional block Q. Left: its four vertex lines are preserved by a $45°$ rotation, although the coordinate and diagonal vertices have different lengths. Right: the rotated octagon RQ extends outside Q, while $cRQ \subset Q$ for $c = 6/(5\sqrt{2})$. The contracted map mixes the two coordinates and sends Q into itself. The drawing shows one coordinate block of the higher-dimensional construction in Appendix I.3.

Their number is

$$J_{m,k} = \sum_{\substack{u,v\geq 0 \\ u+2v=p \\ u+v\leq m}} \frac{m!}{u!v!(m-u-v)!}\, 4^u. \tag{302}$$

For $k = 0$, a sufficient certificate is $q = 3$, $d_3 = 3$. For $1 \leq k \leq 2m-1$, a sufficient certificate is

$$q = 2, \qquad d_2 = \min\{2m, 2(k+1)\}. \tag{303}$$

In particular, for $k \geq 1$, the underdetermined dimension budget holds whenever $2(k+1) \leq M < 2m$.

Proof. **Description of the face spans.** A nonzero linear functional $\mathbf{u}$ has support value $h_{\mathcal{P}_m}(\mathbf{u}) = \max_b h_Q(\mathbf{u}^{(b)}) > 0$. In each block attaining this maximum, it exposes either a vertex or an edge of Q; the exposed face of $\mathcal{P}_m$ is the convex hull of these block faces. The whole octagon cannot be exposed at a positive support value. Let u blocks contribute a vertex and v blocks contribute an edge. Their spans are, respectively, a vertex line and the whole block plane. These spaces are in distinct coordinate blocks, so their sum has dimension $u + 2v$. All selected vertices lie in the same supporting hyperplane with a nonzero right-hand side. Their affine hull consequently has dimension $u + 2v - 1$. One can also see this by summing the dimensions of the block faces and adding $u + v - 1$ for their convex hull across disjoint blocks.

Conversely, choose any block pattern in Equation (301), a vertex on each selected line, and an edge in each selected full block. Scale their exposing functionals to have support value one, and set the functionals on inactive blocks to zero. This exposes a face with precisely the stated span. Every proper face avoids the origin in its affine hull because its support value is positive. Distinct block choices give distinct spans. Selecting u line blocks, v full blocks, and one of four lines in each line block proves Equation (302).

Relation completeness for $k \geq 1$. Suppose a family of linear maps $T_L : L \to W$, $L \in \mathcal{L}_{m,k}$, agrees on pairwise intersections. Because $p \geq 2$, for each b there is an arrangement member containing the whole plane E_b: allocate the remaining $p - 2$ dimensions among the other blocks. Define $T_b : E_b \to W$ as the restriction of any such map. Pairwise compatibility makes this choice independent of the arrangement member. Since $\mathbb{R}^{2m} = \bigoplus_b E_b$, these maps define a linear map $T : \mathbb{R}^{2m} \to W$ by $T\mathbf{x} = \sum_b T_b \mathbf{x}^{(b)}$.

Fix an arrangement member $L = \bigoplus_b L_b$. If $L_b \neq \{0\}$, intersect L with a member containing E_b. Compatibility implies $T_L|_{L_b} = T_b|_{L_b}$. The block spaces L_b span L, so $T_L = T|_L$. The pairwise compatibility criterion proves 2-relation-completeness. This also covers $p = 2m$, when the arrangement has just one member.

Relation completeness for $k = 0$. The arrangement consists of four lines in each block. Its relation kernel is the direct sum of the relation kernels within the individual blocks. In one block, the two relations expressing each diagonal direction in terms of the two coordinate directions have support

three. They are independent and span that block's two-dimensional relation kernel. Thus all relations are generated by relations on at most three lines. Any three lines span at most three dimensions, and this bound is attained by two coordinate lines in one block and a coordinate line in another. Hence $d_3 = 3$.

Dimension budget for $k \geq 1$. Two p-dimensional spans have sum dimension at most $\min\{2m, 2p\}$. Moreover, every coordinate p-subspace occurs in Equation (301): a block contributes zero, one coordinate line, or its whole plane according to how many of its coordinates were selected. Choose two coordinate supports of size p whose union has size $\min\{2m, 2p\}$. Their spans attain the bound, proving Equation (303). □

I.2 Polytope symmetries and feasibility of scattering

The block symmetries allow coordinate sign changes, swaps of the two coordinates within a block, and permutations of the blocks. The unique maximum-volume inscribed ellipsoid must be invariant under these maps. Central symmetry forces its center to be zero, coordinate sign changes force its shape matrix to be diagonal, and the remaining symmetries force all diagonal entries to agree. It is therefore a Euclidean ball.

The polar of Q is

$$Q^\circ = \{\mathbf{u} : \ |u_1|, |u_2| \leq 5/6, \ |u_1| + |u_2| \leq 1\}. \tag{304}$$

Its eight vertices are $(\pm 5/6, \pm 1/6)$ and $(\pm 1/6, \pm 5/6)$, each of squared norm $26/36$. Since $\mathcal{P}_m^\circ = (Q^\circ)^m$, every polar vertex has squared norm $26m/36$. The facet inequalities of $\mathcal{P}_m$ therefore all have supporting hyperplanes at distance $6/\sqrt{26m}$ from zero. It follows that

$$\mathcal{E}_{\mathcal{P}_m} = \{\mathbf{x} : \|\mathbf{x}\|_2 \leq 6/\sqrt{26m}\}, \tag{305}$$

and every facet is tangent to this ellipsoid.

In particular, including all $8m$ vertices among the source columns makes their convex hull equal to $\mathcal{P}_m$ and satisfies the scattering condition. Explicitly, the polar ellipsoid is the ball of radius $\sqrt{26m}/6$ and all polar vertices lie on its boundary. Strict convexity of the Euclidean norm implies that every convex combination of distinct polar vertices lies strictly inside that ball. Hence the intersection of the polar polytope with the polar ellipsoid boundary consists exactly of the polar vertices, as required by the polar part of scattering. This argument concerns scattering; the full-cell population condition requires additional samples, described below.

Every linear automorphism of $\mathcal{P}_m$ preserves its MVIE, so it is orthogonal. The coordinate vertices have norm $6/5$, and the diagonal vertices have norm $\sqrt{2}$. It must therefore permute the coordinate vertices, and is a signed permutation. It must also preserve the collection of coordinate pairs supporting the diagonal vertices. Conversely, every signed permutation preserving this collection maps the vertex set onto itself. Thus

$$\mathcal{A}_{\mathcal{P}_m} = \{\text{signed permutations preserving the partition into pairs}\} \subseteq \mathrm{SP}(2m). \tag{306}$$

I.3 A coordinate-mixing alternative for $k = 1$

Let

$$R = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}, \qquad T = \mathrm{diag}(R, I_2, \ldots, I_2). \tag{307}$$

The rotation R permutes the four vertex lines of Q and preserves its whole block plane. The face-span description Equation (301) therefore gives

$$T \in G_{\mathcal{L}_{m,k}} \quad \text{for every } 0 \leq k \leq 2m - 1. \tag{308}$$

However, $R(1,1)^\top = (0, \sqrt{2})^\top \notin Q$, since $\sqrt{2} > a$. Thus $T \notin \mathcal{A}_{\mathcal{P}_m}$. In particular, the difference between the two groups includes a map mixing source coordinates.

Now set

$$c = \frac{a}{\sqrt{2}} = \frac{6}{5\sqrt{2}}, \qquad B = \mathrm{diag}(cR, I_2, \ldots, I_2), \qquad cR = \frac{3}{5} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}. \tag{309}$$

A diagonal vertex of Q maps under cR to a coordinate vertex of Q. A coordinate vertex maps to a signed diagonal point with both coordinate magnitudes $a^2/2 = 18/25 < 1$, which lies in Q. By convexity, $cRQ \subset Q$, and hence $B\mathcal{P}_m \subseteq \mathcal{P}_m$. Multiplication by c does not change a span. Therefore B preserves every arrangement $\mathcal{L}_{m,k}$ and maps its bounded skeleton into itself.

For $k = 1$, let (H_g, S_g) be a ground-truth factorization with source columns in this skeleton. The pair

$$H_{\text{alt}} = H_g B^{-1}, \qquad S_{\text{alt}} = BS_g \tag{310}$$

has exactly the same observations and feasible source columns. If H_g is injective on every two-cell sum, so is H_{alt}: the map B^{-1} permutes those sums. Thus this is an exact feasible pair under the same local injectivity constraint. It mixes the first two source coordinates and is not a signed-permutation ambiguity.

The relation certificate is $q = 2$, $d_2 = 4$, so any $4 \le M < 2m$ is admissible. An absolutely continuously distributed $M \times 2m$ matrix satisfies the required injectivity almost surely, because there are only finitely many two-cell sums. For source samples satisfying scattering, S_g has full row rank and the alternative cannot coincide with QS_g for any $Q \in \mathcal{A}_{\mathcal{P}_m}$. Moreover,

$$\det B = c^2 = \frac{18}{25}, \qquad \frac{\det R_{S_{\text{alt}}}}{\det R_{S_g}} = (\det B)^2 = \frac{324}{625} = 0.5184. \tag{311}$$

The main identifiability theorem selects the polytope automorphisms among global maximizers. Equation (311) shows explicitly how determinant maximization excludes this incorrect feasible alternative. It is an objective comparison, not a convergence guarantee for a solver. The same construction of B applies at higher k whenever the corresponding budget and the remaining theorem assumptions hold.

The same alternative also has a smaller value of K_ϵ for every $\epsilon > 0$. Both pairs reconstruct exactly, so their residual terms are equal. Moreover, $\mathbf{R}_{\mathbf{S}_g} \succ 0$ and $B^\top B = \operatorname{diag}(c^2 I_2, I_{r-2}) \preceq I_r$, with $B^\top B \neq I_r$. Therefore

$$\begin{aligned}\det(\mathbf{R}_{\mathbf{S}_{\text{alt}}} + \epsilon I_r) &= \det(\mathbf{R}_{\mathbf{S}_g}^{1/2} B^\top B \mathbf{R}_{\mathbf{S}_g}^{1/2} + \epsilon I_r) \\ &< \det(\mathbf{R}_{\mathbf{S}_g} + \epsilon I_r).\end{aligned}$$

The equality follows from $\det(\epsilon I + XX^\top) = \det(\epsilon I + X^\top X)$ with $X = B\mathbf{R}_{\mathbf{S}_g}^{1/2}$; the strict inequality follows because the positive definite matrix in the second determinant exceeds that in the first by a nonzero positive semidefinite matrix. Hence Equation (15) excludes this alternative as well. This comparison does not assert global recovery at fixed ϵ.

For example, with $m = 6$, $r = 12$, and $k = 1$, there are

$$8m = 48 \text{ vertices}, \qquad J_{6,1} = 6 + 16\binom{6}{2} = 246 \text{ distinct cells}. \tag{312}$$

The full-cell population condition in the manuscript requires $J(p-1)+1 = 247$ vectors in linear general position in each cell. One sufficient construction is to draw 247 points with a two-dimensional density in each cell and then append all 48 vertices. The selected continuous samples are in linear general position almost surely; the appended vertices provide scattering and need not belong to those selected subsets. This construction uses $246 \cdot 247 + 48 = 60{,}810$ columns. This is a sufficient theorem-level sample count, not a claimed necessary sample size for numerical recovery. For $r = 12$, the budgets also allow $k = 2, 3, 4$, requiring respectively $M \ge 6, 8, 10$; their population counts use Equation (302).

I.4 Euclidean projection onto the skeleton

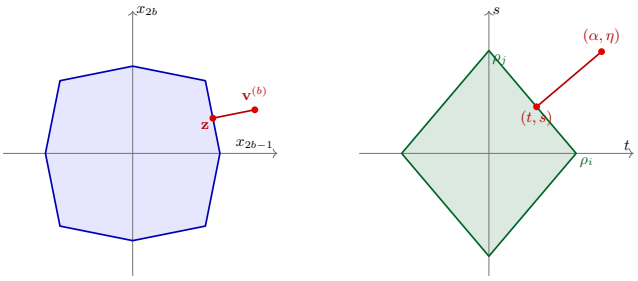
For $k = 1$, the distinct cells have only two forms. In a single full block,

$$C_b = \mathcal{P}_m \cap E_b = Q^{(b)}. \tag{313}$$

For two distinct blocks b, d and directions $i, j \in [4]$, the cell is

$$C_{b,i;d,j} = \left\{ t\mathbf{d}_{b,i} + s\mathbf{d}_{d,j} : \frac{|t|}{\rho_i} + \frac{|s|}{\rho_j} \le 1 \right\}. \tag{314}$$

The latter follows directly from Equation (299), since $\gamma_Q(t\mathbf{d}_i) = |t|/\rho_i$. Directions in different blocks are orthonormal. There are m octagon cells and $16\binom{m}{2}$ diamond cells.



(a) One full block (b) Lines in two different blocks

Figure 8: The two kinds of $k = 1$ span-cells, drawn in orthonormal coordinates on their planes. Left: a full block gives an octagon. Right: a line from each of two different blocks gives a diamond; the illustrated radii are $\rho_i = 6/5$ and $\rho_j = \sqrt{2}$. The shaded regions are the cells, including their interiors. Red points and segments illustrate Euclidean projection of the in-plane input. Components orthogonal to the plane contribute an additional constant squared error.

Projection onto an octagon cell. Let $\mathbf{v}^{(b)}$ be the input block. If $\gamma_Q(\mathbf{v}^{(b)}) \le 1$, retain it. Otherwise, project onto each of the eight edge segments and select the closest point. For an edge with endpoints $\mathbf{v}_0, \mathbf{v}_1$, its segment projection is

$$\mathbf{z} = \mathbf{v}_0 + \operatorname{clip}_{[0,1]}\left(\frac{\langle \mathbf{v}^{(b)} - \mathbf{v}_0, \mathbf{v}_1 - \mathbf{v}_0\rangle}{\|\mathbf{v}_1 - \mathbf{v}_0\|_2^2}\right)(\mathbf{v}_1 - \mathbf{v}_0). \tag{315}$$

The projection of an exterior point onto a polygon lies on its boundary, so this finite comparison is exact. Set all other blocks to zero. Its full squared distance is $\|\mathbf{v}\|_2^2 - \|\mathbf{v}^{(b)}\|_2^2 + \|\mathbf{v}^{(b)} - \mathbf{z}\|_2^2$.

Projection onto a cross-block diamond. Set $\alpha = \langle \mathbf{v}, \mathbf{d}_{b,i}\rangle$ and $\eta = \langle \mathbf{v}, \mathbf{d}_{d,j}\rangle$. The remaining problem is

$$\min_{t,s} \frac{1}{2}(t-\alpha)^2 + \frac{1}{2}(s-\eta)^2 \quad \text{subject to} \quad |t|/\rho_i + |s|/\rho_j \le 1. \tag{316}$$

If the input pair is feasible, take $(t, s) = (\alpha, \eta)$. Otherwise put $u_1 = |\alpha|$, $u_2 = |\eta|$, $w_1 = 1/\rho_i$, and $w_2 = 1/\rho_j$. The solution has matching signs and magnitudes

$$x_\ell = (u_\ell - \lambda w_\ell)_+, \qquad \sum_{\ell=1}^{2} w_\ell x_\ell = 1. \tag{317}$$

These equations follow from stationarity and complementary slackness for the convex problem in the nonnegative magnitudes. There is no iterative search needed: check the three nonempty active sets $A \subseteq \{1, 2\}$, using

$$\lambda_A = \frac{\sum_{\ell \in A} w_\ell u_\ell - 1}{\sum_{\ell \in A} w_\ell^2}. \tag{318}$$

Accept any set for which $\lambda_A \ge 0$, the quantities $u_\ell - \lambda_A w_\ell$ are nonnegative on A and nonpositive outside A. Zero quantities may be assigned to either set. Restore the signs to obtain t, s; a zero input has zero output. The full squared distance is

$$\|\mathbf{v}\|_2^2 - \alpha^2 - \eta^2 + (t-\alpha)^2 + (s-\eta)^2. \tag{319}$$

Select the least-distance candidate over all octagon and diamond cells. This is the exact Euclidean projection onto the $k = 1$ skeleton. Ties may be resolved arbitrarily. Computing the four directional inner products in each block and the input norm takes $O(m)$ operations; every candidate then has constant cost. The complete search takes $O(m^2)$ operations, followed by $O(m)$ to construct the selected output. The threshold step is the usual weighted ℓ_1-ball projection; the contribution here is the description of the cells and their use for this skeleton.

Other values of k. For $k = 0$, compare the $4m$ segment projections $\mathrm{clip}_{[-\rho_j,\rho_j]}(\langle \mathbf{v}, \mathbf{d}_{b,j}\rangle)\mathbf{d}_{b,j}$. For general k, Equation (301) and Equation (299) give a finite convex projection problem in each cell. Let F index its full blocks and S its line blocks, with chosen line direction j_b in each $b \in S$. Write the output as $\mathbf{y}_b$ on full blocks and $t_b\mathbf{d}_{j_b}$ on line blocks. Minimize

$$\frac{1}{2}\sum_{b\in F}\|\mathbf{y}_b - \mathbf{v}^{(b)}\|_2^2 + \frac{1}{2}\sum_{b\in S}\big(t_b - \langle \mathbf{v}^{(b)}, \mathbf{d}_{j_b}\rangle\big)^2 \tag{320}$$

subject to

$$\sum_{b\in F}\gamma_Q(\mathbf{y}_b) + \sum_{b\in S}|t_b|/\rho_{j_b} \le 1. \tag{321}$$

For candidate comparison, add the omitted orthogonal squared errors, namely $\frac{1}{2}\sum_{b\notin F\cup S}\|\mathbf{v}^{(b)}\|_2^2$ and $\frac{1}{2}\sum_{b\in S}(\|\mathbf{v}^{(b)}\|_2^2 - \langle \mathbf{v}^{(b)}, \mathbf{d}_{j_b}\rangle^2)$. The gauge is the maximum of eight linear forms, so introducing one gauge upper-bound variable per full block and one absolute-value upper-bound variable per line block turns this into a convex quadratic program with linear constraints. Comparing all $J_{m,k}$ cells is exact when these subproblems are solved exactly, and otherwise is subject to their numerical tolerance. We do not claim the $O(m^2)$ bound for this general-k search. At $k = 2m - 1$, there is one cell and this is projection onto $\mathcal{P}_m$.

J Online ER-CorInfoMax for Underdetermined PMF

In this section we adapt the correlative information maximization (CorInfoMax) framework of Bozkurt et al. (2023) to the underdetermined regime $M < r$. The resulting online algorithm admits a biologically plausible neural network interpretation as an *expansion network*: M input neurons project to $r > M$ output neurons, mirroring the dimensional expansion observed, for example, in the lateral geniculate nucleus (LGN) to primary visual cortex (V1) pathway.

We start from the source-entropy and reconstruction objective of Appendix C.3 (with $\lambda = 1$),

$$K_\epsilon(\mathbf{H}, \mathbf{S}) = \frac{1}{2}\log\det(\mathbf{R}_S + \epsilon\mathbf{I}_r) - \frac{1}{2}\log\det(\mathbf{R}_E(\mathbf{H}, \mathbf{S}) + \epsilon\mathbf{I}_M), \tag{322}$$

where $\epsilon > 0$ and the moments are *uncentered*:

$$\mathbf{R}_S = \frac{1}{N}\mathbf{S}\mathbf{S}^\mathsf{T}, \qquad \mathbf{R}_E(\mathbf{H}, \mathbf{S}) = \frac{1}{N}(\mathbf{Y} - \mathbf{H}\mathbf{S})(\mathbf{Y} - \mathbf{H}\mathbf{S})^\mathsf{T}.$$

For a prescribed polytope $\mathcal{P} \subset \mathbb{R}^r$ and face dimension $0 \le k_0 < r$, the source constraint is

$$\mathbf{S}_{:,k} \in \mathcal{C}, \qquad \mathcal{C} := \mathrm{Skel}_{k_0}(\mathcal{P}), \quad k = 1, \dots, N. \tag{323}$$

Thus the first term promotes source spread, the second penalizes reconstruction error, and the feasible set includes the face-span restriction in addition to membership in $\mathcal{P}$. This is the ER-CorInfoMax objective K_ϵ; its source-entropy term distinguishes it from symmetric CorInfoMax. We adapt the recurrent inference and Hebbian/anti-Hebbian construction of Bozkurt et al. (2023) to this objective and to $M < r$. We first leave the constraint nonlinearity abstract, then give its neural realization for the nonnegative ℓ_1-ball.

J.1 Online objective and neural drive

For an observation $\mathbf{y}(k) \in \mathbb{R}^M$, hold the current synapses $\mathbf{H}(k) \in \mathbb{R}^{M\times r}$ fixed during inference and write

$$\mathbf{e}(\mathbf{s}) = \mathbf{y}(k) - \mathbf{H}(k)\mathbf{s}. \tag{324}$$

Let $b_s = 1-\zeta_s$ and $b_e = 1-\zeta_e$, with $0 < \zeta_s, \zeta_e < 1$. Replace the batch moments by the exponentially weighted second moments, initialized by positive semidefinite matrices:

$$\begin{aligned}\mathbf{R}_s(k;\mathbf{s}) &= \zeta_s\mathbf{R}_s(k-1) + b_s\mathbf{s}\mathbf{s}^\mathsf{T},\\ \mathbf{R}_e(k;\mathbf{s}) &= \zeta_e\mathbf{R}_e(k-1) + b_e\mathbf{e}(\mathbf{s})\mathbf{e}(\mathbf{s})^\mathsf{T}.\end{aligned} \tag{325}$$

The past moments are fixed during inference; the accepted sample is incorporated once, after the neural response. The conditional online counterpart of Equation (322) is

$$\max_{\mathbf{s}\in\mathcal{C}} K_\epsilon^{(k)}(\mathbf{s}), \qquad K_\epsilon^{(k)}(\mathbf{s}) = \frac{1}{2}\log\det(\mathbf{R}_s(k;\mathbf{s}) + \epsilon\mathbf{I}_r) - \frac{1}{2}\log\det(\mathbf{R}_e(k;\mathbf{s}) + \epsilon\mathbf{I}_M). \quad (326)$$

This is a streaming adaptation: past residuals are not recomputed under each new value of $\mathbf{H}(k)$.

Differentiating the two log-determinants gives

$$\nabla_{\mathbf{s}} K_\epsilon^{(k)} = b_s(\mathbf{R}_s(k;\mathbf{s}) + \epsilon\mathbf{I}_r)^{-1}\mathbf{s} + b_e\mathbf{H}(k)^\mathsf{T}(\mathbf{R}_e(k;\mathbf{s}) + \epsilon\mathbf{I}_M)^{-1}\mathbf{e}(\mathbf{s}). \quad (327)$$

The reconstruction term has a positive sign because the minus sign in the objective combines with $\partial\mathbf{e}/\partial\mathbf{s} = -\mathbf{H}(k)$.

For the compact rate-network realization, let $\mathbf{B}_s(k) \simeq (\mathbf{R}_s(k-1) + \epsilon\mathbf{I}_r)^{-1}$ be the slowly adapting source-precision synapses. Neglecting ridge replenishment in the source-moment recursion, the matrix inversion lemma gives

$$b_s(\mathbf{R}_s(k;\mathbf{s}) + \epsilon\mathbf{I}_r)^{-1}\mathbf{s} \simeq \gamma_s(\mathbf{s})\mathbf{B}_s(k)\mathbf{s}, \qquad \gamma_s(\mathbf{s}) = \frac{b_s}{\zeta_s + b_s\mathbf{s}^\mathsf{T}\mathbf{B}_s(k)\mathbf{s}}. \quad (328)$$

Approximate the error precision by $\epsilon^{-1}\mathbf{I}_M$ when the residual second moment is small relative to $\epsilon\mathbf{I}_M$, and freeze or slowly modulate the source gain. This yields the compact neural drive

$$\widetilde{\mathbf{g}}(k;\iota) = c_s\mathbf{B}_s(k)\mathbf{s}(k;\iota) + c_e\mathbf{H}(k)^\mathsf{T}\mathbf{e}(k;\iota), \qquad c_e = \frac{b_e}{\epsilon}, \quad (329)$$

where $c_s > 0$ approximates γ_s during the response. In the long-memory regime $b_s\mathbf{s}^\mathsf{T}\mathbf{B}_s(k)\mathbf{s} \ll \zeta_s$, one may use $c_s = b_s/\zeta_s$. These approximations give recurrent source input and an error-correction input, rather than evaluating the exact gradient Equation (327) at every neural step.

J.2 Fast inference and slow synaptic learning for a general polytope

Let $\Pi_\mathcal{C}$ denote a selected Euclidean projection onto $\mathcal{C} = \mathrm{Skel}_{k_0}(\mathcal{P})$. Its implementation is left unspecified in the general algorithm. The fast neural response alternates error computation, integration of the drive Equation (329), and this constraint nonlinearity. After the response, $\mathbf{H}$ is updated by the error–source Hebbian product, and $\mathbf{B}_s$ by an anti-Hebbian product of the transformed activities $\mathbf{z} = \mathbf{B}_s\mathbf{s}$. Algorithm 1 summarizes these two timescales without introducing a separate numerical projection routine.

Algorithm 1 Online K_ϵ-based neural learning for a general $\mathcal{P}$

Require: Stream $\{\mathbf{y}(k)\}$; $\mathcal{C} = \mathrm{Skel}_{k_0}(\mathcal{P})$; gains $c_s, c_e > 0$, $\zeta_s \in (0,1)$; step sizes η_s, μ_H.

1: Initialize $\mathbf{H}(1)$ and $\mathbf{B}_s(1) \succ 0$.
2: **for** each observation $\mathbf{y}(k)$ **do**
3: Initialize $\mathbf{s}(k;0) \in \mathcal{C}$; hold the synapses fixed.
4: Run the fast neural dynamics until settling or the iteration limit:
$\mathbf{e}(k;\iota) = \mathbf{y}(k) - \mathbf{H}(k)\mathbf{s}(k;\iota)$,
$\widetilde{\mathbf{g}}(k;\iota) = c_s\mathbf{B}_s(k)\mathbf{s}(k;\iota) + c_e\mathbf{H}(k)^\mathsf{T}\mathbf{e}(k;\iota)$,
$\mathbf{s}(k;\iota+1) = \Pi_\mathcal{C}(\mathbf{s}(k;\iota) + \eta_s(\iota)\widetilde{\mathbf{g}}(k;\iota))$.
5: Accept $\mathbf{s}(k)$ and set $\mathbf{e}(k) = \mathbf{y}(k) - \mathbf{H}(k)\mathbf{s}(k)$.
6: Update generative synapses: $H_{ij}(k+1) = H_{ij}(k) + \mu_H(k)e_i(k)s_j(k)$.
7: Form $\mathbf{z}(k) = \mathbf{B}_s(k)\mathbf{s}(k)$ and $\alpha_s(k) = \dfrac{1-\zeta_s}{\zeta_s + (1-\zeta_s)\mathbf{s}(k)^\mathsf{T}\mathbf{z}(k)}$.
8: Update lateral synapses: $[\mathbf{B}_s(k+1)]_{j\ell} = \dfrac{1}{\zeta_s}\left([\mathbf{B}_s(k)]_{j\ell} - \alpha_s(k)z_j(k)z_\ell(k)\right)$.
9: **end for**

The error-precision approximation makes the current-sample $\mathbf{H}$-gradient a positive scalar multiple of $\mathbf{e}(k)\mathbf{s}(k)^\mathsf{T}$, with its gain absorbed in $\mu_H(k)$. The lateral update is a synaptic scaling plus a

rank-one anti-Hebbian decrement $-\alpha_s z_i z_j$; its normalization denominator is retained. This compact inverse-moment rule neglects the isotropic ridge-replenishment term $(1-\zeta_s)\epsilon\mathbf{I}_r$. It is therefore an approximation to a fixed-ϵ precision update, not an exact identity for that objective. The reduced network maintains only $\mathbf{H}$ and $\mathbf{B}_s$: it requires neither matrix inversions during inference nor a separate error-precision matrix.

The same learned pathways apply for every $\mathcal{P}$; only $\Pi_{\mathcal{C}}$ changes. In particular, replacing it by $\Pi_{\mathcal{P}}$ would omit the structural constraint. We do not assume a universal local circuit for arbitrary skeletons. The following example supplies a concrete constraint circuit without changing either synaptic learning rule.

J.3 Nonnegative ℓ_1-ball: a p-winner, activity-budget circuit

Domain and face-span constraint. For $\rho > 0$, take

$$\mathcal{P}_\rho = \{\mathbf{s} \in \mathbb{R}^r : \mathbf{s} \succeq \mathbf{0},\ \mathbf{1}^\mathsf{T}\mathbf{s} \le \rho\}. \tag{330}$$

The unit nonnegative ℓ_1-ball is obtained with $\rho = 1$. For a face dimension k_0, set $p = k_0 + 1$. A retained face $\mathrm{conv}\{\rho\mathbf{e}_j : j \in T\}$, $|T| = p$, has affine dimension $p-1$ and coordinate span $\mathrm{span}\{\mathbf{e}_j : j \in T\}$. Consequently,

$$\mathcal{C}_{p,\rho} = \mathrm{Skel}_{p-1}(\mathcal{P}_\rho) = \{\mathbf{s} \succeq \mathbf{0} : \mathbf{1}^\mathsf{T}\mathbf{s} \le \rho,\ \|\mathbf{s}\|_0 \le p\}. \tag{331}$$

These are bounded span-cells, not just boundary faces: fewer than p sources may be active and the activity budget need not be exhausted. A single activity-budget interneuron does not enforce the additional support-size bound.

Constraint nonlinearity. For the candidate $\mathbf{v} = \mathbf{s} + \eta_s\widetilde{\mathbf{g}}$, select the p largest strictly positive entries, or all such entries when there are fewer than p. Denote their indices by $T_p(\mathbf{v})$, resolving equal-value ties consistently, and set $a_j = \mathbf{1}\{j \in T_p(\mathbf{v})\}$. The projection used in Algorithm 1 is

$$\boxed{[\Pi_{\mathcal{C}_{p,\rho}}(\mathbf{v})]_j = a_j[v_j - \beta]_+, \qquad [u]_+ := \max(u, 0).} \tag{332}$$

Here $\beta = 0$ if $\sum_j a_j[v_j]_+ \le \rho$; otherwise $\beta > 0$ is determined by

$$\sum_j a_j[v_j - \beta]_+ = \rho. \tag{333}$$

This is the nonnegative sparse projection in Appendix K.1 (Duchi et al., 2008; Kyrillidis et al., 2013). Its two operations have distinct roles: selecting the largest positive candidates gives an optimal allowable support, and the common threshold solves the convex amplitude projection on that support. The exchange argument and fixed-support optimality conditions in Appendix K.1 establish exactness. Selection must precede amplitude inhibition; truncating a full-ball projection is not the same operation.

Winner-control interneuron. The additional interneuron I_p controls a common recruitment threshold $\kappa \ge 0$. For a fixed candidate $\mathbf{v}$, local comparators and the pooled count feedback implement

$$a_j^t = \mathbf{1}\{v_j > \kappa^t\}, \qquad \kappa^{t+1} = \left[\kappa^t + \nu_p^t\left(\sum_j a_j^t - p\right)\right]_+, \tag{334}$$

away from threshold ties. Here t indexes the faster constraint response. The interneuron pools *binary recruitment signals*, compares their sum with the tonic budget p, and inhibits recruitment when too many candidates are selected. The threshold update is projected subgradient descent on the scalar function $D_p(\kappa) = p\kappa + \sum_j [v_j - \kappa]_+$. When more than p candidates are positive and $v_{(p)} > v_{(p+1)}$, any $v_{(p+1)} \le \kappa < v_{(p)}$ selects the top p. With at most p positive candidates, $\kappa = 0$ selects just those candidates. Equal values at the selection boundary require a binary competitive tie rule; a common threshold alone cannot split them.

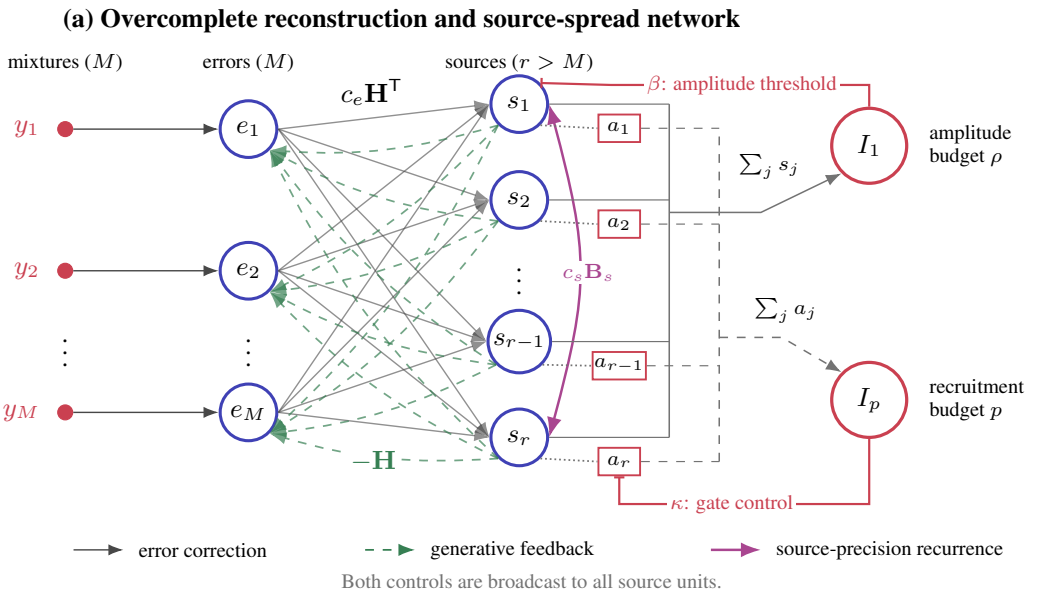


Figure 9: **Nonnegative ℓ_1 face-span-skeleton network for K_ϵ.** (a) Mixture, error, and source layers have dimensions M, M, r, with $r > M$. Black and dashed green paths carry reciprocal error correction and generative prediction; magenta recurrence carries the source-precision drive. Gray pooling and red constraint pathways are fixed. I_1 pools amplitudes and controls β; I_p pools binary recruitment gates and controls κ. The representative return paths are broadcast to all source units. (b) A local gate multiplies the thresholded amplitude, implementing $s_j^+ = a_j[v_j - \beta]_+$. Settled, tie-resolved controls enforce nonnegativity, the budget ρ, and at most $p = k_0 + 1$ active sources. Colors distinguish pathways and plasticity classes, not individual synaptic signs.

Activity-budget interneuron. After recruitment settles, hold the binary gates fixed. A second interneuron I_1 controls the analog threshold through

$$s_j^t = a_j[v_j - \beta^t]_+, \qquad \beta^{t+1} = \left[\beta^t + \nu_1\left(\sum_j s_j^t - \rho\right)\right]_+. \tag{335}$$

This neuron pools *source amplitudes*, rather than gates. Excess total activity increases the common subtractive inhibition; unused budget decreases it. For fixed $\mathbf{v}$ and gates, the update minimizes $D_1(\beta) = \rho\beta + \frac{1}{2}\sum_j a_j[v_j - \beta]_+^2$ over $\beta \geq 0$. At settling,

$$\beta \geq 0, \qquad \sum_j s_j \leq \rho, \qquad \beta\left(\sum_j s_j - \rho\right) = 0, \tag{336}$$

which supplies exactly the threshold in Equation (332). For this fixed-candidate scalar response, $0 < \nu_1 < 2/p$ is a sufficient step-size choice.

Importantly, κ controls the *gate*, not the retained amplitude: the output is $a_j[v_j - \beta]_+$, not $[v_j - \beta - \kappa]_+$. For example, $\mathbf{v} = (0.4, 0.3, 0.2)^\mathsf{T}$ with $p = 2$, $\rho = 1$ gives $(0.4, 0.3, 0)^\mathsf{T}$; recruitment suppresses the

third source without shrinking the other two. A recruited candidate may subsequently be silenced by β, so I_p must count the gates, not the number of nonzero post-inhibition amplitudes.

For each neural candidate, recruitment is resolved first and the activity-budget response second, with $\mathbf{v}$ fixed during these constraint operations. Their settled output replaces the abstract $\Pi_{\mathcal{C}}$ in Algorithm 1. At the next source step, candidates are recomputed for *all* neurons, including silent ones, allowing the support to change. Exact projection refers to settled, tie-resolved binary gates and amplitude control. Finite-time or simultaneous controller dynamics, including finite-step count feedback that may chatter, are approximations to this operation.

Neural architecture and local learning. Figure 9 retains the mixture–error–source organization. Each error neuron subtracts the generative prediction, $e_i = y_i - \sum_j H_{ij} s_j$. Each source neuron integrates

$$v_j = s_j + \eta_s \left[c_e \sum_{i=1}^{M} H_{ij} e_i + c_s \sum_{\ell=1}^{r} (B_s)_{j\ell} s_\ell \right], \tag{337}$$

and applies the gated ReLU in Equation (332). The two red interneurons provide fixed constraint feedback to different local operations: I_p controls recruitment, while I_1 controls amplitude. Their pooling connections are fixed; only their thresholds evolve during inference. Learning still consists solely of the error–source Hebbian update of $\mathbf{H}$ and the transformed-activity anti-Hebbian update of $\mathbf{B}_s$ in Algorithm 1.

Interpretation and scope. The circuit is a rate/compartment-level model with local comparison, gating, rectification, and pooled inhibition. Its activity-product plasticity assumes access to the transformed signals z_i, z_j and a shared scalar gain; the reciprocal $\mathbf{H}, \mathbf{H}^\mathsf{T}$ pathways retain a weight-symmetry assumption, and signed error signals require a signed or opponent-population encoding. These assumptions are separate from constraint enforcement.

K Algorithm and Details on Numerical Experiments

K.1 Full-data projected optimization for synthetic recovery

The synthetic recovery experiments use alternating least-squares and projected-gradient updates for the source-entropy and reconstruction objective of Equation (15).

$$K_{\epsilon,\lambda}(\mathbf{H}, \mathbf{S}) = \frac{\lambda}{2} \log\det(\mathbf{R}_{\mathbf{S}} + \epsilon \mathbf{I}_r) - \frac{1}{2} \log\det(\mathbf{R}_{\mathbf{E}}(\mathbf{H}, \mathbf{S}) + \epsilon \mathbf{I}_M). \tag{338}$$

For a source iterate $\mathbf{S}_t$, the mixing update is

$$\mathbf{H}_t = \mathbf{Y}\mathbf{S}_t^\dagger, \qquad \mathbf{E}_t = \mathbf{Y} - \mathbf{H}_t \mathbf{S}_t. \tag{339}$$

The least-squares residual minimizes the residual second moment in the positive-semidefinite order, so Equation (339) maximizes $K_{\epsilon,\lambda}$ over $\mathbf{H}$. With $\mathbf{H}_t$ fixed, the source gradient and projected update are

$$\mathbf{G}_t = \frac{1}{N} \left[\lambda (\mathbf{S}_t \mathbf{S}_t^\top / N + \epsilon \mathbf{I}_r)^{-1} \mathbf{S}_t + \mathbf{H}_t^\top (\mathbf{E}_t \mathbf{E}_t^\top / N + \epsilon \mathbf{I}_M)^{-1} \mathbf{E}_t \right], \tag{340}$$

$$\mathbf{S}_{t+1} = \Pi_{\mathrm{Skel}_{k_0}(\mathcal{P})}(\mathbf{S}_t + \alpha_t \mathbf{G}_t), \qquad \alpha_t = \mu \gamma^{\max(0, t - t_0)}. \tag{341}$$

All second moments are uncentered and computed from the full set of N samples. The step size is constant through $t = t_0$ and decays geometrically thereafter. Linear solves evaluate the gradient, and a QR-based least-squares solve evaluates Equation (339). Algorithm 2 summarizes one initialization.

Skeleton projection. For the union of cells in Equation (10), projection selects $\Pi_{C_a}(\mathbf{v})$ with the smallest squared distance to the input. The canonical geometries admit sorted candidate searches. For $\mathcal{B}_1^r$ and $\mathcal{B}_{1,+}^r$, retain the $p = k_0 + 1$ largest magnitudes or positive entries, respectively, and project onto the unit ℓ_1 ball on that support. This combines hard support selection (Blumensath & Davies, 2008) with threshold-based simplex/ℓ_1 projection (Duchi et al., 2008). For a

Algorithm 2 Full-data synthetic recovery

Require: Observations $\mathbf{Y} \in \mathbb{R}^{M\times N}$; polytope $\mathcal{P}$; face dimension k_0; $T, \epsilon, \lambda, \mu, t_0, \gamma$.
Draw $\mathbf{U} \in \mathbb{R}^{r\times N}$ with independent uniform entries from the initialization interval in Table 2.
$\mathbf{S}_0 \leftarrow \Pi_{\mathrm{Skel}_{k_0}(\mathcal{P})}(\Pi_{\mathcal{P}}(\mathbf{U}))$; $K_\star \leftarrow -\infty$.
for $t = 0, \dots, T$ **do**
  Fit $\mathbf{H}_t \leftarrow \mathbf{Y}\mathbf{S}_t^\dagger$ and form $\mathbf{E}_t \leftarrow \mathbf{Y} - \mathbf{H}_t\mathbf{S}_t$.
  Evaluate $K_t \leftarrow K_{\epsilon,\lambda}(\mathbf{H}_t, \mathbf{S}_t)$.
  if $K_t > K_\star$ **then**
    $(\widehat{\mathbf{H}}, \widehat{\mathbf{S}}, K_\star) \leftarrow (\mathbf{H}_t, \mathbf{S}_t, K_t)$.
  end if
  if $t < T$ **then**
    Compute $\mathbf{G}_t$ using Equation (340).
    $\alpha_t \leftarrow \mu\gamma^{\max(0,t-t_0)}$.
    $\mathbf{S}_{t+1} \leftarrow \Pi_{\mathrm{Skel}_{k_0}(\mathcal{P})}(\mathbf{S}_t + \alpha_t\mathbf{G}_t)$, columnwise.
  end if
end for
return $(\widehat{\mathbf{H}}, \widehat{\mathbf{S}})$.

cube cell, free coordinates are clipped to their allowed intervals. A tied group B has optimal common value $\mathrm{clip}_{[0,1]}(|B|^{-1}\sum_{j\in B}|v_j|)$ for the signed cube, with the input signs restored, and $\mathrm{clip}_{[0,1]}(|B|^{-1}\sum_{j\in B} v_j)$ for the nonnegative cube. Coordinates fixed to zero remain zero. Sorting reduces the search to candidate tied groups and free-coordinate sets; their squared projection errors determine the selected cell.

K.2 Synthetic experiments for canonical polytopes

Data and experimental design. For each of the four canonical polytopes, source samples are obtained using the inflated maximum-volume inscribed ellipsoid (MVIE) construction of Tatli & Erdogan (2021b), followed here by projection onto the skeleton in Equation (10). Write the MVIE as $\{\mathbf{C}_{\mathcal{P}}\mathbf{u} + \mathbf{g}_{\mathcal{P}} : \|\mathbf{u}\|_2 \le 1\}$. Independently for each sample, generate

$$\begin{aligned} &\rho = 0.85\sqrt{r}, && \mathbf{w} \sim \mathcal{N}(\mathbf{0}, \rho^2\mathbf{I}_r/r), \\ &\mathbf{u} = \mathbf{w}\min\{1, \rho/\|\mathbf{w}\|_2\}, && \mathbf{s}_g = \Pi_{\mathrm{Skel}_{k_0}(\mathcal{P})}(\mathbf{C}_{\mathcal{P}}\mathbf{u} + \mathbf{g}_{\mathcal{P}}). \end{aligned} \tag{342}$$

The MVIE affine parameters (Tatli & Erdogan, 2021b) are

$\mathcal{P}$	$\mathbf{C}_{\mathcal{P}}$	$\mathbf{g}_{\mathcal{P}}$
$\mathcal{B}_1^r$	$r^{-1/2}\mathbf{I}_r$	$\mathbf{0}$
$\mathcal{B}_\infty^r$	$\mathbf{I}_r$	$\mathbf{0}$
$\mathcal{B}_{1,+}^r$	$\dfrac{\mathbf{I}_r + ((r+1)^{-1/2} - 1)\mathbf{1}\mathbf{1}^\top/r}{\sqrt{r(r+1)}}$	$\mathbf{1}/(r+1)$
$\mathcal{B}_{\infty,+}^r$	$\mathbf{I}_r/2$	$\mathbf{1}/2$

The entries of $\mathbf{H}_g$ are independent standard Gaussians. Observations are $\mathbf{Y} = \mathbf{H}_g\mathbf{S}_g + \mathbf{W}$, with independent $\mathcal{N}(0, 10^{-16})$ entries in $\mathbf{W}$; this is an essentially noiseless experiment. Inflation spreads the proposals across the MVIE geometry before the skeleton projection imposes the prescribed source structure.

The source dimension is $r = 25$, with $k_0 \in \{1, 3\}$. The measurement sweep fixes $N = 6000$ and varies M from 2 to 25. Within each data repetition, it uses the same source matrix and nested rows of a single Gaussian mixing matrix. The sample-size sweep fixes $M = 20$ and uses nested source and observation columns for $N \in \{50, 100, 200, 400, 800, 1600, 3200\}$. Each sweep point has ten independent data repetitions and five random initializations per repetition. Multiple initializations reduce sensitivity to the starting point in this nonconvex optimization. Each initialization retains its maximum-objective iterate according to Algorithm 2. Among the five returned pairs, the pair with the

Table 2: Synthetic initialization and step-size settings.

	$\mathcal{B}_1^r$	$\mathcal{B}_{1,+}^r$	$\mathcal{B}_\infty^r$	$\mathcal{B}_{\infty,+}^r$
Initial entry	$\mathcal{U}[-1/4, 1/4]$	$\mathcal{U}[0, 1/4]$	$\mathcal{U}[-1/2, 1/2]$	$\mathcal{U}[0, 1/2]$
μ	10^7	10^8	10^8	10^8
γ	0.99	0.99	0.97	0.97

highest $K_{\epsilon,\lambda}$ is selected. Curves show the mean SIR over the ten data repetitions, with one sample standard deviation shaded.

The shared optimization settings are $T = 3000$, $t_0 = 500$, $\epsilon = 10^{-5}$, and $\lambda = 10^{-3}$. Skeleton projection is applied at every update. Table 2 lists the initialization distributions and schedule parameters.

K.3 Source recovery metric and interpretation

Theorem 3.3 characterizes exact recovery by global optimizers, whereas numerical experiments evaluate the approximate factorizations returned by a finite-iteration solver. The relevant question is therefore how closely the estimated factors agree with the ground truth, rather than whether they are exactly identical. This comparison must account for the admissible ambiguities: the theorem identifies $(\mathbf{H}_g, \mathbf{S}_g)$ up to $(\mathbf{H}_g\mathbf{Q}^{-1}, \mathbf{Q}\mathbf{S}_g)$ with $\mathbf{Q} \in \mathcal{A}_\mathcal{P}$. For the canonical polytopes, $\mathcal{A}_\mathcal{P}$ consists of signed permutations in the symmetric cases and permutations in the nonnegative cases. Direct comparison without component alignment would therefore penalize equivalent solutions. Conversely, a small observation-reconstruction error alone does not establish recovery of the generating factors. Our source-recovery plots assess one aspect of agreement with the ground truth: the suppression of cross-talk between estimated source components.

Signal-to-interference power ratio. Let $\widehat{\mathbf{S}} \in \mathbb{R}^{r\times N}$ denote the estimated source matrix and $\mathbf{S}_g \in \mathbb{R}^{r\times N}$ the ground-truth source matrix. We first form the least-squares source transformation

$$\mathbf{G} := \widehat{\mathbf{S}}\mathbf{S}_g^\dagger, \tag{343}$$

where † denotes the Moore–Penrose pseudoinverse. When $\mathbf{S}_g$ has full row rank, $\mathbf{G}$ is the unique minimizer of $\|\widehat{\mathbf{S}} - \mathbf{A}\mathbf{S}_g\|_F^2$ over $\mathbf{A} \in \mathbb{R}^{r\times r}$. Thus, G_{ij} is the coefficient of true source j in the joint least-squares approximation of estimated source i.

To identify dominant source correspondences, define

$$P_{ij} := \mathbf{1}\left\{|G_{ij}| > 0.99 \max_{1\le\ell\le r} |G_{i\ell}|\right\}, \qquad \mathbf{T} := \mathbf{G}\mathbf{P}^\top. \tag{344}$$

Here $\mathbf{1}\{\cdot\}$ denotes the indicator function. The matrix $\mathbf{P}$ implements rowwise dominant-component matching, including near ties. It is not constrained to be a permutation: different estimated components can select the same true component, and a row can select multiple components whose coefficient magnitudes exceed 99% of its maximum.

We report the coefficient-based *signal-to-interference power ratio* (SIR), expressed in decibels:

$$\mathrm{SIR} := 10\log_{10}\frac{\sum_{i=1}^r T_{ii}^2}{\|\mathbf{T}\|_F^2 - \sum_{i=1}^r T_{ii}^2} = 10\log_{10}\frac{\sum_{i=1}^r T_{ii}^2}{\sum_{i\ne j} T_{ij}^2}. \tag{345}$$

The numerator is the matched diagonal coefficient energy, and the denominator is the off-diagonal coefficient energy after matching. Higher SIR indicates greater dominance of the matched coefficients over cross-talk. For example, an SIR of 20 dB means that the off-diagonal coefficient energy in $\mathbf{T}$ is 1% of its diagonal coefficient energy. These are coefficient energies; for correlated or unequal-power sources, their ratio need not equal the corresponding ratio of source-waveform energies. For Figure 3a in the main text and Figure 11, an interference denominator at or below 10^{-12} is reported at an SIR of 120 dB for numerical stability.

Component ambiguities and repeated matches. If $\mathbf{S}_g$ has full row rank and $\widehat{\mathbf{S}} = \mathbf{Q}\mathbf{S}_g$ for a signed permutation $\mathbf{Q}$, then $\mathbf{G} = \mathbf{Q}$, $\mathbf{T}$ is diagonal, and the ideal SIR is infinite. More generally, the score is invariant to permutations and sign changes of the estimated components and to a common nonzero scaling of $\widehat{\mathbf{S}}$.

Repeated matches are not simply counted as independent successful recoveries. When each matching row selects a single true component, write its index as $c(i)$, so that $T_{ij} = G_{i,c(j)}$. If two estimated components $i \neq j$ select the same true component c, then

$$T_{ij} = T_{ii} = G_{ic}, \qquad T_{ji} = T_{jj} = G_{jc}.$$

Consequently, this pair contributes as much off-diagonal coefficient energy as matched coefficient energy. This single-component argument applies when the threshold rule selects one index per row; near ties are handled by Equation 344.

Scope of the metric. SIR measures separation in the fitted source transformation, rather than absolute gain accuracy. In particular, perfectly separated sources with incorrect nonzero component gains still give a diagonal $\mathbf{T}$ and infinite ideal SIR. The metric therefore does not assess the gain calibration supplied by Det-Max under the theorem's assumptions.

The least-squares decomposition also gives

$$\widehat{\mathbf{S}} = \mathbf{G}\mathbf{S}_g + \mathbf{\Delta}, \qquad \mathbf{\Delta}\mathbf{S}_g^\top = \mathbf{0}, \tag{346}$$

where $\mathbf{\Delta}$ is the part of the estimated sources not represented by the fitted transformation. Equation 345 depends on $\mathbf{G}$ but does not measure this residual. Accordingly, SIR is a normalized source-separation diagnostic, not a total source-estimation error or, by itself, a certificate of joint factor recovery up to $\mathcal{A}_{\mathcal{P}}$.

K.4 Discussion of synthetic SIR results

Effect of sample size. Figure 3a in the main text and Figure 11 report SIR versus sample count for $k_0 = 3$ and $k_0 = 1$, respectively, with $r = 25$ and $M = 20$. Across the four canonical domains, increasing N initially produces substantial improvements, followed by smaller gains or plateaus. The two ℓ_1 models attain higher SIR than the cube models in these experiments, while the latter illustrate separation with non-coordinate face-span structure. Useful separation is obtained with sample counts far below the worst-case population requirement $J[J(p-1)+1]$ used in the identifiability proof. This comparison highlights the distinction between that sufficient requirement and the sample sizes needed for useful empirical performance; it does not establish a sharper exact-recovery sample bound.

Effect of measurement dimension. Figure 10 fixes $r = 25$ and $N = 6000$ and varies M for $k_0 \in \{1, 3\}$. SIR generally increases with the number of measurements, with substantial separation already achieved in the underdetermined regime $M < r$. The locations of the transitions, attained SIR levels, and variability across repetitions depend on the source domain and face dimension. For these settings, the geometric condition $M \geq 2(k_0 + 1)$ gives sufficient local-injectivity budgets of 4 and 8 measurements, respectively. These are not predicted transition points for the numerical solver: Theorem 3.3 additionally requires the population and scattering assumptions and characterizes global optimizers rather than finite-iteration outputs.

Interpretation of saturation. The plateaus at larger N or M characterize performance under the chosen optimization protocol: 3000 projected-gradient updates, fixed positive regularization, and a step size that decays geometrically after the first 500 iterations. Finite optimization accuracy, the solution reached on the nonconvex objective, regularization, and numerical precision may all affect the attained SIR. The present sweeps do not isolate these effects. Consequently, the observed saturation should not be interpreted as an intrinsic separation limit, an identified observation-noise floor, or evidence that additional iterations alone would necessarily improve performance.

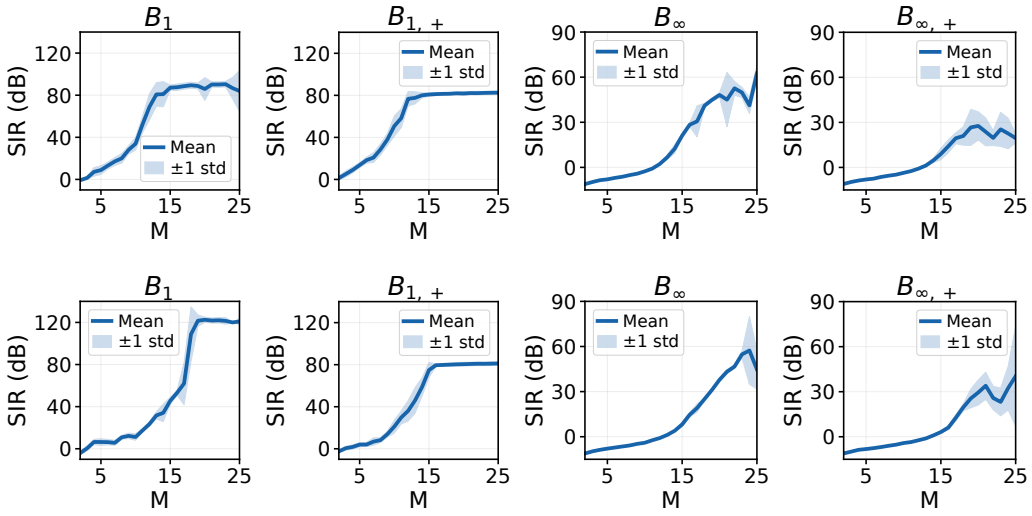


Figure 10: Source SIR versus measurement dimension, with $r = 25$ and $N = 6000$, displayed through $M = 25$. Rows: $k_0 = 1$ (top) and $k_0 = 3$ (bottom). Columns: $\mathcal{B}_1^r$, $\mathcal{B}_{1,+}^r$, $\mathcal{B}_\infty^r$, and $\mathcal{B}_{\infty,+}^r$. Lines and shading show the mean and one standard deviation over ten data repetitions, after the selection procedure described above.

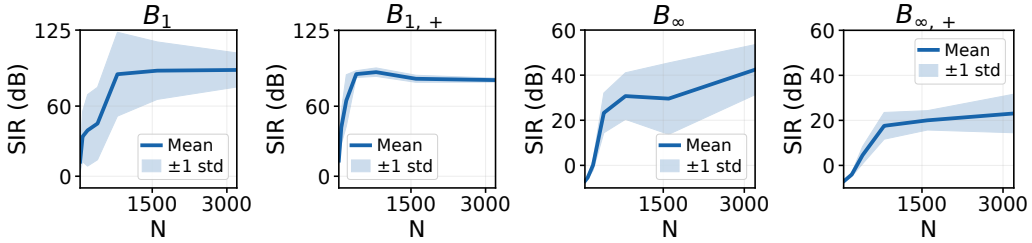


Figure 11: Source SIR versus sample count, with $r = 25$, $M = 20$, $k_0 = 1$, and $N \in \{50, 100, 200, 400, 800, 1600, 3200\}$. The polytope order and averaging convention match Figure 10.

K.5 $B_{1,+}$ DICTIONARY LEARNING FROM IMAGE PATCHES

The image experiment revisits the sparse-coding setting of Olshausen et al. (2009), using the natural-image collection of van Hateren & van der Schaaf (1998).[1] The data comprise 150,000 training and 40,000 validation patches of size 16×16. Each patch is mean-subtracted and divided by its pixel standard deviation plus 10^{-3}. PCA whitening, fitted on the training patches only, retains $M = 100$ components. The dictionary has $r = 1000$ atoms, and the coefficient domain is the scaled nonnegative ball

$$\mathcal{C} = \{\mathbf{s} \geq \mathbf{0} : \mathbf{1}^\top \mathbf{s} \leq 6\}, \qquad \mathcal{C}_{10} = \{\mathbf{s} \in \mathcal{C} : \|\mathbf{s}\|_0 \leq 10\}. \tag{347}$$

Thus $\mathcal{C}_{10} = \mathrm{Skel}_9(6\mathcal{B}_{1,+}^{1000})$ by Equation (10).

Minibatch coefficient inference. Training maximizes K_ϵ in Equation (15), with $\lambda = 1$ and $\epsilon = 10^{-6}$, evaluated on minibatches of $B = 320$ patches. At each dictionary update, a new batch $\mathbf{Y}_B$ is sampled and its coefficients are initialized by $\mathbf{S}_B^{(0)} = \Pi_{\mathcal{C}}(\mathbf{H}^\top \mathbf{Y}_B)$. Holding $\mathbf{H}$ fixed, seven projected source steps use Equation (340) with $N = B$, $\lambda = 1$, and constant step size $\eta_S = 0.12$. During the first 7699 dictionary updates (approximately 35% of training), all seven steps project onto $\mathcal{C}$. Thereafter the first two steps project onto $\mathcal{C}$ and the remaining five onto $\mathcal{C}_{10}$. The sparse projection retains the ten largest positive entries and projects them onto the nonnegative ℓ_1 ball of radius 6.

[1] Dataset downloaded from `https://pirsquared.org/research/vhatdb/`.

Dictionary update. Using the inferred coefficients and the updated batch residual $\mathbf{E}_B = \mathbf{Y}_B - \mathbf{H}\mathbf{S}_B$, take one dictionary step

$$\mathbf{H}^+ = \mathbf{H} + \frac{\eta_H}{B}(\mathbf{E}_B\mathbf{E}_B^\top/B + \epsilon\mathbf{I}_M)^{-1}\mathbf{E}_B\mathbf{S}_B^\top, \qquad \mathbf{h}_j \leftarrow \frac{\mathbf{h}_j^+}{\|\mathbf{h}_j^+\|_2 + 10^{-8}}, \qquad \eta_H = 0.045. \tag{348}$$

The initial dictionary consists of randomly selected whitened training patches, normalized in the same way. Training performs 22,000 dictionary updates with constant learning rates. Dictionary normalization fixes the atom scale; batchwise coefficient inference adapts the sparse representation to the current dictionary. Both updates use second moments computed from the current batch.

Validation and visualization. Validation fixes the final dictionary and uses seven coefficient steps per batch with the sparse projection enabled after the first two steps. For visualization, atoms are mapped back through the inverse PCA transform, reshaped to 16×16, and individually mean-subtracted and contrast-scaled. Figure 12 shows all 1000 atoms.

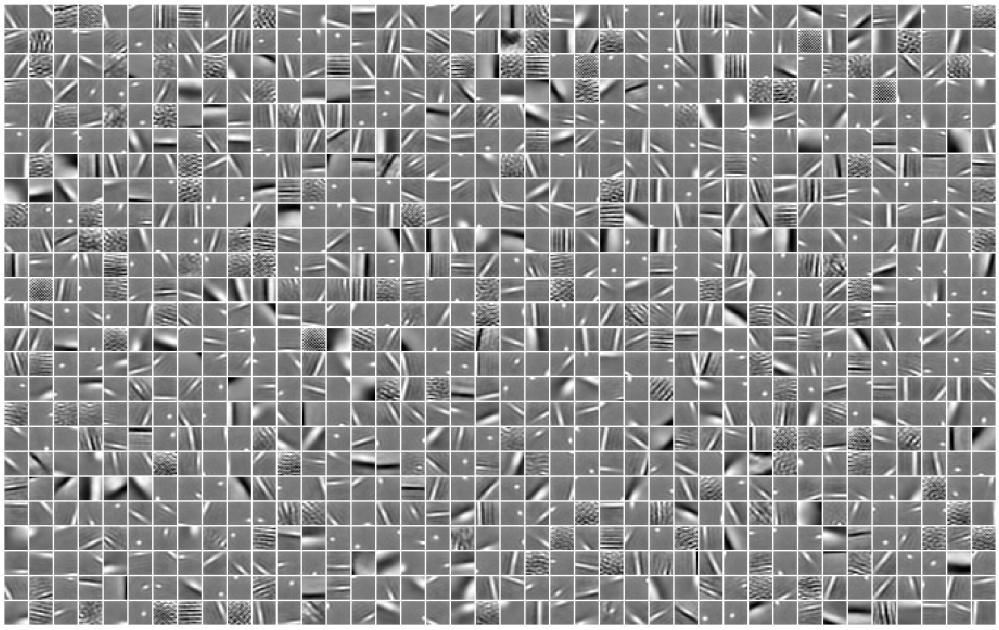


Figure 12: The 1000-atom dictionary learned from van Hateren image patches using minibatch optimization of K_ϵ with nonnegative sparse coefficients. Each tile shows one atom in pixel coordinates, with independent display contrast.

K.6 Synthetic Experiments for Octagon

This example illustrates that Det-Max can do more than calibrate source amplitudes and it can resolve the coordinate mixing ambiguities, and provides synthetic experiments for the numerical demonstration given for the octagon polytope of Appendix I. Some transformations of the octagon preserve the face-span arrangement and admit feasible factorizations with identical predictions, yet genuinely mix source coordinates. We use this geometry to distinguish two questions: *which feasible representation the objective prefers*, and *whether an optimization procedure finds that representation*.

Experimental setting. We use six coordinate pairs ($m = 6$, $r = 12$), $M = 8$ measurements, and $\mathrm{Skel}_1(\mathcal{P}_m)$ ($k_0 = 1$, $p = 2$), with octagon parameter $a = 6/5$. The polytope automorphisms $\mathcal{A}_{\mathcal{P}_m}$ are the pair-preserving signed permutations. Sources are generated by the inflated-MVIE sampler of Equation (342), with $\rho = 0.85\sqrt{r}$, and projected onto the skeleton using Appendix I.4. No vertices were observed or appended. The dictionary $\mathbf{H}_g$ has independent standard Gaussian entries, and $\mathbf{Y} = \mathbf{Y}_0 + \mathbf{W}$, where $\mathbf{Y}_0 = \mathbf{H}_g\mathbf{S}_g$ and $W_{ij} \sim \mathcal{N}(0, 10^{-16})$ independently. We evaluate 20 independent sample pools at nested sizes $N \in \{500, 1000, 2000, 5000\}$, giving 80 dataset–sample-size trials, not 80 independent datasets. Throughout, $\epsilon = 10^{-6}$.

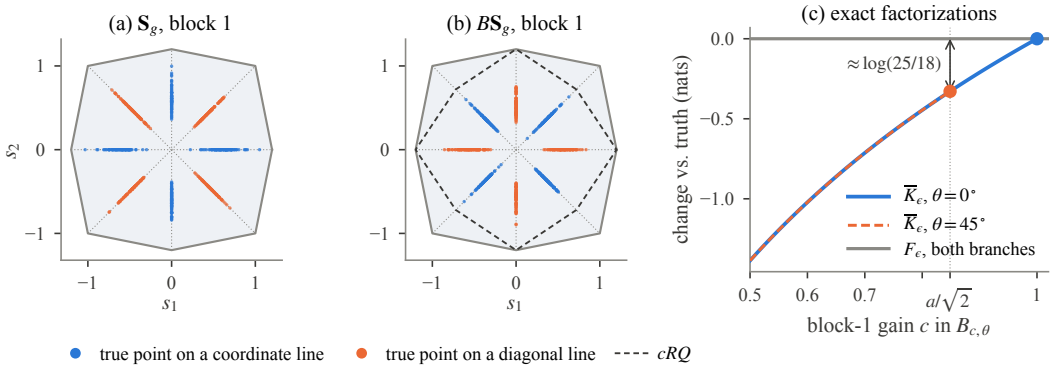


Figure 13: Why reconstruction and Det-Max distinguish different aspects of recovery (one pool, $N = 2000$). (a) True block-one components in the octagon Q. Blue and orange identify coordinate-line and diagonal-line components. (b) Rotation and contraction exchange these line classes while preserving feasibility. (c) Reconstruction is unchanged along both orientation branches. At the same gain c, their Det-Max-consistent objective values also coincide by orthogonal invariance; their *largest feasible gains* differ, producing the gap in Equation (350). "Exact factorizations" refers to the noiseless signal $\mathbf{Y}_0$.

What reconstruction leaves unresolved. Let R be the $45°$ rotation in Equation (307). It exchanges the coordinate and diagonal vertex lines of the octagon, preserving the face-span arrangement, but is not a polytope automorphism. With

$$\mathbf{B} = \operatorname{diag}(cR, I_2, \dots, I_2), \qquad c = \frac{a}{\sqrt{2}} = \frac{6}{5\sqrt{2}},$$

the contraction makes $\mathbf{BS}_g$ skeleton-feasible, while $(\mathbf{H}_g\mathbf{B}^{-1})(\mathbf{BS}_g) = \mathbf{Y}_0$. Thus the original and transformed factors have identical predictions and identical residuals against the observed $\mathbf{Y}$. This ambiguity is not a change of gains alone: it also changes the orientation within a coordinate pair.

To compare the criteria, profile out the dictionary by least squares:

$$\begin{aligned} F_\epsilon(\mathbf{S}) &= K_{\epsilon,0}(\mathbf{Y}\mathbf{S}^\dagger, \mathbf{S}), \\ \overline{K}_\epsilon(\mathbf{S}) &= F_\epsilon(\mathbf{S}) + \tfrac{1}{2}\log\det(\mathbf{R}_\mathbf{S} + \epsilon\mathbf{I}_r) = \Phi_{\epsilon,1}(\mathbf{S}) - \tfrac{M}{2}\log\epsilon. \end{aligned} \tag{349}$$

Here $\Phi_{\epsilon,1}$ is the weighted profile of Equation (44). Since invertible source transformations preserve the source row space, $F_\epsilon(\mathbf{BS}) = F_\epsilon(\mathbf{S})$, even for noisy observations. In contrast, for full-row-rank $\mathbf{S}_g$,

$$\tfrac{1}{2}\log\det\mathbf{R}_{\mathbf{BS}_g} - \tfrac{1}{2}\log\det\mathbf{R}_{\mathbf{S}_g} = \log|\det\mathbf{B}| = -\log(25/18) \approx -0.3285. \tag{350}$$

The rotated block's largest feasible gain was $a/\sqrt{2}$ in every trial, with the other blocks at unit gain. Det-Max detects the volume loss required by the wrong orientation, rather than rotation itself (Figure 13). The measured finite-ϵ objective gap was approximately 0.3285, consistent with Equation (350).

Fitting and evaluation. Stage one fits the data using reconstruction information only ($\lambda = 0$), with projected source updates, dictionary refinement, and restarts. A fit is called *near-exact* when

$$\max_j \|\mathbf{y}_j - \mathbf{H}\mathbf{s}_j\|_2 < 10^{-4}\,\text{median}_j\,\|\mathbf{y}_j\|_2. \tag{351}$$

We then compare two refinements. *Gain calibration and local polishing* jointly maximize the unregularized source-moment determinant over positive block gains subject to skeleton feasibility, followed by projected-gradient polishing of $K_{\epsilon,1}$. *Candidate selection* additionally proposes $45°$ block reorientations, rescales them for feasibility, applies the same gain calibration and polishing, and only accepts improvements in $\overline{K}_\epsilon$. These proposals use the known polytope, not the ground truth, and are additional search steps rather than ordinary projected-gradient updates.

Table 3: Outcomes on the same 76 near-exact stage-one fits. Mixed blocks are counted out of 456; recovery uses Equation (352). The four other trials remained unrecovered and are excluded from this conditional comparison.

Procedure	Mixed blocks	Recovered fits
Reconstruction only	91	$0/76$
Gain calibration and local polishing	91	$18/76$
Additional candidate selection	0	$76/76$

Recovery is measured up to the actual polytope automorphisms, without fitting arbitrary gains:

$$e_{\mathcal{A}}(\widehat{\mathbf{S}}) := \min_{\mathbf{Q}\in\mathcal{A}_{\mathcal{P}_m}} \frac{\|\widehat{\mathbf{S}} - \mathbf{Q}\mathbf{S}_g\|_F}{\|\mathbf{S}_g\|_F}, \qquad \text{recovery: } e_{\mathcal{A}} < 10^{-3}. \tag{352}$$

Results: selection beyond gain calibration. Reconstruction-only fitting reached a near-exact solution in $76/80$ trials, with relative residual at most 3.1×10^{-7}. Nevertheless, 58 of these 76 fits contained coordinate mixing: 91 of their 456 blocks were in the 45° class. Their arrangement-preserving form was verified numerically, not inferred from the theorem. Table 3 shows that gain calibration and local polishing recovered the 18 unmixed fits but left all 91 mixed blocks unchanged in orientation. Adding candidate selection corrected all 91 mixed blocks and recovered all 76 near-exact fits, with $e_{\mathcal{A}} \le 4.05 \times 10^{-6}$. The four stage-one failures remained unrecovered, giving $76/80$ recovery overall.

The *fixed-candidate comparison* isolates the objective's role from the search procedure. For each of the 456 blocks, we compared the current representation with its 45° alternative after applying identical gain calibration to both, without polishing. The reconstruction criterion tied in every comparison. In contrast, $\overline{K}_\epsilon$ favored correcting all 91 mixed blocks and rejected reorienting the other 365, with differences of approximately 0.3285. Thus the source term provides selection information that reconstruction lacks, even when the competing candidates are held fixed.

Interpretation and scope. The example makes the gap between skeleton symmetries $G_{\mathcal{L}_{m,1}}$ and polytope symmetries $\mathcal{A}_{\mathcal{P}_m}$ concrete: fitting the data on the skeleton can leave genuine coordinate mixing unresolved, while Det-Max uses the bounded geometry to prefer the correct orientation. This illustrates an ambiguity reduction beyond gain calibration, not a global-convergence guarantee for projected gradient. The experiments also use far fewer samples than the theorem's sufficient all-cell count (60,762), and sufficient scattering was not verified after synthetic source generation; the observed recoveries are therefore not a test of all theorem assumptions or a sample-complexity bound. Finally, the trace criterion ranked the fixed candidates identically. The experiment illustrates Det-Max's selection role, not its empirical uniqueness among spread criteria; its broader justification is the scattering-based identifiability result.

K.7 Online nonnegative sparse ER-CorInfoMax for image-patch dictionary learning

We trained the online biologically plausible dictionary learning network in Appendix J.3 on the whitened natural-image patch dictionary learning problem. Each observation $\mathbf{x}_t \in \mathbb{R}^{144}$ corresponds to a 12×12 image patch. The image patches are obtained from the link http://www.rctn.org/bruno/sparsenet, which are 312500 pre-whitened imag patches. The overcomplete factor is 10, i.e., the representation dimension is $r = 1440$. The learned dictionary is

$$\mathbf{H} \in \mathbb{R}^{144\times 1440},$$

corresponding to an overcompleteness factor $K = 10$. The source activity vector $\mathbf{s}_t \in \mathbb{R}^{1440}$ was constrained to the nonnegative sparse polytope

$$\mathcal{P}_{+,\ell_1,p} = \left\{\mathbf{s} \in \mathbb{R}^{1440}_+ : \|\mathbf{s}\|_1 \le 1,\ \|\mathbf{s}\|_0 \le p\right\}, \qquad p = 16.$$

Thus, each sample activates at most 16 nonnegative source neurons, with total activity bounded by one.

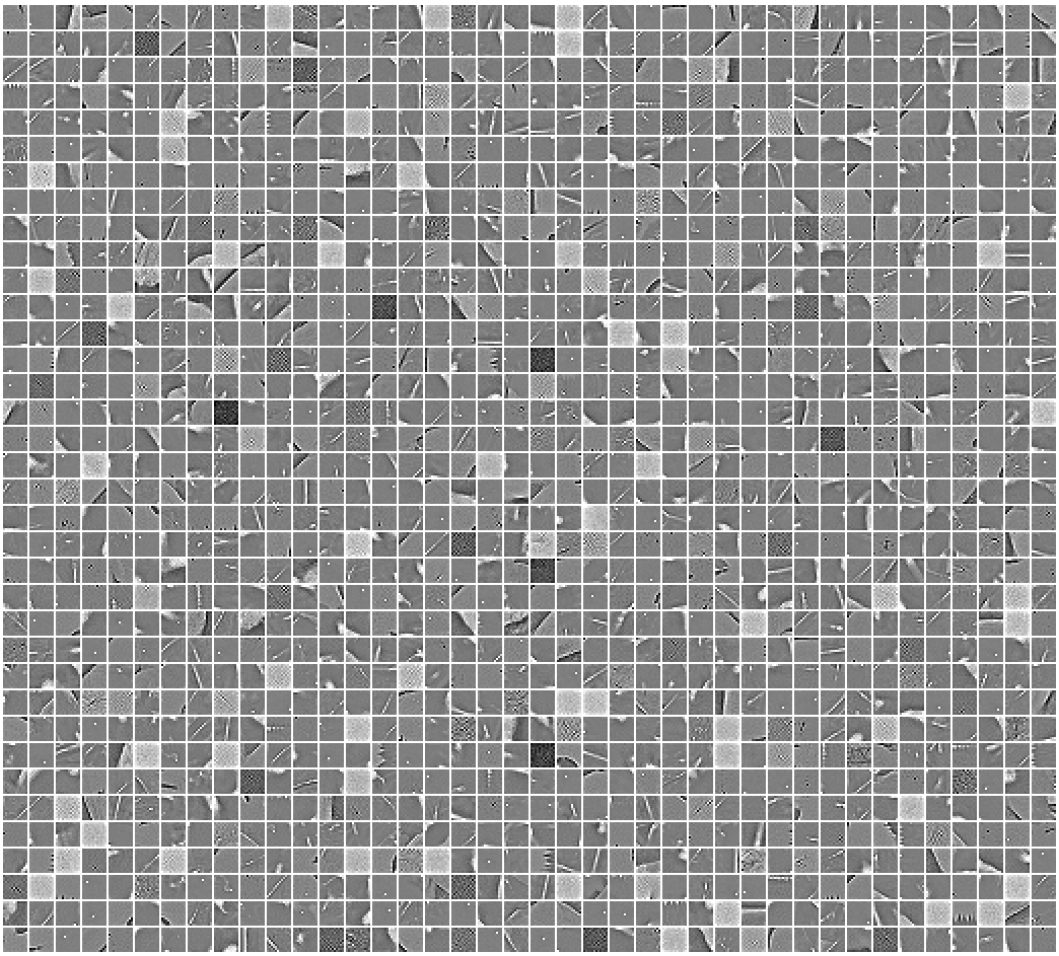


Figure 14: Whole learned dictionary from the online nonnegative sparse ER-CorInfoMax dictionary learning run with $p = 16$ and overcomplete factor of 10. Each tile shows one column of $\mathbf{H}$ reshaped as a 12×12 image patch.

For the representative run shown in Fig. 14, we used the online neural circuit implementation in Appendix J.3 with fading covariance parameters

$$\zeta_s = \zeta_e = 0.999995, \qquad \epsilon = 5 \times 10^{-5}.$$

The neural dynamics used a projected rate update on the feasible face of $\mathcal{P}_{+,\ell_1,p}$, with tolerance 10^{-3}, minimum 3 iterations, and maximum 5000 iterations per input patch. The dictionary update gain used a cosine schedule from

$$\eta_H^{\text{start}} = 0.15 \quad \text{to} \quad \eta_H^{\text{end}} = 0.0015,$$

The dictionary was initialized with column norm scale 2.0 and random seed 0.

The resulting dictionary contains many localized oriented edge-like filters, together with a subset of weak or high-spatial-frequency atoms. For visualization, each atom is displayed as a 12×12 patch using signed symmetric contrast: zero is mapped to mid-gray, positive values to lighter gray, and negative values to darker gray. Thin white lines separate adjacent atoms. The display additionally uses a polarity convention in which the dominant low-pass lobe of each atom is shown as positive; this affects only the visualization and does not modify the learned dictionary matrix.